

DevTreks –social budgeting that improves lives and livelihoods

Social Performance Analysis 3

Last Updated: March 27, 2019; First Released: May 17, 2018

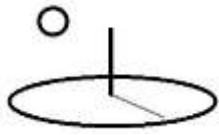
Author: Kevin Boyle, President, DevTreks

Version: DevTreks 2.1.8

Appendix E. Examples 5 to 8

This Appendix uses online datasets to explain how to carry out Social Performance Analysis, primarily related to stakeholder Performance Monitoring, Impact Assessment, and Impact Evaluation. The first 4 examples can be found in in the Social Performance Analysis 2 (SPA2) reference, which should be read first. The RCA Framework can be found in Appendix A of the Social Performance Analysis 1 (SPA1) reference. The CTAP algorithms, covering Disaster Risk Prevention, can be found in the CTA-Prevention reference in the SPA tutorial.

Example	Page
• 5. Rural Stakeholder Resource Conservation Value Accounting (RCA6)	2
• 6. Disaster Stakeholder Resource Conservation Value Accounting (RCA6)	88
• 6A. Stakeholder Abbreviated Resource Conservation Value Accounting (RCA7)	135
• 6B. Correlated Hazards	149
• 7. SDG Stakeholder Resource Conservation Value Accounting	160
• 8A. SDG Stakeholder Resource Conservation Value Accounting using .NetStandard Libraries	180
• 8B. SDG Stakeholder Resource Conservation Value Accounting using R	204
• 8C. SDG Stakeholder Resource Conservation Value Accounting using Python	206



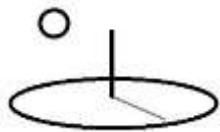
DevTreks –social budgeting that improves lives and livelihoods

• Appendix A. Standard Statistical Impact Evaluation Analysis	208
• Appendix B. Machine Learning Impact Evaluation Analysis	242

Several of the algorithms introduced in this reference employ machine learning techniques that are under active development in their underlying statistical libraries. Like all DevTreks algorithms, they’ll evolve in future releases. Most of the algorithms referenced in this tutorial have been tested using the upgraded 2.1.6 calculator patterns. Although Version 2.1.6’s upgraded security will automatically redirect <http://localhost:5000> URLs to <https://localhost:5001>, the SSL URL should be used with all algorithms.

A video tutorial explaining this reference can be found at:

The Social Performance Analysis tutorial on the DevTreks home page.



DevTreks –social budgeting that improves lives and livelihoods

Example 5. Rural Stakeholder Resource Conservation Value Accounting (RCA6)

Algorithms: RCA algorithms and algorithm1, subalgorithm17

URLs:

<https://www.devtreks.org/greentreks/preview/carbon/resourcepack/Coffee Firm RCA Example 5/1558/none>

<https://www.devtreks.org/greentreks/preview/carbon/output/Coffee Firm RCA5 Stock/2141223484/none>

<http://localhost:5000/greentreks/preview/carbon/resourcepack/Coffee Firm RCA Example 5/543/none>

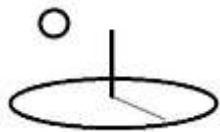
<http://localhost:5000/greentreks/preview/carbon/output/Coffee Firm RCA5 Stock/2141223499/none>

A. Introduction

This example continues with the Coffee Farm, and surrounding community, introduced in the SPA2 reference to demonstrate tying stakeholder socioeconomic characteristics, such as population, age, and gender, to the Social Performance Assessment metrics introduced in that reference. Beside that reference’s emphasis on private sector materiality impact and performance reporting, this example also illustrates how community service organizations, such as Producer Organizations, RCA Districts, Community Sustainability Offices, or National Statistical Offices (NSOs), can report on periodic accomplishments in efforts to conserve scarce resources for targeted stakeholders.

B. Introduction to Sustainable Development Data Systems

The Sustainable Development Solutions Network (Espey et al, 2017) summarize the challenges faced by most countries, and developing countries in particular, in developing the data systems needed to accomplish the UN’s Sustainable Development Goals, or SDGs. They describe the primary challenge as follows:



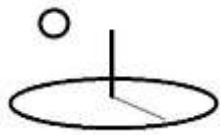
DevTreks –social budgeting that improves lives and livelihoods

“The adoption of the SDGs by the global community in 2015 represents a concerted attempt to ensure sustainable and equitable development over the period from 2015 to 2030. Meeting the goals and their associated targets will be challenging, particularly in the context of the overarching principle that no one should be left behind. However, there is a real risk that the efforts made to meet the goals and targets may be compromised by the inadequacy of the data, particularly in the Global South, to benchmark, monitor and track the 232 unique indicators of progress in meeting the SDGs.”

Using the context of climate change, the UN (2011) summarizes why the SDGs are important – they are designed to help ameliorate conditions endemic to most poor people in most countries. As the climate change-related SDG Indicators make clear, they also will help ameliorate conditions that have the potential to wreak havoc on the majority of the world’s countries and citizens (1*).

“In sum, the people most vulnerable to climate change are usually poor, undernourished, of poor health, live in precarious housing conditions, farm on degraded lands, have low levels of education, lack rights, have little opportunities to influence decision making, work under precarious conditions, and/or reside in countries and regions with non-resilient health systems, limited resources and sometimes poor governance systems. Social, cultural or political circumstances, often including inequalities and discriminatory practices, deprive them of the basic assets and entitlements and the institutional support needed to make a living and ensure their well-being even under normal conditions, let alone for mastering the increased and additional challenges posed by climate change. These non-climatic factors and the socioeconomic context in which climatic problems occur is likely to be as important, if not more so, than climate-related hazards themselves.”

Espey et al (2017) document how National Statistical Offices (NSOs) have primary responsibility for the censuses, household surveys, civil registration and vital statistics systems, administrative systems, and geospatial systems, needed to report civil sustainability data. Although the authors may be right about the central roles played by country-based NSOs in



DevTreks –social budgeting that improves lives and livelihoods

ensuring success with the SDGs (i.e. as contrasted to contemporary alternatives such as international, Cloud Statistical Offices, or CSOs), this tutorial emphasizes the roles of social networks, clubs, stakeholders, producer organizations, community service organizations, companies, investors, end-product consumers, and supply chain participants, in collecting, reporting, and acting on sustainability data (2*).

To anchor these data challenges, and the reality of their moral imperative, the following images (UNEP/SETAC, 2017; UN, 2017) display the SDG Targets and the first few of the 230+ SDG Indicators.

A1.4 Hotspots Analysis and the Sustainable Development Goals

At the United Nations Sustainable Development Summit on 25th September 2015, world leaders adopted the 2030 Agenda for Sustainable

Goal 11 on sustainable cities and communities is implicitly linked to identifying poverty hotspots and addressing these in an inclusive and participatory manner. Goal 12, in particular target 12.8 aimed at ensuring that people everywhere

15 <http://socialhotspot.org/>

16 <http://bookshop.europa.eu/en/social-sustainability-in-trade-and-development-policy-pbLBN26483/>



Figure 19: UN Sustainable Development Goals

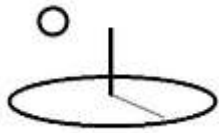


DevTreks –social budgeting that improves lives and livelihoods

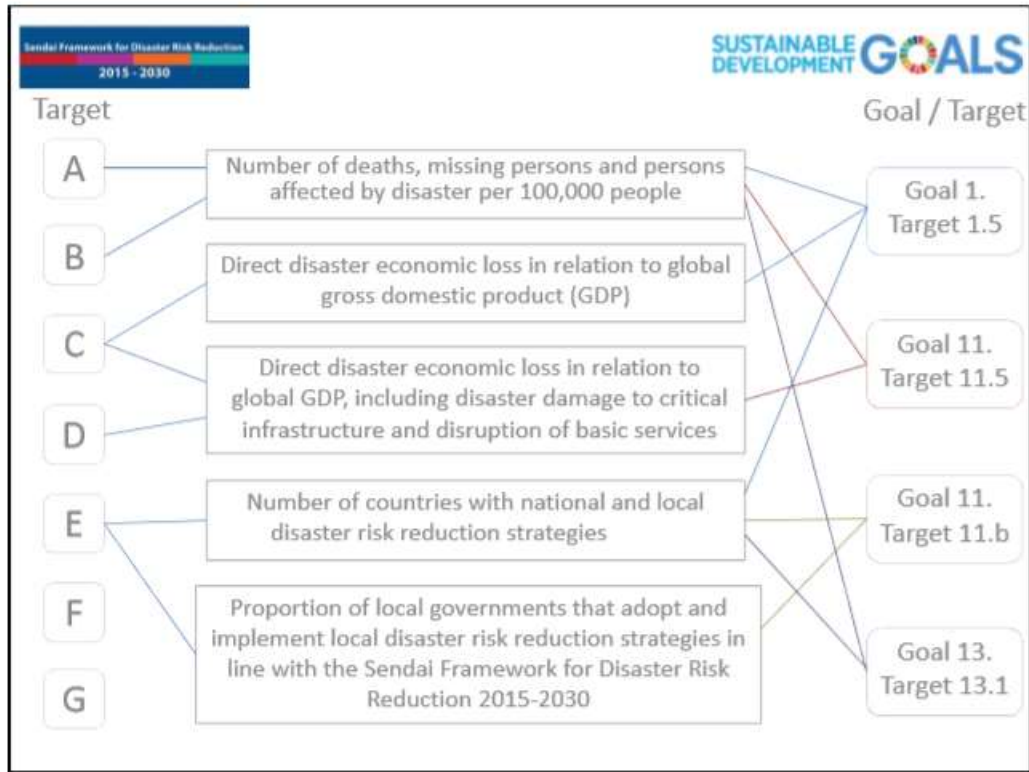
Tier Classification Sheet (as of 15 December 2017)

UNSD Indicator Code	Target	Indicator	Initial Proposed Tier (by Secretariat)	Possible Custodian Agency(ies)	Partner Agency(ies)	Updated Tier Classification (by IAEG-SDG Members)
Goal 1. End poverty in all its forms everywhere						
C010101	1.1 By 2030, eradicate extreme poverty for all people everywhere, currently measured as people living on less than \$1.25 a day	1.1.1 Proportion of population below the international poverty line, by sex, age, employment status and geographical location (urban/rural)	Tier I	World Bank	ILO	Tier I
C010201	1.2 By 2030, reduce at least by half the proportion of men, women and children of all ages living in poverty in all its dimensions according to national definitions	1.2.1 Proportion of population living below the national poverty line, by sex and age	Tier I	World Bank	UNICEF	Tier I
C010202		1.2.2 Proportion of men, women and children of all ages living in poverty in all its dimensions according to national definitions	Tier II	National Gov.	UNICEF, World Bank, UNDP	Tier II
C010301	1.3 Implement nationally appropriate social protection systems and measures for all, including floors, and by 2030 achieve substantial coverage of the poor and the vulnerable	1.3.1 Proportion of population covered by social protection floors/systems, by sex, distinguishing children, unemployed persons, older persons, persons with disabilities, pregnant women, newborns, work-injury victims and the poor and the vulnerable	Tier I	ILO	World Bank	Tier II
C010401	1.4 By 2030, ensure that all men and women, in particular the poor and the vulnerable, have equal rights to economic resources, as	1.4.1 Proportion of population living in households with access to basic services 1.4.2 Proportion of total adult	Tier III	UN-Habitat	UNICEF, WHO	Tier III

Gordon (2016) uses the following image to explain the relation between the SDG and the related Sendai Disaster Risk Reduction (SDRR) Framework. As with the SDG, the SDRR’s targets and indicators have also been adopted by most countries and most countries face similar challenges in collecting and applying the data. Country leaders, or more accurately, informed country leaders, are motivated to comply because they have first-hand experience proving that climate change is wrecking their countries and citizens. Writing billion dollars checks for disaster recovery provides clarity for the informed country leaders. Espey et al (2017) mention that adoption of these global data systems “has created a rare but significant opportunity to build coherence across different but overlapping policy areas”.



DevTreks –social budgeting that improves lives and livelihoods



KPMG (2016, in the SPA1 reference) points out the SDGs contain a specific target (12.6) for private sector companies to integrate sustainability [Indicators] into their reporting cycles. As documented in the SPA1 and SPA2 references, recent business news reports continue to verify the increasing importance of ESG (Environmental, Social, and Good Governance) reporting and investment in the private sector. Investors, supply chain participants, end-product consumers, employees, and local communities, want this evidence, and more companies are responding to the demand (i.e. Footnote 1’s “taking independent action”).

As an example for the agricultural sector, the IFC (2013) confirms that, like many investors in this sector, “[they require their] clients to identify, avoid, mitigate, and manage [Environmental and Social], or E&S, risks and impacts as a way of conducting sustainable business”. They incorporate the SDG throughout their corporate strategy for the years 2018 to 2020 (IFC, 2017). They document the immense role that the private sector can play in accomplishing the SDG and



DevTreks –social budgeting that improves lives and livelihoods

clearly expect their private sector clients to also incorporate the SDG in their corporate strategy. For example, they document how 34 development finance institutions have adopted their Corporate Governance Development Framework, which help companies to fully employ the management accountability systems needed to carry out the SDG. Like many investors and asset managers, they recognize that the SDG Indicators are comprehensive and flexible enough to be incorporated into their client companies' sustainability accounting and reporting systems (3*).

Vanclay et al (2015) mention that the IFC's performance standards (see IFC, 2012), have in fact, become international benchmarks, as suggested by IFC (2013): "Whether or not a company is an IFC client, IFC Performance Standards provide a useful framework for identifying risk assessment variables. Companies may choose to use all eight IFC Performance Standards as the framework to manage risk in their supply chain". The SPA1 reference introduced equivalent sustainable reporting systems, such as EMAS, SASB, GSSB, and B-Lab, that companies also use as sustainability standards. Examples 1 to 4 demonstrate how specific industries have developed industry-specific supplements to these data systems, such as WHO's health care reporting framework and COSA's agricultural smallholder performance framework. The goal of most sustainability standards systems, including the SDG and related standards systems, is to institutionalize "integrated and cross sectoral approaches" for sustainable social, environmental, and economic development.

The Task Force on Climate Related Disclosures (2017) provide professional financial accounting recommendations showing companies how to tie domain-specific sustainability reporting (i.e. related to climate change) to "publication of [domain-specific] financial information in mainstream financial filings". The 17 SDG Goals address many domain-specific topics, including climate change and gender equality, which can directly lead to company materiality impacts and therefore require "publication of [SDG]-related financial information in mainstream financial filings [that] will ensure that appropriate controls govern the production and disclosure of the required information".

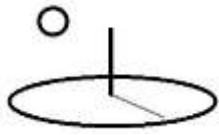


DevTreks –social budgeting that improves lives and livelihoods

Many private sector companies are realizing that, in order to remain competitive with their peers, they must follow suit and adopt sustainable business data systems that assist them to prove the “social soundness” of their business activities. IFC (2015) uses the term Environmental and Social Management System (ESMS) and Vanclay et al (2015) use the term Social Impact Management Plan in similar ways to sustainable business data systems. To assist in the adoption of these systems and their international reporting norms, this reference replaces the globally-accepted term, “Sustainable Development Goals”, and the sustainable value accounting framework term, Resource Conservation Value Accounting (RCA), for the hodgepodge of terms commonly used in past eras for materiality reporting, such as “Economic, Social, and Good Governance”, “Environmental and Social Management System”, or “Corporate Social Responsibility”. As with the SDRR system, adoption of these global data systems “has created a rare but significant opportunity to build coherence across different but overlapping policy areas [, including private sector accountability via SDG materiality impact reporting]”.

Espey et al (2017) imply that no clear data platform has appeared, or will likely appear, that is appropriate, acceptable, and useful for meeting the needs of every country’s sustainable development data systems and every company’s sustainable business data systems. The authors conclude that a broad array of “data actors” have roles to play in collecting, managing, and employing this data (2*). They further speculate about public sector data roles “[morphing] from producer to coordinator of a broad data ecosystem”.

Appendix A in the SPA1 reference points out that an important goal of this tutorial is to harmonize private and public sector sustainability reporting. This example demonstrates how both the public and private sector can monitor, evaluate, and report, their SDG, and general sustainability, accomplishments. In particular, this example demonstrates harmonizing NSO data systems, which document SDG goal accomplishment, with this reference’s emphasis on drilling down to explore the why and how behind the accomplishments –specifically, “what factors caused sustainability to change for different groups of stakeholders? How will sustainability change for particular groups of stakeholders? How did businesses, communities, and public agents contribute to these accomplishments?”



DevTreks –social budgeting that improves lives and livelihoods

Ideally, NSOs and industries institutionalize both sets of data via “administrative data systems”. One-time, “very sound”, “outside expert”, assessments are actively discouraged in favor of ongoing, “reasonably professional”, “local administrative”, assessments. As DevTreks demonstrates, evolution in data development is an ongoing task.

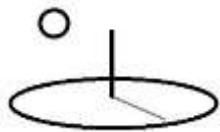
C. Introduction to Stakeholder Impact Assessment (SIA)

Vanclay et al (2015) use The International Principles for Social Impact Assessment to introduce SIA:

“[SIA is defined as being] the processes of analysing, monitoring and managing the intended and unintended social consequences, both positive and negative, of planned interventions (policies, programs, plans, projects) and any social change processes invoked by those interventions”.

This reference extends SIA definitions and purposes beyond special projects and public sector actions to include regular activities carried out by private sector firms and community service organizations (CSO). In addition, SIA is extended beyond community stakeholders to include company. CSO, and local community stakeholders, including employees, supply chain participants, consumers, investors, and asset managers. Although these companies and organizations are not in the public census-taking business often needed for full SIAs, they are, or will be, fully involved in the performance monitoring, impact evaluating, and transparent reporting, business needed for SDG materiality reporting. As this reference will make clearer, if the companies and organizations can’t or won’t fulfill this enhanced reporting, their stakeholders may either do it for them or engage in other types of “ameliorating actions”.

Vanclay et al (2015) use the following statement to further clarify the relation between company and organization activities, SIA, the SDGs, and “sustainable social development”. The statement makes clear that “social soundness”, as proven by SDG materiality impact reporting, involves more than just the charitable activities carried out by many private sector firms and CSOs. A key



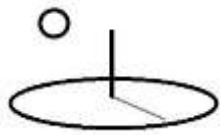
DevTreks –social budgeting that improves lives and livelihoods

SIA goal is to improve the company and community institutions that many people in many countries have come to distrust (as demonstrated by several of the Footnotes) but are often the main avenue for improving local society. Even hard headed business people who question the validity of this statement may agree that modern stakeholders have “modern issues” with current institutions.

“Social development means more than just providing a few jobs and providing funding for a new school building or swimming pool, it requires that the project partner with the local communities in being a force for positive social change and beneficial social development. Social development should be a participatory process of planned social change designed to improve the wellbeing of the community as a whole and especially of the vulnerable, disadvantaged or marginalised groups within a region. Rather than being about benefits to individuals per se, social development is more about facilitating change in institutions and society to reduce social exclusion and fragmentation, to promote social inclusion and democratisation, and to build capacity in institutions and governance.”

Vanclay et al (2015) use the following description and image of “community capitals” to illustrate the relation between the frameworks that underlie SIA and the SDG. The RCA Framework introduced in the SPA1 reference employs the same capitals but uses the terms Physical Capital for Built Capital, Economic Capital for Financial Capital, and Institutional Capital for Political Capital. These frameworks aim to achieve better societal, or public service, outcomes and impacts from private and public sector activities.

“The Sustainable Livelihoods Approach considers the capabilities, livelihood resources (assets, capitals) and livelihood strategies (activities) people undertake to make their living and conduct their way of life. At the heart of the model is the notion that all community resources or assets can be represented as a set of capitals. The assessment of social investment strategies can consider these capitals and how strengthening one or more of these capitals might increase the overall wellbeing in the community.”



DevTrek –social budgeting that improves lives and livelihoods

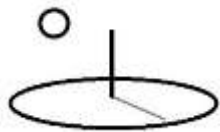


Figure 4: Community Capitals

Source: Cornelia Butler, North Central Regional Center for Rural Development (used with permission).

In terms of private sector materiality impact reporting, Vanclay et al (2015) introduce a glossary of SIA definitions to help companies understand “international reporting norms”, especially related to human rights, which, in this digital age, can quickly impact their companies’ reputations, product sales, and existence (i.e. sexual harassment in the USA movie industry). For example, in the following SIA definition, the term “local impacted communities” can easily be replaced by the term “investors, supply chain participants, end-product consumers, employees, stakeholders, and impacted communities”:

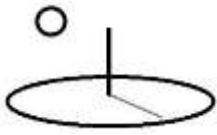
“**Social Licence to Operate** refers to the level of acceptance or approval of the activities of an organization by its stakeholders, especially local impacted communities. Leading corporations



DevTreks –social budgeting that improves lives and livelihoods

now realize that they need to meet more than just the regulatory requirements, they also need to consider, if not meet, the expectations of a wide range of stakeholders, including international NGOs and local communities. If they don't, they risk not only reputational harm and the reduced opportunities that might bring, they also risk being subject to strikes, [, boycotts, disinvestment, whistle blower disclosures, social media campaigns, hacking, inclusion contracts] protests, blockades, sabotage, legal action and the financial consequences of those actions. In some countries, 'social licence' has become an established element of the language of business, actively influencing, if not driving, the business strategy of many companies, and is part of the governance landscape”.

Vanclay et al (2015) and IFC (2013f) use the following images to summarize the business value that SIA, and its applied SDG (i.e. E&S) risk management, offers companies. Although these positive impacts are reason enough for companies to take SIA and the SDG seriously, more companies are realizing that the modern digital age makes Vanclay's (2015) formal SIA instruments, such as Social Licenses to Operate, Impacts and Benefits Agreements, and this reference's algorithms, much more practical, transparent, and verifiable, as well. And even without formal SIA instruments, recent SIA-related social media campaigns (i.e. sexual harassment in the USA) prove that modern digits can quickly impact industries and companies. Finally, TFCD (2017) confirms that the materiality impacts of this increased “social accountability” actually requires companies to carry out formal SDG-related financial reporting. Failure to comply may be grounds for investor lawsuits (i.e. like those recently filed for sexual misconduct in the casino industry) and consequential digital activism (7*).



DevTreks –social budgeting that improves lives and livelihoods

Guidance for assessing and managing the social impacts of projects



Figure 3: Schematic representation of the business value drivers to which SIA adds value



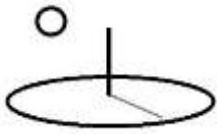
DevTreks –social budgeting that improves lives and livelihoods

The business case for managing E&S risk in agro-commodity supply chains includes value gained through:

- Reputational risk management and value protection;
- access to capital and associated financial services;
- enhanced brand reputation and better market advantage;
- security of supply, supplier competence, and loyalty;
- improving compliance with legislation and “soft” law; and
- cost savings from gains in efficiency and productivity.

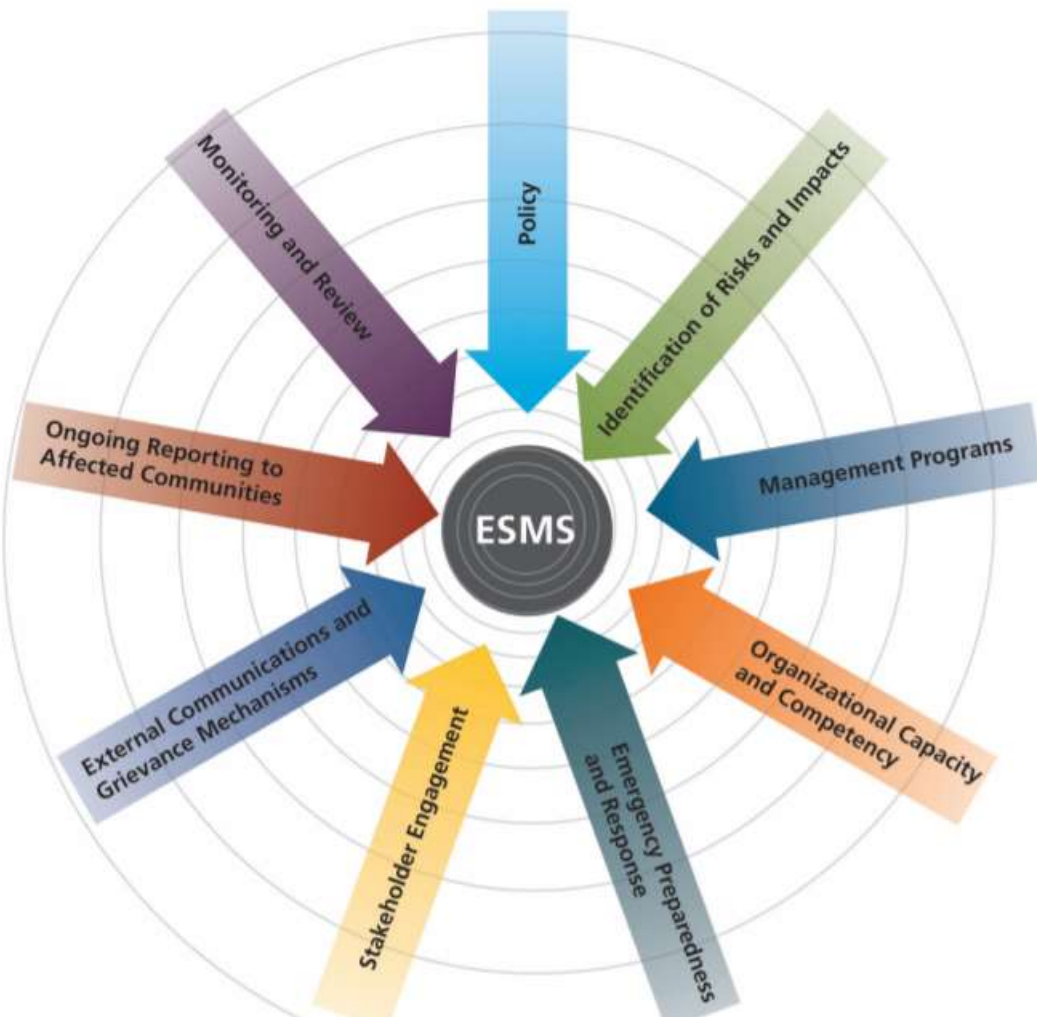
Whether companies like it or not, digits shift power away from conventional institutions, which they have some control over (but which many people distrust), to stakeholders, investors, supply chain participants, consumers, community service organizations, and impacted communities. Many of these stakeholders, or more accurately, informed stakeholders, are realizing their power and recognizing the stark difference between marketing fiction and sustainability evidence. That’s leading them to take independent “ameliorating” actions (i.e. purchases of verifiably sustainable products, protest marches, disinvestment, lawsuits, local activism, app development, social contract riders, and social media campaigns). Bad actor organizations, companies, and their executives have reason for concern. Misinformed employees have reason to become informed (or to find misinformed companies to work for).

The following image (IFC, 2015) demonstrates how companies can integrate sustainable business data systems into their overall management systems. The IFC references provide detailed instructions that companies can follow to implement these systems. They use the term “You can’t improve what you don’t measure” to explain why companies adopt the formal instruments used to apply these systems. Note carefully the emphasis on stakeholders, affected communities, monitoring and evaluating, risk and impact identification, and transparent reporting. Many of the concrete examples used in this IFC reference relate to agricultural firms.



ELEMENTS OF AN ENVIRONMENTAL AND SOCIAL MANAGEMENT SYSTEM (ESMS)

A solid, functioning environmental and social management system (ESMS) is made up of interrelated parts. Take a look at the nine elements of an effective ESMS. Each of these elements is important, because they help you to assess, control and continually improve your environmental and social performance, as part of the Plan-Do-Check-Act cycle. The following section presents step-by-step instructions on how to develop and implement a system using these elements.



Vanclay et al (2015) describe the relationship between the impact pathways introduced in Examples 1 to 4 to this example’s SIAs:

“[S]ocial impacts are rarely singular cause-effect relationships. There are complex patterns of intersecting impact pathways.”



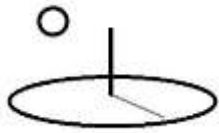
DevTreks –social budgeting that improves lives and livelihoods

This example introduces basic, but professional, SIA. The “impact transition states” explained in this reference and the “impact pathways” used throughout this tutorial will partially address Vanclay’s requirement to understand “complex intersecting patterns”. The algorithm addresses several IFC (2015) recommendations for incorporating SIA and the SDG in formal company management systems (i.e. stakeholder engagement, monitoring and evaluating, managing risks and impacts, and transparent community reporting). The machine learning algorithms introduced in this reference begin to demonstrate fuller SIAs (i.e. because the algorithms are getting better at understanding “complex intersecting patterns”).

D. Stakeholder System Boundary and Stakeholder Engagement

System boundaries are a key initial step in defining the primary stakeholders being impacted by a community service organization intervention, or a company activity. The following images (Vanclay et al, 2015, IFC 2014, UNEP/SETAC 2017) introduce this phase. Although the Vanclay reference relates primarily to capital improvement projects that impact community stakeholders, this reference extends SIA to company and organization activities that impact employees, stakeholders, supply chain participants, consumers, investors, and asset managers.

Example 3 in SPA2 explains that UNEP/SETAC (2017) recognizes stakeholder engagement as a key element in effective LCIA and Hot Spots studies –stakeholders need to be engaged “early and often”. Vanclay et al (2015) discuss using stakeholder engagement to gain stakeholder “trust and respect”. Those authors distinguish between the “statutory” basis for many public participation outreach efforts and the “actual” stakeholder engagement needed to conduct SIA (10*).

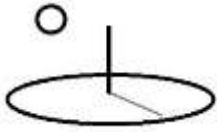


DevTreks –social budgeting that improves lives and livelihoods

Task 3: Identify the preliminary ‘social area of influence’ of the project, likely impacted and beneficiary communities (nearby and distant), and stakeholders.

The concept of a primary area of influence or zone of impact is standard in EIA practice, however, it is not directly transferable to SIA. A ‘social area of influence’ consists of the people potentially impacted by a project. Affected peoples include both ‘communities of place’ and ‘communities of interest’. The location of affected people frequently does not neatly align with the geographic boundaries or the area of influence determined by the environmental impact of a project. In fact, often the buffer zones determined by technical experts are inadequate. Furthermore, downstream water users are often not considered in assessment of impacted peoples. It is worth noting that social impacts do not necessarily decrease in intensity with distance from the project site. People are connected by a vast array of linkages and networks. Projects also can often have a wide logistics corridor and complex value chains (with backwards and forwards linkages). Defining a ‘social area of influence’ does not necessarily require the articulation of a geographic boundary. Instead, the social extent of the project can be determined through a combination of stakeholder analysis and social mapping, and through an iterative process of understanding the social, economic, political and environmental changes induced by the project and the livelihoods and networks of potentially impacted people.

Task 4: Gain a good understanding of the communities likely to be affected by the project by preparing a Community Profile which includes: (a) a thorough stakeholder analysis; (b) a discussion of the socio-political setting; (c) an assessment of the differing needs, interests, values and aspirations of the various subgroups of the affected communities including a gender analysis; (d) an assessment of their impact history, i.e. their experience of past projects and other historical events; (e) a discussion of trends happening in those communities; (f) a discussion of the assets, strengths and weaknesses of the communities; and (g) optionally the results of an opinion survey. This task is typically called profiling.



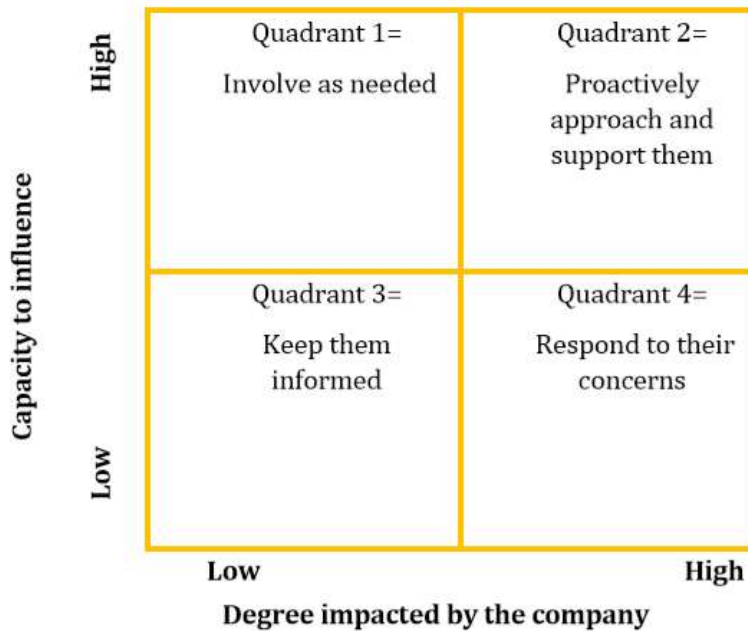
DevTreks –social budgeting that improves lives and livelihoods

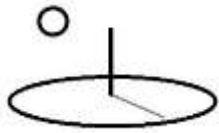
operations). Provide as much detail as possible (i.e. name your main suppliers independently; disaggregate communities by identifying local leaders or other relevant actors).

- Next to each stakeholder group, discuss and list their key concerns, issues, interests.

STAKEHOLDER	ISSUES/CONCERNS/INTERESTS

- Place these stakeholders on the stakeholder map according to the degree to which they are impacted by your operations and their capacity to influence your company operations.





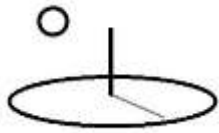
DevTreks –social budgeting that improves lives and livelihoods



Figure 11: Key elements of the stakeholder engagement process undertaken by AHAM

Numbers in parentheses indicate the methodology step during which the stakeholder engagement should occur

IFC (2014) use the following image to illustrate the use of a Stakeholder Engagement Plan as a basis for identifying and working with impacted communities and stakeholders when conducting SIA.



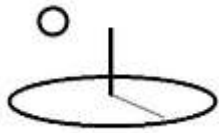
DevTreks –social budgeting that improves lives and livelihoods

ABC FRUIT COMPANY Stakeholder Engagement Plan

Based on the information above, ABC prepared a Stakeholder Engagement Plan. ABC **prioritized** engagement with those groups that are most affected.

STAKEHOLDER ENGAGEMENT PLAN FOR AFFECTED STAKEHOLDERS				
Stakeholder	Concerns	Engagement method	Information to disclose and report back	Most valuable info to obtain
a. Sindanao farmers association (Quadrant 2)	<p>Groundwater supply: possibly contaminated by pesticide residue has caused illness amongst local population</p> <p>Soil deterioration: through overuse of fertilizers and machinery</p> <p>Exploitative land practices</p>	<p>- Grievance mechanism</p> <p>- Quarterly meetings with members of the association and more frequently as demand requires</p> <p>- Participatory monitoring of groundwater and soil quality</p> <p>- Annual perception survey</p>	<p>- Measures to reduce and/or improve safety of storage and handling of pesticides at plantation</p> <p>- Progress on safe disposal of pesticides</p> <p>- Results of groundwater quality monitoring</p> <p>-Measures to use more sustainable cultivation practices</p> <p>-Measures to allow farmers to independently cultivate their land</p>	Exact threat, damage, possible solutions, costs of these
b. Local village population (Quadrant 2)	<p>Concerns of water quality for consumption: contaminated by pesticide residue may have caused illness amongst local population</p> <p>Noise and air pollution from plantation work</p>	<p>- Grievance mechanism</p> <p>- Biannual meetings with community leaders</p> <p>- Annual perception survey</p>	<p>- Progress using alternatives to chemical pesticides</p> <p>- Progress on safe disposal of pesticides and water saving/recycling actions to reduce likelihood of wider spread contamination</p> <p>- Results of groundwater quality monitoring</p> <p>-Control of hours of vehicle traffic to and from plantation</p>	Identification of illnesses associated with use of pesticides
c. Consumers (Quadrant 2)	Food safety	<p>- Grievance mechanism through telephone hotline</p> <p>- Well-briefed</p>	- Progress on actions to improve quality of fruit by reducing use of agrichemicals in their	Consumer concerns Competitors' methods for

Vanclay et al (2015) use the following image to illustrate the use of a Community Engagement Plan as a basis for getting more participation from impacted communities and stakeholders when conducting SIA.



BOX 6: Example of a community engagement continuum

IAP2's Public Participation Spectrum

The IAP2 Federation has developed the Spectrum to help groups define the public's role in any public participation process. The IAP2 Spectrum is quickly becoming an international standard.



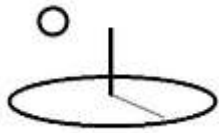
Increasing Impact in the Decision

	Inform	Consult	Involve	Collaborate	Empower
Public Participation Goal	To provide the public with balance and objective information to assist them in understanding the problem, alternatives, opportunities and/or solutions.	To obtain public feedback on analysis alternatives and/or decisions.	To work directly with the public throughout the process to ensure that public concerns and aspirations are consistently understood and considered.	To partner with the public in each aspect of the decision including the development of alternatives and the identification of the preferred solution.	To place final decision making in the hands of the public.
Promise to the Public	We will keep you informed.	We will keep you informed, listen to and acknowledge concerns and aspirations, and provide feedback on how public input influenced the decision. We will seek your feedback on drafts and proposals.	We will work with you to ensure that your concerns and aspirations are directly reflected in the alternatives developed and provide feedback on how public input influenced the decision.	We will work together with you to formulate solutions and incorporate your advice and recommendations into the decisions to maximum extent possible.	We will implement what you decide.

© IAP2 International Federation 2014. All rights reserved.

Source: International Association for Public Participation (<http://www.iap2.org>) (reproduced with permission)

IFC (2015) use the following statement to clarify that SIA is fully applicable to small and medium size enterprises (SMEs). As the previous images demonstrate, standards setting organizations, such as the producer organizations and CSOs introduced in the SPA2 reference, often work at national and industry scale to help SMEs with these systems (5*). SPA2 verifies that many of the SMEs are actually privately run businesses, such as family farms, rather than publicly traded corporations.



DevTreks –social budgeting that improves lives and livelihoods

In the current economic climate, companies are under pressure to perform or even just survive. New initiatives are often met with resistance as people struggle to keep up with their day-to-day responsibilities. Some people think that an environmental and social management system must be big, complicated and expensive. But that is not really true. To be effective, a management system needs to be scaled to the nature and size of the company.

In terms of natural resources conservation, the following watershed management framework (FAO, 2017) confirms that stakeholder boundaries often relate to watersheds, land use areas, ecosystems, and Resource Conservation Districts. These scales are important because the impact of individual private sector, CSO, and local community activities take place at landscape, watershed, and ecosystem, scale. Coordination of multiple stakeholders is required to achieve landscape-wide, and watershed-level, SDG “materiality” impacts.



DevTreks –social budgeting that improves lives and livelihoods

Watershed management is best carried out as a stepwise multistakeholder process. The project review identified a sequence of steps that watershed management projects or programmes should ideally follow. The structure of this publication is based on these steps:

- ▶ promoting an enabling environment (policy, legal and regulatory frameworks, institutional dialogue and cooperation, capacity building, sharing of results, and finance and investment);
- ▶ selecting the area for interventions – the size, number and location of watersheds – based on appropriate criteria;
- ▶ identifying target beneficiaries, mobilizing actors and fostering multistakeholder participation;
- ▶ assessing the state and trends in the watershed – biophysical, socio-economic and institutional – using a variety of scientific and participatory methods and tools, to establish a baseline for defining interventions and monitoring future results;
- ▶ bringing together the stakeholders to analyse the problems in the watershed (using the results from the assessment phase), to identify solutions and to prioritize potential interventions;
- ▶ developing a watershed management plan, formulated through a collaborative process and validated by all stakeholders, to ensure holistic planning, resource allocation and implementation;
- ▶ implementing activities on the ground, establishing a balance among the competing needs, demands and priorities of different stakeholders within the available resources;
- ▶ monitoring to measure the changes in the state or condition of the watershed and to track the performance of the project against its objectives, based on an appropriate set of environmental, social, economic and institutional indicators.

Although companies themselves don't usually lead watershed management teams, they often are stakeholders in, and sponsors of, a community's overall watershed, or ecosystem, management plan (i.e. BLM or Forest Service plans in the USA). FAO (2017) describes the central role played by the SDG in these integrated landscape management approaches and provides practical advice for incorporating the SDG, or "integrated and cross sectoral approaches", in resource conservation planning. Business support for these approaches plays an important role in improving "community institutions", because as FAO states "The private sector has had little role in the mobilization of resources [for watershed planning]", but "Incorporating [SDG-related] projects in existing institutional structures is crucial at both the local and national levels." (10*)

Watershed and landscape scales are also important because tradeoffs are needed to accommodate the needs of impacted stakeholders who hold different value systems. Example 4B in SPA2 discuss the importance of carrying out value assessments, including Cost Effectiveness Analysis, using "perspectives" that reflect the value systems held by different groups of impacted stakeholders. Understanding the value systems, as measured using value assessment "perspectives", offers a better chance of reducing conflict and achieving equity, and thereby

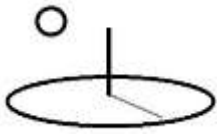


DevTreks –social budgeting that improves lives and livelihoods

accomplishing successful results. Several references endorse allowing each stakeholder group to take ownership of their own sustainability assessment, including the required M&E system. In the latter case, RCA Assessors play consultant roles to each stakeholder group. These experts focus on understanding each stakeholder perspective, identifying potential tradeoffs among stakeholders, and exploring possible sustainability transition states (i.e. low sustainability to high sustainability), that can keep enough of the stakeholders satisfied to achieve the ultimate sustainability goals. In practice, that means carefully considering and recommending refinements to Example 4B-like Natural Capital Care Perspective, Stakeholder Groups Perspectives, and the Societal Perspective.

E. Risk and Impact Identification

The IFC references related to agriculture introduce several agricultural case studies and examples relevant for SIA purposes. For example, the following images (IFC, 2014 and 2015, and TFCD, 2017) demonstrate how to more thoroughly identify crop production risks and impacts. Although similar studies can be found throughout the agricultural sustainability literature, most reflect the era before the SDGs, cloud computing open source apps, and machine learning algorithms.



ELEMENT 2

Identification of Risks and Impacts

5

Mature system, routinely reviewed and updated as part of a continual improvement plan. Internal and external inputs. Procedures extended to contractors, subcontractors, third parties and supply chain as relevant.

4

Systematic, documented identification and prioritization of E&S risks and impacts. Routinely reviewed and updated across existing, new and changing activities. Wide awareness and engagement in company.

3

Awareness and engagement of staff in identification and prioritization of E&S risks and impacts. External experts involved as required.

2

Procedures in place for identification of E&S risks and impacts across all key

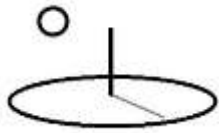
The primary objective of a risk assessment is to identify the potential negative environmental and social risks so that you can develop the appropriate strategies to address the risks and their potential impacts.

In the following pages, we present the key issues that come up across different industries.

GENERAL ENVIRONMENTAL AND SOCIAL RISKS AND IMPACTS

1. Environmental

RISKS	POTENTIAL IMPACTS
Release of air pollutants (air emissions)	Pollution of air, land, and surface water
Release of liquid effluents or contaminated wastewater into local	Surface water pollution

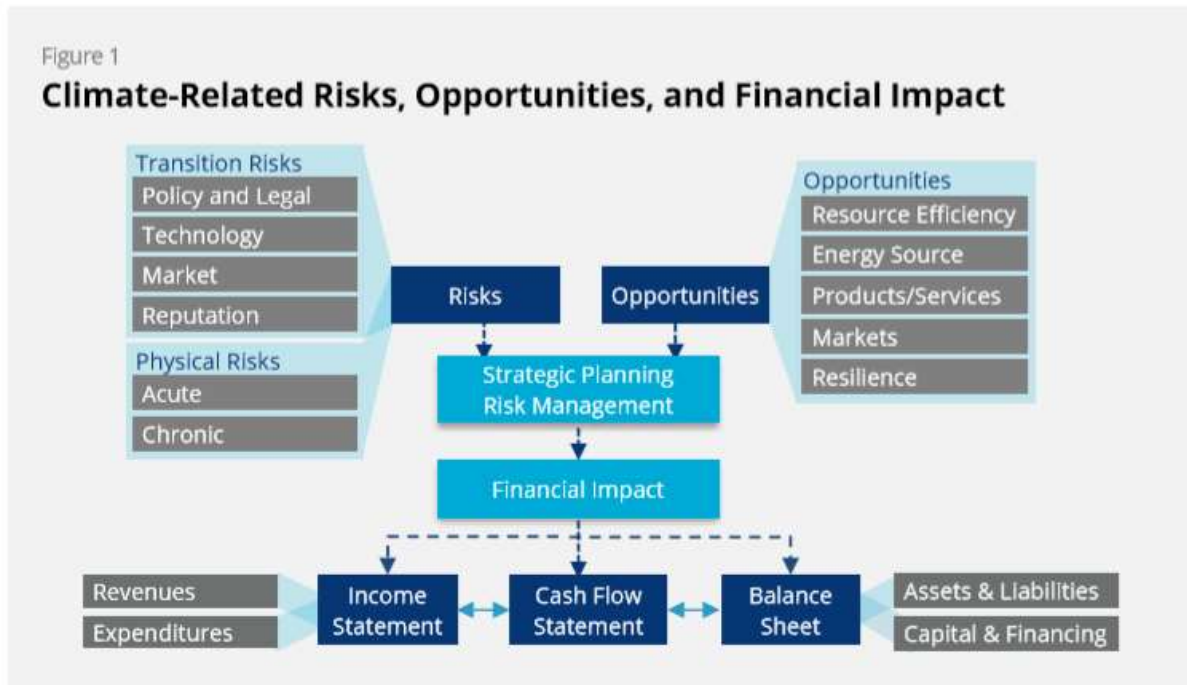
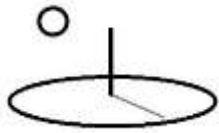


ABC 2. Identification of Risks and Impacts

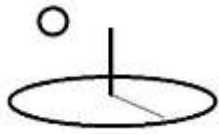
ABC FOOD COMPANY Risk Identification Worksheet

ABC’s ESMS team used the Risk Identification Worksheet below to identify those areas where problems are more likely to happen.

LABOR AND WORKING CONDITIONS RISKS		
RISK FACTORS	My company has the following conditions (circle the appropriate answer)	Potential negative impact (A “yes” response means that there is a potential negative impact)
There is a difference in nationality, race or religion between workers and managers.	Yes/No <input checked="" type="radio"/>	Discrimination. Disciplinary abuse and harassment.
We have an apprentice program that provides young workers with training and work experience.	Yes/No <input checked="" type="radio"/>	Forced labor. Child labor.
Our managers and supervisors are not aware of the workers’ rights under the national labor law or collective agreements.	Yes/No <input checked="" type="radio"/>	Discrimination. Disciplinary abuse and harassment. Excessive overtime. Inadequate Wages. Restriction on freedom of association and collective bargaining.
Children or young workers are employed at the farm.	Yes/No <input checked="" type="radio"/>	Child labor.
We do not employ children but children accompany their parents during work or leisure time.	Yes/No <input checked="" type="radio"/>	Child labor. Exposure of children to workplace hazards.
We do not have a system for recording the “in” and “out” time for agriculture/plantation laborers.	Yes/No <input checked="" type="radio"/>	Excessive working hours. Lack of overtime payment.
Some plantation workers are paid based upon the tasks performed (quota) rather than hours worked.	Yes/No <input checked="" type="radio"/>	Health and safety risks. Inadequate wage payment. Excessive working hours.
Migrant workers or seasonal workers are	Yes/No <input checked="" type="radio"/>	Discrimination.



The Process Map used to identify risks in the following image (IFC, 2014) is similar to the “impact pathways” introduced in SPA2 for conducting short term Performance Monitoring and to DevTreks hierarchical base elements. Example 1, 2, and 3 demonstrate how to use the impact pathways to monitor and evaluate company SDG-related accomplishment. Examples 3, 3A, and 3B, discuss deriving the Resource Inventory Phase of LCAs by using standard Operating and Capital Budgets to define a company’s Inputs, Operations (or Processes, or Activities), Outputs, Outcomes, and Budget (i.e. holding Impacts such as Net Returns) base elements. Example 4 demonstrates using TEXT datasets to define the same base elements. DevTreks applications, Budgets, impact pathways, and this Example’s impact transition states, offer companies and communities several applied tools to thoroughly understand these work processes, to identify and reduce these risks, and to manage the impacts.



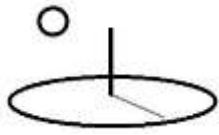
ABC 2. Identification of Risks and Impacts

ABC FOOD COMPANY Process Map

The Process Map presented below covers the activities and operations carried out at ABC Company. However, these would be similar to most of the processes and activities in other chicken processing industries as well.

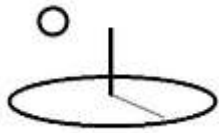
Inputs	Process	Outputs	Potential Negative Impacts - OHS	Potential Negative Impacts - Environment And Community	Opportunity For Waste Reduction/ Energy & Water Savings
Materials, labor, resources	Operational activity	Product, waste, by-product	Injuries, long-term illness	Discharge, contamination, pollution, Shortage	Improved process, re-purposing and recycling by-products
Manual labor	Land preparation (clearing, digging, ploughing and weeding)	Weeds, old pineapple plants	Worker exposure to physical exertion from digging and ploughing trenches, exposure to solar radiation, humidity and to extreme heat Cuts from sharp barbs	Shortage of land for indigenous people and farmers due to long-term lease back arrangements Deforestation and monoculture have altered the biodiversity of the region	Replant pineapple crowns, create wildlife corridors, intercrop (papaya and banana) to reduce erosion and provide shade
	▼				
Manual labor and/or mechanical input, fertilizer	Planting fruit (selecting seeds or seedlings, pineapple crowns or suckers and establishing the plant)	Pineapple plant waste, sediment, fertilizer containers	Worker injuries (back pain/pulled muscles) from repetitive motion and heavy lifting Worker exposure to heat and chemicals (fertilizers)	Surface runoff and leaching of excess crop nutrients leading to contamination of groundwater Eutrophication of surface water resources	Compost and use of plant cover to reduce loss of nutrients, biogas, fertigation, protective buffer zones, organic manure
	▼				
Manual labor	Weed and	Plant waste,	Worker injuries from	Exposure of local	Crop rotation

The following image (Taylor-Powell et al, 2003) of an impact pathway demonstrates that many longer term Impact Evaluations, especially those involving local community interventions or company long term strategy, will be less interested in identifying project or business risks and instead focus on choosing Indicators that can verify results –that can be used to evaluate cause and effect attribution. The primary long term business risks arise from wasting resources on ineffective Activities and failing to achieve acceptable levels of SDG-related Impacts. For developing countries, the IFC documents the importance of identifying and reducing risks that

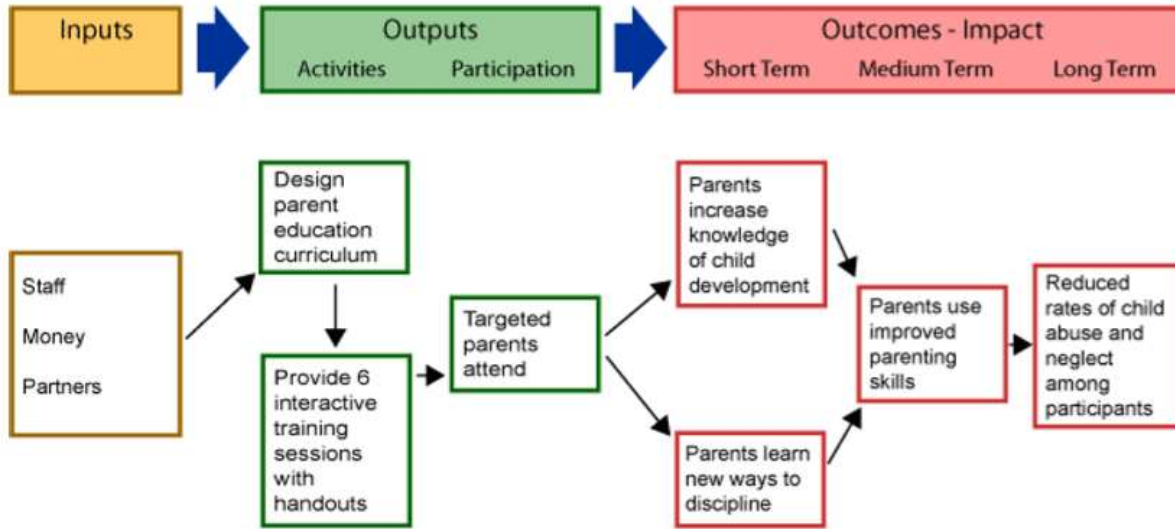


DevTreks –social budgeting that improves lives and livelihoods

are external to a company’s internal activities, such as social discord, disease outbreaks, and institutional corruption. For companies in developed countries, this example identifies external risks related to informed stakeholder actions, such as consequential digital activism. Activities must be broadly defined in these Logic Frames and impact pathways. Example 7 [may] demonstrate using the impact pathways in machine learning algorithms to more fully identify and reduce these risks and evaluate their long term impacts (i.e. causal inference and causal tree algorithms). The image also highlights how activities must be grounded in the socioeconomic root causes of societal problems, with less emphasis on superficial band-aids (i.e. gun violence in the USA).



DevTrek –social budgeting that improves lives and livelihoods



Key Evaluation Questions

Were the inputs sufficient, timely?	Was curriculum developed? Were all 6 sessions delivered?	Did all parents attend that we intended? who did/did not attend? Did they attend all 6 sessions? Why/Why not?	To what extent did knowledge increase? did they learn new approaches? What else happened?	Are parents actually using improved skills? So what? What difference do these skills make?	Has there been a decrease in rates among participants? Were goals met?
-------------------------------------	--	---	---	--	--

Indicators

# staff \$ used # partners When delivered	Curriculum # sessions held	#, % attended per session Certificate of completion	#, % with increased knowledge of... Additional outcomes	#, % using skills Types of differences reported	#, % abusing/neglecting children before - after
--	-------------------------------	--	--	--	---

The following image (Independent Evaluation Group, World Bank, 2017) verifies that impact pathways, or theories of change, make good frameworks for helping companies and local communities to understand, account for, and report on, their own accomplishments, including SDG-related achievements (8*). Section H, Communication, demonstrates how to use Results Chains to communicate this algorithm’s raw Math Results to stakeholders.

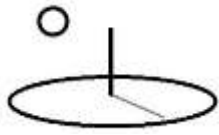
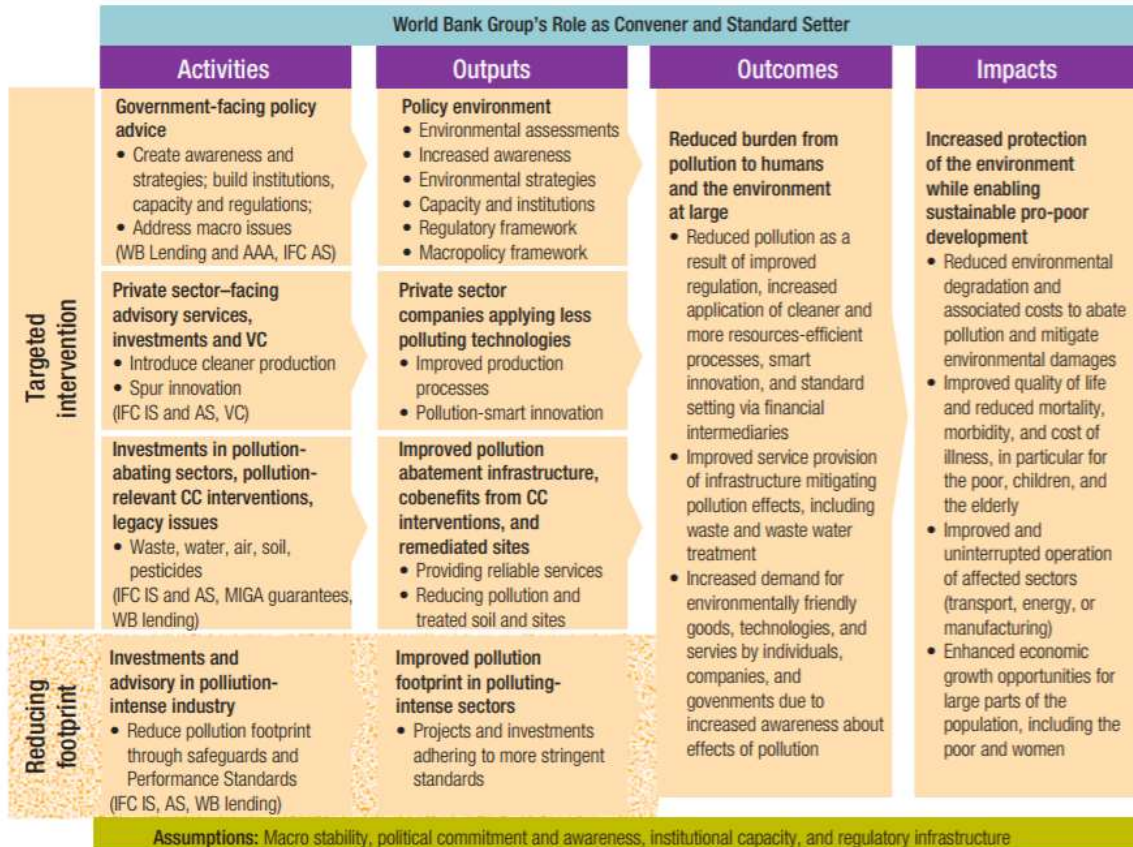


Figure 1.1 | Theory of Change for World Bank Group Pollution Interventions



Note: AAA = analytic and advisory activities; AS = Advisory Services; CP = cleaner production; CC = climate change; IS = Investment Services; WB = World Bank; VC = venture capital.

The RCA framework introduced in Appendix A, SPA2, used the following image to introduce terms used in ecosystem and disaster assessment that are also used to reduce risk, including stressors, drivers, pressures, exposure, vulnerability, and resiliency. These techniques, which will be revisited in Example 6, are especially important when carrying out risk identification and reduction at watershed, landscape, and planet (i.e. climate change) scales.



Box 1. Working definitions.

Urban Climate Resilience: The ability of a city or urban system, through its risk reduction and response capacity capabilities, to reduce exposure and sensitivity to, and recover and learn from, gradual climatic changes or extreme climate events, in order to retain or improve the integrity of its infrastructure and economic systems, vital environmental services and resources, the health and welfare of its populations and communities, and the flexibility and diversity of its institutional and governance structures (adapted from Leichenko, 2011).

Exposure: The presence of people; livelihoods; environmental services and resources; infrastructure; or economic, social, or cultural assets in places that could be affected by climate change stressors (adapted from IPCC, 2012).

Sensitivity: Predisposition of human beings, infrastructure, society, and ecosystems to be affected by exposure to a climate stressor or an effect of that exposure (adapted from IPCC, 2012).

Response Capacity: Intrinsic capacity of a community to recover from alterations in its normal functioning due to gradual changes in the climate or to extreme events that result in adverse human, material, economic, or environmental effects.

Learning: Capability to recognize complex dynamics of socio-ecological systems in order to respond appropriately to risk and make effective adaptation responses, identifying mistakes and shortcoming in those responses following climate stressor events, and evolving as new information becomes available (drawn from IPCC, 2012; Kasperson, 2012). (Learning outcomes are on three levels: reacting, reframing, and transforming [see Figure 1-3, IPCC, 2012]. Examples: reacting—increase levee height; reframing—realizing the need to assess new storm duration frequency distributions; transforming—assessing societal constructs and migrating to a more robust and comprehensive risk management strategy.)

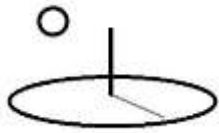
Bridges to Action: Conditions under which unforeseen and huge leaps are made in a community’s ability to respond to and recover from alterations or disruptions in its normal functioning (e.g., due to social or technical innovation).

Risk: A function of the exposure to and severity of the occurrence of a particular type of climate change (gradual or extreme) and the way in which its consequences are likely to be mediated by the social vulnerability of the human system. Risk can be assessed in terms of condition and predictive variables representing factors such as economic well-being; health and education status; and preparedness and coping ability with respect to particular climatic changes.

Risk Reduction Capacity: Ability to reduce risk by reducing exposure and sensitivity or increasing recovery potential and adaptive capacity, to prepare for expected climatic changes or events.

Note: These definitions are considered operational definitions. Therefore, they might not be identical to the definitions in the current literature, but they have been selected for their appropriateness to this application.

The FAO’s (2017) review of 12 watershed management projects found that “capacity gaps identified in the watershed assessments were mainly in the socio-economic disciplines and linked to inadequate analytical skills.” Many natural resource planning groups are bad at incorporating humans and their institutions, and identifying socioeconomic-related risks and impacts, in their work. The instruments, or algorithms, introduced in this reference reinforce the FAO recommendation to employ “standardized format[s] for data and information collection” as a means of institutionalizing the missing socioeconomic and institutional analysis skills. New types of jobs, such as RCA Assessors and RCA Technologists, which focus on applying modern



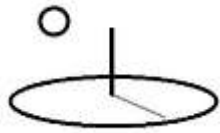
DevTreks –social budgeting that improves lives and livelihoods

IT to understand “value perspectives” and facilitate “stakeholder tradeoffs”, offer a promising means to institutionalize the needed skills at scale and scope.

Dataset 1. Business Resource Conservation Value Accounting (5*)

The following image displays part of a typical, but stylized, TEXT dataset used with this Business Value Report to identify the risks and impacts introduced in Example 1 to 4’s impact pathways and datasets. Many of these socioeconomic characteristics derive from SPA2’s COSA and Sustainable Food Lab references.

1st TEXT dataset. Demographics and SDG Indicators. The following dataset shows that the CIs record socioeconomic characteristics for different stakeholder groups while each CI’s children Indicators measure SDG-related measurements that derive from related subalgorithms, such as Examples 1 to 4. In this dataset, the 2nd CIs are actually measuring land use transition states, rather than population transitions.



DevTreks –social budgeting that improves lives and livelihoods

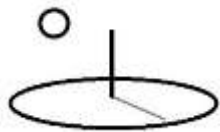
label	location	risks_and_impacts	factor1	factor2	factor3	factor4	factor5	factor6	factor7	factor8	factor9	factor10	factor11
A	0	Farmer Characteristics	hhincome	gender	age	ethnicity	disability	education	experience	hhsize	hhwealth	hhfood index	poverty index
B	0	Farmer Characteristics	tenure	management	size	sublocation	total gross revenue	coffee gross revenue	access to services	trading relationship	farm productivity	visual soil assessment	standard precipitation
A1A	1	Small scale coffee farmers male low income	5500	male	45	hispanic	none	4	10	5	3500	3.5	5
IF1A	1	Non Harassment and Discrimination Index	75	Percent Workers at	37.5	135	lower and upper 80% CI	2	3.00	normal	HCA	production	1/1/2018
IF1B	1	Air Pollution Index	80	daly / kg PM25	60	150.00	Lower and upper	2	3.00	normal	EDE	production	1/1/2018
IF1C	1	Water Consumption Index	60	daly / m3	40	75.00	lower and upper 80% CI	2	3.00	normal	EDF	production	1/1/2018
IF1D	1	Damage Risk Reduction Index	60	daly / extreme	40	75.00	lower and upper 80% CI	2	3.00	normal	PCB	production	1/1/2018
B1A	1	Small scale farms 1000-3000 m slopes <20% 60 SQJ	owned	moderate	2.5	25	5000.00	3000	3.5	5	4	3.5	5
IF2A	1	Biodiversity Index	45	amd	25	65.00	lower and upper 80% CI	2	3.00	normal	EDG	production	1/1/2018
IF2B	1	Climate Change Short Term Index	45	gwp 100	25	65.00	lower and upper 80% CI	2	3.00	normal	ECB	production	1/1/2018
IF2C	1	Climate Change Long Term Index	45	gtp 100	25	65.00	lower and upper 80% CI	2	3.00	normal	ECC	production	1/1/2018
A2A	1	Small scale coffee farmers female low income	4500	female	45	hispanic	none	4	10	5	3500	3.5	5
IF1A	1	Non Harassment and Discrimination Index	75	Percent Workers at	37.5	135	lower and upper 80% CI	2	3.00	normal	HCA	production	1/1/2018
IF1B	1	Air Pollution Index	80	daly / kg PM25	60	150.00	Lower and upper	2	3.00	normal	EDE	production	1/1/2018
IF1C	1	Water Consumption Index	60	daly / m3	40	75.00	lower and upper 80% CI	2	3.00	normal	EDF	production	1/1/2018
IF1D	1	Damage Risk Reduction Index	60	daly / extreme	40	75.00	lower and upper 80% CI	2	3.00	normal	PCB	production	1/1/2018
B2A	1	Small scale farms high elevation moderate slopes	owned	moderate	2.5	25	5000.00	3000	3.5	5	4	3.5	5
IF2A	1	Biodiversity Index	45	amd	25	65.00	lower and upper 80% CI	2	3.00	normal	EDG	production	1/1/2018
IF2B	1	Climate Change Short Term Index	45	gwp 100	25	65.00	lower and upper 80% CI	2	3.00	normal	ECB	production	1/1/2018
IF2C	1	Climate Change Long Term Index	45	gtp 100	25	65.00	lower and upper 80% CI	2	3.00	normal	ECC	production	1/1/2018
SA	1	Small scale poor coffee farmers	0	0	0	0	0.00	none	0	0	0	weights	2
A1B	1	Small scale coffee farmers male medium income	5500	male	45	hispanic	none	4	10	5	3500	3.5	5
IF1A	1	Non Harassment and Discrimination Index	75	Percent Workers at	37.5	135	lower and upper 80% CI	2	3.00	normal	HCA	production	1/1/2018
IF1B	1	Air Pollution Index	80	daly / kg PM25	60	150.00	Lower and upper	2	3.00	normal	EDE	production	1/1/2018
IF1C	1	Water Consumption Index	60	daly / m3	40	75.00	lower and upper 80% CI	2	3.00	normal	EDF	production	1/1/2018
IF1D	1	Damage Risk Reduction Index	60	daly / extreme	40	75.00	lower and upper 80% CI	2	3.00	normal	PCB	production	1/1/2018
B1B	1	Small scale farms 1000-3000 m slopes <20% 60 SQJ	owned	moderate	2.5	25	5000.00	3000	3.5	5	4	3.5	5
IF2A	1	Biodiversity Index	45	amd	25	65.00	lower and upper 80% CI	2	3.00	normal	EDG	production	1/1/2018
IF2B	1	Climate Change Short Term Index	45	gwp 100	25	65.00	lower and upper 80% CI	2	3.00	normal	ECB	production	1/1/2018
IF2C	1	Climate Change Long Term Index	45	gtp 100	25	65.00	lower and upper 80% CI	2	3.00	normal	ECC	production	1/1/2018
A2B	1	Small scale coffee farmers female medium income	4500	female	45	hispanic	none	4	10	5	3500	3.5	5
IF1A	1	Non Harassment and Discrimination Index	75	Percent Workers at	37.5	135	lower and upper 80% CI	2	3.00	normal	HCA	production	1/1/2018
IF1B	1	Air Pollution Index	80	daly / kg PM25	60	150.00	Lower and upper	2	3.00	normal	EDE	production	1/1/2018
IF1C	1	Water Consumption Index	60	daly / m3	40	75.00	lower and upper 80% CI	2	3.00	normal	EDF	production	1/1/2018
IF1D	1	Damage Risk Reduction Index	60	daly / extreme	40	75.00	lower and upper 80% CI	2	3.00	normal	PCB	production	1/1/2018
B2B	1	Small scale farms high elevation moderate slopes	owned	moderate	2.5	25	5000.00	3000	3.5	5	4	3.5	5
IF2A	1	Biodiversity Index	45	amd	25	65.00	lower and upper 80% CI	2	3.00	normal	EDG	production	1/1/2018
IF2B	1	Climate Change Short Term Index	45	gwp 100	25	65.00	lower and upper 80% CI	2	3.00	normal	ECB	production	1/1/2018
IF2C	1	Climate Change Long Term Index	45	gtp 100	25	65.00	lower and upper 80% CI	2	3.00	normal	ECC	production	1/1/2018
SB	1	Small scale medium income coffee farmers	0	0	0	0	0.00	none	0	0	0	weights	2
A1C	1	Small scale coffee farmers male high income	5500	male	45	hispanic	none	4	10	5	3500	3.5	5
IF1A	1	Non Harassment and Discrimination Index	75	Percent Workers at	37.5	135	lower and upper 80% CI	2	3.00	normal	HCA	production	1/1/2018



2nd TEXT dataset. Population transition states and SDG allocations. The following dataset shows that the CIs record uncertain population measurements for the 1st dataset’s stakeholder groups while the CI Indicators allocate the 1st dataset’s SDG-related measurements to these stakeholders. Additional population algorithms may employ alternative population modeling techniques.

label	location	risks_and_impacts	factor1	factor2	factor3	factor4	factor5	factor6	factor7	factor8	factor9	factor10	factor11
A	0												
B	0												
A1A	1	Small scale coffee farmers male low income	1/1/2018	1/1/2020	100000	population	100000	mean	15000	sd	normal	none	0
IF1A	1	Non Harassment and Discrimination Index	100	125	120	100	105	3	4	5	1	weights	3
IF1B	1	Air Pollution Index	25	110	105	100	105	3	4	5	1	weights	3
IF1C	1	Water Consumption Index	25	110	105	100	105	3	4	5	2	weights	3
IF1D	1	Damage Risk Reduction Index	25	110	105	100	105	3	4	5	3	weights	3
B1A	1	Small scale farms 1000-3000 m slopes <20% 60 SQI	1/1/2018	1/1/2020	1000	hectares	1000	mean	100	sd	normal	none	0
IF2A	1	Biodiversity Index	25	110	105	100	105	3	4	5	4	weights	3
IF2B	1	Climate Change Short Term Index	25	110	105	100	105	3	4	5	5	weights	3
IF2C	1	Climate Change Long Term Index	25	110	105	100	105	3	4	5	5	weights	3
A2A	1	Small scale coffee farmers female low income	1/1/2018	1/1/2020	100000	population	100000	mean	15000	sd	normal	none	0
IF1A	1	Non Harassment and Discrimination Index	100	125	120	100	105	3	4	5	1	weights	3
IF1B	1	Air Pollution Index	25	110	105	100	105	3	4	5	1	weights	3
IF1C	1	Water Consumption Index	25	110	105	100	105	3	4	5	2	weights	3
IF1D	1	Damage Risk Reduction Index	25	110	105	100	105	3	4	5	3	weights	3
B2A	1	Small scale farms high elevation moderate slopes	1/1/2018	1/1/2020	1000	hectares	1000	mean	100	sd	normal	none	0
IF2A	1	Biodiversity Index	25	110	105	100	105	3	4	5	4	weights	3
IF2B	1	Climate Change Short Term Index	25	110	105	100	105	3	4	5	5	weights	3
IF2C	1	Climate Change Long Term Index	25	110	105	100	105	3	4	5	5	weights	3
SA	1	Small scale poor coffee farmers	0	0	0	0	0	none	0	0	0	weights	3
A1B	1	Small scale coffee farmers male medium income	1/1/2018	1/1/2020	100000	population	100000	mean	15000	sd	normal	none	0

Vanclay et al (2015) point out that SIA is often conducted concurrently with, or before, fuller sustainability assessments. In many cases, Example 1 to 4’s RCA-style datasets may need to be completed specifically for the impacted populations first identified using this algorithm. For example, once populations have been characterized and aggregated into typical TEXT hierarchies, such as this dataset’s poverty-related CIs, the SPA2’s RCA tools are completed for those specific groups. This algorithm is rerun after the SPA2 data results have been transferred back to these datasets. This technique allows more fined tuned control in identifying SDG-related risks that impact both the companies and their stakeholders. The Monitoring and Evaluation section, below, discusses how companies and local communities carry this out as part of their standard management system



DevTreks –social budgeting that improves lives and livelihoods

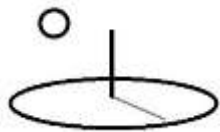
The following features of these 2 datasets highlight how companies use this algorithm for RCA reporting.

1st TEXT file. Demographics and SDG Indicators

Stakeholder Characteristics Titles (3*): In this dataset, the first 2 rows hold the titles for the stakeholder socioeconomic characteristics being recorded. At least 1 row must be used to hold these titles, but the algorithm supports additional rows. These particular characteristics come from several of this tutorial’s references. The Sustainable Food Lab and COSA references found in SPA2, in particular, explain the importance of developing standard characteristics that can be used throughout industries (i.e. agricultural stakeholders in developing countries). Ultimately, this data supports the end goal of cause and effect attribution pursued by Impact Evaluation to explain sustainability and stakeholder impacts.

The specific reason for these 2 particular sets, or rows, of socioeconomic characteristics is that farm sustainability is usually measured in terms of both stakeholder characteristics and supply chain, and in this example, farm, characteristics. Did conservation practices cause farmers’ income go up or down? Were farmworkers protected against safety hazards? Did cultivation practices cause soil erosion rates and farm productivity to go up or down? Did grazing practices result in better wildlife protection? Depending on the purpose of the SIA, additional characteristics may be based on food security, health care, public infrastructure, and specific ecosystems.

Only the most important socioeconomic variables –those needed to explain attribution, should be included in these datasets. Datasets that employ more than 11 to 22 variables are more likely to be “fishing for attribution” rather than “explaining attribution” and should be analyzed using other algorithms prior to being aggregated into these datasets (i.e. refer to Example 7, IITA and COSA, 2016, and Hoebink, 2014, for examples).



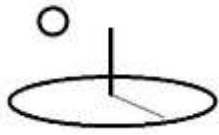
Total Risk Indexes (TR) and Locational Indexes (LI): The examples in this reference will demonstrate a variety of purposes for these Indexes, including further subdivision of socioeconomic data and additional life cycle stages.

Categorical Index (CI) Stakeholder Characteristics: Each CI records measurements for each of the title rows' characteristics. The goal behind this particular set of properties is to support data collection for single stakeholder members or for groups of stakeholders. The 2 CIs for this dataset relate to stakeholder groups, such as small-scale, low income, coffee farmers, and agricultural areas or “sub-ecosystems”, such as moderately sloped, higher elevation, coffee plantations. The Indicators include averages and percentages, such as average poverty index and age, and percent of the population that has at least 1 disability or that has a minority ethnicity. Stakeholder groups that can't be described using averages or percentages may require a separate CI, or a different Indicator property.

The first CI documents demographic characteristics of the CI's stakeholders. These measurements can be based on single stakeholders, straight population counts, or on common numerator/denominator population metrics, such as 1,000 very low income workers per 10,000 total worker population.

The second CI documents land characteristics for single farm plots or for groups of farms. These measurements can be based on single farms, straight area measurement, or on common ecosystem-related metrics, such as land use areas or watersheds.

This example demonstrates using combinations of single demographic characteristics, such as age and gender, with multi-variate indexes, such as poverty indexes and visual soil assessment indexes (refer to IITA and COSA (2016) for an example of how to use index-based evaluations). The decision on whether to use single characteristics or indexes should be based on their causal relationship to sustainability, to international norms for reporting, and to data collection cost. For example, companies that already use SDG M&E systems, or local communities who use watershed simulation models, can take data



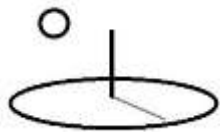
from those systems to develop land use characteristic metrics for soil quality, water quality, air quality, plant quality, and animal quality, indexes.

The RCA-style instruments introduced with Examples 1 to 4 have complementary relationships with these socioeconomic characteristics. For example, the Poverty Index used with this dataset can also be considered a Categorical Index in Examples 1, 1A, 2, and 3B that documents a COSA or SAFA theme, such as Livelihood and Well-Being. This algorithm's socio-economic characteristics serve 2 primary purposes, which help to distinguish them from the RCA impact pathway Indicators (5*):

1. **Disaggregated SDG Reporting.** They support the demographic disaggregation required to fully report SDG and SDRR accomplishments. The UN recommends disaggregating this data by demographic characteristics such as geography, income, sex, age, and disability.
2. **Stakeholder Impact Cause and Effect Attribution.** The socioeconomic characteristics and indexes lead to a better understanding of the cause and effect attribution leading to the RCA impacts, particularly related to stakeholder social impact. This algorithm allocates the general RCA Impacts documented in Examples 1 to 4 to specific groups of stakeholders. Equity, in particular, can be understood better and acted upon.

SDG-related Indicators: The CI or LI math results from subalgorithms 13, 14, 15, or 16, are transferred to separate CI Indicators. As the MathResult section explains, this algorithm will relate these SDG measurements to stakeholder population and/or land use measurements stored in a 2nd TEXT dataset to define stakeholder SDG impact. Custom data derived from other sources can also be used. Example 6 will use this technique with data from the CTAP algorithms to support the SDRR system.

The reason that the natural capital stock Categorical Indexes, Air Pollution and Water Consumption, have been included as Indicators with the stakeholder CI and not the



DevTreks –social budgeting that improves lives and livelihoods

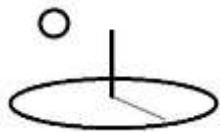
ecosystem CI, is that their units of measurement include Disability Adjusted Life Years (i.e. daly / kg PM25 and daly / m3). Examples 3, 3A, and 3B in the SPA2 reference confirm that many RCA measurements will be based on environmental and socioeconomic impacts that harm humans (1*). This example allocates those damages to particular stakeholder groups and land use areas for subsequent tradeoff analysis (i.e. do some groups experience higher impacts than others? why?). In this tutorial, analyzing equity is a critical part of understanding the term “stakeholder impact” (6*).

2nd TEXT file. Population and Land Use Impact Transition States or Adoption and Diffusion Paths:

A 2nd TEXT dataset stores Categorical Indexes that measure a population and/or land use for the same CI stakeholder group and or land use area defined in the 1st TEXT dataset. CI Indicators are used to allocate the same SDG Indicators from the 1st dataset to the 2nd dataset’s specific population. Additional calculator Indicators (i.e. 2 to 15) can be used to model population and SDG “transition states” that change over time, often as a result of mitigation and adaptation activities.

USHHS (2012) introduces the term “health states” in a similar manner to “impact transition states” and demonstrates how the health care sector carries out Performance Monitoring and Impact Evaluation, or, their preferred term, “Epidemiology”. The transition states are modeled on these “health states” which measure how populations move between states such as healthy, sick, seriously ill, and death (i.e. quality of life states measured as QALYs or DALYs). The Impact Transition states define how the population and their ecosystem move from a baseline, no mitigation action, state, to a final, fully sustainable, impact state. The health care sector has developed several algorithms to analyze and simulate these population transitions (i.e. Markov transition algorithms; the WHO, 2003, reference introduces advanced population models).

In terms of agricultural sustainability, transition states often measure adoption and diffusion “states”, or paths, for new agricultural technologies and policies, such as



DevTreks –social budgeting that improves lives and livelihoods

conservation practices or carbon pricing schemes. IITA and COSA (2016) provide an example for coffee production that show how the poorest stakeholders have significantly lower adoption rates for sustainable practices, with related causal factors, than the richest. A stylized, but timely, example of a “technology adoption and diffusion path” can be found in the definition for Technology Development, Diffusion, and Adoption in SPA1, Appendix D, and below, in Section G, Quality of Life Scenarios.

This example makes arbitrary allocations of the SDG measurements to each stakeholder population and land use area. In practice, better rules are needed for making the allocations. More sophisticated algorithms can be used to codify the rules.

Advanced SIA

Example 7 [may] demonstrate the use of algorithms to train machine learning algorithms to define likely technology diffusion and adoption paths. Once trained, the algorithms predict SDG and SDRR impacts on targeted stakeholders that arise from company activities and community service organization interventions. More advanced algorithms predict likely transition states and advise companies how to dynamically tweak their mitigation and adaptation actions to achieve SDG-related Impacts.

Dataset 2. Community Resource Conservation Value Accounting (5*)

Gertler et al (2016) use the following reference dataset, containing 19,800+ rows of population socioeconomic characteristics, to conduct an Impact Evaluation of a health insurance social intervention. Examples 7 and 8 employ the same dataset and explain more about the data.

	locality_id	household	treatment	promotion	eligible	enrolled	enrolled_r	poverty_index	round	health_ex	age_hh	age_sp	educ_hh	educ_sp	female_hh	indigenous	hhsz	dirtfloor	bathroom	land	hospital_d	hospital
0	26	5	1	1	1	1	1	55.95054245	0	15.18546	24	23	0	6	0	0	4	1	0	1	124.82	0
1	26	5	1	1	1	1	1	55.95054245	1	19.5809	25	24	0	6	0	0	4	1	0	1	124.82	0
2	26	11	1	1	1	1	0	46.05873108	0	13.07626	30	26	4	0	0	0	6	1	0	2	124.82	0
3	26	11	1	1	1	1	0	46.05873108	1	2.398854	31	27	4	0	0	0	6	1	0	2	124.82	1
4	26	13	1	1	1	1	0	54.0958252	1	0	59	57	0	0	0	0	6	1	0	4	124.82	1
5	26	13	1	1	1	1	0	54.0958252	0	15.28635	58	56	0	0	0	0	6	1	0	4	124.82	0
6	26	16	1	1	1	1	1	56.90340042	1	20.02691	36	25	3	0	0	0	7	1	0	2	124.82	0
7	26	16	1	1	1	1	1	56.90340042	0	11.31176	35	24	3	0	0	0	7	1	0	2	124.82	0
8	26	21	1	1	1	1	1	46.90880966	0	11.22391	37	35	0	0	0	0	7	1	0	2	124.82	0
9	26	21	1	1	1	1	1	46.90880966	1	16.66469	39	36	0	0	0	0	7	1	0	2	124.82	0
10	26	22	1	1	1	1	1	48.18392563	0	8.876925	34	41	0	2	1	0	9	1	0	1	124.82	0
11	26	22	1	1	1	1	1	48.18392563	1	0.115546	35	41	0	2	1	0	9	1	0	1	124.82	0
12	26	23	1	1	1	1	0	49.69087982	1	0	31	30	0	6	0	0	7	0	0	1	124.82	0
13	26	23	1	1	1	1	0	49.69087982	0	12.68582	30	29	0	6	0	0	7	0	0	1	124.82	0



DevTreks –social budgeting that improves lives and livelihoods

The goal of this algorithm is not to conduct formal Impact Evaluations for capital projects and social interventions, as demonstrated by Gertler et al (2016). The broader goal is for companies and organizations to achieve sustainable development, as defined by SDG-related targets, by using causal attribution to selfishly identify and reduce risks that impact stakeholders, and to selflessly improve the well-being of targeted stakeholders. Appendix A demonstrates how to use this dataset to conduct a formal Impact Evaluation and then adding those results to this algorithm’s datasets to more thoroughly understand organization and stakeholder long term impacts.

To support the broader SDG goal, Gertler’s reference dataset is aggregated into the following hierarchical TEXT datasets. The Indicators include averages and percentages, such as average poverty index and age, and percent of the population that is indigenous or that has dirt floors. The main purpose of this dataset is to illustrate “harmonizing” existing socioeconomic data, including Gertler’s partial SDG data, with the purposes of this algorithm. Given the transparent nature of all DevTreks data, these aggregated hierarchies also help to address the “misuse of highly personal, private, data” (Espey et al, 2017). A more typical dataset might address SDG-related risks and stakeholders at watershed and landscape level and would be carried out as part of an overall “watershed management plan” (see FAO, 2017).

Indicator 1. Benchmark or Baseline

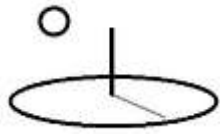
1st TEXT dataset. Demographics and SDG Indicators.



DevTreks –social budgeting that improves lives and livelihoods

label	locatio	risks_and_impacts	factor1	factor2	factor3	factor4	factor5	factor6	factor7	factor8	factor9	factor10	factor11
A	0	HH Characteristics 1	poverty_in dex	age_hh	age_sp	educ_hh	educ_sp	female_hh	indigenous	hhsiz	dirtfloor	bathroom	land
A1A	1	High Poverty Index HHs	44.33	40.69	35.86	2.68	2.52	6.47	43.32	6.20	76.93	52.03	1.68
IF1A	1	Non Harassment and Discrimination Index	75	Percent Workers at Capacity	37.5	135	lower and upper 80%	2	3.00	normal	HCA	production	1/1/2018
IF1B	1	Air Pollution Index	80	daly / kg PM25	60	150.00	Lower and upper estimates	2	3.00	normal	EDE	production	1/1/2018
IF1C	1	Water Consumption Index	60	daly / m3	40	75.00	lower and upper 80%	2	3.00	normal	EDF	production	1/1/2018
IF1D	1	Damage Risk Reduction Index	60	daly / extreme event	40	75.00	lower and upper 80%	2	3.00	normal	PCB	production	1/1/2018
IF2A	1	Biodiversity Index	45	amd	25	65.00	lower and upper 80%	2	3.00	normal	EDG	production	1/1/2018
IF2B	1	Climate Change Short Term Index	45	gwp 100	25	65.00	lower and upper 80%	2	3.00	normal	ECB	production	1/1/2018
IF2C	1	Climate Change Long Term Index	45	gtp 100	25	65.00	lower and upper 80%	2	3.00	normal	ECC	production	1/1/2018
A1B	1	Medium Poverty Index HHs	59.11	48.96	42.25	2.79	2.61	11.25	33.59	4.87	59.00	62.50	1.99
IF1A	1	Non Harassment and Discrimination Index	75	Percent Workers at Capacity	37.5	135	lower and upper 80%	2	3.00	normal	HCA	production	1/1/2018
IF1B	1	Air Pollution Index	80	daly / kg PM25	60	150.00	Lower and upper estimates	2	3.00	normal	EDE	production	1/1/2018
IF1C	1	Water Consumption Index	60	daly / m3	40	75.00	lower and upper 80%	2	3.00	normal	EDF	production	1/1/2018
IF1D	1	Damage Risk Reduction Index	60	daly / extreme event	40	75.00	lower and upper 80%	2	3.00	normal	PCB	production	1/1/2018
IF2A	1	Biodiversity Index	45	amd	25	65.00	lower and upper 80%	2	3.00	normal	EDG	production	1/1/2018
IF2B	1	Climate Change Short Term Index	45	gwp 100	25	65.00	lower and upper 80%	2	3.00	normal	ECB	production	1/1/2018
IF2C	1	Climate Change Long Term Index	45	gtp 100	25	65.00	lower and upper 80%	2	3.00	normal	ECC	production	1/1/2018
A1C	1	Low Poverty Index HHs	78.54	54.50	47.51	3.20	2.90	11.49	15.52	4.72	16.86	79.50	3.88
IF1A	1	Non Harassment and Discrimination Index	75	Percent Workers at Capacity	37.5	135	lower and upper 80%	2	3.00	normal	HCA	production	1/1/2018
IF1B	1	Air Pollution Index	80	daly / kg PM25	60	150.00	Lower and upper estimates	2	3.00	normal	EDE	production	1/1/2018
IF1C	1	Water Consumption Index	60	daly / m3	40	75.00	lower and upper 80%	2	3.00	normal	EDF	production	1/1/2018
IF1D	1	Damage Risk Reduction Index	60	daly / extreme event	40	75.00	lower and upper 80%	2	3.00	normal	PCB	production	1/1/2018
IF2A	1	Biodiversity Index	45	amd	25	65.00	lower and upper 80%	2	3.00	normal	EDG	production	1/1/2018
IF2B	1	Climate Change Short Term Index	45	gwp 100	25	65.00	lower and upper 80%	2	3.00	normal	ECB	production	1/1/2018
IF2C	1	Climate Change Long Term Index	45	gtp 100	25	65.00	lower and upper 80%	2	3.00	normal	ECC	production	1/1/2018
SA	1	Treated Villages	0	0	0	0	0.00	none	0	0	0	weights	2
A1A	1	High Poverty Index HHs	44.33	40.69	35.86	2.68	2.52	6.47	43.32	6.20	76.93	52.03	1.68

2nd TEXT dataset. Populations and Allocated SDG Indicators.



The following dataset was first run unsuccessfully.

label	location	risks_and_impacts	factor1	factor2	factor3	factor4	factor5	factor6	factor7	factor8	factor9	factor10	factor11
A	0												
A1A	1	High Poverty Index HHs	1/1/2018	1/1/2020	1,205	populatio n	1,085	minus 10%	1,325	plus 10%	none	none	none
IF1A	1	Non Harassment and Discrimination Index	100	125	120	100	105	3	4	5	1	weights	3
IF1B	1	Air Pollution Index	25	110	105	100	105	3	4	5	2	weights	3
IF1C	1	Water Consumption Index	25	110	105	100	105	3	4	5	3	weights	3
IF1D	1	Damage Risk Reduction Index	25	110	105	100	105	3	4	5	4	weights	3
IF2A	1	Biodiversity Index	25	110	105	100	105	3	4	5	5	weights	3
IF2B	1	Climate Change, Short Term Index	25	110	105	100	105	3	4	5	0.5	weights	3
IF2C	1	Climate Change, Long Term Index	25	110	105	100	105	3	4	5	10	weights	3
SA	1	High Poverty Index HHs	0	0	0	0	0.00	none	0	0	0	weights	2
A1B	1	Medium Poverty Index HHs	1/1/2018	1/1/2020	3,227	populatio n	2,905	minus 10%	3,549	plus 10%	none	none	none

The following dataset was then run successfully. TEXT csv datasets need careful attention. Good practice is to always run searches for commas in the TEXT files prior to saving them and to inspect the final TEXT dataset in a text editor, such as Notepad, to check for anomalies.

label	location	risks_and_impacts	factor1	factor2	factor3	factor4	factor5	factor6	factor7	factor8	factor9	factor10	factor11
A	0												
A1A	1	High Poverty Index HHs	1/1/2018	1/1/2020	1.21E+03	populatio n	1.09E+03	minus 10%	1.33E+03	plus 10%	none	none	none
IF1A	1	Non Harassment and Discrimination Index	100	125	1.20E+02	100	1.05E+02	3	4.00E+00	5	1	weights	3
IF1B	1	Air Pollution Index	25	110	1.05E+02	100	1.05E+02	3	4.00E+00	5	2	weights	3
IF1C	1	Water Consumption Index	25	110	1.05E+02	100	1.05E+02	3	4.00E+00	5	3	weights	3
IF1D	1	Damage Risk Reduction Index	25	110	1.05E+02	100	1.05E+02	3	4.00E+00	5	4	weights	3
IF2A	1	Biodiversity Index	25	110	1.05E+02	100	1.05E+02	3	4.00E+00	5	5	weights	3
IF2B	1	Climate Change Short Term Index	25	110	1.05E+02	100	1.05E+02	3	4.00E+00	5	0.5	weights	3
IF2C	1	Climate Change Long Term Index	25	110	1.05E+02	100	1.05E+02	3	4.00E+00	5	10	weights	3
A1B	1	Medium Poverty Index HHs	1/1/2018	1/1/2020	3.23E+03	populatio n	2.91E+03	minus 10%	3.55E+03	plus 10%	none	none	none
IF1A	1	Non Harassment and Discrimination Index	100	125	1.20E+02	100	1.05E+02	3	4.00E+00	5	1	weights	3

The following features of these datasets highlight how local communities use this algorithm for RCA reporting.



DevTreks –social budgeting that improves lives and livelihoods

Total Risk Indexes (TRs). This example only uses the TR Indexes to aggregate the children LIs. Other datasets may choose to use the TRs to further subdivide the data.

Locational Indexes (LIs). A total of 200 separate villages are included in the full dataset. The treatment_locality variable is used to aggregate the data into the following 2 LIs.

1. Treatment Villages (SA): the SDG intervention causing impacts
2. No Treatment Villages (SB): the counterfactual used to verify impact

For convenience, the actual dataset uses the same data for both the Treatment and No Treatment Villages.

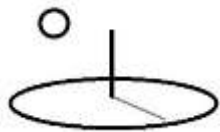
Categorical Indexes (CIs). The poverty index variable is used to aggregate the socioeconomic characteristics of the full population. The Poverty Index was chosen because of its importance in the SDG. As discussed, the actual purpose of the Gertler dataset was to evaluate whether or not the intervention reduced household health care expenditures. Whether these particular Poverty Indexes can really support machine learning algorithms is explored in Example 7.

1. Stakeholder Group 1: Poverty Index range: ≤ 50 : high poverty
2. Stakeholder Group 2: Poverty Index range: 51 to 70: medium poverty
3. Stakeholder Group 3: Poverty Index range: ≥ 71 : low poverty

Indicator 1 Indicators. The “round”, or survey round, variable is used to document the starting and actual, allocated, SDG impacts for the specific stakeholder populations. For this dataset, the target allocation has to be surmised (i.e. from the SDG targets).

Indicator 2 to 15 Indicators and Performance Monitoring. Additional transition states define how the populations’ socioeconomic characteristics and their SDG risks and impacts will change over time. The Resource Stock and M&E calculators support up to 15 Indicators, each of which can define a transition state.

These additional Indicators can be defined using the following 2 techniques.



DevTreks –social budgeting that improves lives and livelihoods

Categorical Index-based Data: Each CI holds updated socioeconomic characteristics for these transition state Indicators. Supports thorough SIAs.

Indicator-based Data: Each CI holds the same socioeconomic characteristics for these transition state Indicators. Supports abbreviated SIAs.

The 2nd TEXT file’s population measurements show that the SDG measurements have been allocated equally to each stakeholder group, even though the number of people in each group varies considerably. In practice, better rules are needed for making the allocations.

With the exception of health care expenditures, Gertler’s socioeconomic characteristics did not change much after the intervention took place. That allows the SDG “Actual” measurements, or “round 1” rows, from the intervention to be realistically included in a 1 Indicator dataset. When the socioeconomic characteristics, such as Poverty Index, change significantly, additional Indicators should be used to document the changes, or transition states. In the latter case, each of the 15 calculator Indicators then serve as short term Performance Monitoring instruments, while the entire group of Indicators, as summarized in the Score, serve as long term, audit-like, Impact Evaluation instruments.

Scores and Impact Evaluation. The Score can be used to hold the same 2 TEXT datasets used by each Indicator, except Score datasets hold data for the full period being evaluated and serve a summary, long term, Impact Evaluation role (i.e. from benchmark to final impact state). Example 1 explained that full Impact Evaluation requires using counterfactual evidence to measure social performance. Example 7 demonstrates these fuller Impact Evaluations.

The aggregation of this type of survey and census data into these hierarchies needs careful thought. For example, if SDG target accomplishment data is needed for each village, Example 7 shows that the LIs can be broken down into the 200 individual villages, and 2 TRs can be used to



DevTreks –social budgeting that improves lives and livelihoods

define the Treatment and No-Treatment villages. Caution must be exercised because the resultant village data could lead to the “misuse of highly personal, private, data” (Espey et al, 2017). In addition, the ultimate destination for this data is likely to be the machine learning algorithms – until those algorithms are fully developed, the best use of this data can’t be determined.

F. Mitigation and Adaptation Actions

The following images (FAO, 2017 and IFC, 2015) demonstrate translating Section E’s risks (or problem areas) into management actions that companies use to mitigate and adapt to risks that reduce “shared value” or that generate negative SDG materiality impacts. One of the main goals of the algorithms in this tutorial is to use the Performance Monitoring and Impact Evaluation techniques to identify sustainable business activities that have been customized to work for specific companies and organizations.



DevTreks –social budgeting that improves lives and livelihoods

WATERSHED MANAGEMENT IN ACTION – LESSONS LEARNED FROM FAO FIELD PROJECTS

TABLE 5. ANALYSIS OF PRIORITY PROBLEMS, THEIR CAUSES AND PROPOSED ACTIVITIES IN THE PAKISTAN PROJECT (FROM THE GULMERA WATERSHED MANAGEMENT PLAN)

Problem area	Causes	Prioritized activities
Crops		
Uncultivated land	Soil erosion Unlevelled land Non-availability of seed Lack of financial resources Lack of irrigation	Field terracing/levelling Retaining walls Introduction of new crops and provision of seed Intercropping Diversion channel
Lack of small-scale vegetable and orchard production	Small landholdings Lack of skills in orchard management Lack of awareness of new cash crops (fruits and vegetables)	Introduction of new cash crops (orchard and vegetables) Skill enhancement training in orchard management and vegetable production Development of market linkages
Low income from crop husbandry	Lack of improved seed Lack of skills in cultural practices Undulated land for cropping	Provision of good quality seed (rainfed) Intercropping Training for skill enhancement Land levelling
Livestock		
Low milk production	Lack of improved breeds Disease attack Non-availability of immediate services Insufficient feed supply Lack of improved skills in animal rearing Lack of financial resources	Regular vaccination and deworming campaign at village level Provision of doorstep veterinary services Introduction of new fodder species Hedge plantation of fodder trees Training of local extension workers
Population decrease	Mortality during earthquake Lack of financial support Local breeds	Provision of good-quality breeds for breeding purposes
Forest		



ABC 3. Management Programs

Risk 3:

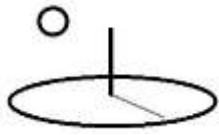
Pesticide/chemical contamination of local drinking water sources from leaching and surface runoffs:

MITIGATION HIERARCHY	ACTIONS	OBJECTIVE AND TARGET	DEADLINE	RESPONSIBLE STAFF	RESOURCES REQUIRED	OPERATIONAL PROCEDURES
Avoid	<ul style="list-style-type: none"> - Substitute pesticides that fall under WHO Hazard Class I or II or listed in Stockholm Convention - Prevent chemical usage including fertilizers, insecticides, fungicides, herbicides/weedicides, growth regulators, etc. by: <ul style="list-style-type: none"> Maintaining and increasing the long-term fertility of the soil by incorporating compost into the soil during the land preparation stage Ensuring proper mulching is done to prevent weeds and conserve moisture Rotating pineapple with other crops (groundnut, beans, rice, vegetables, etc.) Implementing mechanical traps, light and sound to repel pests 	<p>Objective: Prevent/control surface and ground water contamination from agrochemicals</p> <p>Target: 100% Compliance with water quality requirements for selected bore-well and surface water samples</p>	1 year	<ul style="list-style-type: none"> - Operations Manager - Plantation supervisors 	<p>Staff time Intensity: High</p> <p>Capital Intensity: Low</p>	<ul style="list-style-type: none"> - Manual for Integrated Pest Management
Minimize	<ul style="list-style-type: none"> - Use only pesticides whose manufacture is licensed, registered and approved by FAO - Follow manufacturer's instructions and select application technologies/methods that reduce unintentional drift or runoff - Store agrochemicals in dedicated dry, cool and well-ventilated location designed with spill containment measures - Ensure protective clothing worn for pesticide application is carefully disposed - Establish untreated buffer zones along water sources, rivers, streams, ponds, lakes and ditches to protect 		3 months	<ul style="list-style-type: none"> - Operations Manager - Procurement Manager - Plantation Supervisors 	<p>Staff time Intensity: Medium</p> <p>Capital Intensity: Medium</p>	<ul style="list-style-type: none"> - Procedure for selection and procurement of agrochemicals - Procedure for handling, storage and use of agrochemicals - List of banned and restricted substances

G. Quality of Life, or Sustainability, Scenarios

The TFCD (2017) discuss, in detail, the important role that Scenario Analysis plays in helping companies [and local communities] to understand the risks they face from transitioning to higher levels of sustainability. These risks come from highly uncertain, but increasingly likely, SDG target-related materiality impacts, such as the transition risks associated with climate change [and gender equality].

The following image (IFC, 2015) use internal company management system development to show how companies themselves transition between adoption states. Although the IFC case study companies are not small-scale coffee farms, nor small and medium sized enterprises, IFC makes clear that the whole supply chain is accountable for sustainability, including SMEs, producer organizations, cooperatives, regulatory agencies, industry support professionals, input sellers, output buyers, and advisory service companies. Example 3 in SPA2 identified the need to



DevTreks –social budgeting that improves lives and livelihoods

educate consumers about sustainability in ways that result in behavioral changes. Specifically, so that they’ll prefer buying sustainable products and services.

THE SYSTEM MATURITY LEVELS (5 = HIGHEST)	
Level 5	Mature system implemented internally and with key supply chain partners – continual improvement embedded in operations
Level 4	Systems well developed and implemented internally – routine improvement projects
Level 3	Systems approach adopted, but development and implementation is inconsistent – improvement sporadic
Level 2	Limited system development with sporadic implementation – primarily reactive
Level 1	Little systems awareness or repeatable processes
Level 0	No systems awareness or repeatable processes



TFCD (2017) use the following Adoption and Diffusion Path to illustrate how industries can use sustainability reporting systems to accomplish SDG targets, such as those related to climate change [and gender equality].

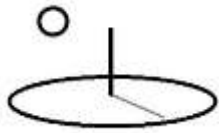
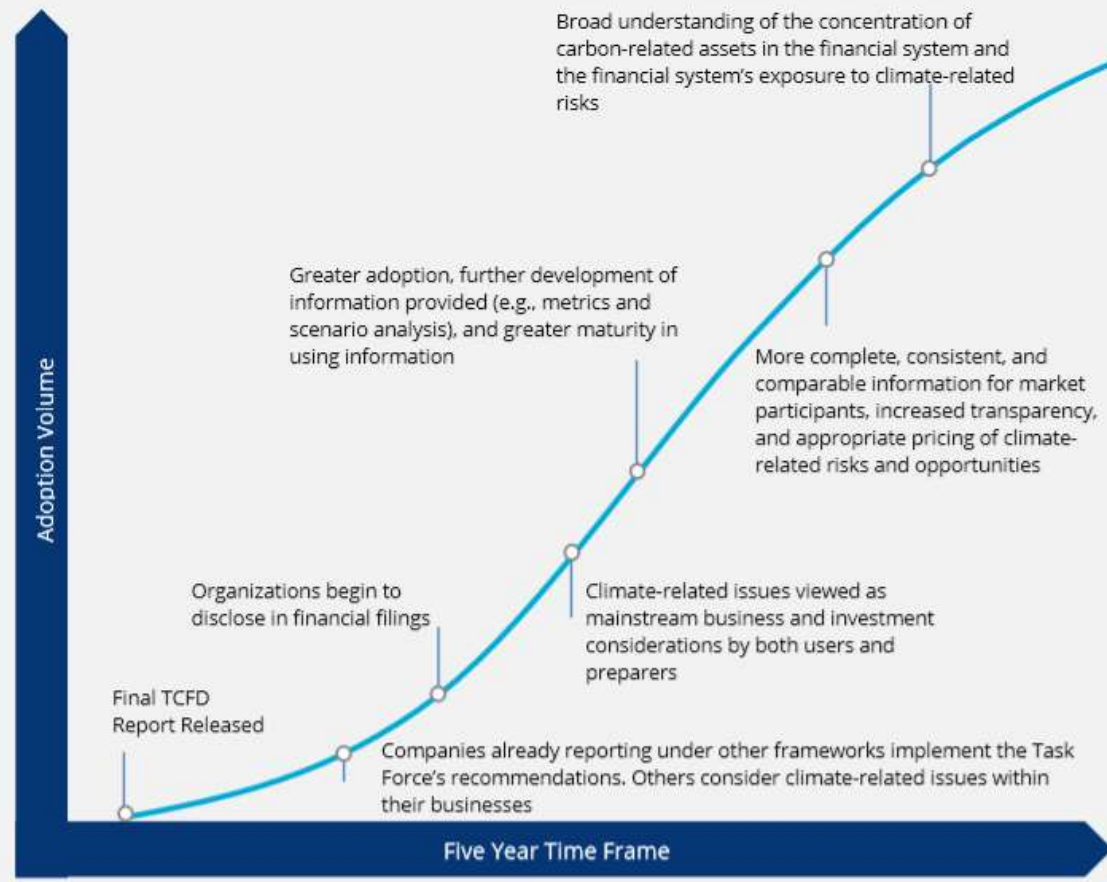


Figure 12

Implementation Path (Illustrative)



The Task Force is not alone in its work. A variety of stakeholders, including stock exchanges, investment consultants, credit rating agencies, and others can provide valuable contributions toward adoption of the recommendations. The Task Force believes that advocacy for these

The following Scenario Analysis is used by companies and local communities to more fully understand the transition risks they face from changes needed in their business activities and social interventions to accomplish the SDG targets.

General Scenario



DevTreks –social budgeting that improves lives and livelihoods

Threatened Quality of Life: High GHG result in 1.5°C temperature increase with higher incidence of biodiversity loss, droughts, severe heat waves, crop and livestock production risks, air pollution, floods, migration, and social discord that leads to the incapacity to achieve the SDG goals (7*).

Targeted Social Performance Risks and Targets: Combinations of the 17 SDG Goals such as climate change and gender equality

Mitigation and Adaptation Actions: Portfolio 1 consists of a) ..., b)...., and c)....

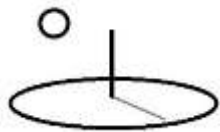
Sustainability Scenarios.

In terms of SDG goal accomplishment, logical sustainability scenarios relate directly to a company’s or local community’s transition to higher levels of sustainability, as illustrated by the following scenarios:

- A. Transition State _A: baseline (i.e. current GHG emissions and lack of full gender equality)
- B. Transition State _B: low sustainability
- C. Transition State _C: medium sustainability
- D. Transition State _D: high sustainability (i.e. fully sustainable GHG emissions and full gender equality)

Simple data conventions, including the use of sibling base elements and Example 1’s “_xx” labelling convention, can be used to model these more comprehensive transition scenarios. Section J, Decisions, demonstrates how to use the Reference Case Cost Effectiveness Results introduced in Example 4B to assist making decisions about the degree of sustainability, or stage of sustainability, appropriate for specific companies and local communities.

Example 7 may demonstrate how machine learning algorithms anticipate transition states and alternative scenarios, given current circumstances, and tailor sustainability recommendations to specific companies and communities, based on their assessment of the company and society’s most probable sustainable development paths.



DevTreks –social budgeting that improves lives and livelihoods

H. Social Performance Score

The following algorithm properties define the socioeconomic characteristics and an “impact transition state” for the stakeholder groups identified in these datasets.

Indicator 1. Benchmark or Baseline

The following images of Indicator 1 displays the “start-target-actual” properties explained more thoroughly in this section that produces the benchmark (i.e. start), target, and actual scores.

These Indicator and Score metadata displayed in this image, and in reports, are calculated as the sum of the TR Indexes for all locations. They are calculated as the sum of the TR’s children normalized and weighted LIs. This dataset uses 2 locations for testing purposes. These scores will be used with the “standard sustainability scorecards” that will be explained in Section I, Decisions.



DevTreks –social budgeting that improves lives and livelihoods

Stakeholder Characteristics Titles. Rows used to describe socioeconomic variables must include a 0 in the 2nd data column, the location, property.

Total Risk and Locational Indexes. The final 11 columns of data for TRs and LIs are not used to change calculations.

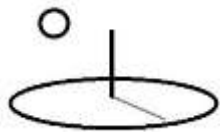
- **factor1 to factor11:** not used – the SDG CI Indicators use normalized data

Categorical Indexes. The following list defines how the final 11 columns of data for Categorical Indexes are used.

- **factor1 to factor11:** socioeconomic characteristic measurements for separate stakeholder groups

SDG-related Indicators. The following list defines the final 11 columns of data that are taken from the CI or LI normalized impacts from subalgos 13, 14, 15, or 16 or custom data. Example 6 demonstrates using the results from the CTAP algorithms.

- **factor1:** Most likely quantity of associated CI.
- **factor2:** Hotspots unit of measurement for most likely quantity CI.
- **factor3:** Low quantity of associated CI.
- **factor4:** High quantity of associated CI.
- **factor5:** Unit of measurement for low and high quantity CI (i.e. upper and lower 80% CI).
- **factor6:** certainty1 of associated CI (i.e. likelihood of CI risk)
- **factor7:** certainty2 of associated CI (i.e. severity of CI risk)
- **factor8:** distribution type (if PRA was used with associated CI)
- **factor9:** Hotspots production process (i.e. farm operation label)
- **factor10:** Hotspots life cycle stage (i.e. production) or Results Chain stage (i.e. outcome)
- **factor11:** date of measurement



Example 3, 3A, and 3B, explain the importance of the Hotspots-related factors. This algorithm is documenting which company, CSO, and local community activities and “stages”, caused the SDG allocations in the 2nd dataset to change. Those activities include mitigation and adaptation activities and work processes that companies adopt to achieve the SDG-related targets. Although this algorithm, by itself, can’t determine cause and effect attribution, when multiple datasets are added to machine learning algorithms, attribution may be found. Example 7 [may] demonstrate how.

2nd TEXT Dataset: Populations and SDG Allocations. This TEXT file is added as the 2nd URL in the Indicator.URL property (using a semicolon delimiter). Additional algorithms may use different properties and Indexes in these datasets to conduct other types of population simulations (i.e. Markov simulations).

Locational Index Measurement. The following properties show that only the last 2 columns of data are used to run calculations. These 2 columns will normalize and sum all of the calculated LIs within each TR Index. This is the same normalization and weighting as used in subalgorithm15.

- **factor1 to factor9:** none
- **factor10:** normalization type
- **factor11:** weight

Categorical Index Population Measurement. These properties support the basic probability density functions demonstrated in Example 2 and throughout the CTA and CTAP tutorials. The data must be put in the exact same CI row index as its related SDG data (i.e. by copying the 1st TEXT and then changing the CIs).

- **factor1:** start date
- **factor2:** end date
- **factor3.** QT population or land use count, or most likely estimate (i.e. QTM)



DevTreks –social budgeting that improves lives and livelihoods

- **factor4:** QT unit
- **factor5:** QTD1 PRA shape (i.e. mean), or lower estimate (i.e. QTL)
- **factor6:** QTD1 unit
- **factor7:** QTD2 PRA scale (i.e. standard deviation), or higher estimate (i.e. QTU)
- **factor8:** QTD2 unit
- **factor9:** PRA distribution type (i.e. normal) ; set to “none” if PRA is not being conducted
- **factor10:** none
- **factor11:** none

Indicator SDG and Population M&E. These properties define how the SDG measurements from the 1st TEXT dataset and the parent CI’s population are expected to change between the CI’s start date and the end date. Indicators that are initially measured using Threshold systems may find it easier to make these allocations based on the 0-100% SAFA Thresholds introduced in Example 1. The data must be put in the exact same Indicator row index as its related SDG dataset.

- **factor1:** start SDG allocation multiplier; a multiplier, expressed as a percentage, for allocating the SDG Indicator measurement to this population; acts as an initial 100% measurement for factor2 and factor3.
- **factor2:** end SDG target multiplier; a multiplier, expressed as a percentage, for defining a target SDG Indicator measurement for the end date;
- **factor3:** end SDG actual multiplier; a multiplier, expressed as a percentage, for defining the actual SDG Indicator measurement at the end date; the target multiplier can be entered as a placeholder in the initial data.
- **factor4:** start Population/land use allocation multiplier; a multiplier, expressed as a percentage, for allocating the CI population measurement to this SDG Indicator; acts as an initial 100% measurement for factor5.
- **factor5:** end Population/land use actual multiplier; a multiplier for defining the actual Population measurement at the end date; an “expected” multiplier can be entered as the initial data.



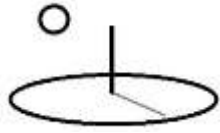
DevTreks –social budgeting that improves lives and livelihoods

- **factor6:** certainty1; severity, and probable consequence, of this SDG risk on this population/land use as of end date
- **factor7:** certainty2; likelihood of this SDG risk on this population/land use as of end date
- **factor8:** certainty3; probable impact of this population/land use on company/CSO/local community from this SDG risk as of end date
- **factor9:** end Cost per Unit SDG Quantity taken from a CEA conducted using the techniques introduced in Examples 4A, 4B, and 4C. Section I, Decisions, show that this property supports the Reference Case Cost Effectiveness Results introduced in those examples.
- **factor10:** normalization type
- **factor11:** weight

Factor 8, certainty3, documents the degree of responsibility by the local community, CSO, or company for the allocated SDG risk. If the risk is caused by internal activities, the degree of responsibility is very high. If the risk is external to company/CSO/local community activities, the degree of responsibility is low, provided the company does not have indirect responsibility for the risk. As explained in Section I, Decisions, objective, evidence-based, 3rd party, standards organizations periodically carry out Impact Evaluations to verify the accuracy of these accounting metrics and to complete their own scorecards.

Indicators 2 to 15. Population Impact Transition States

Exactly the same properties are filled out to define additional population and land use impact transition states. This algorithm supports all 15 Indicators. These “states” can be defined flexibly. For example, they can be used in the same manner as the Results Chains, or impact pathways, used throughout this tutorial. As introduced in Example 1, each Indicator models the elements of the chain (i.e. Indicator 1 = Inputs-> Indicator 2 = Activities-> Indicator 3 = Outputs -> Indicator 4 = Outcomes -> Score = Impacts). Alternatively, instead of using the 1st TEXT dataset’s



factor10 property to define a life cycle stage, that property can be used to define a Results Chain stage.

Business Resource Conservation Value Accounting MathResults. The following table displays the results for the company SIA.

label	locat	risks_and_factor1	factor2	factor3	factor4	factor5	factor6	factor7	factor8	factor9	factor10	factor11	qtmost	percentar	qtlow	qthigh	certainty1	certainty2	certainty3	totalcost	sdgunit	production	lifecyclest
A	0.00	Farmer Ch	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
B	0.00	Farm Char	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
A1A	1.00	Small scal	43101.00	43831.00	99449.37	actual mos	98543.17	lower 80 %	100355.58	upper 80 %	normal	none	0.00	3.39	95.59	1.95	5.73	3.00	4.00	5.00	189.75	0.00	0.00
IF1A	1.00	Non Haras	100.00	125.00	120.00	99450.00	104400.00	3.00	4.00	5.00	1.00	weights	3.00	2.14	96.00	1.07	3.86	3.00	4.00	5.00	90.00	Percent W HCA	production
IF1B	1.00	Air Pollutic	25.00	110.00	105.00	99450.00	104400.00	3.00	4.00	5.00	1.00	weights	3.00	0.50	95.45	0.38	0.94	3.00	4.00	5.00	21.00	daly / kg PI EDE	production
IF1C	1.00	Water Cor	25.00	110.00	105.00	99450.00	104400.00	3.00	4.00	5.00	2.00	weights	3.00	0.38	95.45	0.25	0.47	3.00	4.00	5.00	31.50	daly / m3 EDF	production
IF1D	1.00	Damage Ri	25.00	110.00	105.00	99450.00	104400.00	3.00	4.00	5.00	3.00	weights	3.00	0.38	95.45	0.25	0.47	3.00	4.00	5.00	47.25	daly / extri PCB	production
B1A	1.00	Small scal	43101.00	43831.00	996.33	actual mos	990.29	lower 80 %	1002.37	upper 80 %	normal	none	0.00	0.84	95.45	0.47	1.22	3.00	4.00	5.00	165.38	0.00	0.00
IF2A	1.00	Biodiversit	25.00	110.00	105.00	996.30	1046.00	3.00	4.00	5.00	4.00	weights	3.00	0.28	95.45	0.16	0.41	3.00	4.00	5.00	47.25	amd	EDG
IF2B	1.00	Climate Cf	25.00	110.00	105.00	996.30	1046.00	3.00	4.00	5.00	5.00	weights	3.00	0.28	95.45	0.16	0.41	3.00	4.00	5.00	59.06	gwp 100	ECB
IF2C	1.00	Climate Cf	25.00	110.00	105.00	996.30	1046.00	3.00	4.00	5.00	5.00	weights	3.00	0.28	95.45	0.16	0.41	3.00	4.00	5.00	59.06	gtp 100	ECC
A2A	1.00	Small scal	43101.00	43831.00	99449.37	actual mos	98543.17	lower 80 %	100355.58	upper 80 %	normal	none	0.00	3.39	95.59	1.95	5.73	3.00	4.00	5.00	189.75	0.00	0.00
IF1A	1.00	Non Haras	100.00	125.00	120.00	99450.00	104400.00	3.00	4.00	5.00	1.00	weights	3.00	2.14	96.00	1.07	3.86	3.00	4.00	5.00	90.00	Percent W HCA	production
IF1B	1.00	Air Pollutic	25.00	110.00	105.00	99450.00	104400.00	3.00	4.00	5.00	1.00	weights	3.00	0.50	95.45	0.38	0.94	3.00	4.00	5.00	21.00	daly / kg PI EDE	production
IF1C	1.00	Water Cor	25.00	110.00	105.00	99450.00	104400.00	3.00	4.00	5.00	2.00	weights	3.00	0.38	95.45	0.25	0.47	3.00	4.00	5.00	31.50	daly / m3 EDF	production
IF1D	1.00	Damage Ri	25.00	110.00	105.00	99450.00	104400.00	3.00	4.00	5.00	3.00	weights	3.00	0.38	95.45	0.25	0.47	3.00	4.00	5.00	47.25	daly / extri PCB	production
B2A	1.00	Small scal	43101.00	43831.00	996.33	actual mos	990.29	lower 80 %	1002.37	upper 80 %	normal	none	0.00	0.84	95.45	0.47	1.22	3.00	4.00	5.00	165.38	0.00	0.00
IF2A	1.00	Biodiversit	25.00	110.00	105.00	996.30	1046.00	3.00	4.00	5.00	4.00	weights	3.00	0.28	95.45	0.16	0.41	3.00	4.00	5.00	47.25	amd	EDG
IF2B	1.00	Climate Cf	25.00	110.00	105.00	996.30	1046.00	3.00	4.00	5.00	5.00	weights	3.00	0.28	95.45	0.16	0.41	3.00	4.00	5.00	59.06	gwp 100	ECB
IF2C	1.00	Climate Cf	25.00	110.00	105.00	996.30	1046.00	3.00	4.00	5.00	5.00	weights	3.00	0.28	95.45	0.16	0.41	3.00	4.00	5.00	59.06	gtp 100	ECC
SA	1.00	Small scal	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	8.47	95.52	4.83	13.90	3.00	4.00	5.00	710.25	0.00	0.00
A1B	1.00	Small scal	43101.00	43831.00	99449.37	actual mos	98543.17	lower 80 %	100355.58	upper 80 %	normal	none	0.00	3.39	95.59	1.95	5.73	3.00	4.00	5.00	189.75	0.00	0.00
IF1A	1.00	Non Haras	100.00	125.00	120.00	99450.00	104400.00	3.00	4.00	5.00	1.00	weights	3.00	2.14	96.00	1.07	3.86	3.00	4.00	5.00	90.00	Percent W HCA	production
IF1B	1.00	Air Pollutic	25.00	110.00	105.00	99450.00	104400.00	3.00	4.00	5.00	1.00	weights	3.00	0.50	95.45	0.38	0.94	3.00	4.00	5.00	21.00	daly / kg PI EDE	production
IF1C	1.00	Water Cor	25.00	110.00	105.00	99450.00	104400.00	3.00	4.00	5.00	2.00	weights	3.00	0.38	95.45	0.25	0.47	3.00	4.00	5.00	31.50	daly / m3 EDF	production
IF1D	1.00	Damage Ri	25.00	110.00	105.00	99450.00	104400.00	3.00	4.00	5.00	3.00	weights	3.00	0.38	95.45	0.25	0.47	3.00	4.00	5.00	47.25	daly / extri PCB	production
B2B	1.00	Small scal	43101.00	43831.00	996.33	actual mos	990.29	lower 80 %	1002.37	upper 80 %	normal	none	0.00	0.84	95.45	0.47	1.22	3.00	4.00	5.00	165.38	0.00	0.00
IF2A	1.00	Biodiversit	25.00	110.00	105.00	996.30	1046.00	3.00	4.00	5.00	4.00	weights	3.00	0.28	95.45	0.16	0.41	3.00	4.00	5.00	47.25	amd	EDG
IF2B	1.00	Climate Cf	25.00	110.00	105.00	996.30	1046.00	3.00	4.00	5.00	5.00	weights	3.00	0.28	95.45	0.16	0.41	3.00	4.00	5.00	59.06	gwp 100	ECB
IF2C	1.00	Climate Cf	25.00	110.00	105.00	996.30	1046.00	3.00	4.00	5.00	5.00	weights	3.00	0.28	95.45	0.16	0.41	3.00	4.00	5.00	59.06	gtp 100	ECC
SB	1.00	Small scal	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	8.47	95.52	4.83	13.90	3.00	4.00	5.00	710.25	0.00	0.00
A1C	1.00	Small scal	43101.00	43831.00	99449.37	actual mos	98543.17	lower 80 %	100355.58	upper 80 %	normal	none	0.00	3.39	95.59	1.95	5.73	3.00	4.00	5.00	189.75	0.00	0.00
IF1A	1.00	Non Haras	100.00	125.00	120.00	99450.00	104400.00	3.00	4.00	5.00	1.00	weights	3.00	2.14	96.00	1.07	3.86	3.00	4.00	5.00	90.00	Percent W HCA	production
IF1B	1.00	Air Pollutic	25.00	110.00	105.00	99450.00	104400.00	3.00	4.00	5.00	1.00	weights	3.00	0.50	95.45	0.38	0.94	3.00	4.00	5.00	21.00	daly / kg PI EDE	production
IF1C	1.00	Water Cor	25.00	110.00	105.00	99450.00	104400.00	3.00	4.00	5.00	2.00	weights	3.00	0.38	95.45	0.25	0.47	3.00	4.00	5.00	31.50	daly / m3 EDF	production
IF1D	1.00	Damage Ri	25.00	110.00	105.00	99450.00	104400.00	3.00	4.00	5.00	3.00	weights	3.00	0.38	95.45	0.25	0.47	3.00	4.00	5.00	47.25	daly / extri PCB	production
B2C	1.00	Small scal	43101.00	43831.00	996.33	actual mos	990.29	lower 80 %	1002.37	upper 80 %	normal	none	0.00	0.84	95.45	0.47	1.22	3.00	4.00	5.00	165.38	0.00	0.00
IF2A	1.00	Biodiversit	25.00	110.00	105.00	996.30	1046.00	3.00	4.00	5.00	4.00	weights	3.00	0.28	95.45	0.16	0.41	3.00	4.00	5.00	47.25	amd	EDG
IF2B	1.00	Climate Cf	25.00	110.00	105.00	996.30	1046.00	3.00	4.00	5.00	5.00	weights	3.00	0.28	95.45	0.16	0.41	3.00	4.00	5.00	59.06	gwp 100	ECB
IF2C	1.00	Climate Cf	25.00	110.00	105.00	996.30	1046.00	3.00	4.00	5.00	5.00	weights	3.00	0.28	95.45	0.16	0.41	3.00	4.00	5.00	59.06	gtp 100	ECC
SC	1.00	Small scal	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	8.47	95.52	4.83	13.90	3.00	4.00	5.00	710.25	0.00	0.00
TR	1.00	All small sc	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.82	95.52	1.61	4.63	3.00	4.00	5.00	2130.75	0.00	0.00
A1A	2.00	Small scal	43101.00	43831.00	99449.37	actual mos	98543.17	lower 80 %	100355.58	upper 80 %	normal	none	0.00	3.39	95.59	1.95	5.73	3.00	4.00	5.00	189.75	0.00	0.00
IF1A	2.00	Non Haras	100.00	125.00	120.00	99450.00	104400.00	3.00	4.00	5.00	1.00	weights	3.00	2.14	96.00	1.07	3.86	3.00	4.00	5.00	90.00	Percent W HCA	production

Indicator Math. The following 8 columns of Indicator data are added in the MathResult. These measurements are summed, normalized, and weighted, into their parent CI, LI, and TR Indexes. Existing Properties (updated population allocations)

- **Population Start Count** = (popStartCount * ((popStartAllocation / 100)
- **Population End Count** = (popStartCount * ((popStartAllocation / 100) * (popEndAllocation / 100))



DevTreks –social budgeting that improves lives and livelihoods

New Properties (added as new data columns)

- **qtmmost**: most likely SDG quantity per total population (not per population member because the calculations are difficult to interpret). The mathematical formula used in this calculation is:

$$\text{sdgPerPopulation} = (\text{sdgQuantity} * (\text{sdgStartAllocation} / 100) * (\text{sdgEndActualAllocation} / 100))$$

- **percenttarget**: $(\text{qtmmost} / (\text{sdgQuantity} * (\text{sdgStartAllocation} / 100) * (\text{sdgEndTargetAllocation} / 100))) * 100$
- **qtlow**: low estimate of likely SDG quantity per total population
- **qthigh**: high estimate of likely SDG quantity per total population
- **certainty1**: factor6 averages
- **certainty2**: factor7 averages
- **certainty3**: factor8 average
- **totalcost**: total cost of most likely SDG quantity (qtmmost) (only found in the local community dataset)
 $\text{totalSDGCost} = \text{factor9} (\text{Cost per Unit SDG}) * \text{qtmmost}$ (most likely quantity of SDG prior to normalization and weighting)

The final 3 columns display the following properties from the 1st TEXT dataset.

- **sdgunit (Elementary Flow)**: Hotspots unit of measurement from the 1st TEXT SDG Indicator.
- **productionprocess**: Hotspots production process from the 1st TEXT SDG Indicator (i.e. in terms of results chains, company Operation, Component, or Outcome)
- **lifecyclestage**: Hotspots life cycle stage (i.e. if using Results Chains, can be Inputs, Operations, Components, Outputs, Outcomes, or Impacts)

Categorical Index Math. The first 11 columns display the results of the PRA population calculations. The final 7 columns are calculated by normalizing and weighting all of the



DevTreks –social budgeting that improves lives and livelihoods

Indicators in each parent Locational Index and then summing the normalized and weighted Indicators into their respective CIs, as displayed in the following list.

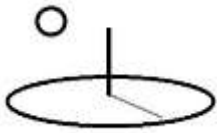
- **qtmmost:** sum of CIs
- **percenttarget:** average of CIs
- **qtlow:** sum of CIs
- **qthigh:** sum of CIs
- **certainty1:** factor6 averages
- **certainty2:** factor7 averages
- **certainty3:** factor8 averages
- **totalcost:** sum of total costs

Locational Index Math. Locational Indexes are calculated as the sum of the normalized and weighted Indicators contained in all of the LI’s children CIs.

TR Index Math. The TR Indexes are calculated as the sum of the normalized and weighted children LIs, and display the same properties as the LIs. The LIs can be normalized and weighted separately from their children Indicators. Section I, Decisions, demonstrates displaying these Scores in uniform, “SDG-ASB Scorecards”.

Indicator Metadata Math. The displayed Indicator properties (i.e. Q1 to Q5, QTMost, QTLow, QTHigh) are summations of all of the final TR Index rows of data, across all locations.

Local Community and CSO Resource Conservation Value Accounting MathResults. The following image displays the results for the local community’s SIA. For testing purposes, the datasets were added to Indicator 15.



QT 15 6.9473	QT Unit 15 target high score
Math Type 15 algorithm1	Math Sub Type 15 subalgorithm17
QT D1 15 3.0000	QT D1 Unit 15 actual certainty1
QT D2 15 4.0000	QT D2 Unit 15 actual certainty2
QT Most 15 4.2366	QT Most Unit 15 actual most score
QT Low 15 2.4154	QT Low Unit 15 actual low score
QT High 15 6.9508	QT High Unit 15 actual high score
Math Expression 15 I15.Q1.factor1 + I15.Q2.factor2 + I15.Q3.factor3 + I15.Q4.fact	
Math Result 15 rca results label,location,risks_and_impacts,factor1,factor2,factor3,factor4,factor5,factor6,factor7,factor8,factor9,factor10,factor11,qtmost,percenttarget,qtlow,qthigh,certainty1,certainty2,certainty3,totalcost,sdgunit,productionprocess,lifecyclestage A,0,HH Characteristics	

Score. The following Score properties show that the Score employs the same 2 TEXT datasets as Indicators, but support longer term Impact Evaluations. Although Scores only partially display the “start-target-actual” properties returned in Indicators, the raw MathResults contain all of that data for reporting. For testing purposes, Indicator 1’s datasets were reused in this Score. Resource Stock and M&E Analyzers, as explained in their respective tutorials, can be used to further analyze all of the “actual” scores for both the Score and the Indicators (i.e. Totals, Statistics, Change Bys, and Progress, Analyzers).



DevTreks –social budgeting that improves lives and livelihoods

“A key point underpinning [...] this [reference] is that rather than seeing SIA as being a cost to business, SIA should be seen as an appropriate, useful management process that reduces risk and brings benefits to companies and to communities – in other words, that operationalises the concept of shared value. There is thus a rock-solid, strong business case for [companies and organizations] to do effective social impact assessment and management.”

The following image (Vanclay et al, 2015) illustrates how to use this algorithm’s SDG Indicator certainty properties, factor6 and factor7, to prioritize and manage business risks and shared value. Those factors derive largely on the 2 certainty factors included in the 1st TEXT dataset. The third certainty factor, factor8, measures the degree of responsibility for the SDG risk by the company, CSO, or local community and can be reported on a 3rd plane, such as a 3D Column graph.

		Consequence Level				
		1	2	3	4	5
Likelihood Level	Descriptor	Insignificant	Minor	Moderate	Major	Catastrophic
A	Almost certain	A1	A2	A3	A4	A5
B	Likely	B1	B2	B3	B4	B5
C	Possible	C1	C2	C3	C4	C5
D	Unlikely	D1	D2	D3	D4	D5
E	Rare	E1	E2	E3	E4	E5

Figure 11: Risk assessment framework.

Risk Rating ■ Low ■ Moderate ■ High ■ Extreme

The following image (IFC, 2013) summarizes a concrete example of assessing SIA risks in agricultural supply chains. DevTreks hierarchical base elements, including Data Services, support the IFC recommendation to use “supplier databases” to document these types of risk ratings. The aggregation of SDG-related risks is especially important for watershed level planning, or for applying “integrated and cross sectoral approaches” to regional resource conservation planning.



DevTreks –social budgeting that improves lives and livelihoods

HOW TO AGGREGATE RISK RATINGS

To achieve a single risk rating associated with each commodity and country combination, the risks can be aggregated across E&S issues. Weighting of indicators can be used, if their importance differs.

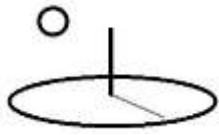
To facilitate decision making on appropriate risk management measures, risk factors can be combined in a matrix to identify an overall risk rating for each supply chain based on the risk ratings of the country of origin and commodity (see Figure 8). It is important to ensure that the outputs (risk levels assigned to agro-commodity supply chains) lead to clearly defined actions.

FIGURE 8. RISK RATING MATRIX

	Child Labor and Forced Labor	Significant Safety Issues Related to Supply Chain Workers	Significant Conversion of Critical and/or Natural Habitat	Aggregate Risk
Commodity/Country A	High risk	High risk	High risk	High risk
Commodity/Country B	High risk	High risk	Low risk	High Risk
Commodity/Country C	High risk	Medium risk	Low risk	Medium Risk
Commodity/Country D	Low risk	Low risk	Low risk	Low risk

The approach described here is typically contained within a stand-alone spreadsheet or database, which is used as a reference tool. As mentioned earlier, to manage E&S risk in supply chains, it is useful to maintain a supplier database. The risk assessment matrix described above can be used to populate the supplier database with risk ratings.

Full SDG Monitoring and Evaluation Report: The following table illustrates how to use this algorithm’s primary decision support variables, including the 3 risk factors, to develop a full M&E Report for decision making purposes. In this report, Indicators 1 to 15 serve as Performance Monitoring instruments (i.e. Indicator 1 = 2018, 2028 = Indicator 6), and the Score acts as a summary Impact Evaluation of total SDG accomplishment. In practice, this image’s SDG Indexes consist of combinations of the 17 SDG Goals (i.e. poverty, climate change, and



DevTreks –social budgeting that improves lives and livelihoods

gender equality). RCA Assessors work at industry, company, and community scale to develop combinations of the SDG Targets relevant to impacted stakeholders, along with recommended targets and completion percentages for each Monitoring period.

Full SDG Monitoring and Evaluation Report								
indicator	1	2	3	4	5	6	7	Score
period	2017-2018	2019-2020	2021-2022	2023-2024	2025-2026	2027-2028	2029-2030	2017-2030
Non Discrimination and Non Harassment Index								
sdg target	x	x	x	x	x	x	x	x
most likely sdg actual	x	x	x	x	x	x	x	x
sdg percent	x	x	x	x	x	x	x	x
low sdg actual	x	x	x	x	x	x	x	x
high sdg actual	x	x	x	x	x	x	x	x
certainty 1	x	x	x	x	x	x	x	x
certainty 2	x	x	x	x	x	x	x	x
certainty 3	x	x	x	x	x	x	x	x
totalcost	x	x	x	x	x	x	x	x
Water Consumption Index								
sdg target	x	x	x	x	x	x	x	x
most likely sdg actual	x	x	x	x	x	x	x	x
sdg percent	x	x	x	x	x	x	x	x
low sdg actual	x	x	x	x	x	x	x	x
high sdg actual	x	x	x	x	x	x	x	x
certainty 1	x	x	x	x	x	x	x	x
certainty 2	x	x	x	x	x	x	x	x
certainty 3	x	x	x	x	x	x	x	x
totalcost	x	x	x	x	x	x	x	x
Biodiversity Index								
sdg target	x	x	x	x	x	x	x	x

Reference Case Cost Effectiveness Results Report. Example 4B used a Reference Case Cost Effectiveness Results to illustrate analyzing the cost effectiveness of alternative sustainability scenarios. That example also completed separate “value perspectives” that supported tradeoff analysis for the value systems held by industry, society, and community, stakeholders. These perspectives support equitable decisions and reduce conflict among impacted stakeholders. The following image illustrates using this algorithm’s decision support data to complete similar Reference Cases, or basic value assessments, for impacted stakeholders (i.e. industry, community, society). The following Reference Case Cost Effectiveness Results is comparing the incremental costs and effectiveness of transitioning to higher states of sustainability, or “impact transition states”, for each value perspective.



Reference Case Cost Effectiveness Results (Time Horizon = 22 years; Costs and SDGs Discount Rate = 3%; Escalation Type = none; CEA Threshold = 100,000; Confidence										
SDG Impacts	Most Likely Total Costs	Uncertainty Index (likelihood)	Most Likely SDGs	Uncertainty Index (consequence)	Uncertainty Index (responsibility)	Incremental Costs	Incremental Effectiveness	NMB	Incremental NMB	ICER (Inc Cost / Inc Effects)
Stakeholder Group A Perspective										
Indicator 1. 2017 to 2018	x	x	x	x	x	x	x	x	x	x
Indicator 2. 2019 to 2020	x	x	x	x	x	x	x	x	x	x
Indicator 3. 2021 to 2022	x	x	x	x	x	x	x	x	x	x
Indicator 4. 2023 to 2024	x	x	x	x	x	x	x	x	x	x
Indicator 5. 2025 to 2026	x	x	x	x	x	x	x	x	x	x
Indicator 6. 2027 to 2028	x	x	x	x	x	x	x	x	x	x
Indicator 7. 2029 to 2030	x	x	x	x	x	x	x	x	x	x
Benchmark Indicator 1 to 2030										
Score. 2017 to 2030	x	x	x	x	x	x	x	x	x	x
Private Sector Perspective										
Indicator 1. 2017 to 2018	21,000.00	4.00	16.50	4.50	3.00			1,629,000.00		
Indicator 2. 2019 to 2020	23,100.00	3.50	18.98	2.50	3.00	21,000.00	18.15	1,874,400.00	245,400.00	1,157.02
Indicator 3. 2021 to 2022	25,200.00	3.00	21.45	2.00	5.00	42,000.00	34.65	2,119,800.00	245,400.00	1,212.12
Indicator 4. 2023 to 2024	26,250.00	4.25	21.45	4.50	1.00	63,000.00	51.15	2,118,750.00	(1,050.00)	1,231.67
Indicator 5. 2025 to 2026	x	x	x	x	x	x	x	x	x	x
Indicator 6. 2027 to 2028	x	x	x	x	x	x	x	x	x	x
Indicator 7. 2029 to 2030	x	x	x	x	x	x	x	x	x	x
Benchmark Indicator 1 to 2030										
Score. 2017 to 2030	x	x	x	x	x	x	x	x	x	x
Societal Perspective										
Indicator 1. 2017 to 2018	x	x	x	x	x	x	x	x	x	x
Indicator 2. 2019 to 2020	x	x	x	x	x	x	x	x	x	x
Indicator 3. 2021 to 2022	x	x	x	x	x	x	x	x	x	x
Indicator 4. 2023 to 2024	x	x	x	x	x	x	x	x	x	x
Indicator 5. 2025 to 2026	x	x	x	x	x	x	x	x	x	x
Indicator 6. 2027 to 2028	x	x	x	x	x	x	x	x	x	x
Indicator 7. 2029 to 2030	x	x	x	x	x	x	x	x	x	x
Benchmark Indicator 1 to 2030										
Score. 2017 to 2030	x	x	x	x	x	x	x	x	x	x

Hotspots Analysis Report. Example 3 demonstrated how Hotspots Analysis can be used to “identify and prioritise potential actions around the most significant economic, environmental and social sustainability impacts or benefits associated with a specific country, city, industry sector, organization, product portfolio, product category or individual product or service”. The following table comes directly from this algorithm’s MathResults. The 3 Hotspots elements in the last 3 columns (i.e. SDG Unit, Production Process, and Life Cycle Stage) readily support ranking mitigation and adaptation actions and sustainability stages, or Production Processes and Life Cycle stages, according to their degree of SDG Target Accomplishment and “fairness” for each stakeholder group.



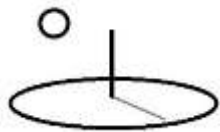
DevTreks –social budgeting that improves lives and livelihoods

label	risks_and_impacts	qtmst	percentar	qtlow	qthigh	certainty1	certainty2	certainty3	totalcost	sdgunit	productior	lifecyclest
A	HH Characteristics 1	0	0	0	0	0	0	0	0	0	0	0
A1A	High Poverty Index HHs	2.82E+00	95.5325	1.61E+00	4.63E+00	3	4	5	425.3438	0	0	0
IF1A	Non Harassment and Discrimina	1.43E+00	96	7.14E-01	2.57E+00	3	4	5	90	Percent W	HCA	productior
IF1B	Air Pollution Index	3.33E-01	95.4545	2.50E-01	6.25E-01	3	4	5	42	daly / kg P	EDE	productior
IF1C	Water Consumption Index	2.50E-01	95.4545	1.67E-01	3.13E-01	3	4	5	47.25	daly / m3	EDF	productior
IF1D	Damage Risk Reduction Index	2.50E-01	95.4545	1.67E-01	3.13E-01	3	4	5	63	daly / extr	PCB	productior
IF2A	Biodiversity Index	1.88E-01	95.4545	1.04E-01	2.71E-01	3	4	5	59.0625	amd	EDG	productior
IF2B	Climate Change Short Term Indi	1.88E-01	95.4545	1.04E-01	2.71E-01	3	4	5	5.9063	gwp 100	ECB	productior
IF2C	Climate Change Long Term Inde	1.88E-01	95.4545	1.04E-01	2.71E-01	3	4	5	118.125	gtp 100	ECC	productior
A1B	Medium Poverty Index HHs	2.82E+00	95.5325	1.61E+00	4.63E+00	3	4	5	425.3438	0	0	0
IF1A	Non Harassment and Discrimina	1.43E+00	96	7.14E-01	2.57E+00	3	4	5	90	Percent W	HCA	productior
IF1B	Air Pollution Index	3.33E-01	95.4545	2.50E-01	6.25E-01	3	4	5	42	daly / kg P	EDE	productior
IF1C	Water Consumption Index	2.50E-01	95.4545	1.67E-01	3.13E-01	3	4	5	47.25	daly / m3	EDF	productior
IF1D	Damage Risk Reduction Index	2.50E-01	95.4545	1.67E-01	3.13E-01	3	4	5	63	daly / extr	PCB	productior
IF2A	Biodiversity Index	1.88E-01	95.4545	1.04E-01	2.71E-01	3	4	5	59.0625	amd	EDG	productior
IF2B	Climate Change Short Term Indi	1.88E-01	95.4545	1.04E-01	2.71E-01	3	4	5	5.9063	gwp 100	ECB	productior
IF2C	Climate Change Long Term Inde	1.88E-01	95.4545	1.04E-01	2.71E-01	3	4	5	118.125	gtp 100	ECC	productior
A1C	Low Poverty Index HHs	2.82E+00	95.5325	1.61E+00	4.63E+00	3	4	5	425.3438	0	0	0
IF1A	Non Harassment and Discrimina	1.43E+00	96	7.14E-01	2.57E+00	3	4	5	90	Percent W	HCA	productior
IF1B	Air Pollution Index	3.33E-01	95.4545	2.50E-01	6.25E-01	3	4	5	42	daly / kg P	EDE	productior
IF1C	Water Consumption Index	2.50E-01	95.4545	1.67E-01	3.13E-01	3	4	5	47.25	daly / m3	EDF	productior
IF1D	Damage Risk Reduction Index	2.50E-01	95.4545	1.67E-01	3.13E-01	3	4	5	63	daly / extr	PCB	productior
IF2A	Biodiversity Index	1.88E-01	95.4545	1.04E-01	2.71E-01	3	4	5	59.0625	amd	EDG	productior
IF2B	Climate Change Short Term Indi	1.88E-01	95.4545	1.04E-01	2.71E-01	3	4	5	5.9063	gwp 100	ECB	productior
IF2C	Climate Change Long Term Inde	1.88E-01	95.4545	1.04E-01	2.71E-01	3	4	5	118.125	gtp 100	ECC	productior
SA	Treated Villages	8.47E+00	95.5325	4.83E+00	1.39E+01	3	4	5	1276.031	0	0	0

The previous 3 images imply that fully sustainable impact transition states, or technology and diffusion paths, can be defined with a level of accuracy needed to establish these initial targets. In practice, that’s unlikely. RCA Assessors can establish the initial targets, but the most appropriate technology for fully accomplishing the targets needs experimentation. In general, with the exception of “low hanging fruit”, or proven sustainable business activities, the business practices causing the SDG scores have to be discovered and then customized for each business and community via algorithms like those introduced in this reference. The purpose of this SDG M&E accounting system is to provide the feedback needed by managers and advisors to identify these practices and then adjust their activities to achieve the necessary targets.

SDG Accounting Standards Board Scorecards (SDG-ASB) (10*)

The institutional improvements recommended throughout this tutorial include the independent, evidence-based, 3rd party, verification of the content in these reports. SPA2 demonstrated the role of standards setting groups, such as the ISEAL Alliance, in setting standards for CSOs who administer commodity and producer organization standards for their members. TFCFD (2017) explains the importance of financial accounting standards, such as those enforced by the



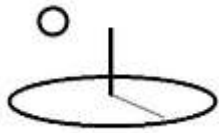
DevTreks –social budgeting that improves lives and livelihoods

International Accounting Standards Board, in ensuring the integrity of private sector financial reporting. These reports require similar enforcement.

CSO certification standards setters (EMAS, SASB, GSSB, B-Lab, FairTrade, Rainforest Alliance, ...), make logical enforcers for their members. In addition, many local communities are adopting a variety of accounting techniques to measure sustainability. For example, many U.S. states now require standard energy efficiency reports for all home sales. These types of reports are completed by 3rd party professionals (i.e. sustainability officers or RCA Assessors) who specialize in sustainability reporting. Local community sustainability offices make logical enforcers for their stakeholders.

The importance and international reality of the SDG, justifies a separate “SDG Accounting Standards Board” (SDG-ASB) to verify and “harmonize” the disparate standard setting CSOs and their reporting requirements. The UN references suggest such an organization may already exist (i.e. refer to the United Nations Forum on Sustainability Standards (UNFSS) reference in SPA1). The following images illustrate several ways to use this algorithm’s metrics to develop uniform Scorecards used by the SDG-ADB, standards setting CSOs, and local community sustainability offices, to evaluate sustainability target accomplishment. One advantage to these uniform reports is that they eliminate the need for hundreds of disparate reports –all of which ultimately have the same purpose. Importantly, their uniform “data standards” support more advanced algorithms, such as machine learning algorithms, which can facilitate SDG Target accomplishment at industry and landscape scale.

1. **Scored Full M&E SDG Report.** The following image shows that Scores have been added to the full M&E SDG Report generated from this algorithm’s results.

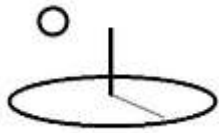


DevTreks –social budgeting that improves lives and livelihoods

SDG-ASB Scorecard								
indicator	1	2	3	4	5	6	7	Score
period	2017-2018	2019-2020	2021-2022	2023-2024	2025-2026	2027-2028	2029-2030	2017-2030
Non Discrimination and Non Harassment Index								
sdg target								
most likely sdg actual								
sdg percent								
low sdg actual								
high sdg actual								
certainty 1								
certainty 2								
certainty 3								
totalcost								
Score								xx
Water Consumption Index								
sdg target								
most likely sdg actual								
sdg percent								
low sdg actual								
high sdg actual								
certainty 1								
certainty 2								
certainty 3								
totalcost								
Score								xx
Biodiversity Index								
sdg target								
...								
Total SDG Score								xx

Scored 1 Page M&E SDG Report. The following image demonstrates using the Full SDG M&E Report directly as a 1 page Scorecard. This instrument recognizes the nature and size of many SMEs do not justify full sustainability accounting, such as the use of sustainability algorithms. These annual, 1 page, “running record”, sustainability reports, are completed by 3rd party sustainability officers, many of whom work for CSOs supported through local, fee-based, sustainability programs.

The use of indexes, such as this image’s SDG Target Indexes, for sustainability reporting is also discussed throughout this tutorial. For example, Example 2 demonstrates how to generate the Indexes using simple MCDA techniques that can employ Most Likely-Low-High Estimates for the Index’s Indicators.



DevTreks –social budgeting that improves lives and livelihoods

SDG-ASB Scorecard									
indicator	1	2	3	4	5	6	7	Score	
period	2017-2018	2019-2020	2021-2022	2023-2024	2025-2026	2027-2028	2029-2030	2017-2030	
SDG Target 4. Gender Equality									
sdg target									
most likely sdg actual									
sdg percent									
low sdg actual									
high sdg actual									
certainty 1									
certainty 2									
certainty 3									
totalcost									
Score									xx
SDG Target 13. Climate Change									
sdg target									
most likely sdg actual									
sdg percent									
low sdg actual									
high sdg actual									
certainty 1									
certainty 2									
certainty 3									
totalcost									
Score									xx
SDG Target 6. Clean Water and Sanitation									
sdg target									
...									
Total SDG Score									xx

2. **Scored Stakeholder Reference Case CEA.** The following image shows that Scores have been added to the full Reference Case Cost Effectiveness Results displayed above. Equity is the primary concern of the SDG-ADB. They use the company, CSO, and local community rating matrixes to independently evaluate stakeholder equity, which can be assessed from these types of reports.

Appendix 1 to Example 4B discusses the important role that QALYs, DALYs, and QASYs can play in these scorecards (i.e. the 4th column holds QASYs rather than SDGs). The overall goal of any sustainability system is to increase human quality of life. Measuring all SDG accomplishment according to that yardstick may be the best metric for any of these accounting systems because they derive from stakeholder preferences.



Example 4B confirms that the health care sector in many countries is already using that yardstick in these CEA instruments.

Reference Case Cost Effectiveness Results (Time Horizon = 22 years; Costs and SDGs Discount Rate = 3%; Escalation Type = none; CEA Threshold = 100,000; Confidence										
SDG Impacts	Most Likely Total Costs	Uncertainty Index (likelihood)	Most Likely SDGs	Uncertainty Index (consequence)	Uncertainty Index (responsibility)	Incremental Costs	Incremental Effectiveness	NMB	Incremental NMB	ICER (Inc Cost / Inc Effects)
Stakeholder Group A Perspective										
Indicator 1. 2017 to 2018	x	x	x	x	x	x	x	x	x	x
Indicator 2. 2019 to 2020	x	x	x	x	x	x	x	x	x	x
Indicator 3. 2021 to 2022	x	x	x	x	x	x	x	x	x	x
Indicator 4. 2023 to 2024	x	x	x	x	x	x	x	x	x	x
Indicator 5. 2025 to 2026	x	x	x	x	x	x	x	x	x	x
Indicator 6. 2027 to 2028	x	x	x	x	x	x	x	x	x	x
Indicator 7. 2029 to 2030	x	x	x	x	x	x	x	x	x	x
Benchmark Indicator 1 to 2030										
Score. 2017 to 2030	x	x	x	x	x	x	x	x	x	x
Private Sector Perspective										
Indicator 1. 2017 to 2018	21,000.00	4.00	16.50	4.50	3.00			1,629,000.00		
Indicator 2. 2019 to 2020	23,100.00	3.50	18.98	2.50	3.00	21,000.00	18.15	1,874,400.00	245,400.00	1,157.02
Indicator 3. 2021 to 2022	25,200.00	3.00	21.45	2.00	5.00	42,000.00	34.65	2,119,800.00	245,400.00	1,212.12
Indicator 4. 2023 to 2024	26,250.00	4.25	21.45	4.50	1.00	63,000.00	51.15	2,118,750.00	(1,050.00)	1,231.67
Indicator 5. 2025 to 2026	x	x	x	x	x	x	x	x	x	x
Indicator 6. 2027 to 2028	x	x	x	x	x	x	x	x	x	x
Indicator 7. 2029 to 2030	x	x	x	x	x	x	x	x	x	x
Benchmark Indicator 1 to 2030										
Score. 2017 to 2030	x	x	x	x	x	x	x	x	x	x
Societal Perspective										
Indicator 1. 2017 to 2018	x	x	x	x	x	x	x	x	x	x
Indicator 2. 2019 to 2020	x	x	x	x	x	x	x	x	x	x
Indicator 3. 2021 to 2022	x	x	x	x	x	x	x	x	x	x
Indicator 4. 2023 to 2024	x	x	x	x	x	x	x	x	x	x
Indicator 5. 2025 to 2026	x	x	x	x	x	x	x	x	x	x
Indicator 6. 2027 to 2028	x	x	x	x	x	x	x	x	x	x
Indicator 7. 2029 to 2030	x	x	x	x	x	x	x	x	x	x
Benchmark Indicator 1 to 2030										
Score. 2017 to 2030	x	x	x	x	x	x	x	x	x	x

3. **Scored Scenario Analysis.** The following image from SPA2 shows that Scores have been added to the full Reference Case Cost Effectiveness Results displayed for alternative scenarios.



DevTreks –social budgeting that improves lives and livelihoods

Reference Case Cost Effectiveness Results (Time Horizon = 22 years; Costs and QASYS Discount Rate = 3%; Escalation Type = none; CEA Threshold =										
Alternative	Most Likely Total Costs	Uncertainty Index (likelihood)	Most Likely SDGs	Uncertainty Index (consequence)	Uncertainty Index (Agent Responsibility)	Incremental Costs	Incremental Effectiveness	NMB	Incremental NMB	ICER (Inc Cost / Inc Effects)
Stakeholder Group A Perspective										
No Mitigation and Adapt.										
AA	to be completed when data becomes available									
AB										
AC										
Private Sector Perspective										
No Mitigation	21,000.00	4.00	16.50	4.50	3.00			1,629,000.00		
AA	23,100.00	3.50	18.98	2.50	3.00	21,000.00	18.15	1,874,400.00	245,400.00	1,157.02
AB	25,200.00	3.00	21.45	2.00	5.00	42,000.00	34.65	2,119,800.00	245,400.00	1,212.12
AC	26,250.00	4.25	21.45	4.50	1.00	63,000.00	51.15	2,118,750.00	(1,050.00)	1,231.67
Societal Perspective										
No Mitigation and Adapt.										
AA	to be completed when data becomes available									
AB										
AC										

4. **Scored Hotspots Analysis.** The following image shows that Scores have been added to rank the effectiveness of coping and resilience work processes and population epidemiology stages in achieving SDG Targets. Example 6 explains that coping capacity and resiliency work processes, and population transition states, make appropriate replacements for Production Processes and Life Cycle Stages in many types of Hotspots-related sustainability accounting.

SDG Hotspots Report	SDG Targets					Score
	No Poverty	Good Health and Well Being	Gender Equality	Affordable and Clean Energy	Climate Change	
Coping and Resilience Process 1	%	%	%	%	%	%
Coping and Resilience Process 2	%	%	%	%	%	%
Coping and Resilience Process 3	%	%	%	%	%	%
Coping and Resilience Process 4	%	%	%	%	%	%
Coping and Resilience Process 5	%	%	%	%	%	%
Population Epidemiology Stage 1 Total	%	%	%	%	%	%
Coping and Resilience Process 1	%	%	%	%	%	%
Coping and Resilience Process 2	%	%	%	%	%	%
Coping and Resilience Process 3	%	%	%	%	%	%
Coping and Resilience Process 4	%	%	%	%	%	%
Coping and Resilience Process 5	%	%	%	%	%	%
Population Epidemiology Stage 2 Total	%	%	%	%	%	%
High Poverty Index HHs Totals	%	%	%	%	%	%

The SPA2 reference documented that many industries are already developing sustainability standards for their member companies. It's not farfetched for an organization such as the SDG-



DevTreks –social budgeting that improves lives and livelihoods

ASB to provide incentives to these industry groups to maintain “Reference Case Scorecards” that document the most probable, cost effective, equitable, practices needed to achieve specific sustainability targets in those industries.

In terms of tackling landscape-wide SDG impacts, many local communities are already mandating the use of “sustainability reporters” for a variety of purposes. These impacted communities and their newly hired, or fee-supported, reporters may welcome having these types of uniform scorecards to carry out their reporting requirements of actual SDG target accomplishment.

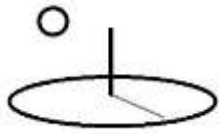
Impacted stakeholders, such as consumers, may also welcome having uniform scorecards so that they can make “informed” purchases, support “informed” stores, judge how well their public officials are delivering public services and cast resultant “informed” votes, and lead meaningful lives. The latter likelihood reinforces the role that objective, evidence-based, science must play in these scorecards.

These independent scorecards also deal with the “gaming the financial accounting system” employed by companies and accounting firms to indefinitely postpone financial disclosures while they conduct poorly defined, opaque, audits. Rather than wait for the audit results, impacted stakeholders conduct their own Social Performance Assessments and communicate the results directly to other “informed” stakeholders.

J. Communication

The following list summarizes the full storyline that companies and local communities communicate to stakeholders using this algorithm’s approach.

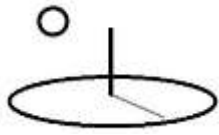
- CSOs work at industry and watershed scale to assist businesses and other CSOs achieve more effective SDG-related activities and to ensure that independent business mitigation and adaptation efforts are coordinated in a way that leads to sector and landscape level



SDG-related results. They maintain the “Scorecards” introduced in Section I, Decisions, which lay out industry-wide and community-wide best sustainable practices and targets.

- Business managers and CSO executives employ impact pathways and theories of change to identify and prioritize the primary SDG risks associated with their activities and work processes and to assess the impacts of those activities on targeted stakeholders.
- Managers take mitigation and adaptation actions based on the SDG-related risks and impacts they identify as relevant to company, CSO, and local community stakeholders.
- Formal Performance Monitoring data systems allow companies and CSOs to account for periodic, short term, SDG-related accomplishments. These instruments are submitted to the sustainable business accounting system to establish “cause and effect” attribution.
- Formal Impact Evaluation data systems allow companies and CSOs account for periodic, SDG-related, long term strategy and to conduct external audits. These instruments are submitted to the sustainable business accounting system to verify the accuracy of the initial “cause and effect” conclusions.
- Full theories of change, impact pathways, and results chains, help companies and CSOs to fully communicate periodic SDG-related accomplishments to relevant stakeholders and to make adjustments to their mitigation and adaptation activities to achieve more efficient and effective results.
- The accounting metrics not only assess company, CSO, and local community performance, they also help stakeholders take ownership of their “own impacts”. That can mean completing their own SIAs, proposing their own mitigation and adaptation actions, monitoring and evaluating their own performance, and using their own “perspective” to report their SDG accomplishments.

Section H, Social Performance Score, demonstrated basing SIA on final SDG Impacts. Examples 1 to 4 demonstrated that Impacts are the final part of the full impact pathways, or Results Chains. The following summary Results Chain use both Example 1 to 4’s results (or direct DevTreks Operating and Capital Budgets) and this algorithm’s results, to communicate final results to



DevTreks –social budgeting that improves lives and livelihoods

decision makers. The report also supports communicating this storyline to impacted stakeholders who hold specific “value systems”, or Example 4B-like “perspectives”.

Business RCA report (8*)

The following report answers the evaluation question

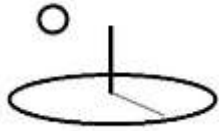
To what extent has the [company] [organization] [community:] been effective and efficient in achieving SDG-related targets that are relevant to our stakeholders?

Example 7 [may] demonstrate the use of machine learning algorithms to codify this report.

Business Value Report	Inputs	Activities	Outputs	Outcomes	Impacts
Societal Perspective					
Non Harassment and Discrimination Index	x	x	x	x	x
Air Pollution Index	x	x	x	x	x
Water Consumption Index	x	x	x	x	x
Disaster Risk Reduction Index	x	x	x	x	x
Biodiversity Index	x	x	x	x	x
Climate Change, Short Term Index	x	x	x	x	x
Climate Change, Long Term Index	x	x	x	x	x
Stakeholder Group A Perspective					

Industry Perspective					

The following image demonstrates using a calculator’s Media View to communicate these reports and scorecards to stakeholders.



DevTrek –social budgeting that improves lives and livelihoods

GreenTrek	Search	Preview	Select
Edit	Pack	Views	Club

Select
PackIt

Edit Linked Views
Make base

Output Stock Calculator--
Get

Media
Mobile
Desktop

Dataset: [Coffee Firm RCA5 Stock IRI](#) This example of Social Performance Stock Assessment can be found in the Social Performance Analysis reference.

RCA-Value1

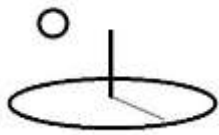
Business Value Report	Inputs	Activities	Outputs	Outcomes	Impacts
Societal Perspective					
Non Harassment and Discrimination Index	x	x	x	x	x
Air Pollution Index	x	x	x	x	x
Water Consumption Index	x	x	x	x	x
Damage Risk Reduction Index	x	x	x	x	x
Biodiversity Index	x	x	x	x	x
Climate Change, Short Term Index	x	x	x	x	x
Climate Change, Long Term Index	x	x	x	x	x
Stakeholder Group A Perspective					
-					
Industry Perspective					

Download

MandE1

Full SDG Monitoring and Evaluation Report	Indicator						
period	1	2	3	4	5	6	7 Score
	2017-2018	2019-2020	2021-2022	2013-2024	2015-2026	2027-2028	2019-2030
Non Discrimination and Non Harassment Index							
sdg target	x	x	x	x	x	x	x
most likely sdg actual	x	x	x	x	x	x	x
sdg percent	x	x	x	x	x	x	x
low sdg actual	x	x	x	x	x	x	x
high sdg actual	x	x	x	x	x	x	x
certainty 1	x	x	x	x	x	x	x
certainty 2	x	x	x	x	x	x	x
certainty 3	x	x	x	x	x	x	x
totalcost	x	x	x	x	x	x	x
Water Consumption Index							
sdg target	x	x	x	x	x	x	x
most likely sdg actual	x	x	x	x	x	x	x
sdg percent	x	x	x	x	x	x	x
low sdg actual	x	x	x	x	x	x	x
high sdg actual	x	x	x	x	x	x	x
certainty 1	x	x	x	x	x	x	x
certainty 2	x	x	x	x	x	x	x

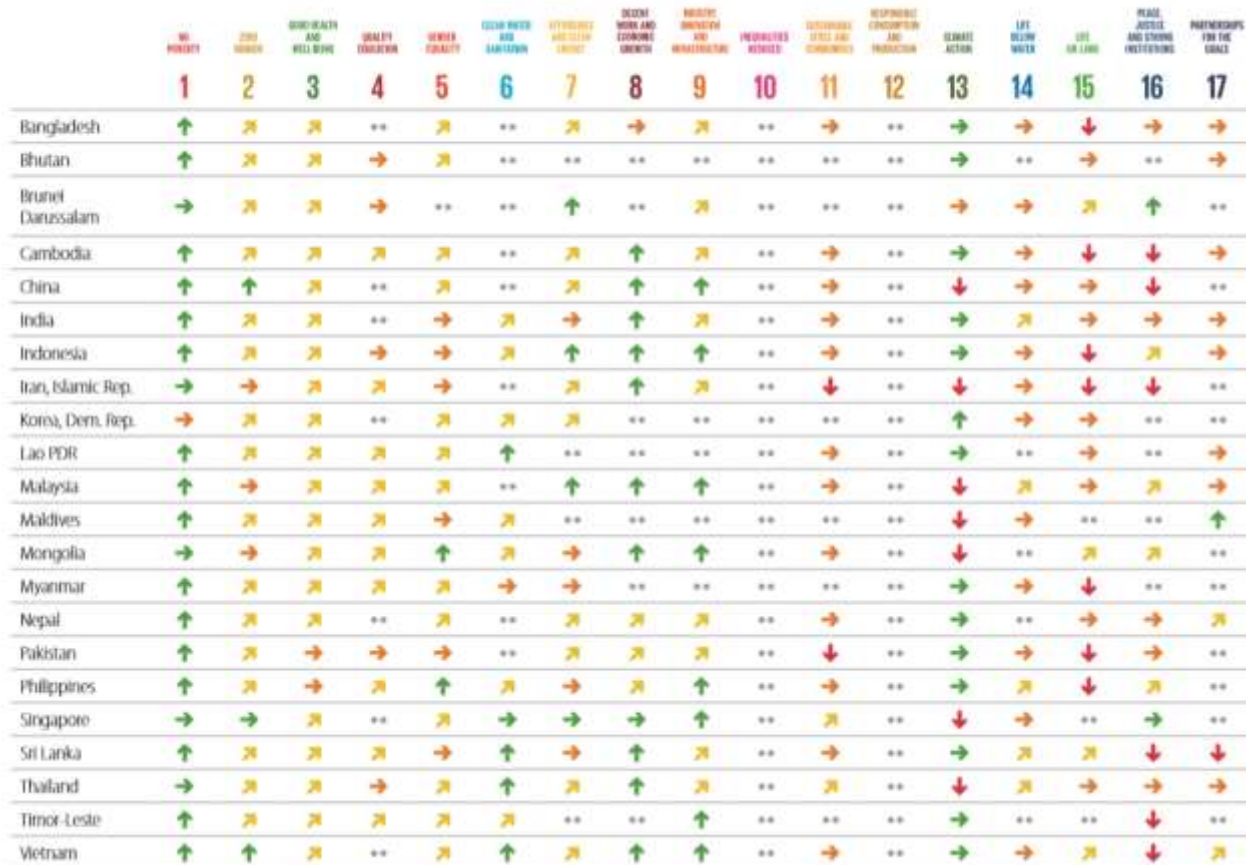
One of the most comprehensive examples of communicating SDG target accomplishment relates to the SDG Index and Dashboards Report project summarized in the following image



DevTreks –social budgeting that improves lives and livelihoods

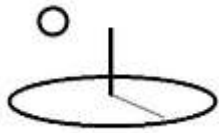
(Bertelsmann Stiftung and SDSN, 2018). Example 9 in the SDG Plan reference, demonstrates using a similar approach for assessing the feasibility of mitigation and adaptation technologies for local communities and local businesses within watersheds or landscapes.

Figure 10 | SDG Trend Dashboard for East and South Asia



K. Performance Monitoring and Impact Evaluation (M&E)

The following image (IFC, 2014) illustrates how the Indicators used in this tutorial’s RCA instruments relate to company monitoring and reporting systems. Unlike the “benchmark-target-actual” M&E approach introduced in Example 1, the “start-target-actual” Indicator properties used in this algorithm support direct M&E (i.e. without the “_xx” suffixes in the labels).



DevTreks –social budgeting that improves lives and livelihoods

IFC (2015) uses the term “audits” in a similar manner to the Impact Evaluations introduced in SPA2 and demonstrated in Example 7.

ABC | 9. Monitoring and Reporting

Monitoring Plan for ABC

ABC’s ESMS team developed a Monitoring Plan based on the Action Plans and their targeted objectives.

Risk 1: Chronic diseases among workers due to prolonged exposure to pesticides and other banned substances

Objective: Prevent use of all banned substances

Target: 100% elimination of banned pesticides and other agrochemicals from farm operations

PERFORMANCE INDICATORS	
Monitoring indicators	Monitoring records
Amount of banned pesticides and agrochemicals purchased	- Procurement records
Amount of banned pesticides and agrochemicals disposed of	- Hazardous waste disposal records
Amount of banned pesticides and agrochemicals in storage	- Inventory records

PROCESS INDICATORS	
Monitoring indicators	Monitoring records
% of banned pesticides and agrochemicals substituted by non-banned products	-Production/Process design checklist

Objective: Minimize workers’ exposure to agrochemicals

Target: Ensure safe work practices and 100% use of PPEs

PERFORMANCE INDICATORS	
Monitoring indicators	Monitoring records
Occupational injury frequency rate attributed to	- Health and safety logs

This reference endorses engaging impacted stakeholder groups to use their own “perspective” to carry out SIA and M&E. These types of “participatory M&E systems”, or “Stakeholder Perspectives”, must be accompanied by an independent “Societal Perspective” that doesn’t lose sight of the overall public service, SDG, objectives.



DevTreks –social budgeting that improves lives and livelihoods

Although the purpose of this example is not to demonstrate full Impact Evaluation, as explained by Gertler et al (2016), the local community dataset employs 2 LIs to document impact results for Treatment and No Treatment villages. These 2 datasets allow basic, but statistically invalid, comparison between the intervention (Treatment villages) and counterfactual (No Treatment villages) used to support full Impact Evaluations. Example 6A demonstrates how to use these datasets to reach initial conclusions about the effectiveness of an intervention or company activity. Example 7 shows how to conduct full, statistically valid, Impact Evaluations.

L. Additional SDG M&E Tools

The following image (IFC, 2013) introduces additional M&E-related tools that companies can use in agricultural supply chains to further accomplish the SDGs. This tutorial introduces tools that support many of these mechanisms. For example, Example 3’s Score Hot Spots analysis can be used as a “supplier score card”. Example 7’s Impact Evaluations can be used in the same manner as “supplier audits”.



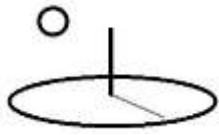
Step 3. Implementation, Monitoring, and Review

Once a strategy outlining the E&S management approach in the supply chain is developed (with priorities and acceptable performance levels), the strategy needs to be implemented. Some tools to do this appear in Table 9.

TABLE 9. TYPES OF IMPLEMENTATION TOOLS

Activity	Type of tool	Potential costs	Time to implement	Key partners in development and/or implementation
Supplier codes of conduct	Process improvements	Low	Short	IFC Performance Standards, UN Global Compact, ILO, roundtables, industry and other sector alliances
Supplier scorecards	Process improvements	Low	Short	IFC Performance Standards, UN Global Compact, ILO, roundtables, industry and other sector alliances
Supplier contract clauses	Process improvements	Low	Medium, recurring	IFC Performance Standards, UN Global Compact, ILO, roundtables, industry and other sector alliances
Supplier communication	Building skills, sharing information	Low	Medium, recurring	Suppliers
Incentivizing performance	Process improvements	Variable	Recurring	Suppliers, senior management, other internal departments
Financing mechanisms	Process improvements	Variable	Ongoing	Domestic financial institutions
Capacity building of suppliers	Building skills	High	Long Recurring	Suppliers, roundtables, consultants
Capacity building of smallholders	Building skills	High	Long Recurring	Suppliers, roundtables, consultants, local service providers, extension offices, local government
Building partnerships	Building skills, information sharing	Medium	Long	Range of potential stakeholders, roundtables, industry and other sector alliances
Supplier audits	Information gathering	Medium (Traders, processors)	Short Recurring	Auditing companies, third-party certification schemes; for primary production, local NGOs, governments, research organizations
		High (Primary production)	Short–Medium Recurring	
Grievance mechanism	Process improvements,	Low	Medium Ongoing	Senior management, other internal departments, communities, NGOs, local

RCA-style instruments can digitize many of these tools and make the consequences of local community and company activity fully transparent. Regular financial disclosures that include



DevTreks –social budgeting that improves lives and livelihoods

SDG accomplishments, disclosures completed by 3rd party RCA Assessors, and new types of SIA-related techniques, such as this reference’s algorithms, SDG-ASB Scorecards, Social Licenses to Operate, social contract riders, and social media campaigns, will make SDG accomplishment much more transparent. Stakeholders, investors, asset managers, supply chain participants, consumers, employees, and impacted communities, can more easily assess the degree of SDG target accomplishment, or “social soundness”, achieved by specific communities, companies, and their executives. Companies and communities can anticipate a large proportion of stakeholders, or more accurately, informed stakeholders, to act accordingly (11*).

Footnotes

1. As mentioned or implied throughout this tutorial, given that the conventional institutions in some countries appear incapable of understanding, let alone ameliorating, these conditions, networks and clubs may need to take independent action.
2. Espey et al (2017) also mention the need for “vetting [of NGO data products] by the international community”. That’s why DevTreks’ source is open and the tutorials plentiful. It’s not necessary for DevTreks itself to be proven right with this approach, nor is it absolutely essential to get “international vetting” (although, by all means, vet), –it’s necessary to teach the next generation about alternative approaches that might help them to be right. As stated in the SPA2 reference, reinforced in Footnote 10, and double-downed in Example 6’s footnotes, too much is at stake to keep repeating the same mistakes over and over again.
3. The SDGs are best viewed as an “integrated and multi-sectoral approach” that companies and communities can use for achieving sustainability goals. For example, although the SDG Indicators use terms associated with country populations or broad ecosystems, many of those terms can be adjusted for use by private sector companies and community service organizations. Examples include, “16.b.1. Proportion of ~~population~~ [employees] reporting having personally felt discriminated against or harassed in the previous 12 months ...”, “12.5.1 ~~National~~ [Company] recycling rate, tons of material recycled”, and



DevTreks –social budgeting that improves lives and livelihoods

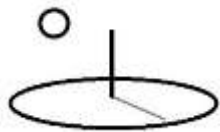
“9.4.1 CO₂ emission per unit of value added [product gross revenue]”. Most progressive sustainable business indicator systems will make, or have made, similar adjustments.

4. The Hoebink et al (2014) reference highlights the challenges of using socioeconomic characteristics to explain cause and effect attribution related to agricultural sustainability. The reference runs through a gamut of feasible socioeconomic explanations –from on-farm income, producer organization support, gender roles, to informal institution development. This tutorial’s consistent advice that networks need to work with experts and information technologists to fine tune these algorithms, is no more specious than researchers continual request to sponsor more research. This tutorial is authored by a technologist who believes applied IT offers a better chance of “ameliorating” than the approaches taken by conventional institutions, including many academic research ones. If you have evidence proving otherwise, take these recommendations with a grain of salt (i.e. but for some reason, the word “futile” comes to mind –where are the reference datasets and source code?).
5. SMEs should recognize that the RCA instruments used throughout this tutorial and the CTAP tutorial employ TEXT csv datasets –even companies that don’t have direct access to computers can probably find a way to complete these instruments. The TEXT files can be submitted to the sustainable business data system when access to an Internet connection is available. Even if the international community develops “international standards to ensure data integrity” (Espey et al, 2017), the simplicity of the TEXT files increases the likelihood of compliance with the new standards. For example, the following image demonstrates harmonizing IFC (2015) Company Risks, related to natural resources, with the SDG Indicators (2017). SPA2 and Footnote 3 demonstrate how to further harmonize disparate sustainability standards.



Harmonization of IFC Company Risks and Impacts and SDG Indicators		
IFC Risks	IFC Impacts	SDG Indicators
Release of air pollutants (air emissions)	Pollution of air, land, and surface water	11.6.2 Annual mean levels of fine particulate matter (e.g. PM2.5 and PM10) in cities
Release of liquid effluents or contaminated wastewater into local	Surface water pollution	6.3.1 Proportion of wastewater safely treated
Generation of large amounts of solid waste and improper waste	Pollution of land, and ground and surface water	12.2.1 Material footprint, material footprint per capita, and material footprint per GDP
Improper management of hazardous substances	Contamination of adjacent land and water	12.4.2 Hazardous waste generated per capita and proportion of hazardous waste treated, by
Excessive energy use	Depletion of local energy sources and release of combustion residuals	7.2.1 Renewable energy share in the total final energy consumption
Excessive water use	Depletion of water resources	6.4.2 Level of water stress: freshwater withdrawal as a proportion of available
High or excessive noise levels	Negative effects on human health and disruption of local wildlife	no SDG
Improper or excessive land use	Soil degradation and biodiversity loss	15.3.1 Proportion of land that is degraded over total land area

6. Example 3 in SPA2 explains that the current UNSETAC LCIA characterization factors (cfs) are most useful at country or regional scale. DevTreks assumes that science and technology will continue to evolve to make local cfs available as well. As stated in SPA2, the take home message is that people need to be getting their hands dirty and gaining experience building these algorithms and applying their tools so that they know when and how to use them properly. And as this tutorial makes clear, science and technology’s end goal is evidence-based societal wealth via sound IT, not marketing-based technologist or company wealth via sound bites. While it may be easy for an IT company to attract investment for this reference’s algorithms, it may be impossible for most of them to supply scientific evidence of the required societal end goals (i.e. stock market returns are not societal end goals).
7. Version 2.1.8 changed the targeted increase in global temperatures from 2°C to 1.5°C because of recent IPCC (2018) recommendations. IPCC documents how the higher target poses too many costly risks to too many local stakeholders.
8. The IFC 2013 reference dealing with agricultural supply chains includes a case study of Malaysian palm oil processing. The authors acknowledge “There are serious concerns regarding the environmental sustainability of palm oil production, particularly in Southeast Asia” and mention how the company addresses those concerns. That doesn’t



DevTreks –social budgeting that improves lives and livelihoods

seem to have stopped the biodiversity loss and reduction in carbon sequestration.

DevTreks emphasizes the importance of institutional improvement because that's where these broader issues have to be addressed. Specifically, real private sector SDG materiality impact accountability only works when SDG-related institutions are developed that can independently monitor and evaluate the accounting metrics and then have the power to influence private sector behavior. SPA2 confirms that product certification standards organizations, with some international scrutiny, have the potential to fulfill this role and those standards systems are expanding. This example explains that many local communities are expanding their sustainability accounting requirements and devising new incentives to pay for improved sustainability reporting. These local sustainability efforts also have the potential to fulfill this role. "Consequential digital activism" that helps CSOs, companies, and local communities, to further achieve the SDG has the potential to be a powerful institutional improvement.

9. The author has assisted U.S. Resource Conservation Districts in Southern California to report on their annual accomplishments. He employed abbreviated Theories of Change, or impact pathways, to structure those annual reports. Similar to the IEG evaluation (2017) of pollution reduction activities, RCA Districts face the annual evaluation question "To what extent has the RCA District been relevant, effective, and efficient in addressing SDG-related natural resources concerns for our stakeholders?". Unlike the qualitative World Bank Results Chain, the Districts preferred quantified numbers in the reports, especially related to money received as Inputs, number of Activities carried out with the money, and quantified natural resource and socioeconomic Impacts for each Activity. The original motivation for this reference came from the desire to institutionalize, by automating, the annual RCD [RCA] reports.
10. Many of the references used in this tutorial emphasize the importance of improving "existing [or conventional] institutions". That poses a quandary to a technologist/economist who has worked for and with many of those institutions. Personal experience suggests that many of these institutions are ill equipped to make innovative changes and even less equipped to handle the technological changes required by this tutorial. Even well intentioned institutions can't succeed when their leaders simply don't



understand public service (or this reference). Many are stuck carrying out statutory, contract-related, special interest group-influenced, or peer-siloed, work. In many cases, they're good at spending money, accommodating special interests, or pandering to peers, because the special interests and peers like the money or the rigged rules, but not accounting for how well the money is spent. Consequential digital activism may require more of a start-up, moral accountability, culture and less of an existing, rote bureaucratic, special interest-infested or peer-siloed, institution. FAO (2017) provides practical guidance for achieving institutional improvement at national scale. The “next generation” needs to experiment and find the exact nature of the needed institutional improvements. They better hurry up (i.e. by reading Footnote 11).

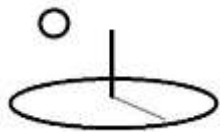
11. After completing the first draft of the SDG-ASB Scorecards, the author felt that companies and communities might find the idea of “universal scorecards” far-fetched or hubristic. Until he read of a popular web application used throughout the U.S.A. that allows employees to rate their companies for how well the company contributes to their “personal quality of life”. “Informed stakeholders” should recognize that rating companies and communities for their contribution to “societal quality of life” may be inevitable, if more complicated and consequential. This reference recommends basing the ratings on “international SDG reporting norms”, “evidence-based” science, and concrete, open-source, IT that employs “next generation”, M&E-based machine learning algorithms. Refer to the Version 2.1.8 SDG Plan reference for additional guidance.

Case Study References

Bertelsmann Stiftung and the Sustainable Development Solutions Network (SDSN). SDG Index and Dashboards Report 2018. Implementing the Goals. Global Responsibilities. 2018

Espey, Jessica et al. Counting on the World. Building Modern Data Systems for Sustainable Development. Sustainable Development Solutions Network. Thematic Research Network on Data and Statistics. 2017

FAO. 2017. Watershed management in action – lessons learned from FAO field projects. Rome.



DevTreks –social budgeting that improves lives and livelihoods

Gertler, Paul J., Sebastian Martinez, Patrick Premand, Laura B. Rawlings, and Christel M. J. Vermeersch. 2016. *Impact Evaluation in Practice*, second edition. Washington, DC: Inter-American Development Bank and World Bank. doi:10.1596/978-1-4648-0779-4. License: Creative Commons Attribution CC BY 3.0 IGO

Marc Gordon, United Nations Office for Disaster Risk Reduction (UNISDR). *Monitoring progress in disaster risk reduction in the Sendai Framework for Action 2015-2030 and the 2030 Agenda for Sustainable Development*, slide presentation, 2016.

Hoebink, Paul; Ruben, Ruerd; Elbers, Willems (eds). *The Impact of Coffee Certification on Smallholder Farmers in Kenya, Uganda and Ethiopia*. Centre for International Development Issues Nijmegen (CIDIN), Radboud University. Nijmegen, The Netherlands, 2014

Intergovernmental Panel on Climate Change. *IPCC Special Report on Global Warming of 1.5°C*. 2018

International Finance Corporation. *Environmental and Social Management System Implementation Handbook*. General. 2015

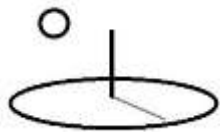
International Finance Corporation. *Environmental and Social Management System Toolkit and Case Studies*. Crop Production. 2014

International Finance Corporation. *Good Practice Handbook*. *Assessing and Managing Environmental and Social Risks in an Agro-Commodity Supply Chain*. 2013

International Finance Corporation. *IFC Performance Standards on Environmental and Social Sustainability*. 2012

International Finance Corporation. *Strategy and Business Outlook*. FY18-FY20. *Creating Markets and Mobilizing Private Capital*. 2017

International Institute for Tropical Agriculture (IITA) and Committee on Sustainable Agriculture (COSA). *Impacts of Certification on Organized Small Coffee Farmers in Kenya*. Baseline, 2016.



DevTreks –social budgeting that improves lives and livelihoods

Task Force on Climate-related Financial Disclosure (TCFD). Recommendations of the Task Force on Climate-related Financial Disclosure. Final Report. 2017

Ellen Taylor-Powell, Larry Jones, and Ellen Henert. Enhancing Program Performance with Logic Models, University of Wisconsin-Extension, Feb. 2003

United Nations (20 organizations collaborated). The Social Dimensions of Climate Change. Discussion Draft. 2011

United Nations. Global indicator framework for the Sustainable Development Goals and targets of the 2030 Agenda for Sustainable Development. 2017

United Nations. SDG Indicators 2017. Available from: <https://unstats.un.org/sdgs/indicators/indicators-list/> [last accessed: September, 2017].

UNEP/SETAC. Hotspots Analysis. An overarching methodological framework and guidance for product and sector level application. 2017

U.S. Department of Health and Human Services. Centers for Disease Control and Prevention (CDC). Principles of Epidemiology in Public Health Practice. Third Edition. An Introduction to Applied Epidemiology and Biostatistics. 2012

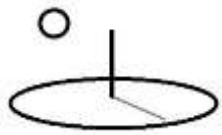
Vanclay, Frank, Ana Maria Esteves, Ilse Aucamp, Daniel M. Franks. Social Impact Assessment: guidance for assessing and managing the social impacts of projects. Fargo, ND: International Association for Impact Assessment. 2015

World Bank. Data for Development. An Evaluation of World Bank Support for Data and Statistical Capacity. Independent Evaluation Group. Washington DC, 2017

World Bank. Toward a Clean World for All. An IEG Evaluation of the World Bank Group's Support to Pollution Management. Independent Evaluation Group. Washington DC, 2017

World Health Organization. Guide to Cost-Effectiveness Analysis. 2003

References Note



DevTreks –social budgeting that improves lives and livelihoods

We try to use references that are open access or that do not charge fees.



DevTreks –social budgeting that improves lives and livelihoods

Example 6. Disaster Stakeholder Resource Conservation Value Accounting

Algorithms: CTAP algorithms and algorithm1, subalgorithm17

URLs

Monitoring and Evaluation Output Calculator

<https://www.devtreks.org/greentreks/preview/carbon/output/Disaster Risk Management, Example 6/2141223485/none>

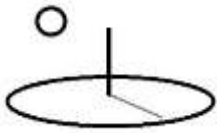
<https://www.devtreks.org/greentreks/preview/carbon/resourcepack/Disaster Accounting, Example 6/1560/non>

<http://localhost:5000/greentreks/preview/carbon/resourcepack/Disaster Accounting, Example 6/545/none>

<http://localhost:5000/greentreks/preview/carbon/output/Disaster Risk Management, Example 6/2141223500/none>

A. Introduction to Disaster Risk Management

As with the SDG, the SDRR's targets and indicators have also been adopted by most countries and most countries face similar challenges in collecting and maintaining the data. In particular, this example, like the CTAP examples, support local DRR efforts carried out by companies, producer organizations, community service organizations, investors, supply chain participants, consumers, and impacted communities. The following image (UNISDR, 2017) verifies that most countries do not have local DRR practices in place that can support the SDRR or their national DRR efforts.



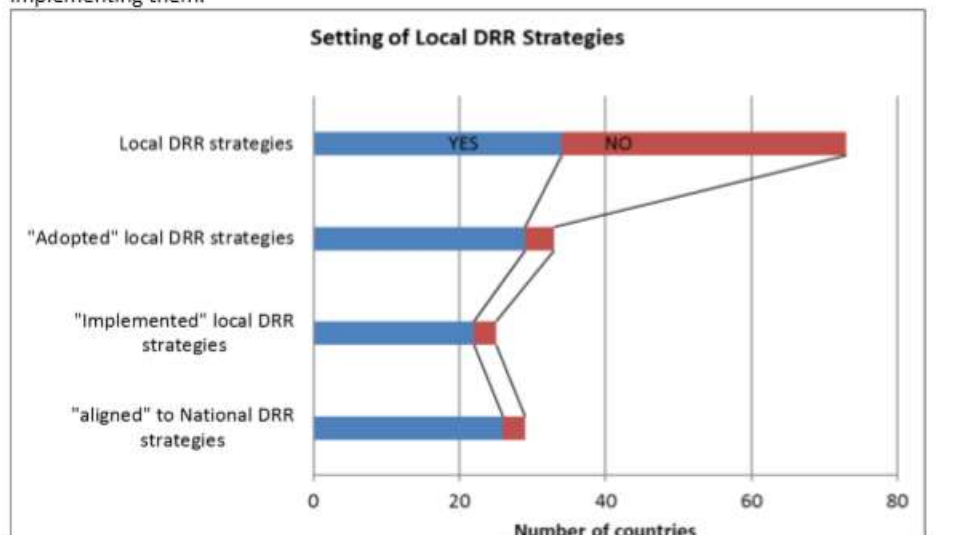
Local disaster risk reduction strategies.

Compared to national strategies, **local disaster risk reduction strategies** are far more heterogeneous, vary across countries and local administrative units, and change over time. Local governments, again with highly heterogeneous characteristics and capabilities, are normally responsible for their development. In general, *national disaster risk reduction strategies* serve a normative function, providing, inter alia guiding principles and an overarching framework for disaster risk reduction. Local strategies, aligned with the national strategy, are generally more specific, reflecting local context and hazard profile, and tend to focus on planning and implementation with clear roles and tasks assigned at local level.

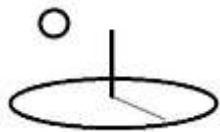
Given these considerations, the alignment of **local disaster risk reduction strategies** with respective **national disaster risk reduction strategies** is considered imperative. The outcome of the Data Readiness Review shows that about a half of the reporting countries have local DRR strategies and that 26 countries out of 29 countries reporting this questionnaire have ones in line with national

135

strategies. The outcome also shows a discrepancy between having local DRR strategies and implementing them.



Poljanšek et al (2017) use the following quote from the Sendai Framework 2015 to reinforce the relationship between the SDG “drivers” of sustainability and the SDRR “drivers” of disaster risk.



DevTreks –social budgeting that improves lives and livelihoods

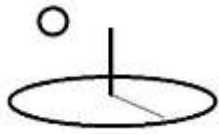
“more dedicated action needs to be focused on tackling underlying disaster risk drivers, such as the consequences of poverty and inequality, climate change and variability, unplanned and rapid urbanization, poor land management and compounding factors such as demographic change, weak institutional arrangements, non-risk-informed policies, lack of regulation and incentives for private disaster risk reduction investment, complex supply chains, limited availability of technology, unsustainable uses of natural resources, declining ecosystems, pandemics and epidemics”.

The following paragraph (UNIDSR, 2015) relates these drivers to government policies that fail to manage risk adequately, such as the failure to price carbon properly. TFCD (2017) discuss the equivalent risk to the financial system posed by climate change. The 2 statements explain why both global data systems need concurrent action and why socioeconomic and institutional analysis skills, or “disaster [and sustainable business] risk assessment capacities”, must be learned and applied to tackle these drivers, causes, and needed policies.

“The continuous mispricing of risk means that consequences are rarely attributed to the decisions that generate the risks. This lack of attribution and accountability creates perverse incentives for continued risk-generating behaviour, as those who gain from risk rarely bear the costs.”

Example 1 includes the Physical Capital category, Flood Control, which qualitatively addresses the coffee firm’s concern about reducing the risk of losses from floods. That capital stock can be expanded to include additional categories that identify additional risks from disasters. In addition, the UNEP (2016) recommendation to include biodiversity and ecosystem services in disaster loss estimates, can be accommodated in a manner similar to Examples 1, 2, and 3. The most straightforward way for companies, CSOs, and local communities to carry out Social Performance Assessments for local disaster risks, is to include more Physical Capital Stock categories and Indicators when using subalgorithms 13, 14, 15, 16, and 17.

The following image (IFC, 2015) verifies that the importance of SDRR, or CTA-Prevention, justifies separate Social Performance Assessments. The International Recovery Platform (2016) refers to these CTAP-related instruments as business continuity plans while IFC uses the term,



DevTreks –social budgeting that improves lives and livelihoods

emergency preparedness plans. In these types of specialized instruments, the RCA-style algorithms focus on the environmental and socioeconomic impacts of disasters. Example 1's company, CSO, and local community activities measure how those actions change the disaster-related impacts. This example demonstrates using Example 5's population transition states to record impacts related to population epidemiology. These algorithms can be used flexibly –some disaster accounting systems may also need to use the transition states to record impacts related to specific disaster hazards or specific hazard events.

Example 6A's dataset uses Example 1 to 4's semi-qualitative, and quantitative, company data to complete these CTAP plans (i.e. "business continuity plans" or "emergency preparedness plans"). Unlike the SPA2 approaches, the Indicators and Indexes used in these reports relate directly to disasters and hazards that impact specific businesses and their stakeholders.



ELEMENT 5

Emergency Preparedness and Response

5

Regular engagement with local community and government for onsite and offsite emergency plan. Formal resource-sharing agreements with neighboring companies.

4

Senior management and all units and shifts, including contract and temporary workers, participate in emergency risk assessment, preparedness planning and mock drills. Continual improvement.

3

All onsite and off-site emergency issues have been identified and an effective preparedness plan is in place. The plan meets the local regulatory requirements and the local industry best practices.

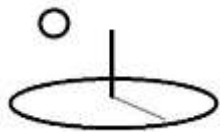
2

The emergency preparedness plan is in place, but there is no evidence of consistent implementation. Some trainings are

Even when you have considered all the risks and put the appropriate management programs in place, accidents and emergency situations can happen.

Your business is a dynamic operation, and many things change from day to day – people go in and out of your workforce, materials and suppliers enter and exit your supply chain, facilities and equipment are added to and removed from your production line. A management system will help to maintain continuity and consistency throughout these changes. However, there may be a momentary lapse or gaps in the system (e.g. someone not properly trained, someone not following the procedures, a machine breakdown), or an external force (e.g. natural disaster) that can lead to an accident or

The CTAP algorithms demonstrate that an important aspect of disaster loss estimating is to correctly identify the Hazards, Exposures, and Vulnerabilities that generate Loss Exceedance Probability distributions (i.e. disaster impact pathways). The CTAP reference addressed disaster risk analysis for single hazards, such as typhoons, droughts, floods, and earthquakes, and human-induced hazards, such as bad management of the drivers of disaster risk (i.e. weakened environmental rules).



DevTreks –social budgeting that improves lives and livelihoods

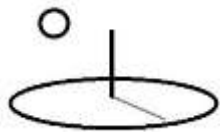
This example’s first dataset, Watershed Disaster Risk Management Plans, demonstrates taking the data from multiple, single hazard, CTAP, or CTAP-like, analyses and adding their results to Example 5’s datasets for specific impacted stakeholders and/or land use areas and using multiple “impact transition states” to understand the resultant population epidemiology. This approach begins to address the UNISDR (2017b) recommendation for estimating “sequential, simultaneous, cascading and interrelated effects of some hazards” and for thoroughly understanding the resultant “population epidemiology”. Individual business continuity and emergency preparedness plans must support this type of overall community disaster management plan.

Although incentives to complete these instruments are not yet offered by companies, insurers, or locally impacted communities, they add another potential revenue stream to support the uniform sustainability reporting carried out by RCA Assessors and RCA Technologists. Local sustainability offices and industry standards organizations can develop fee-supported reporting requirements for these purposes. Given the potential savings from losses and business disruption and the strong possibility of reputational enhancement, informed businesses may become strong supporters of this approach (i.e. especially when evidence starts coming out from disaster areas that either confirm the approach or point to ways to improve the approach).

B. Stakeholder System Boundary and Stakeholder Engagement

The following statement (UNISDR, 2017a) summarizes the UN’s recommendation for data disaggregation by hazard, location, [CTAP] hazard events, and Example 5’s impact transition states. USHHS (2012) confirms that the health care sector uses health states, or impact transition states, to analyze and simulate population “epidemiology” (1*):

“Temporal aspects for attribution and cut-off for data collection: Countries may choose to have different timeframes for each type of hazard, because they have different epidemiology. If so decided, timeframes for each hazard should be based on the epidemiology of injury and illness rates during the event and the feasibility of recording those injuries and diseases.”



DevTreks –social budgeting that improves lives and livelihoods

FAO (2017) uses the following statement to confirm the importance of carrying out these assessments in the larger context of integrated landscape management, including watershed planning (2*).

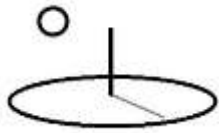
“[The goals of] the Sendai Framework for Disaster Risk Reduction [...] should be seen as a call for watershed management to play its part and take on a stronger role in risk management and resilience building. Institutional strengthening and capacity development for risk management, coordination and contingency planning will be crucial in this regard.”

The CTAP reference introduced using the Exposure and Vulnerability elements of CTAP algorithms to identify and define stakeholder boundaries and impacted stakeholders. All of the references used in this example emphasize the importance using those elements for the same purpose. Serious disaster risk management requires CTAP-like instruments, with their full disaster impact pathways, to be completed prior to this algorithm.

C. Disaster Risk Management Planning

This example customizes Example 5’s techniques for disaster risk management, or DevTreks IT-rooted term, Conservation Technology Assessment-Prevention (CTAP). Rather than repeat the same techniques again, the following list summarizes the prominent differences used when assessing disaster risks.

Disaster Risk Management Pathways (for cause and effect). The CTAP reference introduced algorithms that employ a disaster impact pathway consisting of Hazards -> Exposure -> Vulnerability -> Mitigation Actions -> Impacts. Poljanšek et al (2017) confirm that this is similar to the SDRR impact pathway. UNIDSR (2017b) defines a related, typical impact pathway for disaster risks: Drivers -> Hazards -> Exposure -> Vulnerability -> Capacity -> Impacts. Version 2.1.4 upgraded several subalgorithms by allowing them to be used in 6 or more separate calculator Indicators so that they can used to define 6 or more of these “paths”.



DevTreks –social budgeting that improves lives and livelihoods

This example follows Example 5’s approach of identifying drivers of disaster risk in related algorithms’ impact pathway stages (i.e. see Example 5’s risk and impact identification stage). Poljanšek et al (2017) identify additional examples of risk reduction pathways, such as HAZOP and FMEA, which can also be modeled using this example’s approach. The following image (UNISDR, 2017b) illustrates how these impact pathways support “holistic disaster risk management”.

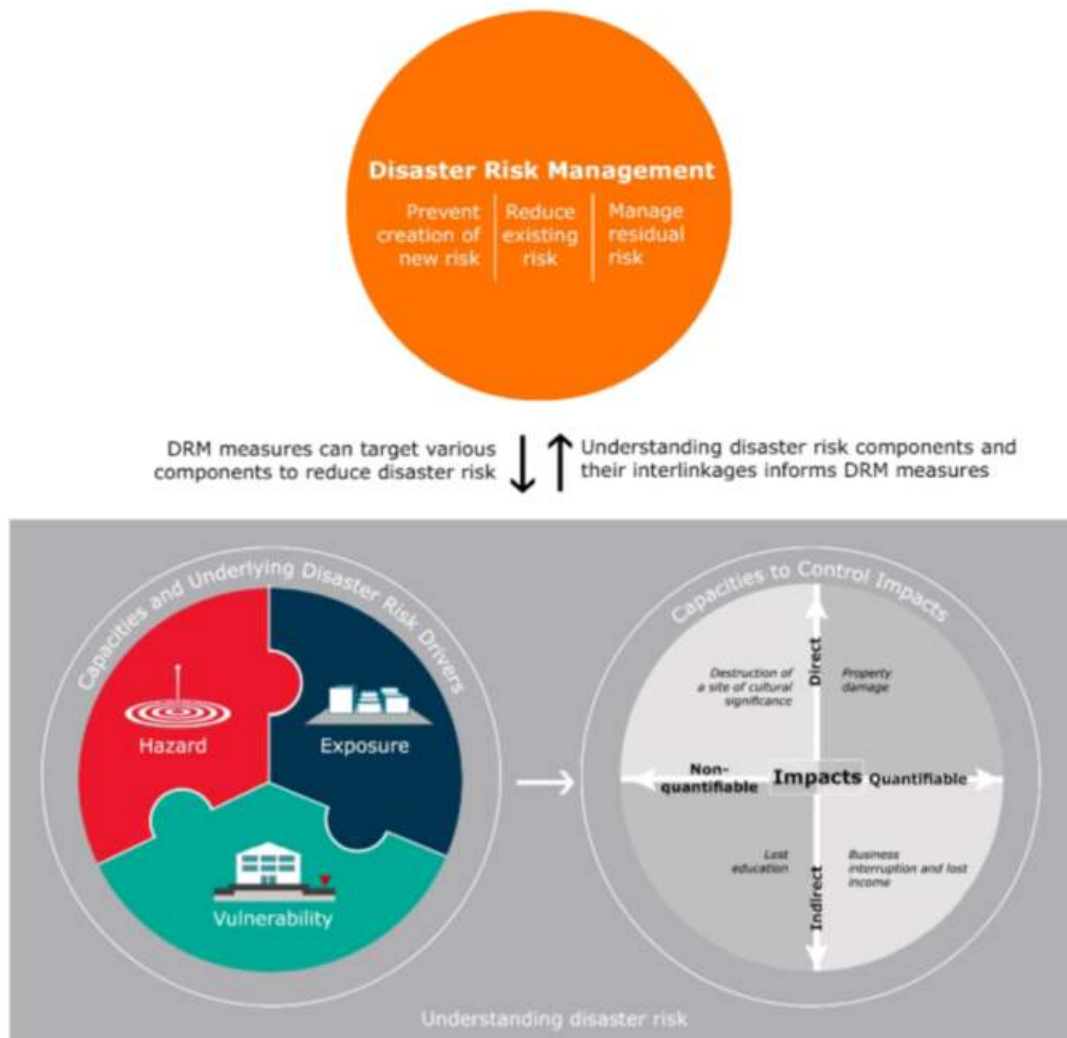


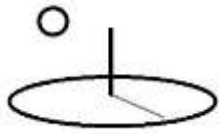
Figure i - Holistic understanding of disaster risk empowers effective and comprehensive disaster risk management (source: UNISDR)



DevTreks –social budgeting that improves lives and livelihoods

Disaster Risk Management Indicators and Indexes: UNISDR (2017a) provides “minimum standards” for SDRR Indicators that “allow for consistent measurement of progress towards the global targets across countries and over the duration of the Sendai Framework and Sustainable Development Goals”. They also mention that individual governments have the latitude to develop alternative guidelines and Indicators that are consistent with these minimum standards.

Example 3B in SPA2 demonstrated the use of “ecopoints” as the final, uniform, score generated from multi-indexed sustainable Indicator systems. Examples 8 and 9 in CTAP demonstrated the use of Multi Criteria Decision Analysis (MCDA or MCA) for the same purpose. This example’s references (UNISDR, 2017; Poljanšek et al 2017) provides similar examples, such as INFORM and the World Risk Index, for multi-index disaster risk management systems. The following image (UNISDR, 2017b) confirms that INFORM employs a basic disaster impact pathway of Hazards & Exposures -> Vulnerability -> Lack of Coping Capacity. INFORM (2018) demonstrates how a fully developed, mature, MCDA, system can support disaster risk management at global scale. In the context of this reference, the authors explain how their 1 Total Risk Score, 4 Indexes, 6 Categories, 14 Components, and 53 Indicators, can be “harmonized”, to a partial degree, with the SDG and SDRR Targets and Indicators (7*).



Risk	INFORM																
Dimensions	Hazard & Exposure					Vulnerability				Lack of Coping Capacity							
Categories	Natural			Human		Socio-Economic		Vulnerable Groups		Institutional	Infrastructure						
Components	Earthquake	Tsunami	Flood	Tropical cyclone	Drought	Current Conflict Intensity	Projected Conflict Intensity	Development & Deprivation (50%)	Inequality (25%)	Aid Dependency (25%)	Uprooted People	Other Vulnerable Groups	DRR	Governance	Communication	Physical Infrastructure	Access to Health System

Figure 11 - INFORM model.

Source: INFORM methodology document (accessed March 2017)

This tutorial will refine the existing “ecopoints” and MCDA approaches to demonstrate how these MCDA systems can support disaster risk management at industry, company, community, and stakeholder group, scale. The fact that many of INFORM’s, CTAP’s, and related comprehensive disaster risk management Indexes and Indicators, extend well beyond the SDRR Indexes and Indicators suggest that comprehensive disaster and crisis management must also look beyond just the UN “metadata” system.

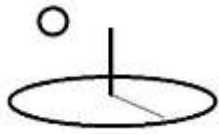
Multiple Disaster Hazards. Disaster risk management requires understanding how multiple hazards impact populations, which is another way of saying understanding “population epidemiology”. The first image (UNIDSR, 2017a) shows a Hazard Classification system that includes earthquakes, typhoons, droughts, floods, biological epidemics, and human-caused hazards (i.e. terrorism, hazardous waste spills, weakened environmental laws, weakened protected area boundaries). Many of the drivers of human-caused hazards can be traced to bad public service management and corrupted institutions which must be addressed as “drivers of disaster risk” in any serious sustainability accounting system.



ANNEX I

Proposed Hazard Classification

FAMILY	Hazards
Geophysical	Earthquake, Tsunami
	Mass movement mostly triggered by geophysical events (overlaps with hydrological category) Landslide, avalanche, rock fall, mud flow, debris flow, subsidence)
	Volcanic activity, eruption, lava flow, ash fall, pyroclastic flow, lahar
Hydrological	Flood, riverine flood, coastal flood, ponding flood, urban flood, flash flood,
	Mass movement mostly triggered by hydrological events (overlaps with geophysical category) Landslide, avalanche, rock fall, mud flow, debris flow, subsidence
	Wave Action, coastal erosion, shoreline change.
Meteorological	Convective Storm, rain, wind, hail, snow, ice, blizzard, lightning, sand, dust, derecho, tornado
	Extra-tropical storm
	Extreme temperature, cold wave, heat wave, frost, freeze
	Fog
	Tropical cyclone, cyclonic wind, cyclonic rain, cyclone surge
Climatological	Drought
	Glacial lake outburst (GLOF)
	Wildfire
Extra-terrestrial	Impact
	Space weather
Environment degradation	Erosion
	Deforestation

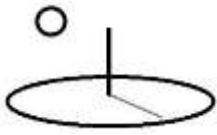


The following image’s (Poljanšek et al 2017) correlation matrix demonstrates a basic approach for understanding the interaction of multiple hazards. Examples 1b to 1e in the CTA 01 reference begins to demonstrate formal techniques that can be used in more advanced risk assessments to account for correlated Indicators, such as interacting hazards (i.e. see Example 6B).

Slides (H4)	2	2	0–No interaction 1–Weak interaction 2–Medium interaction 3–Strong interaction	Slides (H4)	Deposits only	Cut off a flow in a water course
0	Debris flows (H5)	2		No interaction	Debris flows (H5)	Change of river bed morphology
1	1	River floods (H6)		Erosion / saturation of deposits	Re-mobilisation of deposits	River floods (H6)

107

Multiple Hazard Events. The CTAP reference introduced algorithms that employ typical hazard event periods for disasters, such as 2 year, 5 year, 10 year, 25 year, 50 year, and 100 year, events. In a related fashion, Example 5 demonstrates how impact transition states model population epidemiology over time. The following image (Poljanšek et al 2017) shows that the aggregation of multi-hazard event losses can be carried out using data derived from the PRA techniques introduced in CTAP. UNIDSR (2017b) explains how Average Annual Loss estimates or MCDA qualitative ratings, both of which are demonstrated in CTAP, also make appropriate

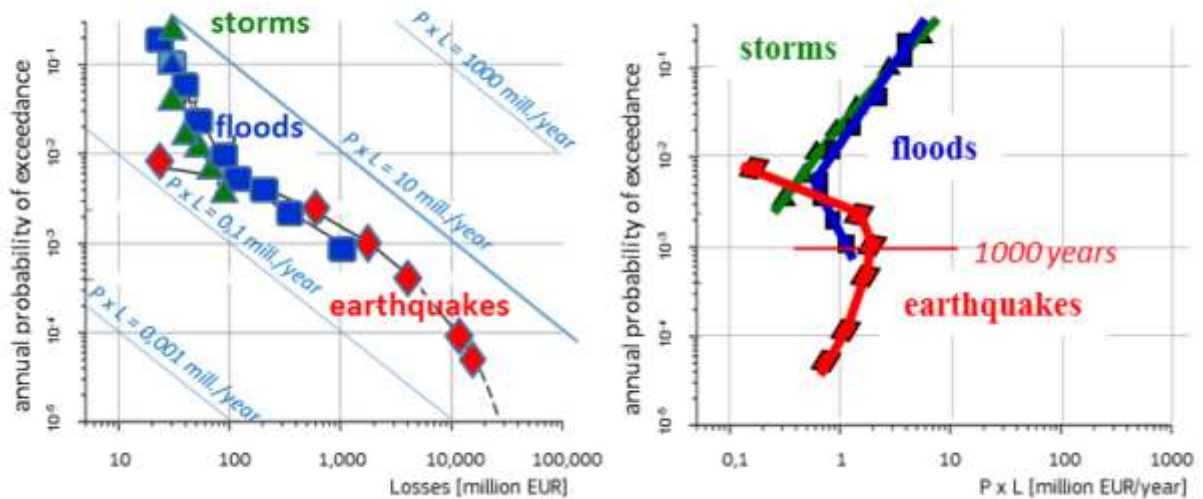


metrics for aggregating losses from multi hazards and their events.

FIGURE 2.24

Risk curves and P x L - curves for the city of Cologne (Exceedance probability versus loss (left) and versus its product with loss (right))

Source: courtesy of author



105

Mitigation and Adaptation Activities and Life Cycle Stages. The mitigation and adaptation activities “causing” changes in impact scores must be defined more broadly for disaster risk management. Much of a community’s ability to recover from disasters relates to their resilience and coping capacity (Poljanšek et al 2017 define these terms in the following image; UNIDSR, 2017b defines these terms in Annex 3). Example 5’s Hotspots Work Processes can be expanded to include improvements to community resilience and coping capacity.



BOX 2.1

Resilience and capacities

Besides the notion of 'vulnerability' there are other terms and concepts addressing the possibility of harm to a system, people or specific objects by certain events and processes. Vulnerability – understood as a holistic and systemic concept – is closely related to and partly overlaps, for example, with the concepts of resilience and of coping and adaptive capacity.

'Resilience' is a term that has been widely used over the last years to describe characteristics related to the ability to absorb stresses, to respond to changes and to recover from shocks. Some authors see resilience as the positive flipside of vulnerability. A broader under-

standing of resilience incorporates the ability and willingness to learn, to reorganise and to undertake critical self-reflection (Alexander 2013; Kelman et al., 2016). Climate resilience has emerged into a new doctrine under the umbrella of which communities define the activities to combat the impending implications of climate change.

There are numerous related activities within Europe, for example the RESIN project is investigating climate resilience in European cities, the European Commission's FP7 project emBRACE has focused on community resilience and developed a set of key indicators for assessing it, and the Commission's

Horizon 2020 project 'resilens' is scrutinising the resilience of European critical infrastructure.

Just as the term 'resilience', the concept of capacities relates to the possibilities and abilities to reduce harm under hazardous conditions. Hereby, 'coping capacity' rather deals with the short-term conservation and protection of the current system and institutional settings, whilst 'adapting capacity' denotes a longer-term and constantly unfolding process of learning (Birkmann et al., 2013).

In addition, the following image (Poljanšek et al 2017) shows that several disaster stages require mitigation, including prevention, preparedness, response, recovery, and reconstruction. Example 5's Hotspots Life Cycle Stages can be replaced with these stages for community disaster management. Alternatively, some damage assessments may prefer to use multiple hazards as replacements.

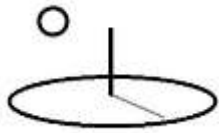
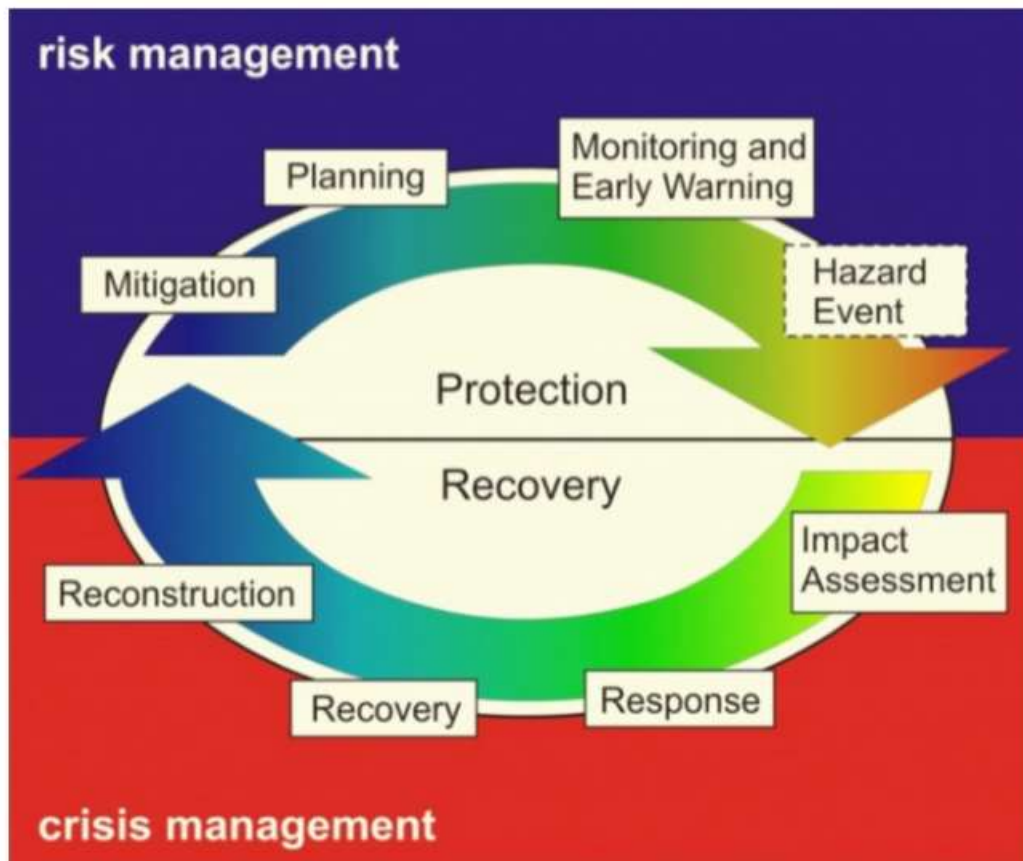


FIGURE 3.45

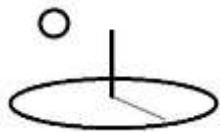
The Cycle of Disaster Management illustrating the importance of risk management in reducing future drought-related impacts.

Source: National Drought Mitigation Center, University of Nebraska-Lincoln.



Stakeholder Impact Analysis (SIA). UNIDSR (2015) discuss the relation between Example 5’s SIA approach and the need for a similar approach for ensuring the fairness of disaster risk management. UNIDSR (2017b) discuss the centrality of understanding “socially constructed risks” in a related manner.

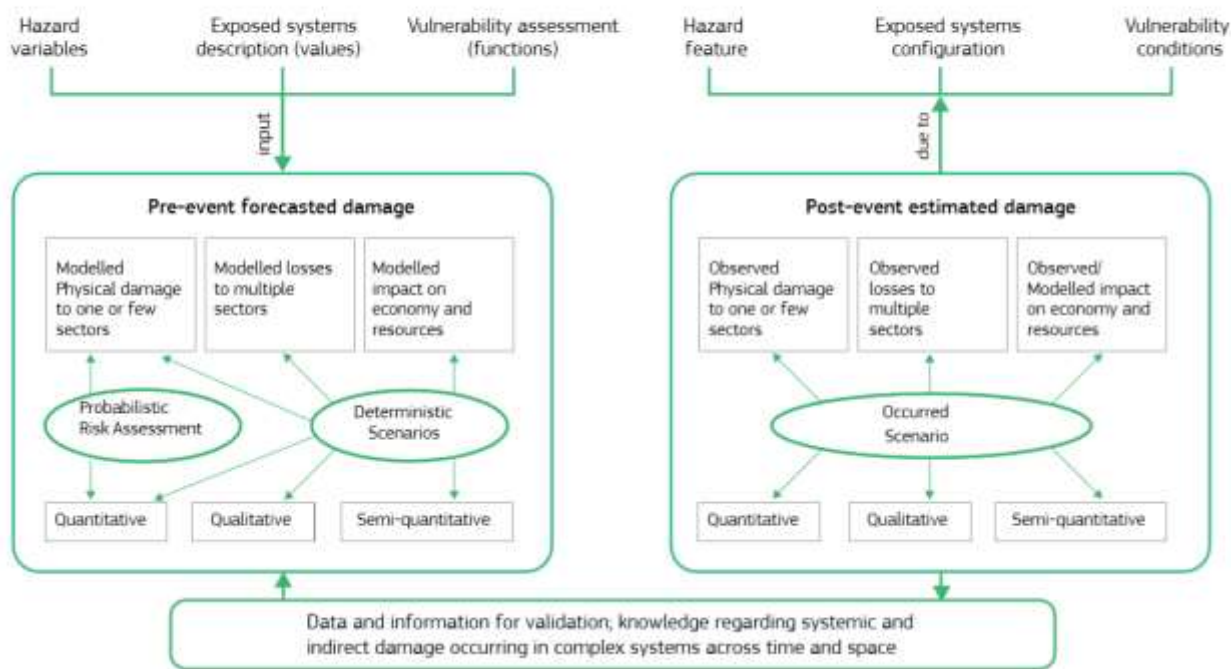
“Fewer countries addressed the need for a social impact analysis (SIA) even though SIA is important because the scale of disasters differ depending on the vulnerability of the community. Poor people, children, the elderly and the disabled are more vulnerable to hazards. SIA is an important tool for supporting social policy planning and requires disaggregated data (e.g. age and gender) to identify the vulnerable segments of society that need support.”



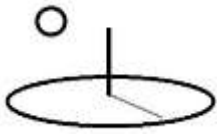
Actual, or Post-Disaster, Impacts: Poljanšek et al (2017) use the following image to discuss the role that post-disaster data collection can play in improving the accuracy of the pre-disaster risk assessment models and techniques. The “start-target-actual” properties used with this algorithm’s Indicators require this approach. The “start and target” properties can still be generated from CTAP-style pre-disaster risk modelling and the “actual”, or post-event, properties support further improvement in those algorithms.

FIGURE 2.14

Pre - and post - disaster damage assessments
Source: courtesy of authors



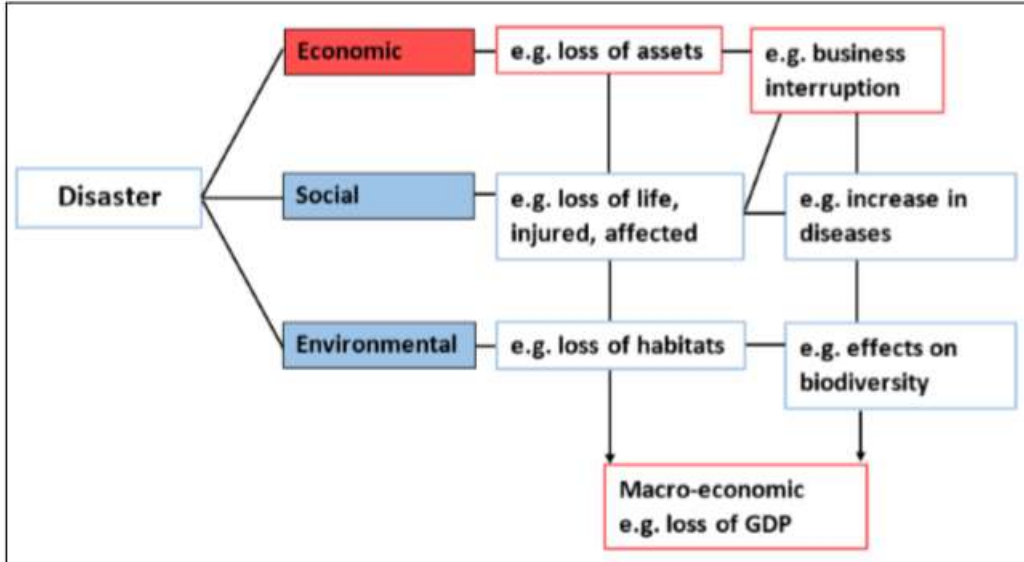
Indirect and Direct, Tangible and Intangible, Benefits and Costs: The SDG impacts being evaluated must relate to both physical losses (i.e. loss of life) and monetary losses (i.e. loss of property). The following images (UNISDR, 2015, and V. Meyer et al, 2014) summarize the full benefits and costs that must be evaluated using “Reference Case Socioeconomic Results”. The CTAP reference introduced algorithms (i.e. subalgorithm10) that address several of these dimensions of costs and benefits. Disaster risk management requires expanded documentation of costs and benefits. Specifically, Example 6A will demonstrate using socioeconomic SDG



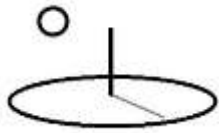
DevTreks –social budgeting that improves lives and livelihoods

measurements, such as QASYs, QALYs, and DALYs, rather than relying solely on physical SDG measurements. Example 4B confirms that they measure stakeholder preferences better.

Figure 7: Impact of Disaster

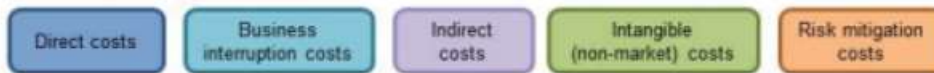


Source: Author



		Tangible costs	Intangible (non-market) costs
Damage costs	Direct	<ul style="list-style-type: none"> Physical damage to assets: <ul style="list-style-type: none"> – buildings – contents – infrastructure 	<ul style="list-style-type: none"> Loss of life Health effects Loss of environmental goods
	Business interruption	<ul style="list-style-type: none"> Production interruption because of destroyed machinery 	<ul style="list-style-type: none"> Ecosystem services interrupted
	Indirect	<ul style="list-style-type: none"> Induced production losses of suppliers and customers of companies directly affected by the hazard 	<ul style="list-style-type: none"> Inconvenience of post-flood recovery Increased vulnerability of survivors
Risk mitigation costs	Direct	<ul style="list-style-type: none"> Set-up of infrastructure Operation & maintenance costs 	<ul style="list-style-type: none"> Environmental damage <ul style="list-style-type: none"> - due to the development of mitigative infrastructure - or due to a change in agricultural practices
	Indirect	<ul style="list-style-type: none"> Induced costs in other sectors 	

Cost categories applied in this article:



Disaster Impact Scenario Analysis: Example 5’s scenario analysis addressed alternative scenarios for SDG Target-related Impacts. The following image (UNIDSR, 2017b) summarize the impacts to analyze in disaster scenario analysis. These authors also discuss the use of logarithmic scales, related to the probability of disaster events, for this algorithm’s 3 certainty factors in disaster scenario analysis.



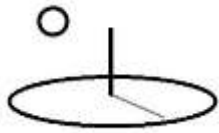
Box 11– Examples multi-criteria impact analysis

The impact analysis in the scenario-based approach should take into account all different kinds of societal impacts. As defined in the Sendai Framework, the societal impacts to be considered are impacts on human health and safety, the economy, the environment, the social and political stability, cultural heritage and education, as well as negative consequences for SDG's.

The impact on each of these societal values has to be operationalized in one or more impact criteria, so the impact can be estimated in a transparent and comparable way for each risk. Examples of criteria in existing NDRA scenario based methods are²⁵:

Human impact (health and safety)	<ul style="list-style-type: none">• fatalities• severely injured or ill people• permanently displaced people• people with lack of basic necessities
Economic impact	<ul style="list-style-type: none">• fatalities• severely injured or ill people• permanently displaced people• people with lack of basic necessities
Environmental impact	<ul style="list-style-type: none">• disruption of ecosystems• environmental pollution• loss of ecological value
Social and political impact	<ul style="list-style-type: none">• public outrage and anxiety/social-psychological impact• disruption of daily life• disruption of the education system• encroachment of the territory• infringement of the international position• violation of the democratic system• impact on public order and safety

Stakeholder Communication: Unlike most DRR references, Poljanšek et al (2017) devote several chapters to this topic. The authors address public perception of risk, communication of uncertainty, the role of modern social media, and the importance of interactive communication between experts and audience. UNIDSR (2017b) describe the ultimate purpose of these communication aids: “The end result of the risk evaluation is a decision by the authorities [...] after stakeholder consultations and [...] public participation, on the “prioritization of risk””. The following image (Michellini et al, 2015) demonstrates the use of standard communication aids proposed for an EU-wide hazard reduction system.



ARISTOTLE EMERGENCY REPORT

EARTHQUAKE IN BALANGONAN, PHILIPPINES

POSSIBLE SMALL TSUNAMI

MAIN DETAILS

Area	ASIA and OCEANIA	Operation mode	Proactive
Event start	28 April 2017, 20:23 UTC	Event end	-
Report created	28 April 2017, 21:48 UTC	Report finalized	28 April 2017, 23:45 UTC

EXECUTIVE SUMMARY

- A STRONG earthquake with magnitude 6.8 occurred on Fri Apr 28 20:23:17 2017 (UTC) with latitude 5.52°N, longitude 125.06°E and depth of 26.0 km. This is a offshore event at a distance of 26.3 km from the nearest coast.
- There is nothing to evidence for the multi-hazard assessment since the weather forecast does not show any deterioration within the next five days, there is no volcano erupting and only small <0.3 m variations of the tide level might be generated. This is supported also by numerical modeling (0.2 m max).
- The population exposed to the largest earthquake shaking (VI-VII degree Mercalli Intensity) along the coast amounts to 12k people. Degree VI-VII MMI can result in generally in minor damage although it depends on the vulnerability of the constructions.
- The closest villages that experienced the largest shaking are Baliton (MMI VII, 6k) Burinas (MMI VII, 4k) and Pangyan (MMI VII, 5k). The largest cities above 100k are General Santos (680k, 67 km), Koronadal (125k, 111 km). There are no dams and nuclear plants
- Given the distance (see table in the additional material) of the largest cities (i.e., >100k people) and the level of ground shaking experienced, it is unlikely any major damage and as a result casualties
- Orange alert level for Aristotle

GEOGRAPHICAL LOCATION

MINDANAO, PHILIPPINES: 5.52 125.06
Magnitude 6.8, depth 26 km

OVERALL IMPACT

Medium

LACK OF COPING CAPACITY

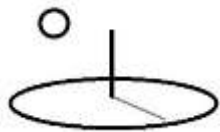
Philippines 6.8

ALERT LEVEL

Likelihood of major impact	High			
	Medium			X
	Low			
Required Resources	Sub-national	National	International	

D. Risk and Impact Identification

The following 2 datasets demonstrate one approach that RCA Assessors can use to carry out these assessments. This approach focuses on harmonizing the approaches taken with the CTAP



DevTreks –social budgeting that improves lives and livelihoods

algorithms with the most important features of this algorithm –population epidemiology or stakeholder impact, and equity.

Dataset 1. Watershed Disaster Risk Management Plans

The CTAP reference datasets (i.e. typhoons, earthquakes, floods, droughts, related risk management indexes) are summarized in related SDRR-related Targets and Indicators that impact specific stakeholder groups. For this example, Example 5's 2nd dataset is stylistically modified for these purposes (i.e. disaster losses are being illustrated but not measured). Key differences from Example 5 include:

1st TEXT dataset. Demographics and SDRR Indicators (CTAP-like indexes): The Indicators derive from CTAP-like Categorical Indexes. Many of the following examples of SDRR Indexes and Indicators from UNIDSR (2017a) can be derived directly from the CTAP datasets. However some CTAP or CTAP-like datasets may need further harmonization, or modification, to make them fully compatible with the SDRR (i.e. INFORM (2018) demonstrates how to harmonize existing datasets with SDG and SDRR targets). The CTAP and SPA2 references demonstrate that several algorithms related to MCDA (i.e. subalgorithms 10, 11, 12 and 13) also support direct measurement of these Indexes and Indicators.

Target A. Mortality. Substantially reduce global disaster mortality by 2030, measured as the number of deaths and missing persons attributed to disasters, per 100,000 population.

Indicator 1. Number of deaths attributed to disasters, per 100,000 population.

Indicator 2. Number of missing persons attributed to disasters, per 100,000 population.

Target B. Affected People. Substantially reduce the number of affected people globally by 2030, measured as the number of directly affected people attributed to disasters, per 100,000 population.

Indicator 1. Number of injured or ill people attributed to disasters, per 100,000 population, per 100,000 population.

Indicator 2 to 4: ...



DevTreks –social budgeting that improves lives and livelihoods

Target C. Economic Loss. Reduce direct disaster economic loss in relation to global gross domestic product (GDP) by 2030, measured as the direct economic loss attributed to disasters in relation to global gross domestic product

Indicator 1. Direct agricultural loss attributed to disasters

Indicator 2 to 6: ...

Target D. Infrastructure Damage. Substantially reduce disaster damage to critical infrastructure and disruption of basic services, among them health and educational facilities, including through developing their resilience by 2030, measured as the damage to critical infrastructure attributed to disasters.

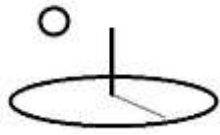
Indicator 1. Number of destroyed or damaged health facilities attributed to disasters.

Indicator 2 to 8: ...

Target F. Total International Support. Total official international support, (official development assistance (ODA) plus other official flows), for national disaster risk reduction actions.

Target F is an example of Sendai Targets and Indicators that must be adapted to work at landscape and business scale. For example, a better Target for watershed planning might be Total State or Province Support.

- **Categorical Indexes (CIs for specific stakeholder groups):** This plan uses exactly the same stakeholder groups and land use types as Example 5's 1st dataset.
- **Location Indexes (LIs for specific hazards) and Total Risk Indexes:** Annex 1 in UNIDSR (2017a) demonstrates the use of a Hazard Classification System to fully identify all of the hazards that can impact targeted stakeholders. This example uses this classification system for Total Risk Indexes (i.e. Geophysical, Hydrological) and Locational Indexes (Earthquakes, Tsunamis). Alternatively, the Hotspots Life Cycle Stages recorded in the 2nd dataset can be used for this purpose.
- **Indicator Metadata:** Most of the properties displayed in each Indicator (i.e. Q1 ... QTMost) are generated automatically and calculated as the summation of the all of the TR Indexes in a dataset.



DevTrek –social budgeting that improves lives and livelihoods

label	locatio	risks_and_impacts	factor1	factor2	factor3	factor4	factor5	factor6	factor7	factor8	factor9	factor10	factor11
A	0	HH Characteristics 1	poverty_in dex	age_hh	age_sp	educ_hh	educ_sp	female_hh	indigenous	hhsz	dirfloor	bathroom	land
A1A	1	High Poverty Index HHS	44.33	40.69	35.86	2.68	2.52	6.47	43.32	6.20	76.93	52.03	1.68
IF1A	1	Target A. Mortality	75	number of deaths and missing person	37.5	135	lower and upper 80% CI	2	3.00	normal	HCA	prevention	1/1/2018
IF1B	1	Target B. Affected People	80	number of directly affected people	60	150.00	Lower and upper estimates	2	3.00	normal	EDE	prevention	1/1/2018
IF1C	1	Target C. Economic Loss	60	direct economic loss	40	75.00	lower and upper 80% CI	2	3.00	normal	EDF	prevention	1/1/2018
IF1D	1	Target D. Infrastructure Damage	60	damage to critical infrastructure	40	75.00	lower and upper 80% CI	2	3.00	normal	PCB	prevention	1/1/2018
IF2A	1	Target E. Disaster Risk Reduction Strategy	45	degree of adoption and implementation of DRR strategies	25	65.00	lower and upper 80% CI	2	3.00	normal	EDG	preparedness	1/1/2018
IF2B	1	Target F. International Support	45	degree of state support for DRR actions	25	65.00	lower and upper 80% CI	2	3.00	normal	ECB	preparedness	1/1/2018
IF2C	1	Target G. Multi-hazard Disaster Information	45	degree of multi-hazard disaster information	25	65.00	lower and upper 80% CI	2	3.00	normal	ECC	preparedness	1/1/2018
A1B	1	Medium Poverty Index HHS	59.11	48.96	42.25	2.79	2.61	11.25	33.59	4.87	59.00	62.50	1.99
IF1A	1	Target A. Mortality	75	number of deaths and missing person	37.5	135	lower and upper 80% CI	2	3.00	normal	HCA	prevention	1/1/2018
IF1B	1	Target B. Affected People	80	number of directly affected people	60	150.00	Lower and upper estimates	2	3.00	normal	EDE	prevention	1/1/2018
IF1C	1	Target C. Economic Loss	60	direct economic loss	40	75.00	lower and upper 80% CI	2	3.00	normal	EDF	prevention	1/1/2018
IF1D	1	Target D. Infrastructure Damage	60	damage to critical infrastructure	40	75.00	lower and upper 80% CI	2	3.00	normal	PCB	prevention	1/1/2018
IF2A	1	Target E. Disaster Risk Reduction Strategy	45	degree of adoption and implementation of DRR strategies	25	65.00	lower and upper 80% CI	2	3.00	normal	EDG	preparedness	1/1/2018
IF2B	1	Target F. International Support	45	degree of state support for DRR actions	25	65.00	lower and upper 80% CI	2	3.00	normal	ECB	preparedness	1/1/2018
IF2C	1	Target G. Multi-hazard Disaster Information	45	degree of multi-hazard disaster information	25	65.00	lower and upper 80% CI	2	3.00	normal	ECC	preparedness	1/1/2018
A1C	1	Low Poverty Index HHS	78.54	54.50	47.51	3.20	2.90	11.49	15.52	4.72	16.86	79.50	3.88
IF1A	1	Target A. Mortality	75	number of deaths and missing person	37.5	135	lower and upper 80% CI	2	3.00	normal	HCA	prevention	1/1/2018
IF1B	1	Target B. Affected People	80	number of directly affected people	60	150.00	Lower and upper estimates	2	3.00	normal	EDE	prevention	1/1/2018
IF1C	1	Target C. Economic Loss	60	direct economic loss	40	75.00	lower and upper 80% CI	2	3.00	normal	EDF	prevention	1/1/2018
IF1D	1	Target D. Infrastructure Damage	60	damage to critical infrastructure	40	75.00	lower and upper 80% CI	2	3.00	normal	PCB	prevention	1/1/2018
IF2A	1	Target E. Disaster Risk Reduction Strategy	45	degree of adoption and implementation of DRR strategies	25	65.00	lower and upper 80% CI	2	3.00	normal	EDG	preparedness	1/1/2018
IF2B	1	Target F. International Support	45	degree of state support for DRR actions	25	65.00	lower and upper 80% CI	2	3.00	normal	ECB	preparedness	1/1/2018
IF2C	1	Target G. Multi-hazard Disaster Information	45	degree of multi-hazard disaster information	25	65.00	lower and upper 80% CI	2	3.00	normal	ECC	preparedness	1/1/2018
SA	1	Earthquakes Treated Villages	0	0	0	0	0.00	none	0	0	0	weights	2
A1A	1	High Poverty Index HHS	44.33	40.69	35.86	2.68	2.52	6.47	43.32	6.20	76.93	52.03	1.68
IF1A	1	Target A. Mortality	75	number of deaths and missing person	37.5	135	lower and upper 80% CI	2	3.00	normal	HCA	prevention	1/1/2018
IF1B	1	Target B. Affected People	80	number of directly affected people	60	150.00	Lower and upper estimates	2	3.00	normal	EDE	prevention	1/1/2018
IF1C	1	Target C. Economic Loss	60	direct economic loss	40	75.00	lower and upper 80% CI	2	3.00	normal	EDF	prevention	1/1/2018
IF1D	1	Target D. Infrastructure Damage	60	damage to critical infrastructure	40	75.00	lower and upper 80% CI	2	3.00	normal	PCB	prevention	1/1/2018
IF2A	1	Target E. Disaster Risk Reduction Strategy	45	degree of adoption and implementation of DRR strategies	25	65.00	lower and upper 80% CI	2	3.00	normal	EDG	preparedness	1/1/2018
IF2B	1	Target F. International Support	45	degree of state support for DRR actions	25	65.00	lower and upper 80% CI	2	3.00	normal	ECB	preparedness	1/1/2018
IF2C	1	Target G. Multi-hazard Disaster Information	45	degree of multi-hazard disaster information	25	65.00	lower and upper 80% CI	2	3.00	normal	ECC	preparedness	1/1/2018
A1B	1	Medium Poverty Index HHS	59.11	48.96	42.25	2.79	2.61	11.25	33.59	4.87	59.00	62.50	1.99
IF1A	1	Target A. Mortality	75	number of deaths and missing person	37.5	135	lower and upper 80% CI	2	3.00	normal	HCA	prevention	1/1/2018
IF1B	1	Target B. Affected People	80	number of directly affected people	60	150.00	Lower and upper estimates	2	3.00	normal	EDE	prevention	1/1/2018
IF1C	1	Target C. Economic Loss	60	direct economic loss	40	75.00	lower and upper 80% CI	2	3.00	normal	EDF	prevention	1/1/2018
IF1D	1	Target D. Infrastructure Damage	60	damage to critical infrastructure	40	75.00	lower and upper 80% CI	2	3.00	normal	PCB	prevention	1/1/2018
IF2A	1	Target E. Disaster Risk Reduction Strategy	45	degree of adoption and implementation of DRR strategies	25	65.00	lower and upper 80% CI	2	3.00	normal	EDG	preparedness	1/1/2018
IF2B	1	Target F. International Support	45	degree of state support for DRR actions	25	65.00	lower and upper 80% CI	2	3.00	normal	ECB	preparedness	1/1/2018
IF2C	1	Target G. Multi-hazard Disaster Information	45	degree of multi-hazard disaster information	25	65.00	lower and upper 80% CI	2	3.00	normal	ECC	preparedness	1/1/2018
A1C	1	Low Poverty Index HHS	78.54	54.50	47.51	3.20	2.90	11.49	15.52	4.72	16.86	79.50	3.88
IF1A	1	Target A. Mortality	75	number of deaths and missing person	37.5	135	lower and upper 80% CI	2	3.00	normal	HCA	prevention	1/1/2018
IF1B	1	Target B. Affected People	80	number of directly affected people	60	150.00	Lower and upper estimates	2	3.00	normal	EDE	prevention	1/1/2018
IF1C	1	Target C. Economic Loss	60	direct economic loss	40	75.00	lower and upper 80% CI	2	3.00	normal	EDF	prevention	1/1/2018
IF1D	1	Target D. Infrastructure Damage	60	damage to critical infrastructure	40	75.00	lower and upper 80% CI	2	3.00	normal	PCB	prevention	1/1/2018
IF2A	1	Target E. Disaster Risk Reduction Strategy	45	degree of adoption and implementation of DRR strategies	25	65.00	lower and upper 80% CI	2	3.00	normal	EDG	preparedness	1/1/2018
IF2B	1	Target F. International Support	45	degree of state support for DRR actions	25	65.00	lower and upper 80% CI	2	3.00	normal	ECB	preparedness	1/1/2018
IF2C	1	Target G. Multi-hazard Disaster Information	45	degree of multi-hazard disaster information	25	65.00	lower and upper 80% CI	2	3.00	normal	ECC	preparedness	1/1/2018
SB	1	Earthquakes Non Treated Villages	0	0	0	0	0.00	none	0	0	0	weights	2
TRG	1	Geophysical Hazards	0	0	0	0	0.00	none	0	0	0	weights	2



...

IF2C	1	Target G. Multi-hazard Disaster Information	45	degree of multi-hazard disaster information	25	65.00	lower and upper 80% CI	2	3.00	normal	ECC	preparedness	1/1/2018
A1C	1	Low Poverty Index HHs	78.54	54.50	47.51	3.20	2.90	11.49	15.52	4.72	16.86	79.50	3.88
IF1A	1	Target A. Mortality	75	number of deaths and missing person	37.5	135	lower and upper 80% CI	2	3.00	normal	HCA	prevention	1/1/2018
IF1B	1	Target B. Affected People	80	number of directly affected people	60	150.00	Lower and upper estimates	2	3.00	normal	EDE	prevention	1/1/2018
IF1C	1	Target C. Economic Loss	60	direct economic loss	40	75.00	lower and upper 80% CI	2	3.00	normal	EDF	prevention	1/1/2018
IF1D	1	Target D. Infrastructure Damage	60	damage to critical infrastructure	40	75.00	lower and upper 80% CI	2	3.00	normal	PCB	prevention	1/1/2018
IF2A	1	Target E. Disaster Risk Reduction Strategy	45	degree of adoption and implementation of DRR strategies	25	65.00	lower and upper 80% CI	2	3.00	normal	EDG	preparedness	1/1/2018
IF2B	1	Target F. International Support	45	degree of state support for DRR actions	25	65.00	lower and upper 80% CI	2	3.00	normal	ECB	preparedness	1/1/2018
IF2C	1	Target G. Multi-hazard Disaster Information	45	degree of multi-hazard disaster information	25	65.00	lower and upper 80% CI	2	3.00	normal	ECC	preparedness	1/1/2018
SB	1	Droughts Non Treated Villages	0	0	0	0	0.00	none	0	0	0	weights	2
TRC	1	Climatological Hazards	0	0	0	0	0.00	none	0	0	0	weights	2

The Treatment and Non Treatment Villages in the previous table supports the use of counterfactual evidence in Impact Evaluations. In the case of disasters, the ethical consequences of that approach has to be balanced with the need to understand what works and what doesn't work. Gertler et al (2016) discuss the use of review boards and ethics committees for these purposes.

2nd TEXT dataset. Population transition states and SDRR allocations. The following dataset employs the same techniques as introduced with Example 5 -the CIs record uncertain population measurements for the 1st dataset's stakeholder groups while the CI Indicators allocate the 1st dataset's SDRR Targets to these stakeholders. Additional population algorithms may employ alternative population modeling techniques.

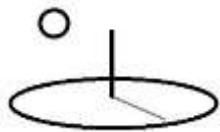


DevTrek –social budgeting that improves lives and livelihoods

label	location	risks_and_impacts	factor1	factor2	factor3	factor4	factor5	factor6	factor7	factor8	factor9	factor10	factor11
A	0												
A1A	1	High Poverty Index HHs	1/1/2018	1/1/2020	1.21E+03	population	1.09E+03	minus 10%	1.33E+03	plus 10%	none	none	none
IF1A	1	Target A. Mortality	100	125	1.20E+02	100	1.05E+02	3	4.00E+00	5	1	weights	3
IF1B	1	Target B. Affected People	25	110	1.05E+02	100	1.05E+02	3	4.00E+00	5	2	weights	3
IF1C	1	Target C. Economic Loss	25	110	1.05E+02	100	1.05E+02	3	4.00E+00	5	3	weights	3
IF1D	1	Target D. Infrastructure Damage	25	110	1.05E+02	100	1.05E+02	3	4.00E+00	5	4	weights	3
IF2A	1	Target E. Disaster Risk Reduction Strategy	25	110	1.05E+02	100	1.05E+02	3	4.00E+00	5	5	weights	3
IF2B	1	Target F. International Support	25	110	1.05E+02	100	1.05E+02	3	4.00E+00	5	0.5	weights	3
IF2C	1	Target G. Multi-hazard Disaster Information	25	110	1.05E+02	100	1.05E+02	3	4.00E+00	5	10	weights	3
A1B	1	Medium Poverty Index HHs	1/1/2018	1/1/2020	3.23E+03	population	2.91E+03	minus 10%	3.55E+03	plus 10%	none	none	none
IF1A	1	Target A. Mortality	100	125	1.20E+02	100	1.05E+02	3	4.00E+00	5	1	weights	3
IF1B	1	Target B. Affected People	25	110	1.05E+02	100	1.05E+02	3	4.00E+00	5	2	weights	3
IF1C	1	Target C. Economic Loss	25	110	1.05E+02	100	1.05E+02	3	4.00E+00	5	3	weights	3
IF1D	1	Target D. Infrastructure Damage	25	110	1.05E+02	100	1.05E+02	3	4.00E+00	5	4	weights	3
IF2A	1	Target E. Disaster Risk Reduction Strategy	25	110	1.05E+02	100	1.05E+02	3	4.00E+00	5	5	weights	3
IF2B	1	Target F. International Support	25	110	1.05E+02	100	1.05E+02	3	4.00E+00	5	0.5	weights	3
IF2C	1	Target G. Multi-hazard Disaster Information	25	110	1.05E+02	100	1.05E+02	3	4.00E+00	5	10	weights	3
A1C	1	Low Poverty Index HHs	1/1/2018	1/1/2020	5.22E+02	population	4.70E+02	minus 10%	5.74E+02	plus 10%	none	none	none
IF1A	1	Target A. Mortality	100	125	1.20E+02	100	1.05E+02	3	4.00E+00	5	1	weights	3
IF1B	1	Target B. Affected People	25	110	1.05E+02	100	1.05E+02	3	4.00E+00	5	2	weights	3
IF1C	1	Target C. Economic Loss	25	110	1.05E+02	100	1.05E+02	3	4.00E+00	5	3	weights	3
IF1D	1	Target D. Infrastructure Damage	25	110	1.05E+02	100	1.05E+02	3	4.00E+00	5	4	weights	3
IF2A	1	Target E. Disaster Risk Reduction Strategy	25	110	1.05E+02	100	1.05E+02	3	4.00E+00	5	5	weights	3
IF2B	1	Target F. International Support	25	110	1.05E+02	100	1.05E+02	3	4.00E+00	5	0.5	weights	3
IF2C	1	Target G. Multi-hazard Disaster Information	25	110	1.05E+02	100	1.05E+02	3	4.00E+00	5	10	weights	3
SA	1	Earthquakes Treated Villages	0	0	0.00E+00	0	0.00E+00	none	0.00E+00	0	0	weights	2
A1A	1	High Poverty Index HHs	1/1/2018	1/1/2020	1.21E+03	population	1.09E+03	minus 10%	1.33E+03	plus 10%	none	none	none

...

IF2C	1	Target G. Multi-hazard Disaster Information	25	110	1.05E+02	100	1.05E+02	3	4.00E+00	5	10	weights	3
A1C	1	Low Poverty Index HHs	1/1/2018	1/1/2020	5.22E+02	population	4.70E+02	minus 10%	5.74E+02	plus 10%	none	none	none
IF1A	1	Target A. Mortality	100	125	1.20E+02	100	1.05E+02	3	4.00E+00	5	1	weights	3
IF1B	1	Target B. Affected People	25	110	1.05E+02	100	1.05E+02	3	4.00E+00	5	2	weights	3
IF1C	1	Target C. Economic Loss	25	110	1.05E+02	100	1.05E+02	3	4.00E+00	5	3	weights	3
IF1D	1	Target D. Infrastructure Damage	25	110	1.05E+02	100	1.05E+02	3	4.00E+00	5	4	weights	3
IF2A	1	Target E. Disaster Risk Reduction Strategy	25	110	1.05E+02	100	1.05E+02	3	4.00E+00	5	5	weights	3
IF2B	1	Target F. International Support	25	110	1.05E+02	100	1.05E+02	3	4.00E+00	5	0.5	weights	3
IF2C	1	Target G. Multi-hazard Disaster Information	25	110	1.05E+02	100	1.05E+02	3	4.00E+00	5	10	weights	3
SA	1	Droughts Non Treated Villages	0	0	0.00E+00	0	0.00E+00	none	0.00E+00	0	0	weights	2
TRC	1	Climatological Hazards	0	0	0.00E+00	0	0.00E+00	none	0.00E+00	0	0	weights	2



DevTreks –social budgeting that improves lives and livelihoods

Indicators 2 to 7. Population Transition States using Performance Monitoring Approach:

The separate calculator Indicators measure population epidemiology using the same M&E approach introduced in Example 5. The CTAP Average Annual Loss (AAL) and risk metrics can be used for this purpose. Using Example 5’s approach, the 1st Indicator measures AALs and risks for the period 2017 to 2019 and subsequent Indicators measure the trend in losses and risks during 2020 to 2030 as new mitigation and adaptation efforts are undertaken. This approach requires including all of the probable hazards in each separate Indicator’s datasets. Example 1A in SPA2 demonstrates using SAFA-like Total Risk Indexes to report each separate hazard in each Indicator’s dataset (i.e. TRA = earthquakes, TRB = landslides, TRC = droughts, TRD = human-induced hazards caused by weakened environmental rules, ...).

Alternatively, some assessments may need to use CTAP-like hazards or hazard events in each Indicator. With the latter approach, each Indicator holds the loss data associated with specific hazards or hazard events.

Score using Impact Evaluation Approach: The Score serves as an Impact Evaluation by summarizing the losses and risks (i.e. for the period 2017 to 2030) in the 15 Performance Monitoring Indicators (Indicator 1 = 2017 to 2018, Indicator 2 = 2019 to 2020, ...).

Dataset 2. Example 6A. Business Continuity and Emergency Preparedness Plans

The Watershed Disaster Management Plan becomes the “reference plan” needed to complete separate plans for individual firms and local communities within the watershed. The plans demonstrated in Example 6A complement the watershed plan and may contain the exact same Indicators and Indexes, except they use abbreviated SIAs and measurements related to individual firms, local communities, and impacted stakeholders. Example 6A modifies Example 5’s 1st dataset for these purposes.

Advanced Disaster Risk Management

This reference [may] demonstrate the use of algorithms to train machine learning algorithms to define likely disaster impact paths. Once trained, the algorithms predict SDG and SDRR impacts on targeted stakeholders that arise from company and community service organization disaster



DevTreks –social budgeting that improves lives and livelihoods

risk reduction, and CTA-Prevention, activities. More advanced algorithms predict likely transition states and advise companies and communities how to dynamically tweak their mitigation and adaptation actions to achieve SDRR-related Targets.

E. Mitigation and Adaptation Activities and Stages

The first paragraph (Poljanšek et al, 2017) explain why the definition for CTAP in the second paragraph focuses on disaster prevention. The role of disaster risk management is to prevent disasters from causing costly losses by “adopting and diffusing” mitigation and adaptation activities that increase the resiliency and coping capacity of communities and companies.

“Based on an analysis of the benefits arising from avoided losses, mitigation and prevention measures are widely considered more cost-effective than ex post disaster interventions. An increase in mitigation investment has occurred in some ... countries, but the lack of public and therefore political interest in prevention and mitigation remains a problem.”

“**CTA-Prevention (CTAP)** is the numeric assessment of the costs and benefits of a portfolio of mitigation and adaptation interventions that prevent or correct resource stock damages.

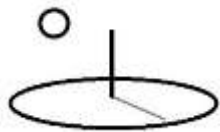
Assessments use relevant Conservation Technology Assessment (CTA) algorithms to quantify the risk and uncertainty associated with resource stock measurement and valuation.”

F. Quality of Life, or Sustainability. Scenarios

The following Scenario Analysis is used by communities and companies to more fully understand the transition risks they face from changes needed in their disaster risk management activities and social interventions to accomplish the SDRR targets.

General Scenario

Threatened Quality of Life: High GHG result in 1.5°C temperature increase with higher incidence of biodiversity loss, droughts, severe heat waves, crop and livestock production



DevTreks –social budgeting that improves lives and livelihoods

risks, air pollution, floods, migration, and social discord that leads to the incapacity to achieve the SDRR targets.

Targeted Social Performance Risks and Targets: Combinations of the SDRR targets, adapted to local stakeholders, including (1*):

- [Local] target A: Substantially reduce [local] disaster mortality by 2030
- [Local] target B: Substantially reduce the number of affected people [locally] by 2030

Mitigation and Adaptation Actions: Portfolio 1 consists of a) ..., b)...., and c)....

Sustainability Scenarios.

In terms of SDRR goal accomplishment, logical sustainability scenarios relate directly to a community’s or company’s transition to higher levels of sustainability, as illustrated by the following scenarios:

- A. Transition State _A: baseline (i.e. current disaster mortality and high economic losses from disasters)
- B. Transition State _B: low sustainability
- C. Transition State _C: medium sustainability
- D. Transition State _D: high sustainability (i.e. substantial reduction in disaster mortality and affordable economic losses from disasters)

Simple data conventions, including the use of sibling base elements and Example 1’s “_xx” labelling convention, can be used to model these more comprehensive transition scenarios. Section J, Decisions, demonstrates how to use the Reference Case Cost Effectiveness Results introduced in Example 4B to assist making decisions about the degree of sustainability, or stage of sustainability, appropriate for specific communities and companies.

Example 7 may demonstrate how machine learning algorithms anticipate transition states, given current circumstances, and tailor sustainability recommendations to specific scenarios, based on their assessment of the company and community’s most probable sustainable development paths.

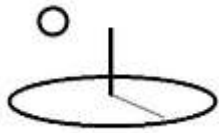


DevTreks –social budgeting that improves lives and livelihoods

E. Social Performance Score

Dataset 1. Watershed Disaster Risk Management Plans

The following images show the properties of the 1st Indicator. The SDRR Indicators used in this dataset derive from results summarized from CTAP or CTAP-like algorithms. In this example, the first Total Risk Index, or TRG, measures disaster risks associated with earthquakes while the second, or TRC, measures droughts. The MathResults properties show the correct convention of storing the results in TEXT files identified by URLs derived from the Resources Application.



Monitoring and Evaluation 2 Output Calculator

M and E Indicators

Indicator 1

DRM 2017 to 2018

Indicator 1 Description
Example 6 in Social Performance Analysis 3.

Indicator 1 URL
http://localhost:5000/resources/network_carbon/resourcepack_545/resource_1939/Ind1-Watershed2.csv;http://localhost:5000/resources/network_carbon/resourcepack_545/resource_1938/Ind1-Watershed1.csv

Label 1	Rel Label 1
DRM1	none
Date 1	Dist Type 1
03/07/2018	none
Q1 1	Q1 Unit 1
7.9256	benchmark most score
Q2 1	Q2 Unit 1
4.5186	benchmark low score
Q3 1	Q3 Unit 1
13.0032	benchmark high score
Q4 1	Q4 Unit 1
8.8694	target most score
Q5 1	Q5 Unit 1
5.0567	target low score
Math Operator 1	BaseIO 1
equalto	none

QT 1	QT Unit 1
14.5517	target high score
Math Type 1	Math Sub Type 1
algorithm1	subalgorithm17
QT D1 1	QT D1 Unit 1
3.0000	actual certainty1
QT D2 1	QT D2 Unit 1
4.0000	actual certainty2
QT Most 1	QT Most Unit 1
8.4732	actual most score
QT Low 1	QT Low Unit 1
4.8308	actual low score
QT High 1	QT High Unit 1
13.9016	actual high score
Math Expression 1	
I1.Q1.factor1 + I1.Q2.factor2 + I1.Q3.factor3 + I1.Q4.factor4 +	
Math Result 1	
http://localhost:55009/resources/network_carbon/resourcepack_545/resource_1944/Ind1-Watershed-Results.csv	

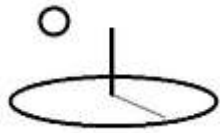
Indicator 2

DRM 2019 to 2020

Indicator 2 Description
Example 6 in Social Performance Analysis 3.

Indicator 2 URL
http://localhost:5000/resources/network_carbon/resourcepack_545/resource_1939/Ind1-Watershed2.csv;http://localhost:5000/resources/network_carbon/resourcepack_545/resource_1938/Ind1-Watershed1.csv

The MathResults appear as follows. The data is stylistic and has no meaning other than to debug the algorithm.



DevTreks –social budgeting that improves lives and livelihoods

label	locatr_risks_and_imp	factor1	factor2	factor3	factor4	factor5	factor6	factor7	factor8	factor9	factor10	factor11	qtmost	percentar	qtlow	qthigh	certainty1	certainty2	certainty3	totalcost	sdgunit	production	lifecyclestr
A	0.00 HH Characteris	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
A1A	1.00 High Poverty In	1/1/2018	1/1/2020	1210.00	population	1090.00	minus 10%	1330.00	plus 10%	none	none	none	2.82	95.53	1.61	4.63	3.00	4.00	5.00	425.34	0.00	0.00	0.00
IF1A	1.00 Target A. Mort	100.00	125.00	120.00	1210.00	1271.00	3.00	4.00	5.00	1.00 weights	3.00	1.43	96.00	0.71	2.57	3.00	4.00	5.00	90.00	number of HCA	prevention		
IF1B	1.00 Target B. Affec	25.00	110.00	105.00	1210.00	1271.00	3.00	4.00	5.00	2.00 weights	3.00	0.33	95.45	0.25	0.63	3.00	4.00	5.00	42.00	number of EDE	prevention		
IF1C	1.00 Target C. Econ	25.00	110.00	105.00	1210.00	1271.00	3.00	4.00	5.00	3.00 weights	3.00	0.25	95.45	0.17	0.31	3.00	4.00	5.00	47.25	direct ecoi EDF	prevention		
IF1D	1.00 Target D. Infra	25.00	110.00	105.00	1210.00	1271.00	3.00	4.00	5.00	4.00 weights	3.00	0.25	95.45	0.17	0.31	3.00	4.00	5.00	63.00	damage to PCB	prevention		
IF2A	1.00 Target E. Disas	25.00	110.00	105.00	1210.00	1271.00	3.00	4.00	5.00	5.00 weights	3.00	0.19	95.45	0.10	0.27	3.00	4.00	5.00	59.06	degree of i EDG	preparedn		
IF2B	1.00 Target F. Interr	25.00	110.00	105.00	1210.00	1271.00	3.00	4.00	5.00	0.50 weights	3.00	0.19	95.45	0.10	0.27	3.00	4.00	5.00	5.91	degree of i ECB	preparedn		
IF2C	1.00 Target G. Multi	25.00	110.00	105.00	1210.00	1271.00	3.00	4.00	5.00	10.00 weights	3.00	0.19	95.45	0.10	0.27	3.00	4.00	5.00	118.13	degree of i ECC	preparedn		
A1B	1.00 Medium Pove	1/1/2018	1/1/2020	3230.00	population	2910.00	minus 10%	3550.00	plus 10%	none	none	none	2.82	95.53	1.61	4.63	3.00	4.00	5.00	425.34	0.00	0.00	0.00
IF1A	1.00 Target A. Mort	100.00	125.00	120.00	3230.00	3392.00	3.00	4.00	5.00	1.00 weights	3.00	1.43	96.00	0.71	2.57	3.00	4.00	5.00	90.00	number of HCA	prevention		
IF1B	1.00 Target B. Affec	25.00	110.00	105.00	3230.00	3392.00	3.00	4.00	5.00	2.00 weights	3.00	0.33	95.45	0.25	0.63	3.00	4.00	5.00	42.00	number of EDE	prevention		
IF1C	1.00 Target C. Econ	25.00	110.00	105.00	3230.00	3392.00	3.00	4.00	5.00	3.00 weights	3.00	0.25	95.45	0.17	0.31	3.00	4.00	5.00	47.25	direct ecoi EDF	prevention		
IF1D	1.00 Target D. Infra	25.00	110.00	105.00	3230.00	3392.00	3.00	4.00	5.00	4.00 weights	3.00	0.25	95.45	0.17	0.31	3.00	4.00	5.00	63.00	damage to PCB	prevention		
IF2A	1.00 Target E. Disas	25.00	110.00	105.00	3230.00	3392.00	3.00	4.00	5.00	5.00 weights	3.00	0.19	95.45	0.10	0.27	3.00	4.00	5.00	59.06	degree of i EDG	preparedn		
IF2B	1.00 Target F. Interr	25.00	110.00	105.00	3230.00	3392.00	3.00	4.00	5.00	0.50 weights	3.00	0.19	95.45	0.10	0.27	3.00	4.00	5.00	5.91	degree of i ECB	preparedn		
IF2C	1.00 Target G. Multi	25.00	110.00	105.00	3230.00	3392.00	3.00	4.00	5.00	10.00 weights	3.00	0.19	95.45	0.10	0.27	3.00	4.00	5.00	118.13	degree of i ECC	preparedn		
A1C	1.00 Low Poverty In	1/1/2018	1/1/2020	522.00	population	470.00	minus 10%	574.00	plus 10%	none	none	none	2.82	95.53	1.61	4.63	3.00	4.00	5.00	425.34	0.00	0.00	0.00
IF1A	1.00 Target A. Mort	100.00	125.00	120.00	522.00	548.10	3.00	4.00	5.00	1.00 weights	3.00	1.43	96.00	0.71	2.57	3.00	4.00	5.00	90.00	number of HCA	prevention		
IF1B	1.00 Target B. Affec	25.00	110.00	105.00	522.00	548.10	3.00	4.00	5.00	2.00 weights	3.00	0.33	95.45	0.25	0.63	3.00	4.00	5.00	42.00	number of EDE	prevention		
IF1C	1.00 Target C. Econ	25.00	110.00	105.00	522.00	548.10	3.00	4.00	5.00	3.00 weights	3.00	0.25	95.45	0.17	0.31	3.00	4.00	5.00	47.25	direct ecoi EDF	prevention		
IF1D	1.00 Target D. Infra	25.00	110.00	105.00	522.00	548.10	3.00	4.00	5.00	4.00 weights	3.00	0.25	95.45	0.17	0.31	3.00	4.00	5.00	63.00	damage to PCB	prevention		
IF2A	1.00 Target E. Disas	25.00	110.00	105.00	522.00	548.10	3.00	4.00	5.00	5.00 weights	3.00	0.19	95.45	0.10	0.27	3.00	4.00	5.00	59.06	degree of i EDG	preparedn		
IF2B	1.00 Target F. Interr	25.00	110.00	105.00	522.00	548.10	3.00	4.00	5.00	0.50 weights	3.00	0.19	95.45	0.10	0.27	3.00	4.00	5.00	5.91	degree of i ECB	preparedn		
IF2C	1.00 Target G. Multi	25.00	110.00	105.00	522.00	548.10	3.00	4.00	5.00	10.00 weights	3.00	0.19	95.45	0.10	0.27	3.00	4.00	5.00	118.13	degree of i ECC	preparedn		
SA	1.00 Earthquakes T	0.00	0.00	0.00	0.00	0.00	none	0.00	0.00	0.00 weights	2.00	8.47	95.53	4.83	13.90	3.00	4.00	5.00	1276.03	0.00	0.00	0.00	
A1A	1.00 High Poverty In	1/1/2018	1/1/2020	1210.00	population	1090.00	minus 10%	1330.00	plus 10%	none	none	none	2.82	95.53	1.61	4.63	3.00	4.00	5.00	425.34	0.00	0.00	0.00
IF1A	1.00 Target A. Mort	100.00	125.00	120.00	1210.00	1271.00	3.00	4.00	5.00	1.00 weights	3.00	1.43	96.00	0.71	2.57	3.00	4.00	5.00	90.00	number of HCA	prevention		

...

A1C	1.00 Low Poverty In	43101.00	43831.00	522.00	population	470.00	minus 10%	574.00	plus 10%	none	none	none	2.82	95.53	1.61	4.63	3.00	4.00	5.00	425.34	0.00	0.00	0.00
IF1A	1.00 Target A. Mort	100.00	125.00	120.00	522.00	548.10	3.00	4.00	5.00	1.00 weights	3.00	1.43	96.00	0.71	2.57	3.00	4.00	5.00	90.00	number of HCA	prevention		
IF1B	1.00 Target B. Affec	25.00	110.00	105.00	522.00	548.10	3.00	4.00	5.00	2.00 weights	3.00	0.33	95.45	0.25	0.63	3.00	4.00	5.00	42.00	number of EDE	prevention		
IF1C	1.00 Target C. Econ	25.00	110.00	105.00	522.00	548.10	3.00	4.00	5.00	3.00 weights	3.00	0.25	95.45	0.17	0.31	3.00	4.00	5.00	47.25	direct ecoi EDF	prevention		
IF1D	1.00 Target D. Infra	25.00	110.00	105.00	522.00	548.10	3.00	4.00	5.00	4.00 weights	3.00	0.25	95.45	0.17	0.31	3.00	4.00	5.00	63.00	damage to PCB	prevention		
IF2A	1.00 Target E. Disas	25.00	110.00	105.00	522.00	548.10	3.00	4.00	5.00	5.00 weights	3.00	0.19	95.45	0.10	0.27	3.00	4.00	5.00	59.06	degree of i EDG	preparedn		
IF2B	1.00 Target F. Interr	25.00	110.00	105.00	522.00	548.10	3.00	4.00	5.00	0.50 weights	3.00	0.19	95.45	0.10	0.27	3.00	4.00	5.00	5.91	degree of i ECB	preparedn		
IF2C	1.00 Target G. Multi	25.00	110.00	105.00	522.00	548.10	3.00	4.00	5.00	10.00 weights	3.00	0.19	95.45	0.10	0.27	3.00	4.00	5.00	118.13	degree of i ECC	preparedn		
SB	1.00 Droughts Non	0.00	0.00	0.00	0.00	0.00	none	0.00	0.00	0.00 weights	2.00	8.47	95.53	4.83	13.90	3.00	4.00	5.00	1276.03	0.00	0.00	0.00	
TRC	1.00 Climatological	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	4.24	95.53	2.42	6.95	3.00	4.00	5.00	2552.06	0.00	0.00	0.00

Indicators 2 to 7. Population Transition States using Performance Monitoring Approach.

The following image displays the results for the 2 year Performance Monitoring periods

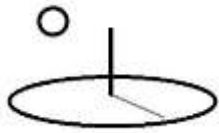
contained in 6 additional Indicators (6*). For convenience, the same datasets were used for the 7 Indicators and the Score.



DevTreks –social budgeting that improves lives and livelihoods

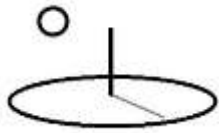
Name (N)		Label	Date	Rel Label	Math Type	Dist Type	Base IO	Math Operator	Math Sub Type
Q1 Amount	Q1 Unit	Q2 Amount	Q2 Unit	Q3 Amount	Q3 Unit	Q4 Amount	Q4 Unit	Q5 Amount	Q5 Unit
QT Amount	QT Unit	QT D1 Amount	QT D2 Amount	QT Most Amount	QT Most Unit	QT Low Amount	QT Low Unit	QT High Amount	QT High Unit
DRM Evaluation		Score1	03/07/2018	none	algorithm1	none	none	equalto	subalgorithm17
14.5517	target high score	3.0000	4.0000	8.4732	actual most score	4.8308	actual low score	13.9016	actual high score
I1.QTM									
DRM 2017 to 2018		DRM1	03/07/2018	none	algorithm1	none	none	equalto	subalgorithm17
7.9256	benchmark most score	4.5186	benchmark low score	13.0032	benchmark high score	8.8694	target most score	5.0567	target low score
14.5517	target high score	3.0000	4.0000	8.4732	actual most score	4.8308	actual low score	13.9016	actual high score
Example 6 in Social Performance Analysis 3.									
I1.Q1.factor1 + I1.Q2.factor2 + I1.Q3.factor3 + I1.Q4.factor4 + I1.Q5.factor5 + I1.Q6.factor6 + I1.Q7.factor7 + I1.Q8.factor8 + I1.Q9.factor9 + I1.Q10.factor10 + I1.Q11.factor11									
DRM 2019 to 2020		DRM2	03/07/2020	none	algorithm1	none	none	equalto	subalgorithm17
7.9256	benchmark most score	4.5186	benchmark low score	13.0032	benchmark high score	8.8694	target most score	5.0567	target low score
14.5517	target high score	3.0000	4.0000	8.4732	actual most score	4.8308	actual low score	13.9016	actual high score
Example 6 in Social Performance Analysis 3.									
I2.Q1.factor1 + I2.Q2.factor2 + I2.Q3.factor3 + I2.Q4.factor4 + I2.Q5.factor5 + I2.Q6.factor6 + I2.Q7.factor7 + I2.Q8.factor8 + I2.Q9.factor9 + I2.Q10.factor10 + I2.Q11.factor11									
DRM 2021 to 2022		DRM3	03/07/2022	none	algorithm1	none	none	equalto	subalgorithm17
7.9256	benchmark most score	4.5186	benchmark low score	13.0032	benchmark high score	8.8694	target most score	5.0567	target low score
14.5517	target high	3.0000	4.0000	8.4732	actual	4.8308	actual	13.9016	actual high

Score using Impact Evaluation Approach. The following Score properties show that the Score employs the same TEXT dataset as Indicators, but support longer term Impact Evaluations.



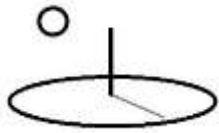
DevTreks –social budgeting that improves lives and livelihoods

Although Scores only partially display the “target-actual” properties returned in Indicators, the raw MathResults contain all of that data for reporting. For testing purposes, Indicator 1’s datasets were reused in this Score. Resource Stock and M&E Analyzers, as explained in their respective tutorials, can be used to further analyze all of the “actual” scores for both the Score and the Indicators (i.e. Totals, Statistics, Change Bys, and Progress, Analyzers).



Score		Most Likely	Most Unit
DRM Evaluation		8.4732	actual most score
Score Math Expression		Low Estimate	Low Unit
I1.QTM		4.8308	actual low score
Label	Rel Label	High Estimate	High Unit
Score1	none	13.9016	actual high score
Total Score	Score Unit	Math Type	Math Sub Type
14.5517	target high score	algorithm1	subalgorithm17
D1	D1 Unit	Math Result	
3.0000	actual certainty1	http://localhost:55009/resources/network_carbon/resourcepack_545/resource_1951/Score-Watershed-Results.csv	
D2	D2 Unit	Score URL	
4.0000	actual certainty2	none	
Date	Dist Type	Calculations Description	
03/07/2018	none	Example 6 in SPA3.v214b	
Iterations	Confidence Interval	Media URL	
500	80	http://localhost:5000/resources/network_carbon/resourcepack_537/resource_1940/UNISDR1.PNG	
Random Seed	BaseIO	Data URL	
6	none	http://localhost:5000/resources/network_carbon/resourcepack_545/resource_1943/Ind1-	
Math Operator		Output Group: RCA Output Examples	
equalto		Output : Disaster Risk Management, Example 6	
Most Likely	Most Unit	Indicators Details	
8.4732	actual most score	Indic 0 Name: DRM Evaluation Label: Score1	
Low Estimate	Low Unit	Date: 03/07/2018 Rel Label: none	
4.8308	actual low score	Math Type: algorithm1 Dist Type: none	
High Estimate	High Unit	Math Sub Type: subalgorithm17 Base IO: none	
13.9016	actual high score	Math Express: I1.QTM Math Operator: equalto	
Math Type	Math Sub Type		

Example 6A demonstrates additional approaches, including the use of related algorithms, for using Scores to support disaster risk management decisions.



DevTreks –social budgeting that improves lives and livelihoods

F. Decisions

The following image (UNISDR, 2017b) introduces the use of scoring criteria in national risk assessment efforts. The image shows that UNISDR (2017b) recommends using the 3rd risk factor to describe the “level of uncertainty” of the first 2 risk factors.

Box 16 – Risk matrix or multi-criteria impact and likelihood analysis

A risk matrix is the common template presenting the impacts, likelihoods of risks and confidence levels. This matrix projects the outcomes of the impact and likelihood analysis onto two axes. The visual tool helps decision makers see a clear distinction between impact and likelihood levels.

The actual impact and likelihood levels of different national risks represented in the risk matrix are the result of a single-hazard risk analysis.

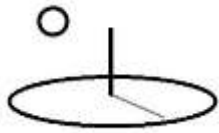
Likelihood	Consequence level				
	Insignificant	Minor	Moderate	Major	Catastrophic
Almost certain	Medium	Medium	High	Extreme	Extreme
Likely	Low	Medium	High	Extreme	Extreme
Unlikely	Low	Low	Medium	High	Extreme
Rare	Very low	Low	Medium	High	High
Very rare	Very low	Very low	Low	Medium	High
Extremely rare	Very low	Very low	Low	Medium	High

Another aspect that needs to be presented is the confidence level or level of uncertainty. The confidence level is an additional level of information supporting decision-making. Various methods to add this element to the risk matrix have been used in different countries and national guidelines.

Likelihood	Consequence				
	Insignificant	Minor	Moderate	Major	Catastrophic
Almost certain	4	4	3	2	1
Likely	5	4	4	2	2
Unlikely	5	5	4	3	2
Rare	5	5	5	3	3
Very rare	5	5	5	4	3
Extremely rare	5	5	5	4	4

Likelihood	Consequence				
	Insignificant	Minor	Moderate	Major	Catastrophic
Almost certain	3	2	1	1	1
Likely	3	3	2	1	1
Unlikely	4	3	2	1	1
Rare	4	4	3	2	1
Very rare	5	4	3	2	2
Extremely rare	5	5	4	3	2

The image above shows two risk matrices with priority levels. The darker colour and number one denote the highest priority and the lighter colour and number 5 the lowest. On top the



DevTreks –social budgeting that improves lives and livelihoods

Example 5 used the 3rd “certainty”, or risk factor, for measuring the degree of responsibility for the disaster risk by a “driver” of hazards, such as a company or community. This risk factor focuses on further drivers, or causes, of human-induced disasters. For example, community leaders, or more accurately, misinformed leaders, who fail to adequately manage disasters (i.e. by weakening environmental laws or allowing special interest groups to pillage public lands) must be held accountable for whatever carnage ensues (7*).

SDG and SDRR Accounting Standards Board Scorecards (SDG-ASB)

The following 4 Scorecards illustrate the use of SDRR metrics in the same manner as Example 5’s SDG metrics.

Full Sendai DRR Monitoring and Evaluation Report							
indicator	1	2	3	4	5	6	7 Score
period	2017-2018	2019-2020	2021-2022	2023-2024	2025-2026	2027-2028	2029-2030 2017-2030
Sendai DRR Target 1. Mortality							
sdg target							
most likely sdg actual							
sdg percent							
low sdg actual							
high sdg actual							
certainty 1							
certainty 2							
certainty 3							
totalcost							
Score							xx
Sendai DRR Target 2. Affected People							
sdg target							
most likely sdg actual							
sdg percent							
low sdg actual							
high sdg actual							
certainty 1							
certainty 2							
certainty 3							
totalcost							
Score							xx
Sendai DRR Target 3. Economic Loss							
sdg target							
...							
Total SDG Score							xx



DevTreks –social budgeting that improves lives and livelihoods

The CTAP algorithms demonstrate direct measurement of monetary benefits and costs. The following Reference Case CEA is based on these benefit and cost metrics. Benefit metrics for many disaster-related assessments, as documented in Column 4 and explained in Example 4B, are based on QALYs, DALYs, or QASYs.

Reference Case Cost Effectiveness Results (Time Horizon = 22 years; Costs and Sendai DRRs Discount Rate = 3%; Escalation Type = none; CEA Threshold = 100,000;										
Sendai DRR Impacts	Most Likely Total Costs	Uncertainty Index (likelihood)	Most Likely Sendai DRRs	Uncertainty Index (consequence)	Uncertainty Index (responsibility)	Incremental Costs	Incremental Effectiveness	NMB	Incremental NMB	ICER (Inc Cost / Inc Effects)
Stakeholder Group A Perspective										
Indicator 1. 2017 to 2018	x	x	x	x	x	x	x	x	x	x
Indicator 2. 2019 to 2020	x	x	x	x	x	x	x	x	x	x
Indicator 3. 2021 to 2022	x	x	x	x	x	x	x	x	x	x
Indicator 4. 2023 to 2024	x	x	x	x	x	x	x	x	x	x
Indicator 5. 2025 to 2026	x	x	x	x	x	x	x	x	x	x
Indicator 6. 2027 to 2028	x	x	x	x	x	x	x	x	x	x
Indicator 7. 2029 to 2030	x	x	x	x	x	x	x	x	x	x
Benchmark Indicator 1 to 2030										
Score. 2017 to 2030	x	x	x	x	x	x	x	x	x	x
Private Sector Perspective										
Indicator 1. 2017 to 2018	21,000.00	4.00	16.50	4.50	3.00			1,629,000.00		
Indicator 2. 2019 to 2020	23,100.00	3.50	18.98	2.50	3.00	21,000.00	18.15	1,874,400.00	245,400.00	1,157.02
Indicator 3. 2021 to 2022	25,200.00	3.00	21.45	2.00	5.00	42,000.00	34.65	2,119,800.00	245,400.00	1,212.12
Indicator 4. 2023 to 2024	26,250.00	4.25	21.45	4.50	1.00	63,000.00	51.15	2,118,750.00	(1,050.00)	1,231.67
Indicator 5. 2025 to 2026	x	x	x	x	x	x	x	x	x	x
Indicator 6. 2027 to 2028	x	x	x	x	x	x	x	x	x	x
Indicator 7. 2029 to 2030	x	x	x	x	x	x	x	x	x	x
Benchmark Indicator 1 to 2030										
Score. 2017 to 2030	x	x	x	x	x	x	x	x	x	x
Societal Perspective										
Indicator 1. 2017 to 2018	x	x	x	x	x	x	x	x	x	x
Indicator 2. 2019 to 2020	x	x	x	x	x	x	x	x	x	x
Indicator 3. 2021 to 2022	x	x	x	x	x	x	x	x	x	x
Indicator 4. 2023 to 2024	x	x	x	x	x	x	x	x	x	x
Indicator 5. 2025 to 2026	x	x	x	x	x	x	x	x	x	x
Indicator 6. 2027 to 2028	x	x	x	x	x	x	x	x	x	x
Indicator 7. 2029 to 2030	x	x	x	x	x	x	x	x	x	x
Benchmark Indicator 1 to 2030										
Score. 2017 to 2030	x	x	x	x	x	x	x	x	x	x

Reference Case Cost Effectiveness Results (Time Horizon = 22 years; Costs and Sendai DRRs Discount Rate = 3%; Escalation Type = none; CEA Threshold = 100,000;											
Sendai DRR Impacts	Most Likely Total Costs	Uncertainty Index (likelihood)	Most Likely Sendai DRRs	Uncertainty Index (consequence)	Uncertainty Index (responsibility)	Incremental Costs	Incremental Effectiveness	NMB	Incremental NMB	ICER (Inc Cost / Inc Effects)	Score
Stakeholder Group A Perspective											
No Mitigation and Adaptation	x	x	x	x	x	x	x	x	x	x	x
Scenario AA	x	x	x	x	x	x	x	x	x	x	x
Scenario AB	x	x	x	x	x	x	x	x	x	x	x
Scenario AC	x	x	x	x	x	x	x	x	x	x	x
Private Sector Perspective											
No Mitigation and Adaptation	x	x	x	x	x	x	x	x	x	x	x
Scenario AA	x	x	x	x	x	x	x	x	x	x	x
Scenario AB	x	x	x	x	x	x	x	x	x	x	x
Scenario AC	x	x	x	x	x	x	x	x	x	x	x
Societal Perspective											
No Mitigation and Adaptation	x	x	x	x	x	x	x	x	x	x	x
Scenario AA	x	x	x	x	x	x	x	x	x	x	x
Scenario AB	x	x	x	x	x	x	x	x	x	x	x
Scenario AC	x	x	x	x	x	x	x	x	x	x	x



DevTreks –social budgeting that improves lives and livelihoods

Sendai DRR Hotspots Report	Sendai DRR Targets							Score
	Mortality	Affected People	Economic Loss	Infrastructure Damage	Disaster Risk Reduction Strategy	State Support	Multi-hazard Disaster Information	
Coping and Resilience Process 1	%	%	%	%	%	%	%	%
Coping and Resilience Process 2	%	%	%	%	%	%	%	%
Coping and Resilience Process 3	%	%	%	%	%	%	%	%
Coping and Resilience Process 4	%	%	%	%	%	%	%	%
Coping and Resilience Process 5	%	%	%	%	%	%	%	%
Population Epidemiology Stage 1 Total	%	%	%	%	%	%	%	%
Coping and Resilience Process 1	%	%	%	%	%	%	%	%
Coping and Resilience Process 2	%	%	%	%	%	%	%	%
Coping and Resilience Process 3	%	%	%	%	%	%	%	%
Coping and Resilience Process 4	%	%	%	%	%	%	%	%
Coping and Resilience Process 5	%	%	%	%	%	%	%	%
Population Epidemiology Stage 2 Total	%	%	%	%	%	%	%	%
High Poverty Index HHs Totals	%	%	%	%	%	%	%	%

The evaluation question being answered by the following Business Value Report is:

To what extent has the [company] [organization] [community] been effective and efficient in achieving SDRR-related targets that are relevant to our stakeholders?

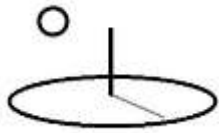
Example 7 [may] demonstrate the use of machine learning algorithms to codify this report.

Business Value Report	Drivers	Hazards	Exposure	Vulnerability	Capacity	Impacts	Score
Societal Perspective							
Mortality	x	x	x	x	x	x	x
Affected People	x	x	x	x	x	x	x
Economic Loss	x	x	x	x	x	x	x
Disaster Risk Reduction Index	x	x	x	x	x	x	x
Disaster Risk Reduction S	x	x	x	x	x	x	x
State Support	x	x	x	x	x	x	x
Multi-hazard Disaster Inf	x	x	x	x	x	x	x
Stakeholder Group A Perspective							

Industry Perspective							

G. Communication

The following image demonstrates using a calculator’s Media View to communicate the Math Results to keep impacted stakeholders informed.



DevTreks –social budgeting that improves lives and livelihoods

GreenTreks	Search	Preview	Select
Edit	Pack	Views	Club

← Select PackIt →

← Edit Linked Views Make base →

M and E Output 2 Calcs--- Get

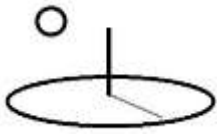
Media Mobile Desktop

Dataset: [Disaster Risk Management, Example 6 IRI](#) Example 6 in the Social Performance Analysis 3 reference.

UNISDR1

Download

One of the most comprehensive examples of communicating disaster risk information relates to the INFORM project summarized in the following image (INFORM, 2018) (4*). This tutorial



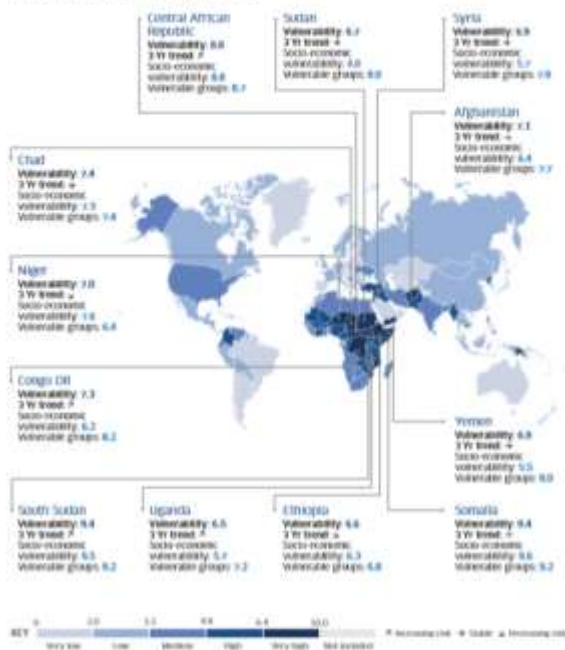
DevTrek – social budgeting that improves lives and livelihoods

demonstrates that INFORM-like MCDA techniques can be adapted to work in a similar fashion at landscape and community, or “sub-national”, scale. Example 10 in the SDG Plan reference, demonstrates using a similar approach for assessing the feasibility of mitigation and adaptation technologies for local communities and local businesses within watersheds or landscapes.

VULNERABILITY

This dimension of INFORM measures the susceptibility of people to potential hazards. It is made up of two categories – socio-economic vulnerability and vulnerable groups. This map shows details for the 12 countries with the highest values in the vulnerability dimension.

INFORM 2018 Vulnerability dimension



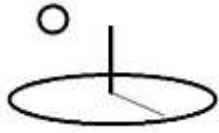
LACK OF COPING CAPACITY

This dimension of INFORM measures the lack of resources available that can help people cope with hazardous events. It is made up of two categories – institutions and infrastructure. This map shows details for the 12 countries with the highest values in the lack of coping capacity dimension.

INFORM 2018 Lack of coping capacity dimension



The following image (UNISDR, 2017b) summarizes the types of communication aids commonly used to communicate disaster management information to stakeholders. CTAP introduces algorithms that contain the raw data needed by several of these communication aids.



The following tools and formats can be used to report the results for different purposes:

- Geospatial tools and mapping for hazards and risk information. For further information, see Module G on the use of geospatial data.
- Risk matrices for comparison and showcasing prioritization of risk levels to decision makers are a common format to use for all risks. See Box 16 in Element 6.
- Scenario information depicted in maps and infographics, as well as preparation of exercises,
- to raise awareness among the general public.
- Brief snapshots of risk values, trends, with and without possible disaster risk reduction policies (if identified in the scoping step) and main findings for communicating with high-level decision makers.

³¹ Risk communication is a two-way process. This is reflected in element 9. Element 8 focuses only on the preparation of the results for two-way communication.

-
- Exceedance probability curves to communicate the concept of risk layers for disaster risk reduction actions.

H. Performance Monitoring and Impact Evaluation (M&E)

The references largely discuss disaster risk management M&E systems in terms of Internet of Things (IoT) device metrics, such as ocean wave height measurements, earthquake tremors, snow quantity, and soil moisture levels. The references discuss how many of these systems are highly developed and used currently for real life disaster planning purposes (i.e. global tsunami



DevTreks –social budgeting that improves lives and livelihoods

monitoring systems, radiation discharge networks). These IoT systems add an important, missing, dimension to the M&E accounting systems introduced in this tutorial. They suggest that additional interfaces are needed that allow both M&E systems to communicate dynamically with one another to advance the overall, disaster risk management, objective. This reference [may] introduce additional algorithms for this purpose.

This reference focuses on disaster risk management Performance Monitoring and Impact Evaluation accounting systems. CTAP introduced 3 examples showing how some of these systems work for disaster risk management. For example, the following image (Khazai et al, 2015) introduces 3 Indexes related to these purposes.

CHAPTER

01 General notions of urban risk and resilience

02 Urban Disaster Risk Index (UDRI)

A risk communication tool which provides a holistic view of disaster risk by capturing through indices both the direct physical damages of buildings and infrastructure, as well as considering social vulnerability or lack of resilience that can aggravate the physical effects.

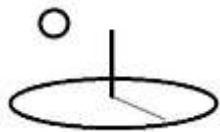
03 Risk Management Index (RMI)

A risk management tool which brings together a group of indicators that measure a city's risk management performance, reflecting organizational, development, capacity and institutional actions taken to reduce risk, to prepare for crisis and to recover efficiently from disasters.

04 Disaster Resilience Index (DRI)

A monitoring and evaluation tool for benchmarking and measuring progress (or lack of progress) along a city's key development policies and processes for mainstreaming of risk reduction and increasing resilience.

I. Conclusions



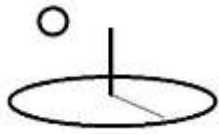
DevTreks –social budgeting that improves lives and livelihoods

The World Bank (2017) use the following statement to summarize the role played by Performance Management frameworks, such as the RCA Framework that underlies this tutorial, in improving government accountability and helping voters to become “informed”. The citizens being impacted by disasters may welcome having accountable governments in place that know how to prevent disasters from taking place, and, in the event of their occurrence, can take care of them, affordably and equitably, after the disasters strike (8*).

“The more citizens hold their governments accountable, the greater the demand and use for data will be for measuring government performance against indicators and targets. One way to make government agencies, [legislators, and executives,] more accountable and more efficient is to widely publicize data about their achievements and shortfalls, and then [to cast informed votes that will lead to adjustments in the] funds delivered in the next budget cycle to reward strong performers [and chasten bad performers].”

Footnotes

1. As mentioned or implied throughout this tutorial, given that the conventional institutions in some countries appear incapable of understanding, let alone ameliorating, these conditions, networks and clubs may need to take independent action.
2. The Poljanšek et al (2017) reference, in particular, highlight how much more advanced IT technologies, many involving satellite imagery, IoT technologies, and emerging GIS software, are improving disaster risk identification and reduction at national and international scales. The reference also highlights how thousands of separate, peer-reviewed, silos are carrying out the work, with little to no coordination. Section 7.8 in UNISDR (2015) further discusses the drawbacks to this conventional approach. The reference offers no reference dataset that demonstrates alternative ways to carry out each chapter’s algorithms and recommendations (i.e. see Gertler et al 2016, or any of DevTreks tutorials for examples). The reference makes passing mention of the fundamental role of a standard IT data platform[s] to actually accomplish all of the



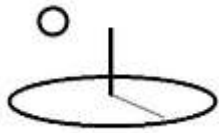
DevTreks –social budgeting that improves lives and livelihoods

needed objectives, partnerships, and “innovative approaches” (i.e. see Section K and L of Appendix A in SPA1 for examples –keep in mind that the author is a technologist).

DevTreks emphasizes the importance of “doing it right” – by starting with the basics: community capital improvements, cause and effect impact pathways, population transition states, multi-stakeholder value assessments, standardized TEXT datasets accessed via URIs, uniform reporting, scientific reference datasets applied in online tutorials, simple software object models, mathematical algorithms derived from open source software libraries, networks, clubs, generic-open source-IT platforms, and morally-grounded \$28 donations.

The “next generation” needs to critically assess whether, although this type of reference represents a valuable legacy from the “last generation”, their generation can, and must, do better. They can start by doing better than the algorithms introduced in these tutorials. Then proceed to automate all of the best algorithms in the Poljanšek reference and include complete reference TEXT datasets, URLs, tutorials, and open source code. Demonstrate how the quality of life of impacted stakeholders is improving. That is, do it right.

3. Although this GIS-style map can make a lot of complex information understandable to stakeholders, the author has found that the use of “pretty pictures and maps” can distort reality, foster audience complacency, and lead to misinformation. UNISDR (2015) use the phrase “decision makers need numbers not maps” to make a similar point. That’s why this reference uses a lot of images of tabular TEXT datasets. The TEXT files can still be inaccurate and boring, but raw data has less tendency to be misleading.
4. Prior to becoming a technologist, the author worked “in the field” for more than 15 years. Although this reference makes a good faith effort to follow “best disaster risk management practices”, that experience raises warning flags about data intensive efforts like this. On the other hand, the technologist experience suggests these practices, which follow UNISDR recommendations, are inevitable because they’re necessary. Footnote 14 in CTAP is worth repeating. “These are all tools that have the potential to assist decision



making. Assess which ones make sense for local contexts, experiment with them, adapt, but ultimately figure out how to make affordable and transparent decisions for [improving peoples’ lives and livelihoods].”

5. A new algorithm for automating INFORM was decided against because the goal of this reference is not global reporting using existing international datasets. The goal is to use impact pathways to understand cause and effect attribution at national, landscape, and business scale. This reference encourages national sustainability efforts to investigate duplicating INFORM’s approach using national datasets (i.e. census, SDG, and SDRR). Given that this tutorial’s algorithms can use Indicators derived from these national datasets, priority is placed on further development of the SPA2 and SPA3 algorithms.
6. This dataset could not, at first, be run with subalgorithm17 because of a bug displayed in the following image’s function:

```
        sb.AppendLine(rb.ToString());
        rb = new StringBuilder();
    }
}
iCount++;
if (iCount == (DataSet2[0].Count * 3))
{
    iCount = 0;
    iRateIndex = 0;
}
else if (iCount == 3)
{
    iRateIndex = 1;
}
else if (iCount == 6)
```

The bug highlighted several problems: a. this was completely the wrong function to use with subalgorithm17, b. even though 17 was under active debugging, several sloppy areas of coding also surfaced associated with previous releases, c. DataSet2[0] has no error checking and is the immediate bug (but also the symptom of another bug), and d. this code only supported a minor part of row headers but had the potential to completely



prevent the results of a CTAP algorithm from being printed. Even if 99% of an algorithm’s code is basically sound, or even if an algorithm appears to work for 99% of datasets, that’s not good enough for algorithms with these potential consequences.

Footnote 10 in CTAP is worth repeating: “Serious [local], national, and international, IT shops must plan on having larger IT staffs.” And adding “Serious local, national, and international, SDG and SDRR efforts must divert staff from last generation jobs to next generation jobs.”

7. For example, many western states in the USA now require the persons responsible for causing fires to pay for the ensuing maelstroms. Impacted communities and stakeholders can require similar accountability for the persons, past and present, responsible for causing losses attributed to their actions and policies that contributed to the human-induced hazards. Modern digital evidence supports such expanded accountability. Consequential digital activism can help to deal with related institutional “issues”, such as rigged judicial systems.

12. “Informed stakeholders” should recognize that rating government agencies, legislators, judicial bodies, and public executives for their contribution to “societal quality of life” may be inevitable and necessary. This reference recommends basing the ratings on “international SDG and SDRR reporting norms”, “evidence-based” science, and concrete, open-source, IT that employs “next generation”, M&E-based machine learning algorithms. Refer to the Version 2.1.8 SDG Plan reference for additional guidance.

Case Study References

Example 5, plus the following.

INFORM Global Risk Index Results 2018 (last accessed February 29, 2018: www.inform-index.org) (see the related reference by Martin-Ferrer et al, 2017)

International Recovery Platform. Guidance Note on Recovery. Private Sector. 2016



DevTreks –social budgeting that improves lives and livelihoods

Khazai, Bijan; Bendimerad, Fouad; Cardona, Omar Dario; Carreno, Martha-Lilliana; Barbat, Alex H.; Burton, Christopher G. A Guide to Measuring Urban Risk Resilience. Principle, Tools and Practice of Urban Indicators (Prerelease Draft). Earthquakes and Megacities Initiative - EMI. 2015

Marin-Ferrer, M. Vernaccini, L. Poljansek, K. INFORM. Index for Risk Management Concept and Methodology Version 2017

V. Meyer, N. Becker, V. Markantonis, R. Schwarze, J. C. J. M. van den Bergh, L. M. Bouwer, P. Bubeck, P. Ciavola, E. Genovese, C. Green, S. Hallegatte, H. Kreibich, Q. Lequeux, I. Logar, E. Papyrakis, C. Pfurtscheller, J. Poussin, V. Przyluski, A. H. Thieken, and C. Viavattene. Review article: Assessing the costs of natural hazards – state of the art and knowledge gaps. Nat. Hazards Earth Syst. Sci., 13, 1351–1373, 2013

Alberto Michelini, Gerhard Wotawa, Delia Arnold-Arias, Daniela Pantosti, et al. EU ARISTOTLE Project. 2017

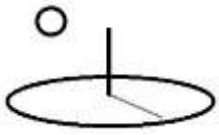
Poljanšek, K., Marín Ferrer, M., De Groeve, T., Clark, I. (Eds.), 2017. Science for disaster risk management 2017: knowing better and losing less. EUR 28034 EN, Publications Office of the European Union, Luxembourg, ISBN 978-92-79-60678-6, doi: 10.2788/688605, JRC102482.

United Nations Office for Disaster Risk Reduction. Technical Guidance for Monitoring and Reporting on Progress in Achieving the Global Targets of the Sendai. Framework for Disaster Risk Reduction 2015-2030. 2017a

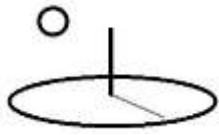
United Nations Office for Disaster Risk Reduction. National Disaster Risk Assessment Words into Action Guidelines. Governance System, Methodologies, and Use of Results, 2017b

United Nations Office for Disaster Risk Reduction. Global Assessment Report (GAR). 2015

United Nations Environment Programme (UNEP) Loss and Damage: The Role of Ecosystem Services. 2016



DevTreks –social budgeting that improves lives and livelihoods



DevTreks –social budgeting that improves lives and livelihoods

Example 6A and 6B. Stakeholder Abbreviated Resource Conservation Value Accounting (RCA7)

Algorithms: algorithm1, subalgorithm18; CTA algorithms

URLs:

Example 6A. Resource Stock Output Calculator

<https://www.devtreks.org/greentreks/preview/carbon/output/Disaster Risk Management, Example 6A/2141223486/none>

<https://www.devtreks.org/greentreks/preview/carbon/resourcepack/Disaster Accounting, Example 6A/1559/none>

<http://localhost:5000/greentreks/preview/carbon/resourcepack/Disaster Accounting, Example 6A/546/none>

<http://localhost:5000/greentreks/preview/carbon/output/Disaster Risk Management, Example 6A/2141223501/none>

Example 6B. Correlated Indicators

<https://www.devtreks.org/greentreks/preview/carbon/resourcepack/Disaster Accounting, Example 6B/1561/none>

<https://www.devtreks.org/greentreks/preview/carbon/output/SDRR, Example 6B/2141223487/none>

<http://localhost:5000/greentreks/preview/carbon/resourcepack/Disaster Accounting, Example 6B/547/none>

<http://localhost:5000/greentreks/preview/carbon/output/SDRR, Example 6B/2141223502/none>

A. Introduction

This example offers an abbreviated alternative to Example 5 and 6's fuller SIA techniques (1*).

The primary differences from those examples include:

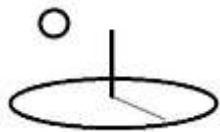
1. **SDG Allocations and Socioeconomic Characteristics in 1 TEXT dataset:** Example 5 and 6's 1st and 2nd datasets are combined into 1 TEXT dataset. This algorithm serves as



an abbreviated SIA compared to Example 5 and 6. Rather than allocate the 1st datasets SDG measurements to the 2nd dataset, this algorithm requires the direct SDG allocated measurements. RCA Assessors must explain the reason for the allocated measurements in separate documentation. Either CTAP-like data can be used as the background data for making the allocations or rules can be developed for making the allocations from Example 6's watershed plans.

2. **QASY, QALY, or DALY Measurements:** The preferred SDG measurements are based on human quality of life-related measurements, not physical or social Indicator measurements. Example 4B explained the reason for preferring these metrics in socioeconomic decision making. That is, they relate directly to stakeholder preferences which do a better job of measuring stakeholder perspectives and understanding equity and tradeoffs. Example 4B also discusses the need to institutionalize these measurements at local, national, and international, scales.
3. **Hotspots Categorical Indexes:** Example 5 and 6's Hotspots-related factors move from the Indicators to the Categorical Index. This approach implies that the original CI measurements taken from related algorithms, must be grouped together in this dataset based on both Hotspots factors, Production Processes and Life Cycle stages. In the context of this example, these processes and stages should focus on Coping Capacity and Resiliency Processes and Disaster Risk Management Stages. The goal is to understand the most effective and equitable ways to achieve SDRR targets at individual business scale (2*).
4. **CEAs use the SPA2 techniques.** This algorithm does not support direct CEA measurement because no costs are added to the datasets. CEA and Reference Cost Effectiveness Results must be completed using the techniques introduced in Examples 4A, 4B, and 4C.
5. **Additional DRM support from additional algorithms.** Example 6B demonstrates how to supplement the Example 6A decision making by using additional risk assessment algorithms.

B. Social Performance Score

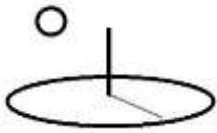


DevTreks –social budgeting that improves lives and livelihoods

The following algorithm properties define the socioeconomic characteristics and an “impact transition state” for the stakeholder groups identified in these datasets.

Indicators

The following images of Indicator 1 displays the “target-actual” properties explained more thoroughly in this section that produces the target and actual scores. Unlike Example 6, these properties don’t include “starting”, or benchmark, properties. These “scores” are calculated as the sum of the TR Indexes for all locations. They are calculated as the sum of the TR’s children normalized and weighted LIs. This dataset uses 2 locations for testing purposes. These scores will be used with the “standard sustainability scorecards” that will be explained in Section I, Decisions.



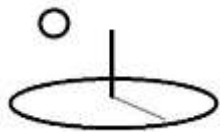
Stock Indicators	
Indicator 1	
DRM 2017 to 2018	
Indicator 1 Description	
Example 6A in the Social Performance Analysis 3 reference.	
Indicator 1 URL	
http://localhost:5000/resources/network_carbon/resourcepack_545/resource_1941/Ind1-RCA7-Impacts.csv	
Label 1	Rel Label 1
DRM1	none
Date 1	Dist Type 1
03/09/2018	none
Q1 1	Q1 Unit 1
0.0000	benchmark most score
Q2 1	Q2 Unit 1
0.0000	benchmark low score
Q3 1	Q3 Unit 1
0.0000	benchmark high score
Q4 1	Q4 Unit 1
35.7367	target most score
Q5 1	Q5 Unit 1
22.4930	target low score
Math Operator 1	BaseIO 1
equalto	none
QT 1	QT Unit 1
53.8169	target high score
Math Type 1	Math Sub Type 1
algorithm1	subalgorithm18
53.8169	target high score
Math Type 1	Math Sub Type 1
algorithm1	subalgorithm18
QT D1 1	QT D1 Unit 1
3.5000	actual certainty1
QT D2 1	QT D2 Unit 1
4.5000	actual certainty2
QT Most 1	QT Most Unit 1
32.6916	actual most score
QT Low 1	QT Low Unit 1
20.5764	actual low score
QT High 1	QT High Unit 1
49.2312	actual high score
Math Expression 1	
I1.Q1.factor1 + I1.Q2.factor2 + I1.Q3.factor3 + I1.Q4.factor4 +	
Math Result 1	
http://localhost:55009/resources/network_carbon/resourcepack_546/resource_1954/Ind1-RCA7-Results.csv	
+ Indicator 2	
+ Indicator 3	
- Indicator 4	
DRM 2023 to 2024	

TEXT dataset. Demographics and SDG Allocated Indicators. This stylized dataset applies Example 5’s coffee farm dataset with 2 hazards, earthquakes and droughts, and 2 location ids.



DevTreks –social budgeting that improves lives and livelihoods

label	location	risks_and_impacts	factor1	factor2	factor3	factor4	factor5	factor6	factor7	factor8	factor9	factor10	factor11
A	0	Farmer Characteristics	hhincome	gender	age	ethnicity	disability	education	population count	population unit	hotspots1	hotspots2	date
B	0	Farm Characteristics	tenure	management	size	sublocation	total gross revenue	coffee gross revenue	land use count	land use unit	hotspots1	hotspots2	date
A1A	1	Small scale coffee farmers male low income	5500.00	male	45.00	hispanic	none	4.00	100000.00	people	HCA	prevention	1/1/2018
IF1A	1	Target A. Mortality	75.00	daly	37.50	135.00	lower and upper 80% CI	80.00	3.00	4.00	2.00	weights	2.00
IF1B	1	Target B. Affected People	80.00	daly	60.00	150.00	Lower and upper	85.00	3.00	4.00	2.00	weights	3.00
IF1C	1	Target E. Disaster Risk Reduction Strategy	60.00	daly	40.00	75.00	lower and upper 80% CI	65.00	3.00	4.00	2.00	weights	4.00
IF1D	1	Target F. International Support	60.00	daly	40.00	75.00	lower and upper 80% CI	65.00	3.00	4.00	2.00	weights	5.00
IF1E	1	Target G. Multi-hazard Disaster Information	60.00	daly	40.00	75.00	lower and upper 80% CI	65.00	3.00	4.00	2.00	weights	6.00
B1A	1	Small scale farms 1000-3000 m slopes <20% 60 SQI	owned	moderate	2.50	25.00	5000.00	3000.00	1000.00	hectare	HCA	prevention	1/1/2018
IF2A	1	Target C. Economic Loss	45.00	US dollar	25.00	65.00	lower and upper 80% CI	50.00	4.00	5.00	1.00	weights	2.50
IF2B	1	Target D. Infrastructure Damage	45.00	US dollar	25.00	65.00	lower and upper 80% CI	50.00	4.00	5.00	1.00	weights	3.50
A2A	1	Small scale coffee farmers female low income	4500.00	female	45.00	hispanic	none	4.00	100000.00	people	HCA	prevention	1/1/2018
IF1A	1	Target A. Mortality	75.00	daly	37.50	135.00	lower and upper 80% CI	80.00	3.00	4.00	2.00	weights	2.00
IF1B	1	Target B. Affected People	80.00	daly	60.00	150.00	Lower and upper	85.00	3.00	4.00	2.00	weights	3.00
IF1C	1	Target E. Disaster Risk Reduction Strategy	60.00	daly	40.00	75.00	lower and upper 80% CI	65.00	3.00	4.00	2.00	weights	4.00
IF1D	1	Target F. International Support	60.00	daly	40.00	75.00	lower and upper 80% CI	65.00	3.00	4.00	2.00	weights	5.00
IF1E	1	Target G. Multi-hazard Disaster Information	60.00	daly	40.00	75.00	lower and upper 80% CI	65.00	3.00	4.00	2.00	weights	6.00
B2A	1	Small scale farms high elevation moderate slopes	owned	moderate	2.50	25.00	5000.00	3000.00	1000.00	hectare	HCA	prevention	1/1/2018
IF2A	1	Target C. Economic Loss	45.00	US dollar	25.00	65.00	lower and upper 80% CI	50.00	4.00	5.00	1.00	weights	2.50
IF2B	1	Target D. Infrastructure Damage	45.00	US dollar	25.00	65.00	lower and upper 80% CI	50.00	4.00	5.00	1.00	weights	3.50
SA	1	Earthquakes low income	0.00	0.00	0.00	0.00	0.00	none	0.00	0.00	0.00	weights	2.00
A1B	1	Small scale coffee farmers male medium income	5500.00	male	45.00	hispanic	none	4.00	100000.00	people	HCA	prevention	1/1/2018
IF1A	1	Target A. Mortality	75.00	daly	37.50	135.00	lower and upper 80% CI	80.00	3.00	4.00	2.00	weights	2.00
IF1B	1	Target B. Affected People	80.00	daly	60.00	150.00	Lower and upper	85.00	3.00	4.00	2.00	weights	3.00
IF1C	1	Target E. Disaster Risk Reduction Strategy	60.00	daly	40.00	75.00	lower and upper 80% CI	65.00	3.00	4.00	2.00	weights	4.00
IF1D	1	Target F. International Support	60.00	daly	40.00	75.00	lower and upper 80% CI	65.00	3.00	4.00	2.00	weights	5.00
IF1E	1	Target G. Multi-hazard Disaster Information	60.00	daly	40.00	75.00	lower and upper 80% CI	65.00	3.00	4.00	2.00	weights	6.00
B1B	1	Small scale farms 1000-3000 m slopes <20% 60 SQI	owned	moderate	2.50	25.00	5000.00	3000.00	1000.00	hectare	HCA	prevention	1/1/2018
IF2A	1	Target C. Economic Loss	45.00	US dollar	25.00	65.00	lower and upper 80% CI	50.00	4.00	5.00	1.00	weights	2.50
IF2B	1	Target D. Infrastructure Damage	45.00	US dollar	25.00	65.00	lower and upper 80% CI	50.00	4.00	5.00	1.00	weights	3.50
A2B	1	Small scale coffee farmers female medium income	4500.00	female	45.00	hispanic	none	4.00	100000.00	people	HCA	prevention	1/1/2018
IF1A	1	Target A. Mortality	75.00	daly	37.50	135.00	lower and upper 80% CI	80.00	3.00	4.00	2.00	weights	2.00
IF1B	1	Target B. Affected People	80.00	daly	60.00	150.00	Lower and upper	85.00	3.00	4.00	2.00	weights	3.00
IF1C	1	Target E. Disaster Risk Reduction Strategy	60.00	daly	40.00	75.00	lower and upper 80% CI	65.00	3.00	4.00	2.00	weights	4.00
IF1D	1	Target F. International Support	60.00	daly	40.00	75.00	lower and upper 80% CI	65.00	3.00	4.00	2.00	weights	5.00
IF1E	1	Target G. Multi-hazard Disaster Information	60.00	daly	40.00	75.00	lower and upper 80% CI	65.00	3.00	4.00	2.00	weights	6.00
B2B	1	Small scale farms high elevation moderate slopes	owned	moderate	2.50	25.00	5000.00	3000.00	1000.00	hectare	HCA	prevention	1/1/2018
IF2A	1	Target C. Economic Loss	45.00	US dollar	25.00	65.00	lower and upper 80% CI	50.00	4.00	5.00	1.00	weights	2.50
IF2B	1	Target D. Infrastructure Damage	45.00	US dollar	25.00	65.00	lower and upper 80% CI	50.00	4.00	5.00	1.00	weights	3.50
SB	1	Earthquakes medium income	0.00	0.00	0.00	0.00	0.00	none	0.00	0.00	0.00	weights	2.00
A1C	1	Small scale coffee farmers male high income	5500.00	male	45.00	hispanic	none	4.00	100000.00	people	HCA	prevention	1/1/2018



...

IF2B	2	Target D. Infrastructure Damage	45.00	US dollar	25.00	65.00	lower and upper 80% CI	50.00	4.00	5.00	1.00	weights	3.50
A2C	2	Small scale coffee farmers female high income	4500.00	female	45.00	hispanic	none	4.00	100000.00	people	HCA	prevention	1/1/2018
IF1A	2	Target A. Mortality	75.00	daly	37.50	135.00	lower and upper 80% CI	80.00	3.00	4.00	2.00	weights	2.00
IF1B	2	Target B. Affected People	80.00	daly	60.00	150.00	Lower and upper estimates	85.00	3.00	4.00	2.00	weights	3.00
IF1C	2	Target E. Disaster Risk Reduction Strategy	60.00	daly	40.00	75.00	lower and upper 80% CI	65.00	3.00	4.00	2.00	weights	4.00
IF1D	2	Target F. International Support	60.00	daly	40.00	75.00	lower and upper 80% CI	65.00	3.00	4.00	2.00	weights	5.00
IF1E	2	Target G. Multi-hazard Disaster Information	60.00	daly	40.00	75.00	lower and upper 80% CI	65.00	3.00	4.00	2.00	weights	6.00
B2C	2	Small scale farms high elevation moderate slopes	owned	moderate	2.50	25.00	5000.00	3000.00	1000.00	hectare	HCA	prevention	1/1/2018
IF2A	2	Target C. Economic Loss	45.00	US dollar	25.00	65.00	lower and upper 80% CI	50.00	4.00	5.00	1.00	weights	2.50
IF2B	2	Target D. Infrastructure Damage	45.00	US dollar	25.00	65.00	lower and upper 80% CI	50.00	4.00	5.00	1.00	weights	3.50
SC	2	Droughts high income	0.00	0.00	0.00	0.00	0.00	none	0.00	0.00	0.00	weights	2.00
TRC	2	Climatological Hazards	0.00	0.00	0.00	0.00	0.00	none	0.00	0.00	0.00	weights	2.00

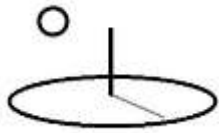
Stakeholder Characteristics Titles. Rows used to describe socioeconomic variables must include a 0 in the 2nd data column, the location, property.

Locational Index Measurement. The following properties show that only the last 2 columns of data are used to run calculations. These 2 columns will normalize and sum all of the calculated LIs within each TR Index. This is the same normalization and weighting as used in subalgorithm15.

- **factor1 to factor9:** none
- **factor10:** normalization type
- **factor11:** weight

Categorical Indexes. The following list defines how the final 11 columns of data for Categorical Indexes are used.

- **factor1 to factor6:** socioeconomic characteristic measurements for separate stakeholder groups; most of the UN references identify the most important characteristics for reporting: age, gender, race, disability, income, and education.
- **factor7:** population/land use count at end date
- **factor8:** population/land use unit of measurement
- **factor9:** Hotspots Coping and Resiliency Mechanism



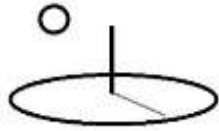
DevTreks –social budgeting that improves lives and livelihoods

- **factor10:** Hotspots Disaster Risk Management Stage
- **factor11:** end date of measurement

Indicators. These properties measure the actual and target SDRR and CTAP-like measurements at the parent CI.end date. QASY, QALY, or DALY measurements are preferred for these allocated measurements.

- **factor1:** actual most likely, allocated, SDRR and CTAP-like measurement count at end date;
- **factor2.** unit of measurement (i.e. QASY)
- **factor3:** actual low, allocated, CTAP count at end date;
- **factor4:** actual high, allocated, CTAP count at end date;
- **factor5:** unit of measurement for low and high estimates;
- **factor6:** target most likely, allocated, CTAP measurement count at end date;
- **factor7.** certainty1; severity, and probable consequence, of this SDG risk on this population/land use as of end date
- **factor8:** certainty2; likelihood of this CTAP risk on this population/land use as of end date
- **factor9:** certainty3; probable impact of this population/land use on company/CSO from this CTAP risk as of end date
- **factor10:** normalization type
- **factor11:** weight

Business Resource Conservation Value Accounting MathResults. The following table displays the partial results for this abbreviated SIA.

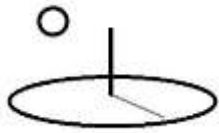


DevTrek –social budgeting that improves lives and livelihoods

label	locatic	risks_and_factor1	factor2	factor3	factor4	factor5	factor6	factor7	factor8	factor9	factor10	factor11	qtmost	percentar	qtlow	qthigh	certainty1	certainty2	certainty3	totalcost	sdgunit	producer	lifecyclesta	
A	0.00	Farmer Ch	hincome	gender	age	ethnicity	disability	education	population	population	hotspots1	hotspots2	date	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
B	0.00	Farm Char	tenure	managem	size	sublocatio	total gross	coffee	land use	co	land use	ur	hotspots1	hotspots2	date	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
A1A	1.00	Small scale	5500.00	male	45.00	hispanic	none	4.00	100000.00	people	HCA	preventior	43101.00	6.44	92.96	4.18	9.81	3.00	4.00	2.00	0.00	0.00	0.00	
IF1A	1.00	Target A. #	75.00	daly	37.50	135.00	lower and	80.00	3.00	4.00	2.00	weights	2.00	1.44	93.75	0.72	2.60	3.00	4.00	2.00	0.00	daly	HCA	prevention
IF1B	1.00	Target B. #	80.00	daly	60.00	150.00	Lower and	85.00	3.00	4.00	2.00	weights	3.00	1.54	94.12	1.15	2.89	3.00	4.00	2.00	0.00	daly	HCA	prevention
IF1C	1.00	Target E. C	60.00	daly	40.00	75.00	lower and	65.00	3.00	4.00	2.00	weights	4.00	1.15	92.31	0.77	1.44	3.00	4.00	2.00	0.00	daly	HCA	prevention
IF1D	1.00	Target F. #	60.00	daly	40.00	75.00	lower and	65.00	3.00	4.00	2.00	weights	5.00	1.15	92.31	0.77	1.44	3.00	4.00	2.00	0.00	daly	HCA	prevention
IF1E	1.00	Target G. #	60.00	daly	40.00	75.00	lower and	65.00	3.00	4.00	2.00	weights	6.00	1.15	92.31	0.77	1.44	3.00	4.00	2.00	0.00	daly	HCA	prevention
B1A	1.00	Small scale	owned	moderate	2.50	25.00	5000.00	3000.00	1000.00	hectare	HCA	preventior	43101.00	1.73	90.00	0.96	2.50	4.00	5.00	1.00	0.00	0.00	0.00	
IF2A	1.00	Target C. E	45.00	US dollar	25.00	65.00	lower and	50.00	4.00	5.00	1.00	weights	2.50	0.87	90.00	0.48	1.25	4.00	5.00	1.00	0.00	US dollar	HCA	prevention
IF2B	1.00	Target D. I	45.00	US dollar	25.00	65.00	lower and	50.00	4.00	5.00	1.00	weights	3.50	0.87	90.00	0.48	1.25	4.00	5.00	1.00	0.00	US dollar	HCA	prevention
A2A	1.00	Small scale	4500.00	female	45.00	hispanic	none	4.00	100000.00	people	HCA	preventior	43101.00	6.44	92.96	4.18	9.81	3.00	4.00	2.00	0.00	0.00	0.00	
IF1A	1.00	Target A. #	75.00	daly	37.50	135.00	lower and	80.00	3.00	4.00	2.00	weights	2.00	1.44	93.75	0.72	2.60	3.00	4.00	2.00	0.00	daly	HCA	prevention
IF1B	1.00	Target B. #	80.00	daly	60.00	150.00	Lower and	85.00	3.00	4.00	2.00	weights	3.00	1.54	94.12	1.15	2.89	3.00	4.00	2.00	0.00	daly	HCA	prevention
IF1C	1.00	Target E. C	60.00	daly	40.00	75.00	lower and	65.00	3.00	4.00	2.00	weights	4.00	1.15	92.31	0.77	1.44	3.00	4.00	2.00	0.00	daly	HCA	prevention
IF1D	1.00	Target F. #	60.00	daly	40.00	75.00	lower and	65.00	3.00	4.00	2.00	weights	5.00	1.15	92.31	0.77	1.44	3.00	4.00	2.00	0.00	daly	HCA	prevention
IF1E	1.00	Target G. #	60.00	daly	40.00	75.00	lower and	65.00	3.00	4.00	2.00	weights	6.00	1.15	92.31	0.77	1.44	3.00	4.00	2.00	0.00	daly	HCA	prevention
B2A	1.00	Small scale	owned	moderate	2.50	25.00	5000.00	3000.00	1000.00	hectare	HCA	preventior	43101.00	1.73	90.00	0.96	2.50	4.00	5.00	1.00	0.00	0.00	0.00	
IF2A	1.00	Target C. E	45.00	US dollar	25.00	65.00	lower and	50.00	4.00	5.00	1.00	weights	2.50	0.87	90.00	0.48	1.25	4.00	5.00	1.00	0.00	US dollar	HCA	prevention
IF2B	1.00	Target D. I	45.00	US dollar	25.00	65.00	lower and	50.00	4.00	5.00	1.00	weights	3.50	0.87	90.00	0.48	1.25	4.00	5.00	1.00	0.00	US dollar	HCA	prevention
SA	1.00	Earthquaki	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	16.35	91.48	10.29	24.62	3.50	4.50	1.50	0.00	0.00	0.00	
A1B	1.00	Small scale	5500.00	male	45.00	hispanic	none	4.00	100000.00	people	HCA	preventior	43101.00	6.44	92.96	4.18	9.81	3.00	4.00	2.00	0.00	0.00	0.00	
IF1A	1.00	Target A. #	75.00	daly	37.50	135.00	lower and	80.00	3.00	4.00	2.00	weights	2.00	1.44	93.75	0.72	2.60	3.00	4.00	2.00	0.00	daly	HCA	prevention
IF1B	1.00	Target B. #	80.00	daly	60.00	150.00	Lower and	85.00	3.00	4.00	2.00	weights	3.00	1.54	94.12	1.15	2.89	3.00	4.00	2.00	0.00	daly	HCA	prevention
IF1C	1.00	Target E. C	60.00	daly	40.00	75.00	lower and	65.00	3.00	4.00	2.00	weights	4.00	1.15	92.31	0.77	1.44	3.00	4.00	2.00	0.00	daly	HCA	prevention
IF1D	1.00	Target F. #	60.00	daly	40.00	75.00	lower and	65.00	3.00	4.00	2.00	weights	5.00	1.15	92.31	0.77	1.44	3.00	4.00	2.00	0.00	daly	HCA	prevention
IF1E	1.00	Target G. #	60.00	daly	40.00	75.00	lower and	65.00	3.00	4.00	2.00	weights	6.00	1.15	92.31	0.77	1.44	3.00	4.00	2.00	0.00	daly	HCA	prevention
B2B	1.00	Small scale	owned	moderate	2.50	25.00	5000.00	3000.00	1000.00	hectare	HCA	preventior	43101.00	1.73	90.00	0.96	2.50	4.00	5.00	1.00	0.00	0.00	0.00	
IF2A	1.00	Target C. E	45.00	US dollar	25.00	65.00	lower and	50.00	4.00	5.00	1.00	weights	2.50	0.87	90.00	0.48	1.25	4.00	5.00	1.00	0.00	US dollar	HCA	prevention
IF2B	1.00	Target D. I	45.00	US dollar	25.00	65.00	lower and	50.00	4.00	5.00	1.00	weights	3.50	0.87	90.00	0.48	1.25	4.00	5.00	1.00	0.00	US dollar	HCA	prevention
SB	1.00	Earthquaki	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	16.35	91.48	10.29	24.62	3.50	4.50	1.50	0.00	0.00	0.00	
A1C	1.00	Small scale	5500.00	male	45.00	hispanic	none	4.00	100000.00	people	HCA	preventior	43101.00	6.44	92.96	4.18	9.81	3.00	4.00	2.00	0.00	0.00	0.00	
IF1A	1.00	Target A. #	75.00	daly	37.50	135.00	lower and	80.00	3.00	4.00	2.00	weights	2.00	1.44	93.75	0.72	2.60	3.00	4.00	2.00	0.00	daly	HCA	prevention
IF1B	1.00	Target B. #	80.00	daly	60.00	150.00	Lower and	85.00	3.00	4.00	2.00	weights	3.00	1.54	94.12	1.15	2.89	3.00	4.00	2.00	0.00	daly	HCA	prevention
IF1C	1.00	Target E. C	60.00	daly	40.00	75.00	lower and	65.00	3.00	4.00	2.00	weights	4.00	1.15	92.31	0.77	1.44	3.00	4.00	2.00	0.00	daly	HCA	prevention
IF1D	1.00	Target F. #	60.00	daly	40.00	75.00	lower and	65.00	3.00	4.00	2.00	weights	5.00	1.15	92.31	0.77	1.44	3.00	4.00	2.00	0.00	daly	HCA	prevention
IF1E	1.00	Target G. #	60.00	daly	40.00	75.00	lower and	65.00	3.00	4.00	2.00	weights	6.00	1.15	92.31	0.77	1.44	3.00	4.00	2.00	0.00	daly	HCA	prevention
B2C	1.00	Small scale	owned	moderate	2.50	25.00	5000.00	3000.00	1000.00	hectare	HCA	preventior	43101.00	1.73	90.00	0.96	2.50	4.00	5.00	1.00	0.00	0.00	0.00	
IF2A	1.00	Target C. E	45.00	US dollar	25.00	65.00	lower and	50.00	4.00	5.00	1.00	weights	2.50	0.87	90.00	0.48	1.25	4.00	5.00	1.00	0.00	US dollar	HCA	prevention
IF2B	1.00	Target D. I	45.00	US dollar	25.00	65.00	lower and	50.00	4.00	5.00	1.00	weights	3.50	0.87	90.00	0.48	1.25	4.00	5.00	1.00	0.00	US dollar	HCA	prevention
SC	1.00	Earthquaki	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	16.35	91.48	10.29	24.62	3.50	4.50	1.50	0.00	0.00	0.00	
TRG	1.00	Geophysic	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	8.17	91.48	5.14	12.31	3.50	4.50	1.50	0.00	0.00	0.00	
A1A	2.00	Small scale	5500.00	male	45.00	hispanic	none	4.00	100000.00	people	HCA	preventior	43101.00	6.44	92.96	4.18	9.81	3.00	4.00	2.00	0.00	0.00	0.00	
IF1A	2.00	Target A. #	75.00	daly	37.50	135.00	lower and	80.00	3.00	4.00	2.00	weights	2.00	1.44	93.75	0.72	2.60	3.00	4.00	2.00	0.00	daly	HCA	prevention

...

B2C	2.00	Small scale	owned	moderate	2.50	25.00	5000.00	3000.00	1000.00	hectare	HCA	preventior	43101.00	1.73	90.00	0.96	2.50	4.00	5.00	1.00	0.00	0.00	0.00	
IF2A	2.00	Target C. E	45.00	US dollar	25.00	65.00	lower and	50.00	4.00	5.00	1.00	weights	2.50	0.87	90.00	0.48	1.25	4.00	5.00	1.00	0.00	US dollar	HCA	prevention
IF2B	2.00	Target D. I	45.00	US dollar	25.00	65.00	lower and	50.00	4.00	5.00	1.00	weights	3.50	0.87	90.00	0.48	1.25	4.00	5.00	1.00	0.00	US dollar	HCA	prevention
SC	2.00	Droughts	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	16.35	91.48	10.29	24.62	3.50	4.50	1.50	0.00	0.00	0.00	
TRC	2.00	Climatolog	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	8.17	91.48	5.14	12.31	3.50	4.50					



DevTreks –social budgeting that improves lives and livelihoods

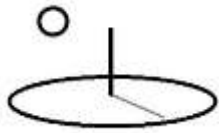
the following list. These properties differ from Example 6 in not having population PRA calculations in the CIs (i.e. no high and low confidence interval) and requiring the totalcosts property to be filled in manually (i.e. to support the Reference Case CEA reports).

- **qtmmost:** sum of CIs
- **percenttarget:** average of CIs
- **qtlow:** sum of CIs
- **qthigh:** sum of CIs
- **certainty1:** factor6 averages
- **certainty2:** factor7 averages
- **certainty3:** factor8 averages
- **totalcost:** 0 (fill in the quantity manually)

Locational Index Math. Locational Indexes are calculated as the sum of the normalized and weighted Indicators contained in all of the LI’s children CIs.

TR Index Math. The TR Indexes are calculated as the sum of the normalized and weighted children LIs, and display the same properties as the LIs. The LIs can be normalized and weighted separately from their children Indicators. Section I, Decisions, demonstrates displaying these Scores in uniform, “SDG-ASB Scorecards”.

Indicators 2 to 15. Hazard Impact Transition States. The following image demonstrates a similar M&E approach for documenting population epidemiology as introduced in Examples 5 and 6. The Indicator metadata displayed in these reports results from summing all of the TR Index data across all locations in each dataset. Note that this algorithm’s datasets do not include starting, or benchmark, SDRR allocations, resulting in benchmark Indicator scores of zero.



RCA Output Examples

Output

Disaster Risk Management, Example 6A

Score	Score Unit	Score D1 Amount	Score D1 Unit	Score D2 Amount	Score D2 Unit	Iterations	Confid Int	Random Seed	Base IO
Score Most Amount	Score Most Unit	Score Low Amount	Score Low Unit	Score High Amount	Score High Unit	Distribution Type	Math Type	Math Sub Type	Observations
53.8169	target high score	3.5000	actual certainty1	4.5000	actual certainty2	500	75	5	none
32.6916	actual most score	20.5764	actual low score	49.2312	actual high score	none	algorithm1	subalgorithm18	1

I1.QTM

http://localhost:55009/resources/network_carbon/resourcepack_546/resource_1962/Score-RCA7-Results.csv

Name (N)	Label	Date	Rel Label	Math Type	Dist Type	Base IO	Math Operator	Math Sub Type	
Q1 Amount	Q1 Unit	Q2 Amount	Q2 Unit	Q3 Amount	Q3 Unit	Q4 Amount	Q4 Unit	Q5 Amount	Q5 Unit
QT Amount	QT Unit	QT D1 Amount	QT D2 Amount	QT Most Amount	QT Most Unit	QT Low Amount	QT Low Unit	QT High Amount	QT High Unit
DRM 2017 to 2018	DRM1	03/09/2018	none	algorithm1	none	none	equalto	subalgorithm18	
0.0000	benchmark most score	0.0000	benchmark low score	0.0000	benchmark high score	35.7367	target most score	22.4930	target low score
53.8169	target high score	3.5000	4.5000	32.6916	actual most score	20.5764	actual low score	49.2312	actual high score

Example 6A in the Social Performance Analysis 3 reference.

I1.Q1.factor1 + I1.Q2.factor2 + I1.Q3.factor3 + I1.Q4.factor4 + I1.Q5.factor5 + I1.Q6.factor6 + I1.Q7.factor7 + I1.Q8.factor8 + I1.Q9.factor9 + I1.Q10.factor10 + I1.Q11.factor11

DRM 2019 to 2020	Label	Date	Rel Label	Math Type	Dist Type	Base IO	Math Operator	Math Sub Type	
Q1 Amount	Q1 Unit	Q2 Amount	Q2 Unit	Q3 Amount	Q3 Unit	Q4 Amount	Q4 Unit	Q5 Amount	Q5 Unit
QT Amount	QT Unit	QT D1 Amount	QT D2 Amount	QT Most Amount	QT Most Unit	QT Low Amount	QT Low Unit	QT High Amount	QT High Unit
DRM 2019 to 2020	DRM2	03/09/2018	none	algorithm1	none	none	equalto	subalgorithm18	
0.0000	benchmark most score	0.0000	benchmark low score	0.0000	benchmark high score	35.7367	target most score	22.4930	target low score
53.8169	target high score	3.5000	4.5000	32.6916	actual most score	20.5764	actual low score	49.2312	actual high score

Score using Multi-Index, MCDA-style, approach. Example 9 in CTAP demonstrates using MCDA to support disaster risk management. INFORM (2018) demonstrates using a uniform 10 point, low risk = 0, high risk =10, MCDA scoring system for disaster risk management at global

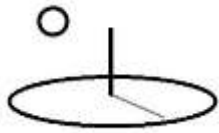


DevTreks –social budgeting that improves lives and livelihoods

scale. They demonstrate using advanced MCDA techniques to rank countries by degree of risk, to cluster countries together according to risk threshold levels, and to assess the trend of each country’s risk. Importantly, they also demonstrate how Indicator-based decision support systems can take several years to mature and require substantial effort and evolution to get the Indicators and Indexes “right”.

Rather than explain again the Score properties explained in Example 6A that are used for these purposes, the following images demonstrates using the Resource Stock and M&E Analyzers to conduct further analyses of the metadata Indicators and Scores. The Scenarios sections of this reference recommends using sibling base elements, which can be analyzed using these techniques, to document sustainability transition states or scenarios.

Totals Output Stock Analysis. Three children Output Series were added to this Output and then the Stock Calculator was copied into the 3 children to produce these results.



DevTreks –social budgeting that improves lives and livelihoods

GreenTreks	Search	Preview	Select
Edit	Pack	Views	Club

Select
PackIt

Edit Linked Views
Make base

Output Stock Totals
Get

Media
Mobile
Desktop

Intro	1	2	3	Help
-------	---	---	---	------

Resource Stock Output Analyzer View

DRM Totals Analysis

Introduction
This tool generates a variety of basic resource stock statistics for outputs.

Calculation View Description
v214b

Version: 1.9.0

[Feedback About carbon/output/Disaster Risk Management, Example 6A/2141223501/none](#)

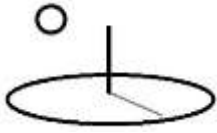
Output Group: RCA Output Examples

Output : Disaster Risk Management, Example 6A

Indicators

<p>Math Expression: I1.QTM</p> <p>Score Amount: 161.4507</p> <p>Score D1 Amount: 10.5000</p> <p>Score D2 Amount: 13.5000</p> <p>Distribution Type: none</p> <p>Score Most Amount: 98.0748</p> <p>Score Low Amount: 61.7292</p> <p>Score High Amount: 147.6936</p> <p>Iterations: 500</p> <p>Score Math Result: http://localhost:55009/resources/network_carbon/resourcepack_546/resource_1962/Score-RCA7-Results.csv</p>	<p>Observations: 3.0</p> <p>Score Unit: target high score</p> <p>Score D1 Unit: actual certainty1</p> <p>Score D2 Unit: actual certainty2</p> <p>Math Type: algorithm1</p> <p>Score Most Unit: actual most score</p> <p>Score Low Unit: actual low score</p> <p>Score High Unit: actual high score</p> <p>Score Math Sub Type: subalgorithm18</p>
--	---

Statistical Output Stock Analysis. The image shows that Statistical Analyzers produce basic total, mean, median, variance, and standard deviation, statistics for the 3 children observations.



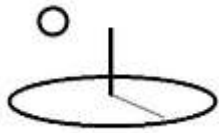
Intro	1	2	3	Help
Resource Stock Output Analyzer View				
DRM Stats				
<p>Introduction This tool generates a variety of basic resource stock statistics for outputs.</p> <p>Calculation View Description v214b</p> <p>Version: 1.9.0</p> <p>Feedback About carbon/output/Disaster Risk Management, Example 6A/2141223501/none</p>				

Output Group : RCA Output Examples

Output : Disaster Risk Management, Example 6A

Indicators	
Observations 1: 3.0	
Score Most Likely: 98.0748	Unit: actual most score
Score Mean: 32.6916	Score Median: 32.6916
Score Variance: 0.0000	Score Std Dev: 0.0000
Score Low: 61.7292	Unit: actual low score
Score Low Mean: 20.5764	Score Low Median: 20.5764
Score Low Variance: 0.0000	Score Low Std Dev: 0.0000
Score High: 147.6936	Unit: actual high score
Score High Mean: 49.2312	Score High Median: 49.2312
Score High Variance: 0.0000	Score High Std Dev: 0.0000
Name 1: DRM 2017 to 2018	Label 1: DRM1
Observations 1: 3.0	
Total 1: 98.0748	Unit 1: actual most score
Mean 1: 32.6916	Median 1: 32.6916
Variance 1: 0.0000	Std Dev 1: 0.0000
Name 2: DRM 2019 to 2020	Label 2: DRM2
Observations 2: 3.0	
Total 2: 98.0748	Unit 2: actual most score
Mean 2: 32.6916	Median 2: 32.6916
Variance 2: 0.0000	Std Dev 2: 0.0000
Name 3: DRM 2021 to 2022	Label 3: DRM3
Observations 3: 3.0	
Total 3: 98.0748	Unit 3: actual most score
Mean 3: 32.6916	Median 3: 32.6916
Variance 3: 0.0000	Std Dev 3: 0.0000
Name 4: DRM 2023 to 2024	Label 4: DRM4
Observations 4: 3.0	

Change By Id Output Stock Analysis. The zero changes in the following image correctly result from copying the same Stock Calculator into the 3 children Output Series, including the same URLs to the same datasets. Changes will be shown when each of the children datasets are changed.



DevTreks –social budgeting that improves lives and livelihoods

Output Group : RCA Output Examples ; none				
Output	All	Alt. 0		
Name		Disaster Risk Management, Example 6A		
Label		RCA6A		
Output Series	All	Alt. 0	Alt. 1	Alt. 2
Name		Disaster Risk Management, Example 6A	DRM, Example 6A, Stats 2	DRM, Example 6A, Stats 3
Label		RCA6A	RCA6A	RCA6A
Indicators	All	Alt. 0	Alt. 1	Alt. 2
Alternative				
Date		03/09/2018	03/09/2018	03/09/2018
Score Observations		1.0	1.0	1.0
Score Amount		32.6916	32.6916	32.6916
Score Unit		actual most score	actual most score	actual most score
Score Amount Change		0.0000	0.0000	0.0000
Score Percent Change		0.0000	0.0000	0.0000
Score Base Change		0.0000	0.0000	0.0000
Score Base Percent Change		0.0000	0.0000	0.0000
Score Low Amount		20.5764	20.5764	20.5764
Score Low Unit		actual low score	actual low score	actual low score
Score Low Amount Change		0.0000	0.0000	0.0000
Score Low Percent Change		0.0000	0.0000	0.0000
Score Low Base Change		0.0000	0.0000	0.0000
Score Low Base		0.0000	0.0000	0.0000



DevTreks –social budgeting that improves lives and livelihoods

Example 6B. Score using Correlated Indicator Approach. This example demonstrates using related algorithms to conduct more advanced disaster risk assessment. INFORM (2018) uses the following image to explore the correlation between the Categories, or Categorical Indexes, in their MCDA disaster risk management system. The authors explain how to use more advanced statistical analysis techniques, including correlation analysis and factor analysis, to ensure that Indicator-based Indexes, as demonstrated throughout this tutorial, are “well-structured and balanced”. OECD (2008) also introduce several statistical techniques, referred to as algorithms in DevTreks, for improving the applied use of composite indicators and indexes.



DevTreks –social budgeting that improves lives and livelihoods

introduced an example of studying the interrelationship of multiple hazards. Poljanšek et al (2017) discuss why it’s important to study how hazards interact, as follows: “various kinds of interactions between hazards that often lead to significantly more severe negative consequences for the society than when they act separately.”

The following example demonstrates the basic approach introduced in CTA01 for analyzing correlated Indicators. In this example, the correlated Indicators are final, aggregated, fictitious, Total Risk Indexes taken from 11 locations for 3 earthquakes that are correlated in some manner with 3 droughts (i.e. the earthquakes destroyed reservoirs). This example’s Score is measuring some type of aggregated risk being measured by aggregating the separate Indicators, such as aggregated, multiplicative, costs, benefits, net benefits, or, in this example, qualitative, MCDA-style, disaster risk TR Indexes.

The following 2 dataset URLs are added to the Score and used with algorithm1, subalgorithm3, to generate the final calculations. That algorithm uses a correlated multivariate analysis technique known as spearman correlation. In preparation for the machine learning algorithms, Version 2.1.6 upgraded older calculator patterns to make them compatible with newer patterns that are compatible with R, Python, and Machine Learning (i.e. the pattern found in the DevTreksStatsAPI app).

Dataset 1. Script File or Training Dataset. The following dataset specifies the type of correlation technique to use, such as pearson or spearman, and the Indicators that are correlated (i.e. E1, earthquake 1 is correlated with D1, or drought 1).

Stock Calculator using Score.JointDataURL[0]

spearman					
E1	D1	E2	D2	E3	D3

M&E Calculator using Score.URL[0]

spearman					
1	2	3	4	5	6



Dataset 2. Full Data File or Test Data File. The following dataset generates the correlation matrix for the interacting Indicators. The assumption is that separate datasets were collected from 11 different entities, such as locations or businesses.

Stock Calculator using Score.JointDataURL[1]

label	locatic	risks_and_in none	qtmost	
E1	1	earthquakes	0	22
E1	2	earthquakes	0	21
E1	3	earthquakes	0	21.25
E1	4	earthquakes	0	21.5
E1	5	earthquakes	0	21.75
E1	6	earthquakes	0	23
E1	7	earthquakes	0	22.75
E1	8	earthquakes	0	22.5
E1	9	earthquakes	0	22.25
E1	10	earthquakes	0	21.9
E1	11	earthquakes	0	22.1
D1	1	droughts	0	11.5
D1	2	droughts	0	10.5
D1	3	droughts	0	10.75
D1	4	droughts	0	11
D1	5	droughts	0	11.25
D1	6	droughts	0	12.5
D1	7	droughts	0	12.25
D1	8	droughts	0	12
D1	9	droughts	0	11.75
D1	10	droughts	0	11.6
D1	11	droughts	0	11.4
E2	1	earthquakes	0	20
E2	2	earthquakes	0	19

...



DevTrek –social budgeting that improves lives and livelihoods

E3	11	earthquakes	0	19.75
D3	1	droughts	0	10
D3	2	droughts	0	8
D3	3	droughts	0	7
D3	4	droughts	0	9
D3	5	droughts	0	10.5
D3	6	droughts	0	11
D3	7	droughts	0	12.5
D3	8	droughts	0	9.5
D3	9	droughts	0	10.25
D3	10	droughts	0	13.5
D3	11	droughts	0	7.5

M&E Calculator using Score.URL[1]

label	location	risks_and_	none	qtmost
1	1	earthquakes	0	22
1	2	earthquakes	0	21
1	3	earthquakes	0	21.25
1	4	earthquakes	0	21.5
1	5	earthquakes	0	21.75
1	6	earthquakes	0	23
1	7	earthquakes	0	22.75
1	8	earthquakes	0	22.5
1	9	earthquakes	0	22.25
1	10	earthquakes	0	21.9
1	11	earthquakes	0	22.1
2	1	droughts	0	11.5
2	2	droughts	0	10.5

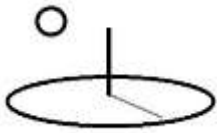
...



DevTreks –social budgeting that improves lives and livelihoods

5	10	earthquak	0	23
5	11	earthquak	0	19.75
6	1	droughts	0	10
6	2	droughts	0	8
6	3	droughts	0	7
6	4	droughts	0	9
6	5	droughts	0	10.5
6	6	droughts	0	11
6	7	droughts	0	12.5
6	8	droughts	0	9.5
6	9	droughts	0	10.25
6	10	droughts	0	13.5
6	11	droughts	0	7.5

Score. The following Score properties display the final aggregated risk generated from the 3 correlated Indicator datasets. The main purpose of this example is to introduce additional statistical techniques that can be automated in algorithms and then used to improve MCDA-style Indicators and Indexes. This image also shows the simpler calculator patterns used in Version 2.1.6.



Score Math Expression	
<input type="text" value="((I1.QT*I2.QT)+(I3.QT*I4.QT)+(I5.QT*I6.QT))/3"/>	
Score	Score Unit
<input type="text" value="253.0195"/>	<input type="text" value="total most likely risk"/>
Score D1	Score D1 Unit
<input type="text" value="0.0000"/>	<input type="text" value="low"/>
Score D2	Score D2 Unit
<input type="text" value="0.0000"/>	<input type="text" value="high"/>
Score Dist Type	Iterations
<input type="text" value="none"/>	<input type="text" value="500"/>
Confidence Interval	Random Seed
<input type="text" value="80"/>	<input type="text" value="7"/>
Score BaseIO	
<input type="text" value="none"/>	
Score Most Likely	Score Most Unit
<input type="text" value="240.9119"/>	<input type="text" value="most likely total risk"/>
Score Low Estimate	Score Low Unit
<input type="text" value="239.1666"/>	<input type="text" value="lower 80 % ci"/>
Score High Estimate	Score High Unit
<input type="text" value="242.6572"/>	<input type="text" value="upper 80 % ci"/>
Score Math Type	Score Math Sub Type
<input type="text" value="algorithm1"/>	<input type="text" value="subalgorithm3"/>
Score Math Result	
<input type="text" value="original spearman
DenseMatrix 6x6-Double
1.000000000 0.000000000 0.000000000 0.000000000 0.000000000 0.000000000"/>	
Joint Data	
<input type="text" value="https://devtreks1.blob.core.windows.net/resources/network_carbon/resourcepack_1561/resource_9207/Score-Corr-Inds4.csv"/>	
Calculations Description	
<input type="text" value="Example 6B in the Social Performance Analysis 3 reference. v214i"/>	
Media URL	
<input type="text"/>	
Data URL	
<input type="text" value="https://devtreks1.blob.core.windows.net/resources/network_carbon/resourcepack_1561/resource_9206/Score-Corr-Data4.csv"/>	
Score Math Result	
Output Group: RCA Output Examples	



DevTreks –social budgeting that improves lives and livelihoods

The reason that the previous images TEXT datasets have names that end with “4” (...Inds4.csv) is that three more complex datasets were unsuccessfully tested with this algorithm before defaulting back to the simpler, CTA01 introductory datasets (confirming that these introductory algorithms need dedicated staff working on the next generation algorithms introduced in Example 7).

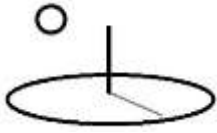
C. Reporting

The differences from Example 5 and 6’s reports and scorecards include:

1. CEAs measurements come directly from techniques in 4A, 4B, and 4C.
2. Starting SDG count comes from the “ending” SDG count from the previous Indicator.

The first Indicator measures starting benchmark, or baseline, conditions.

Advanced risk analysis, such as illustrated in the following chart from Example 6B’s Score.MathResults, support the manual completion of the following types of decision support aids. The CTA01 and CTAP reference demonstrates adding monetary benefit and loss metrics, such as supported by the Reference Case Cost Effectiveness Results reports, to these decision support aids. INFORM (2018) demonstrates how standardizing these scoring systems to a 10 point scale supports comparisons over time and place.

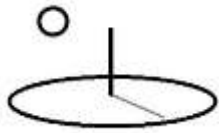


Total Risk Score	163.9182	201.8495	215.1461	223.325	231.8897	240.5363	247.8784	255.0034	266.4874	281.3504	346.3898
Probability	0	10	20	30	40	50	60	70	80	90	100



Footnotes

1. Although supporting abbreviated SIAs, this algorithm still does not target “lowest common denominator” sustainability officers, or RCA Assessors. The author has found that professional work is best accomplished with professional workers. Training those professionals for the work is better than assuming the workers aren’t capable of advancement. Workers may appreciate being trained for “next generation” jobs, rather than being led to believe that “last generation” jobs have a future on a sustainable planet.
2. The only way to fully proof the validity of these algorithms is field testing and running controlled experiments with businesses. It’s likely that all of the algorithms in these references need further refinement as well as the development of related algorithms that offer alternative approaches for achieving the desired targets. DevTreks doesn’t raise funds for these purposes because the author dislikes administrative work and because equating success with money sends the wrong message to any potentially “wiser generation”.



DevTreks –social budgeting that improves lives and livelihoods

3. The following error message was generated when the Make Base command element was first clicked. The Calculators and Analyzers tutorial explains that, since only children Input and Output Series are used in Budgets, each Output and Input must have at least 1 child series before calculations can be run. That tutorial also explains that this contended error message has been retained because DevTreks is not consumer software –it requires study and hard work (and because clicking on an Add Output Series command element is not part of the hard work).

GreenTreks	Search	Preview	Select
Edit	Pack	Views	Club

Point of contention: linkedviews
This contention message can be erased by clicking on any tab.
Application: prices
SubApplication: outputprices
Node Name: output
Id: 2141223502
Service Id: 3449
Account Id: DevTreks.Models.Account
Admin Message: System.Data.SqlClient.SqlException (0x80131904): The offset specified in a OFFSET clause may not be negative. at System.Data.SqlClient.SqlConnection.OnError(SqlException exception, Boolean breakConnection, Action`1 wrapCloseInAction) at System.Data.SqlClient.SqlInternalConnection.OnError(SqlException exception, Boolean breakConnection, Action`1 wrapCloseInAction) at

References

Example 5, plus the following.

OECD. Handbook on Constructing Composite Indicators. Methodology and User Guide. 2008



Example 7. SDG Stakeholder Resource Conservation Value Accounting using Machine Learning (ML)

A. Introduction to Impact Evaluation

Gertler et al (2016) explain how to conduct formal Impact Evaluation for social interventions. The authors provide concrete examples that use standard statistical analysis techniques, such as regression analysis, to evaluate, or verify, the impact of the interventions. The authors summarize the “fundamental equation” underlying impact evaluation in the following images.

the change in an outcome. To establish causality between a program and an outcome, we use impact evaluation methods to rule out the possibility that any factors other than the program of interest explain the observed impact.

The answer to the basic impact evaluation question—what is the impact or causal effect of a program (P) on an outcome of interest (Y)?—is given by the basic impact evaluation formula:

$$\Delta = (Y | P = 1) - (Y | P = 0).$$

This formula states that the causal impact (Δ) of a program (P) on an outcome (Y) is the difference between the outcome (Y) *with* the program (in other words, when $P = 1$) and the same outcome (Y) *without* the program (that is, when $P = 0$).

...

The basic impact evaluation formula is valid for any unit that is being analyzed—a person, a household, a community, a business, a school, a hospital, or other unit of observation that may receive or be affected by a program. The formula is also valid for any outcome (Y) that is related to the program at hand. Once we measure the two key components of this formula—the outcome (Y) both with the program and without it—we can answer any question about the program’s impact.

Appendix A introduce automated versions of Gertler’s standard statistical analysis techniques. Examples 8A, 8B, and 8C introduce complementary Machine Learning (ML) techniques. These



DevTreks –social budgeting that improves lives and livelihoods

techniques will be applied during the Impact Evaluation stage of the sustainable M&E accounting system demonstrated throughout this tutorial.

B. Introduction to RCA Technologists and RCA Assessors Roles, or Automated Statistical Analysis

This example differs from Gertler's, and most conventional statistical analyses, techniques in several IT-based ways.

1. **Developers or Data Scientists, such as RCA Technologists, build online algorithms. Analysts, such as RCA Assessors, apply the online algorithms.** Current statistical packages require every user to learn the statistical languages and data conventions and to work on a local server. That's inefficient and wasteful. While the developers of algorithms need to understand the intricacies of algorithms, the majority of analysts only need to understand the logic behind an algorithm. For example, is it really necessary for every analyst to understand every scripting language convention demonstrated in Examples 8A to 8C, or can 1 IT team build the scripts and then have thousands of analysts employ the packaged algorithms? Once a Social Performance Analysis, or sustainable accounting, data system has been set up, RCA Assessors consult with the RCA Technologists on the best algorithms that will achieve the desired targets. The Assessors focus on the mechanics of using the algorithms: determining the base element data context, facilitating data collection, setting up TEXT datasets needed by the algorithm, understanding stakeholder perspectives, assessing tradeoffs, forecasting scenarios, and completing the final communication aids that support decision making.
2. **Developers use APIs to Automate Everything.** DevTreks started with `algorithm1` because it employs an open-source mathematical library, MathNet or Numerics, with a straightforward API that can be used as a standard C# module, and easily compiled into each released version of DevTreks. Standard statistical analysis packages, including R, Python, or Stata, require separate deployment and awkward coding, such as the use of batch files or DLL syntax (**1***) (i.e. see the CTA02 and CTA03 references). Most statistical analysis software is evolving in ways that might make them easier to code and



DevTreks –social budgeting that improves lives and livelihoods

deploy. The machine learning algorithms will employ the latest developments in this field.

3. **Occam’s Rule supports imperfect socioeconomic decision making.** DevTreks applies Occam’s Rule for data management because it supports the final, imperfect decision making, objective in 95% of all cases: a) Basic algorithms should be used over complex algorithms, b) When possible, TEXT datasets should be limited to 10 columns of explanatory variables, and c) Indicator metadata should be limited to a minimal number of uniform quantitative properties (i.e. Q1-Q5, QTM-QTL). Occam appears to have recognized that the data, the analyst, the algorithms, the statistical techniques, and the decision makers, are all imperfect. Chakraborty and Joseph (2017) provide a formal explanation about under-fitting and over-fitting of algorithms and discuss ML and Occam further.
4. **Qualitative analysis complements quantitative analysis.** This tutorial emphasizes “mixed methods analyses” techniques that include qualitative techniques such as informant interviews, case studies, multimedia story-telling, and review of historical context. Besides triangulation of evidence and better understanding of actual stakeholder needs, they also recognize that quantitative techniques are imperfect and have a tendency to exaggerate claims of proof.

C. Introduction to Machine Learning (ML) Impact Evaluation

The following images (Chakraborty and Joseph, 2017; Varian, 2016) explain some of the primary differences between how ML and standard statistical algorithms are used. Physical scientists may note that the validation and testing of algorithms closely matches their approach for building simulation models (i.e. of climate change).

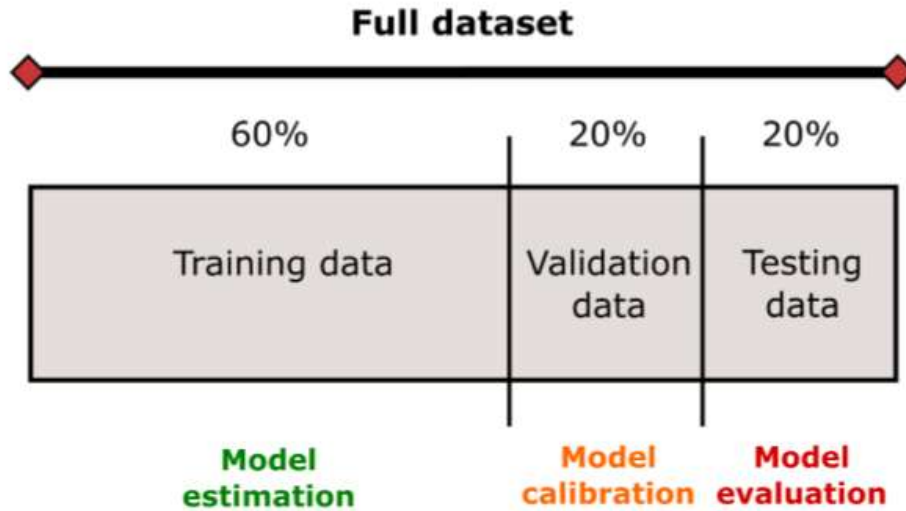
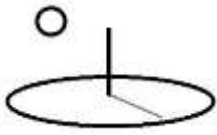


Figure 1: Schematic representation of the three stages of model generation in machine learning: training (model fit), validation (calibration) and testing (evaluation). The given data fraction offer a starting point if no cross-validation is involved (see Section 4.3).

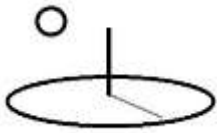
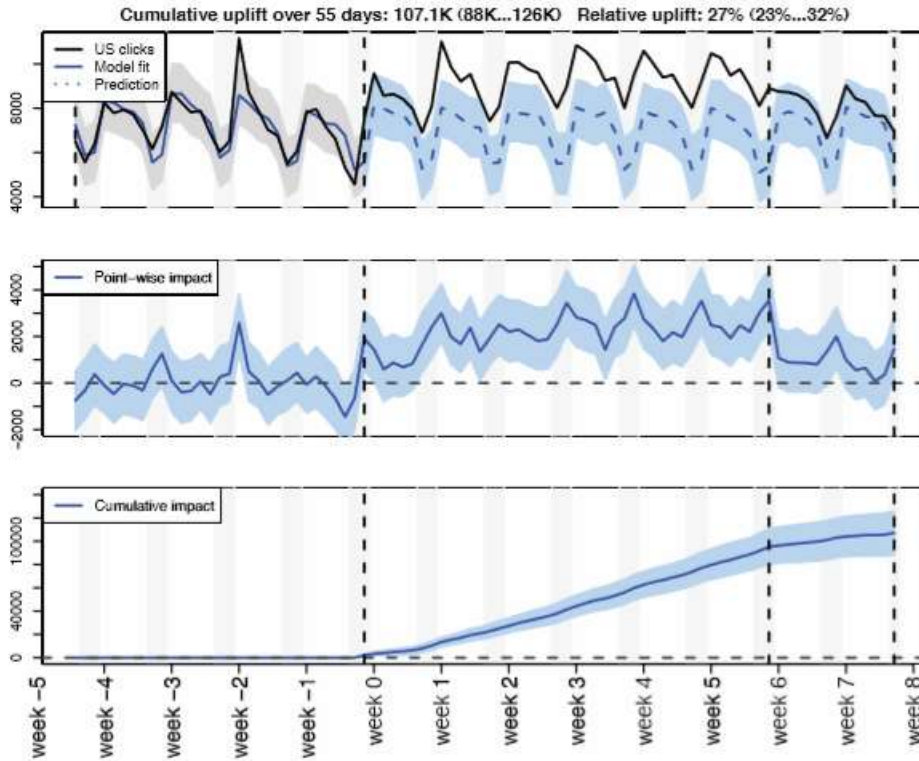
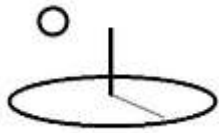


Figure 8
Actual and Predicted Website Visits

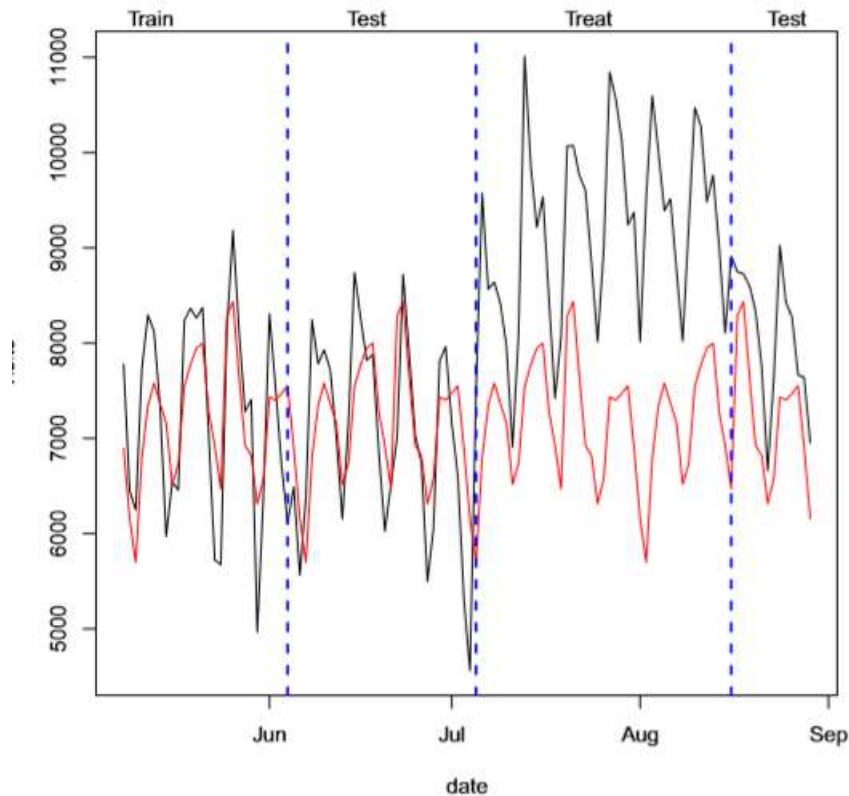


Source: This example is based on the approach proposed in Brodersen, Gallusser, Koehler, Remy, and Scott (2013), but the covariates are chosen automatically from Google Trends categories using Bayesian Structural Time Series (BSTS).

Notes: Suppose a given company wants to determine the impact of an advertising campaign on its website visits. Panel A shows the actual visits and the prediction of what the visits would have been without the campaign based on the BSTS forecasting model. Panel B shows the difference between actual and predicted visits, and Panel C shows the cumulative difference.



DevTreks –social budgeting that improves lives and livelihoods

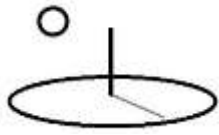


g. 1. Hypothetical TTTC process. The model is estimated during the training period and its predictive performance is assessed during the test period. The extrapolation of the model during the treat period (red line) serves as a counterfactual. This counterfactual is compared with the actual outcome (black line), and the difference is the estimated treatment effect. When the treatment ended, the outcome returns to something close to the original level.

However, where does the counterfactual come from? Answer: it is a predictive model developed using data from before the experiment was run.

In the classic experiment design described earlier, we compare treated and untreated subjects. Here, we treat all of the subjects for a limited time and measure their aggregate response. Our counterfactual is a prediction of what would have happened during the limited period of spend increase.

Athey (2018) highlights the major difference between ML's emphasis on using algorithms and the data itself to make predictions and statistics emphasis on using statistical models to infer causal attribution. The author suggests that the differences between the 2 approaches are gradually becoming complements. The following image (Mullainathan and J. Spiess, 2017)



highlight the approach taken by many analysts to discern the differences between standard statistical techniques, in this case ordinary least squares regression, and alternative ML approaches. Most of these authors conclude that ML complements, and in some cases, surpasses, conventional statistical techniques, such as regression.

Table 1
Performance of Different Algorithms in Predicting House Values

Method	Prediction performance (R^2)		Relative improvement over ordinary least squares by quintile of house value				
	Training sample	Hold-out sample	1st	2nd	3rd	4th	5th
Ordinary least squares	47.3%	41.7% [39.7%, 43.7%]	-	-	-	-	-
Regression tree tuned by depth	39.6%	34.5% [32.6%, 36.5%]	-11.5%	10.8%	6.4%	-14.6%	-31.8%
LASSO	46.0%	43.3% [41.5%, 45.2%]	1.3%	11.9%	13.1%	10.1%	-1.9%
Random forest	85.1%	45.5% [43.6%, 47.5%]	3.5%	23.6%	27.0%	17.8%	-0.5%
Ensemble	80.4%	45.9% [44.0%, 47.9%]	4.5%	16.0%	17.9%	14.2%	7.6%

Note: The dependent variable is the log-dollar house value of owner-occupied units in the 2011 American Housing Survey from 150 covariates including unit characteristics and quality measures. All algorithms are fitted on the same, randomly drawn training sample of 10,000 units and evaluated on the 41,808 remaining held-out units. The numbers in brackets in the hold-out sample column are 95 percent bootstrap confidence intervals for hold-out prediction performance, and represent measurement variation for a fixed prediction function. For this illustration, we do not use sampling weights. Details are provided in the online Appendix at <http://e-jep.org>.

The following image (self-referenced) summarizes prominent differences between the techniques employed by statistics and ML.



ML versus Statistics

Statistics:

- Hypothesis testing
- Experimental design
- Anova
- Linear regression
- Logistic regression
- GLM
- PCA

Machine Learning:

- Decision trees
- Rule induction
- Neural Networks
- SVMs
- Clustering method
- Association rules
- Feature selection
- Visualization
- Graphical models
- Genetic algorithm

<http://statweb.stanford.edu/~jhf/ftp/dm-stat.pdf>

D. Stakeholder System Boundary and Stakeholder Engagement

Examples 5 and 6 explain the importance of ensuring that individual business and CSO activity takes place as part of a larger, integrated landscape management, or watershed planning, approach. Applying that principle to Gertler's dataset means that improved health care delivery is 1 target needed in an overall sustainable development end goal, or SDG target. Those targets become the basis for the measurements being made in the sustainable accounting M&E business systems introduced in this tutorial.

Several references discuss basing Impact Evaluations on the needs and expectations of the impacted stakeholders. These references point out that, prior to conducting a full scale evaluation, smaller scale, mixed methods, qualitative and quantitative studies, such as case



studies, key stakeholder interviews, preparation of operating and capital budgets, can help to ensure that the evaluation is really addressing the greatest concerns of the stakeholders. The point is that, as Gertler et al (2016) explain, Impact Evaluation must take place as part of an ongoing, comprehensive, societal improvement context. They must also be included at the start of societal improvement interventions. This reference argues that the SDG and SDRR make good contexts for societal improvement.

Datasets

Dataset A. Standard Statistics. Gertler et al (2016) use the following reference dataset, containing 19,800+ rows of population socioeconomic characteristics, to conduct an Impact Evaluation of a health insurance CSO social intervention. This dataset is the primary dataset used by all algorithms introduced in this example.

	locality_id	household	treatment	promotion_eligible	enrolled	enrolled_r	poverty_index	round	health_ex	age_hh	age_sp	educ_hh	educ_sp	female_hh	indigenous	hhsz	dirtfloor	bathroom	land	hospital_d	hospital
0	26	5	1	1	1	1	55.95054245	0	15.18546	24	23	0	6	0	0	4	1	0	1	124.82	0
1	26	5	1	1	1	1	55.95054245	1	19.5809	25	24	0	6	0	0	4	1	0	1	124.82	0
2	26	11	1	1	1	1	46.05873108	0	13.07626	30	26	4	0	0	0	6	1	0	2	124.82	0
3	26	11	1	1	1	1	46.05873108	1	2.398854	31	27	4	0	0	0	6	1	0	2	124.82	1
4	26	13	1	1	1	1	54.0958252	1	0	59	57	0	0	0	0	6	1	0	4	124.82	1
5	26	13	1	1	1	1	54.0958252	0	15.28635	58	56	0	0	0	0	6	1	0	4	124.82	0
6	26	16	1	1	1	1	56.90340042	1	20.02691	36	25	3	0	0	0	7	1	0	2	124.82	0
7	26	16	1	1	1	1	56.90340042	0	11.31176	35	24	3	0	0	0	7	1	0	2	124.82	0
8	26	21	1	1	1	1	46.90880966	0	11.22391	37	35	0	0	0	0	7	1	0	2	124.82	0
9	26	21	1	1	1	1	46.90880966	1	16.66469	39	36	0	0	0	0	7	1	0	2	124.82	0
10	26	22	1	1	1	1	48.18392563	0	8.876925	34	41	0	2	1	0	9	1	0	1	124.82	0
11	26	22	1	1	1	1	48.18392563	1	0.115546	35	41	0	2	1	0	9	1	0	1	124.82	0
12	26	23	1	1	1	1	49.69087982	1	0	31	30	0	6	0	0	7	0	0	1	124.82	0
13	26	23	1	1	1	1	49.69087982	0	12.68582	30	29	0	6	0	0	7	0	0	1	124.82	0

The following images (Gertler, 2016) define these socioeconomic characteristics.

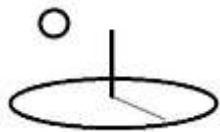


The dataset includes the following variables (columns):

Outcome variables	
health_expenditures	Out of pocket health expenditure (per capita per year)
Control variables	
age_hh	Age of the head of the household (in years)
Age_sp	Age of the spouse (years)
educ_hh	Education of the head of household (completed years of schooling)
educ_sp	Education of the spouse (completed years of schooling)
indigenous	Head of household speaks an indigenous language (0=no, 1=yes)
female_hh	Head of the household is a woman (0=no, 1=yes)
hhsiz	Number of household members (at baseline)
dirtfloor	Home has a dirt floor at baseline (0=no, 1=yes)
bathroom	Home with private bathroom at baseline (0=no, 1=yes)
land	Number of hectares of land owned by household at baseline
hospital_dist	Distance to closest hospital
Other variables	
locality_identifier	Locality identifier
household_identifier	Unique household identifier
round	Survey round (0 = baseline; 1 = follow-up)
enrolled	Household enrolled in HISP (0=no, 1=yes)
enrolled_rp	Household enrolled in HISP under the randomized promotion

	scenario (0=no, 1=yes)
eligible	Household eligible to enroll in HISP (0=no, 1=yes)
treatment_locality	Household is located in treatment community (0=no, 1=yes)
promotion_locality	Household is located in locality randomly assigned promotion of HISP (0=no, 1=yes)
poverty_index	Poverty index 1-100 (eligible \leq 58)
hospital	HH member visited hospital in the past year (0=no, 1=yes), used in power calculations

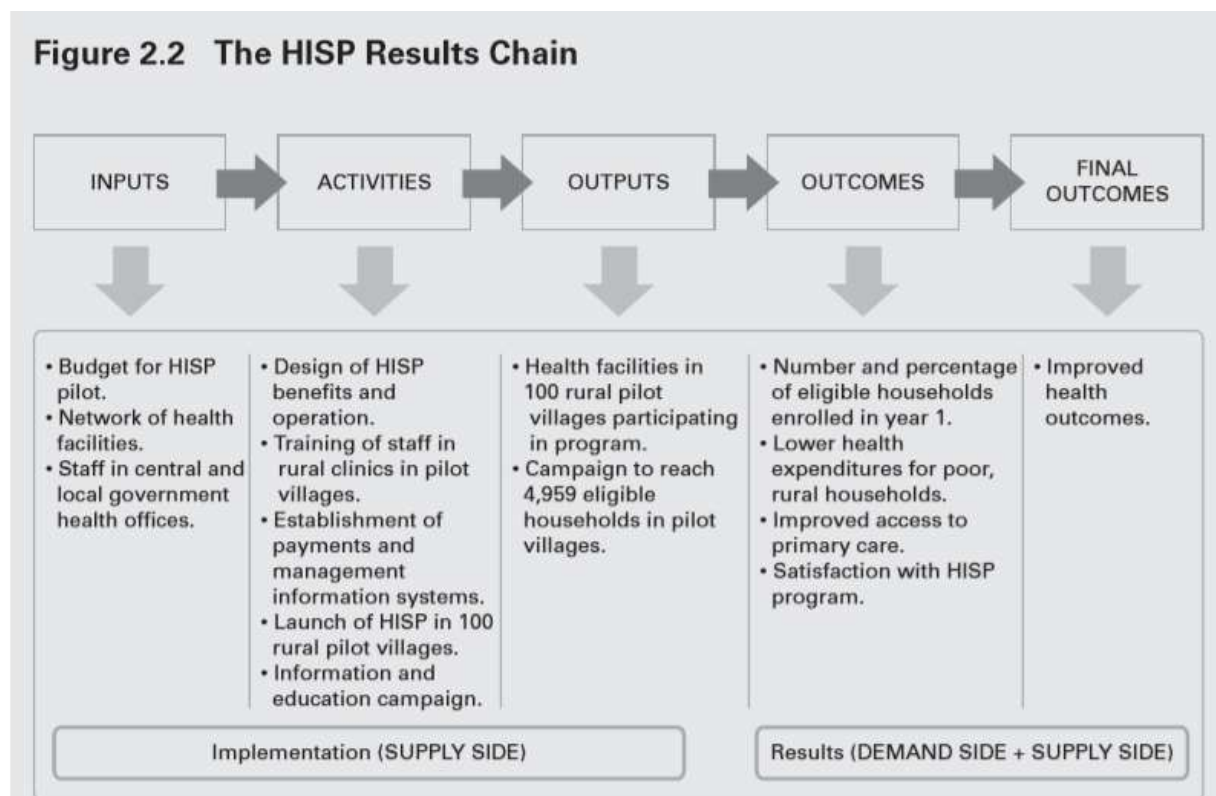
[Dataset B. Big Data Statistics. An additional dataset will introduce a simplified “big data” technique that breaks Gertler’s dataset into separate datasets, derived from randomly changing each of Gertler’s village datasets, using Example 5 and 6’s data structures. This dataset explores potential ML techniques that aggregate the individual datasets from this tutorial’s M&E sustainable accounting data system.]



DevTreks –social budgeting that improves lives and livelihoods

E. Risk and Impact Identification

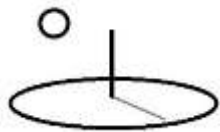
Gertler et al (2016) employ the following image of their results chain to summarize the logic for an Impact Evaluation being conducted for a CSO social intervention. These pathways serve as longer term risk identification and reduction mechanisms, or Impact Evaluations, than the shorter term Performance Monitoring techniques introduced in Examples 1 through 6.



The authors use the following statement to describe the purpose of this CSO intervention. The statement confirms that Impact Evaluation focuses on the “Final Outcomes” (i.e. Impacts) part of the results chain. The key evaluation question addressed by Gertler is:

“What is the impact of HISP on poor households’ out-of-pocket health expenditures?”

The algorithms introduced in this reference supplement Gertler’s evaluation question with the following types of additional impact evaluation questions (2*). Given that Gerler’s dataset may not have been designed for these purposes, additional datasets may be introduced. Setting up



DevTreks –social budgeting that improves lives and livelihoods

these types of evaluation questions at the start of the social intervention (i.e. setting up the underlying M&E system) helps to identify the data that needs to be collected.

Evaluation Question 1. Which villages have the highest risk of not achieving targeted savings in poor households’ health expenditures (i.e. or more generally, SDG-SDRR targets)?

This evaluation is similar to how the INFORM (2018) global index of disaster risk assessment in Example 6 identifies countries that face the greatest risk of negative impacts from disasters. Chakraborty and Joseph (2017) provide a similar case study demonstrating how to use ML to detect banks in financial trouble that need to be alerted of the central bank’s concerns.

Evaluation Question 2. How accurately can the impact of recommended M&E actions be predicted for individual village, poor households’ health expenditures (i.e. or more generally, SDG-SDRR targets)?

This evaluation demonstrates a typical use of ML techniques –making predictions. The example also explores the relation between prediction and “cause and effect” attribution.

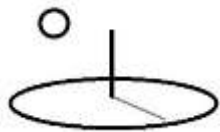
Evaluation Question 3. What are the principal factors causing improvements in individual village, poor households’ health expenditures (i.e. or more generally, SDG-SDRR targets)?

This evaluation explores the relation between prediction and “cause and effect” attribution at greater depth.

Evaluation Question 4. What are the trends in poor households’ health expenditures (i.e. or more generally, SDG-SDRR targets)?

This evaluation is similar to how the INFORM (2018) global index of disaster risk assessment in Example 6 identifies trends in individual country social risks. Chakraborty and Joseph (2017) provide a similar case study demonstrating how to use ML to forecast trends in short term inflation.

F. Quality of Life, or Sustainability, Scenarios



DevTreks –social budgeting that improves lives and livelihoods

Typical scenarios for assessing alternative ways to achieve the SDG and SDRR targets include:

General Scenario

Threatened Quality of Life: High GHG result in 1.5°C temperature increase with higher incidence of health care cost increases, biodiversity loss, droughts, severe heat waves, crop and livestock production risks, air pollution, floods, migration, and social discord that leads to the incapacity to achieve the SDG and SDRR targets.

Targeted Social Performance Risks and Targets: Combinations of the SDG and SDRR targets, adapted to local stakeholders, including (2*):

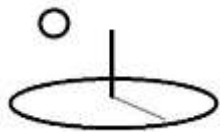
- [Local] target A: Substantially reduce [local] household expenditures on health care while increasing health outcomes
- [Local] target B: Substantially reduce the number of people [locally] who do not have adequate health care

Mitigation and Adaptation Actions: Portfolio 1 consists of a) ..., b)...., and c)....

G. Introduction to Advanced Machine Learning Impact Evaluation

Examples 1 to 5 demonstrate that full SDG-related accountability involves considerably more complexity. Watersheds, ecosystems, the 7 “community capitals”, demographics, stakeholder perspectives, and land use management, are melded together into business and organization sustainable accounting data systems. The impact pathways, impact transition states, scenarios, and hierarchical base elements, used with these accounting systems and applied through algorithms, help companies and organization to identify and reduce risks. Gertler’s dataset, and most of the Impact Evaluation references, have little or nothing to do with private and public sector sustainable accounting systems.

The real value of machine learning comes from turning all of that complexity into actionable decision support. Specifically, companies and organizations want to know the most effective and efficient ways, such as concrete mitigation and adaptation actions, or management policies, needed to achieve the SDG and SDRR targets while advancing the company’s financial performance goals. The holy grail of sustainable development is to more fully understand the



DevTreks –social budgeting that improves lives and livelihoods

cause and effect attribution that leads to better knowledge of what works and what doesn't work. This reference [may] demonstrate additional machine learning algorithms that turn sustainable accounting system complexity into actionable decision support results.

Advanced machine learning algorithms will answer evaluation questions similar to the following (3*):

- 1. M&E:** Given the changing circumstances identified by the current Performance Monitoring System, how can long term SDG and SDRR Targets be achieved, given the constraints associated with the socioeconomic and physical characteristics present in this ecosystem and community?
- 2. Business Value Reports:** To what extent has the [company] [organization] [community] been effective and efficient in achieving SDG and SDRR targets that are relevant to our stakeholders? [examine ML causal chain algorithms]
- 3. Tradeoff Analysis:** Given the limited resources this company or community can muster, what are the most effective, efficient, and equitable, mitigation and adaptation activities that will further the SDG and SDRR Targets for targeted stakeholder perspectives?
- 4. Scenario Analysis:** Given the predicted impact transition states for the scenarios relevant to this company or community, which alternatives have the greatest probability of achieving cost effective, equitable, SDG and SDRR Targets for targeted groups of stakeholders?
- 5. Public Official and Company Executive Accountability:** Given the known costs and negative societal impacts associated with human-induced hazards caused by bad public service or private sector management, what are the best mechanisms for holding the transgressors accountable for their past irresponsibility?

H. Introduction to Standard Impact Evaluation using Statistical Software

The following image (Varian, 2016) summarizes most of the statistical techniques employed by Gertler et al and their relation to machine learning.



random, then we may be able to interpret the results as the impact of the treatment on the population.

There are many examples of difference in differences in the economics literature. For a recent application to online search advertising, see ref. 19.

Summary

We have described four techniques for estimating causal effects from observational data. In each case, techniques for predictive modeling from machine learning may be useful.

Experiments. With either a designed or natural experiment, it is important to have an estimate of the counterfactual—what would have happened in the absence of the experiment. This is essentially a problem of predictive modeling, an area where machine learning offers several powerful techniques.

Regression Discontinuity. When a treatment is applied depending on some threshold, one can estimate causal effects by comparing outcomes for experimental units on each side of the threshold. In this case, one wants to build a predictive model for behavior near the threshold. We can then use the predicted outcome for the treated group estimated using the training data from the untreated units.

Instrumental Variables. The first stage in an IV model involves predicting treatment as a function of instrumental variables (variables that are thought to be independent of potential confounders) along with other helpful covariates. There are good reasons why the instruments should enter the predictive model linearly, but the other covariates could easily be nonlinear. See ref. 5, pp. 190–192, for a discussion of the problem and some recommendations.

Difference in Differences. In this case, we have two groups, the treated and the untreated, and two time periods, before treatment and after treatment. We also have a number of predictors that may

affect the observed values of the outcome for each group. The goal is to estimate a predictive model of what the outcome would be for the treated group if it were not treated. To accomplish this goal, one can use a model, possibly nonlinear, of the observed outcomes of the untreated group in the posttreatment period.

In each of these cases, building a predictive model is a key step in identifying the causal impact. Machine-learning tools offer powerful methods for predictive modeling that may prove useful in this context.

Guide to Further Reading

If you know nothing about machine learning and would like to read an elementary introduction to it written for economists, see ref. 20. The references below are for those familiar with machine learning who want to learn more about the econometric approach to causal inference.

A more advanced approach to causal modeling involves “structural equation modeling,” which involves building a specific model of the data generating behavior. For example, in the Honolulu–Fargo example, we might build a model of how marketing managers choose to allocate ad spend across cities and estimate the behavioral effects along with the responses. See ref. 21 for a detailed survey.

There is also a large literature on propensity scores, which are estimates of the probability of treatment as a function of “observed” characteristics. This approach can be contrasted with the confounding variables framework described earlier, which involves selection on “unobservables.” See ref. 4 for an up-to-date review.

With respect to the econometrics literature, ref. 5 provides a very accessible introduction and ref. 8 provides a somewhat more advanced description of the methods outlined here. Ref. 23 describes the historical development of these methods.

Finally, there are graphical methods pioneered by Pearl (23, 24) that allow one to analyze complex models to determine when and how various causal effects can be identified.

Appendix A demonstrates carrying out Impact Evaluation using some of the standard statistical analysis techniques introduced in the Gertler et al (2016) reference. The Appendix modifies Gertler’s Stata scripts (see the Technical Companion, Version 1, September, 2016) to related statistical software scripts.

I. Introduction to Machine Learning Impact Evaluation (4*)

Appendix B introduces several ML algorithms that address specific impact evaluation questions. Wikipedia URLs provide the definition for each algorithm. Bontempi (2017) provides the mathematical basis for each algorithm. Smola and Vishwanathan (2008) provide the actual algorithm. Examples 8A, 8B, and 8C demonstrate using alternative statistical software to run each algorithm.

Appendix B introduces the following algorithms:



DevTreks –social budgeting that improves lives and livelihoods

Example a. Naïve Bayes. Predicts the poverty index classification of households and villages in the Gertler dataset.

Example b. Deep Neural Network. Predicts a household health care expenditure category for households and villages in the Gertler dataset.

Example c. Time Series Deep Learning. Predicts health care expenditures for upcoming years for fictitious households and villages in the Gertler dataset.

J. Decisions

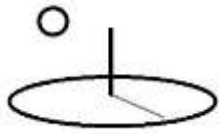
Athey (2018) makes the following predictions about how ML will change the role of applied socioeconomic experts, including economists, in assisting decision making.



5 Broader Predictions About the Impact of Machine Learning on Economics

My prediction is that there will be substantial changes in how empirical work is conducted; indeed, it is already happening, and so this prediction already can be made with a high degree of certainty. I predict that a number of changes will emerge, summarized as follows:

1. Adoption of off-the-shelf ML methods for their intended tasks (prediction, classification, and clustering, e.g. for textual analysis)
2. Extensions and modifications of prediction methods to account for considerations such as fairness, manipulability, and interpretability
3. Development of new econometric methods based on machine learning designed to solve traditional social science estimation tasks
4. No fundamental changes to theory of identification of causal effects
5. Incremental progress to identification and estimation strategies for causal effects that exploit modern data settings including large panel datasets and environments with many small experiments
6. Increased emphasis on model robustness and other supplementary analysis to assess credibility of studies
7. Adoption of new methods by empiricists at large scale
8. Revival and new lines of research in productivity and measurement
9. New methods for the design and analysis of large administrative data, including merging these sources and privacy-preserving methods
10. Increase in interdisciplinary research
11. Changes in organization, dissemination, and funding of economic research
12. Economist as engineer engages with firms, government to design and implement policies in digital environment
13. Design and implementation of digital experimentation, both one-time and as an ongoing process, including “multi-armed bandit” experimentation algorithms, in collaboration with firms and government



DevTreks –social budgeting that improves lives and livelihoods

14. Research on developing high-quality metrics that can be measured quickly, in order to facilitate rapid incremental innovation and experimentation
15. Increased use of data analysis in all levels of economics teaching; increase in interdisciplinary data science programs
16. Research on the impact of AI and ML on economy

K. Conclusions

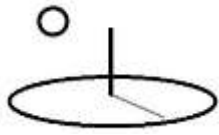
Gertler et al (2016), IITA and COSA (2016), and Hoebink (2014) prove that standard statistical analyses techniques have tremendous value for understanding the cause and effect attribution that helps to achieve sustainable social performance. Example 7 and 8 demonstrate that all of those techniques can be automated as online algorithms, run from subnational, national, or cloud computing data centers, and used by greater numbers of sustainability workers to monitor social performance, evaluate social impacts, report transparently, keep governments and companies aligned with societal objectives, and improve people’s quality of life.

Athey (2018) use the following statement to predict a prominent change in the role of socioeconomic experts, including applied sustainability workers. That is, they’ll become akin to RCA Technologists and RCA Assessors (5*).

“As digitization spreads across application areas and sectors of the economy, it will bring opportunities for [sustainability workers] to develop and implement policies that can be delivered digitally.”

Footnotes

1. The DevTreksStatsApi web application (DTSA) was built because standard statistical packages, such as R and Python, can only be deployed in cloud computing data centers using dedicated servers, or custom data science virtual servers. DTSA exposes a simple API that allows remote software developers to access the remote statistical packages by sending standard http request commands, containing this example’s Script URL and Data URL, and receiving http responses that return the URL holding the statistical results.



These types of applications often support long running remote calculations using server-side techniques such as data queues. DevTreks will clearly document whether or not the cloud computing app will support long running calculations, and if not, DTSA will be upgraded to support that purpose. In the latter case, the remote client processes the Stat Result URL after it returns (i.e. in an email message).

2. As mentioned or implied throughout this tutorial, given that the conventional institutions in some countries appear incapable of understanding, let alone ameliorating, these conditions, networks and clubs may need to take independent action.
3. An appropriate question for “informed” software developers to ask is:
Will algorithm 350, subalgorithm 2154, really help to solve this societal problem?
4. Larger datasets take longer to load and run, which online algorithms must take into account. The first 3 examples of regression were added to 3 consecutive Indicators and run concurrently. The results were displayed after a noticeable delay, suggesting: 1) don’t run an excessive number of algorithms concurrently, 2) break the dataset into the subset data used by the algorithm prior to running the calculation, 3) do not run children series algorithms from their parents –run them separately when conducting long running calculations, and 4) use the next evolution of the DevTreksStatsAPI app to run long running calculations on remote data sciences machines.
5. This reference continually points out “doing it right” will only occur if substantial changes take place in existing, conventional, institutions. Software development is 1 means towards an institutional reform, or replacement, end goal. For example, the primary goal of RCA Technologists and RCA Assessors is evidence-based societal improvement through sound IT. In this context, open source repository publishing is far more important than academic publishing. Conventional institutions may have skewed incentives (or at least, pre-IT era performance targets). This reference predicts that consequential digital activism will find its own path.

Case Study References

Example 5, plus the following.



DevTreks –social budgeting that improves lives and livelihoods

Susan Athey. *The Impact of Machine Learning on Economics*. January, 2018. last accessed: April, 2018: <http://www.nber.org/chapters/c14009.pdf>

Gianluca Bontempi, *Handbook Statistical foundations of machine learning*. Machine Learning Group Computer Science Department Universite Libre de Bruxelles, ULB Belgique. June 2, 2017

Chiranjit Chakraborty and Andreas Joseph. Staff Working Paper No. 674 *Machine learning at central banks*. Bank of England. 2017. last accessed April 18, 2018:

<https://www.bankofengland.co.uk/-/media/boe/files/working-paper/2017/machine-learning-at-central-banks.pdf?la=en&hash=EF5C4AC6E7D7BDC1D68A4BD865EEF3D7EE5D7806>

S. Mullainathan and J. Spiess. *Machine learning: an applied econometric approach*. *Journal of Economic Perspectives*, 31(2):87–106, 2017. Online Appendix. April, 2017

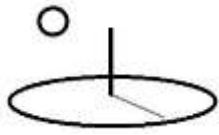
Smola, Alex and S.V.N. Vishwanathan. *Introduction to Machine Learning*. Yahoo! Labs Santa Clara –and– Departments of Statistics and Computer Science Purdue University –and– College of Engineering and Computer Science Australian National University. Cambridge University Press. 2008

Varian, Hal R. *Big Data: New Tricks for Econometrics*. *Journal of Economic Perspectives*—Volume 28, Number 2—Spring 2014—Pages 3–28

Varian, Hal R. *Causal inference in economics and marketing*. pgs. 7310–7315, *PNAS*, July 5, 2016, vol. 113, no. 27

World Bank. *Miscellaneous blogposts*.

<http://www.worldbank.org/en/news/video/2018/02/27/machine-learning-future-of-poverty-prediction>



DevTreks –social budgeting that improves lives and livelihoods

Example 8A. Introduction to Machine Learning (ML) Algorithms using .NetStandard Libraries

Algorithms: algorithm1, multiple subalgorithms;

A. ML NetStandard Examples (algorithm1, subalgorithm_xx)

This example uses the .NetStandard mathematical libraries, MathNet, Accord, Cognitive Toolkit, and System.Math (1*), to demonstrate examples of basic ML algorithms. All of these libraries are compiled and run directly within DevTreks with no need for separate statistical or mathematical software (2*).

Version 2.1.4 began distinguishing machine learning (ML) subalgorithms from standard statistical subalgorithms for algorithm1 using the following conventions:

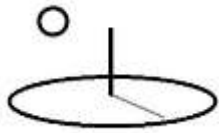
- a. **Subalgorithm Names:** Names must use a “_xx” suffix, such as “subalgorithm_01”.
- b. **Indicator.URL and Score.URL Training and Testing Datasets:** ML datasets must use 2 separate TEXT files –the first is for a ML training dataset, the second for a ML test dataset. The URLs to these datasets must be stored using the Indicator.URL property for Indicators and Score.URL for Scores. Most algorithms will return a new testing dataset that holds row by row calculations. This testing dataset will be stored in either an Indicator.MathResult or Score.MathResults URL.
- c. **Datasets broken into subsets.** Example 7, Footnote 4, explained that online algorithms must be treated differently than desktop algorithms. Specifically, they should be 1) broken into subsets of data prior to running the algorithms, 2) not used to run children algorithms, and c) run using the DevTreksStatsAPI app for long running, remote, calculations.
- d. **Meta MathExpression determines dataset columns.** The Indicator.MathExpression and Score.MathExpression must define the columns of data to include in the analysis.
- e. **Score.Iterations determine number of training rows to use.** Certain machine learning algorithms take a long time to process and not all algorithms need every row



to adequately train the model. The Score.Iterations property determines the number of rows to use in the training dataset.

- f. **Data Transformation:** Version 2.1.4 started using the 2nd row of the training and test datasets and the Indicator and Score metadata to pass instructions to ML algorithms. The following table defines the data transformation commands available in this release.

2 nd Row and Meta	Command	Purpose
columns 1 to 3	varied	Algorithm-specific instructions
columns 4+	none	Do not transform the data column
columns 4+	index	Transform the data column to integer categories
columns 4+	text	Do not transform the data column. Treat the data as a string datatype.
columns 4+ and Meta	qindex	Transform the data column to integer categories derived from index positions for thresholds defined in Indicator.Q1 to Indicator.Q5. At least 2 thresholds must be defined.
columns 4+ and Meta	qcategory	Transform the data column to double categories derived from double thresholds (i.e. midpoints) used in Indicator.Q1 to Indicator.Q5. At least 2 thresholds must be defined.
columns 4+ and Meta	qtext	Transform the double data column to double categories derived from double thresholds (i.e. midpoints)

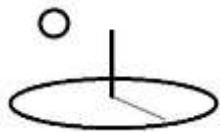


		used in Indicator.Q1 to Indicator.Q5. Then transform the double categories to their corresponding Indicator.Q1Unit to IndicatorQ5Unit text categories. At least 2 thresholds must be defined.
columns 4+	minmax, zscore, logistic, logit, tanh	Normalize the data column according to the command.
Meta: Indicator.RelatedLabel	Alternative statistical library name (under further planning)	Run the algorithm using the library specified

- g. **Automatic Indicator Meta Completion:** The final 3 rows of the Test TEXT file will be used to automatically fill in Indicator.meta properties.
- h. **Manual Indicator Meta Completion:** The quantitative Indicator.meta properties (i.e. Q1, Q2, Q3, Q4, Q5, QT, QTD1, QTD2, QTM, QTL, and QTU) are only replaced by the algorithms when their starting values = 0 or “none”. These algorithms assume that when their starting values are not 0, the analyst has made manual calculations based on the mathematical results (i.e. manual predicted values, confidence intervals, or algorithm instructions). In the same vein, the text Indicator.meta properties (i.e. Q1Unit, Q2Unit, Q3Unit, Q4Unit, Q5Unit, QTUnit, QTD1Unit, QTD2Unit, QTMUnit, QTLUnit, and QTUUnit) are only replaced when their values are string.empty, null, or “none”.
- i. **TEXT csv Math Results:** When possible, statistical results are returned as row-column TEXT, csv, datasets (i.e. see Examples 1 to 5).

Example a. Naïve Bayes (algorithm1, subalgorithm_01) (3*)

System.Math and MathNet URL:



DevTreks –social budgeting that improves lives and livelihoods

Many of the machine learning datasets exceed the 500KB db storage limit. They are not stored in the db and therefore can't be automatically stored in file systems and blobs by previewing them. They must be manually uploaded to their referenced Resource. The Source Code reference has a URL to a zip file where these files can be obtained.

[https://www.devtreks.org/greentreks/preview/carbon/resourcepack/Resource Conservation Value Accounting Example 8A/1563/none](https://www.devtreks.org/greentreks/preview/carbon/resourcepack/Resource%20Conservation%20Value%20Accounting%20Example%208A/1563/none)

[https://www.devtreks.org/greentreks/preview/carbon/output/Machine Learning 8A/2141223488/none](https://www.devtreks.org/greentreks/preview/carbon/output/Machine%20Learning%208A/2141223488/none)

[http://localhost:5000/greentreks/preview/carbon/resourcepack/Resource Conservation Value Accounting Example 8A/549/none](http://localhost:5000/greentreks/preview/carbon/resourcepack/Resource%20Conservation%20Value%20Accounting%20Example%208A/549/none)

[requires first uploading the dataset manually because of its size]

[http://localhost:5000/greentreks/preview/carbon/output/Machine Learning 8A/2141223503/none](http://localhost:5000/greentreks/preview/carbon/output/Machine%20Learning%208A/2141223503/none)

McCaffrey (MSDN, February, 2013) uses System.Math to demonstrate using this algorithm to make predictions about how a dependent variable, or label, can be classified based on the independent variables, or features. The following image introduces McCaffrey's code for this algorithm. The code is available as open access from the referenced issue. McCaffrey cautions that the purpose of these algorithms is to teach software developers how to build better algorithms. All of McCaffrey's code has been modified for this tutorial's M&E sustainability accounting system.

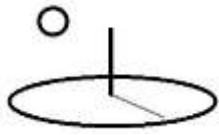


Figure 2 Naive Bayes Program Structure

```
using System;
namespace NaiveBayes
{
    class Program
    {
        static Random ran = new Random(25); // Arbitrary
        static void Main(string[] args)
        {
            try
            {
                string[] attributes = new string[] { "occupation", "dominance",
                    "height", "sex"};
                string[][] attributeValues = new string[attributes.Length][];
                attributeValues[0] = new string[] { "administrative",
                    "construction", "education", "technology" };
                attributeValues[1] = new string[] { "left", "right" };
                attributeValues[2] = new string[] { "short", "medium", "tall" };
                attributeValues[3] = new string[] { "male", "female" };
                double[][] numericAttributeBorders = new double[1][];
                numericAttributeBorders[0] = new double[] { 64.0, 71.0 };
                string[] data = MakeData(40);
                for (int i = 0; i < 4; ++i)
                    Console.WriteLine(data[i]);
                string[] binnedData = BinData(data, attributeValues,
                    numericAttributeBorders);
                for (int i = 0; i < 4; ++i)
                    Console.WriteLine(binnedData[i]);
                int[][][] jointCounts = MakeJointCounts(binnedData, attributes,
                    attributeValues);
                int[] dependentCounts = MakeDependentCounts(jointCounts, 2);
                Console.WriteLine("Total male = " + dependentCounts[0]);
                Console.WriteLine("Total female = " + dependentCounts[1]);
                ShowJointCounts(jointCounts, attributeValues);
            }
        }
    }
}
```

The following images show that the dataset and code have been modified to run as a standard DevTreks machine learning algorithm. The data conventions for ML subalgorithms for algorithm1 include:

- Statistical library name specified using the Indicator.RelatedLabel property



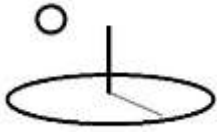
- separate training and test datasets stored in Indicator.URL,
- 2nd row of datasets pass instructions to the algorithm (in this image, the 1st column tells the algorithm whether or not to include Laplacian data transformation, and columns 4 to 7 retain their text categorical data). The 1st source code image show that all of the initial data goes through data normalization prior to being used in the analysis.
- 3 columns of row identifiers,
- dependent variable in the 4th column.
- remaining columns store the independent variables.
- MathExpression defines the columns of independent variables to include in the analysis
- Score.Iterations determines number of training dataset rows to use
- Categories of dependent variable data optionally defined by the Indicator.Q1 to Indicator.Q5 properties

label	locality_id	household	sex	occupation	dominance	height
laplacian=true	none	none	text	text	text	text
0	26	5	male	administra	left	short
1	26	5	male	administra	left	medium
2	26	11	male	constructi	left	medium
3	26	11	male	constructi	left	medium

In the following table, the 2nd row data instructions include the command, qcategory, which converts the double data type to a) double category thresholds defined by Indicator.Q1 to Q5 for analysis and b) text thresholds defined by Indicator.Q1Unit to Indicator.Q5Unit for reporting.

The poverty_index and hhsiz columns are normalized by their respective commands.

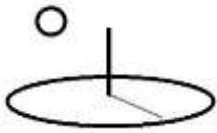
label	locality	househ	health	treatme	poverty	female_hh	hhsiz	dirtfloor
none	none	none	qcategory	none	zscore	none	minmax	none
1	26	5	19.5809	1	55.95054	0	4	1
3	26	11	2.398854	1	46.05873	0	6	1
4	26	13	0	1	54.09583	0	6	1
6	26	16	20.02691	1	56.9034	0	7	1
9	26	21	16.66469	1	46.90881	0	7	1
11	26	22	0.115546	1	48.18393	1	9	1
12	26	23	0	1	49.69088	0	7	0
14	26	25	17.2237	1	57.10974	0	9	0
17	26	27	1.284901	1	49.69088	0	7	1



1 reference | kpbayle1, 5 days ago | 1 author, 3 changes

```
private async Task<StringBuilder> Classify(List<List<string>> trainData,
    List<List<string>> rowNames, List<List<string>> testData)
{
    StringBuilder sb = null;
    try
    {
        //instructions in both row names and datasets
        List<string> dataInstructs = Shared.GetAlgoInstructs(rowNames);
        // prevent joint counts with 0
        bool withLaplacian
            = dataInstructs[0].ToLower().Contains("true") ? true : false;
        //converts rows to columns with normalized data
        List<List<string>> trainDB = Shared.GetNormalizedSData(trainData,
            this.IndicatorQT, _colNames, _depColNames, "F0");
        List<List<string>> testDB = Shared.GetNormalizedSData(testData,
            this.IndicatorQT, _colNames, _depColNames, "F0");
        int iColCount = testDB.Count;
        if (_subalgorithm == MATHML_SUBTYPES.subalgorithm_01.ToString().ToString())
        {
            //subalgo02 needs qtm and percent probability of accuracy, mse, qtm, low ci,
            iColCount = testDB.Count + 2;
        }
        //row count comes from original testdata to account for the instructions row
        DataResults = CalculatorHelpers.GetList(testData.Count, iColCount);
        // trainData columns define number of rows (depcolumns.Length + 1)
        string[][] attributeValues = new string[trainDB.Count][];
        //for each column of trainDB, fill in the unique attribute names (i.e. gender = 2
        for (int i = 0; i < trainDB.Count; i++)
        {
```

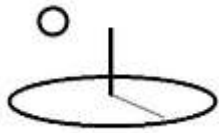
...



```
for (int i = 0; i < trainDB.Count; i++)
{
    attributeValues[i] = Shared.GetAttributeGroups(i, trainDB, this.IndicatorQT);
}
int[][][] jointCounts = MakeJointCounts(trainDB, attributeValues);
int[] dependentCounts = MakeDependentCounts(jointCounts, attributeValues[0].Length);
//classify everything in test dataset and add result to new columns in test dataset
List<string> predictors = new List<string>();
for (int r = 0; r < DataResults.Count - 1; r++)
{
    if (r == 0)
    {
        Shared.FillMLAlgoInstructions(r, DataResults, testData);
    }
    predictors = new List<string>();
    //cols have separate set of predictors
    for (int j = 0; j < testDB.Count; j++)
    {
        //prepare mathresults
        DataResults[r+1][j] = testDB[j][r];
        if (j > 0)
        {
            //going down the rows (j) in the column (r)
            predictors.Add(testDB[j][r]);
        }
    }
    int d = await Classify(r+1, attributeValues, predictors.ToArray(),
        jointCounts, dependentCounts, withLaplacian, attributeValues.Length - 1);
    for (int l = 0; l < attributeValues[0].Length; l++)
    {
        if (d == 1)
        {
            int i = ... DataResults[r+1].Count;
```

The following image shows that the algorithm is predicting the poverty index classification of households in the Gertler dataset. These poverty_index categories were introduced in Example 5. This example used Indicator.Q1 to Indicator.Q3 to categorize the poverty_index label into 3 poverty categories (high, medium, and low). The low probabilities reflect the large number of categories in both the household size and land size independent variables. Example 7 and Appendix B discuss the importance of testing and validation to fine tune these algorithms.

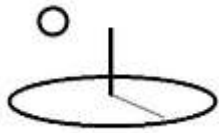
The best way to fill in the Indicator.meta properties using classifying algorithms is not clear yet (i.e. what is a confidence interval in this type of algorithm?). This example manually set the Indicator.QTM, QTL, and QTU, properties.



<p>Indicator 1</p> <p>Naive Bayes</p> <p>Indicator 1 Description</p> <p>Example 8A in the Social Performance Analysis 3 reference.</p> <p>Indicator 1 URL</p> <p>http://localhost:5000/resources/network_carbon/resourcepack_549/resource_1986/gertler1train.csv;http://localhost:5000/resources/network_carbon/resourcepack_549/resource_1985/gertler1test.csv</p> <p>Label 1: ML8A Rel Label 1: none</p> <p>Date 1: 05/02/2018 Dist Type 1: none</p> <p>Q1 1: 51.0000 Q1 Unit 1: high poverty</p> <p>Q2 1: 71.0000 Q2 Unit 1: medium poverty</p> <p>Q3 1: 101.0000 Q3 Unit 1: low poverty</p> <p>Q4 1: 0.0000 Q4 Unit 1: none</p> <p>Q5 1: 0.0000 Q5 Unit 1: none</p> <p>Math Operator 1: equalto BaseIO 1: none</p>		<p>Math Type 1: algorithm1</p> <p>Math Sub Type 1: subalgorithm_01</p> <p>QT D1 1: 0.0000 QT D1 Unit 1: none</p> <p>QT D2 1: 0.0000 QT D2 Unit 1: none</p> <p>QT Most 1: 50.0000 QT Most Unit 1: most poverty index</p> <p>QT Low 1: 25.0000 QT Low Unit 1: lower 80 % ci</p> <p>QT High 1: 68.0000 QT High Unit 1: upper 80 % ci</p> <p>Math Expression 1: I1.Q1.hhsize + I1.Q2.dirtfloor + I1.Q3.land + I1.Q4.fem;</p> <p>Math Result 1:</p> <pre>ml results label,locality_identifier,household_identifier,poverty_index,hhsize,dirtfloor,land,female_hh,label,probability laplacian=true,none,none,qcategory,none,none,none,none,none,none 0,26,5,61,4,1,1,0,medium poverty,0.0158 1,26,5,61,4,1,1,0,medium poverty,0.0158 2,26,11,26,6,1,2,0,high poverty,0.0255 3,26,11,26,6,1,2,0,high poverty,0.0255 4,26,13,61,6,1,4,0,high poverty,0.0342 5,26,13,61,6,1,4,0,high poverty,0.0342 6,26,16,61,7,1,2,0,high poverty,0.0158 7,26,16,61,7,1,2,0,high poverty,0.0158 8,26,21,26,7,1,2,0,high poverty,0.0158 9,26,21,26,7,1,2,0,high poverty,0.0158 10,26,22,26,9,1,1,1,medium poverty,0.0164 11,26,22,26,9,1,1,1,medium poverty,0.0164</pre>
---	--	---

[The Accord Nuget packages were removed from Version 2.1.6 due to their alpha status].

Accord URL:



DevTreks –social budgeting that improves lives and livelihoods

http://accord-framework.net/docs/html/T_Accord_MachineLearning_Bayes_NaiveBayes.htm

http://accord-framework.net/docs/html/T_Accord_Statistics_Filters_Codification.htm

<https://www.devtreks.org/greentreks/preview/carbon/output/ML 8A Accord/2141223489/none>

The following images show that Version 2.1.4 allows multiple statistical libraries to be used to carry out most machine learning algorithms by setting the Indicator.RelatedLabel to the name of the library (or in the source code, Label2). The default uses standard .netstandard libraries. This feature will not be fully functioning until R and Python ML algorithms are fully supported and documented in the CTA02 and CTA03 references in the Technology Assessment tutorial.

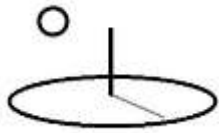
Indicator 1

Naive Bayes

Indicator 1 Description
Example 8A in the Social Performance Analysis 3 reference.

Indicator 1 URL
`http://localhost:5000/resources/network_carbon/resourcepack_549/resource_1986/gertler1train.csv;http://localhost:5000/resources/network_carbon/resourcepack_549/resource_1985/gertler1test.csv`

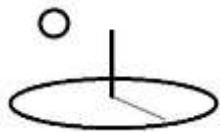
Label 1	Rel Label 1
ML8A	accord
Date 1	Dist Type 1
05/02/2018	none
Q1 1	Q1 Unit 1
51.0000	high poverty
Q2 1	Q2 Unit 1
71.0000	medium poverty
Q3 1	Q3 Unit 1
101.0000	low poverty
Q4 1	Q4 Unit 1



```
if (_subalgorithm == MATHML_SUBTYPES.subalgorithm_01.ToString())
{
    //use sb for feedback when training and testing;
    //but use MathResult for reporting
    StringBuilder sb = new StringBuilder();
    //classify testdata and return new dataset
    if (IndicatorQT.Label2.ToLower() == "accord")
    {
        //use the accord ml library
        sb = await ClassifyAccord(trainData, rowNames, testData);
    }
    else
    {
        //use standard dotnet libraries
        sb = await Classify(trainData, rowNames, testData);
    }
    bHasCalculations = await SetMathResult(rowNames);
}
```

This is the first algorithm built using Accord and that machine learning library is still being critically assessed. The following 2 tables contrast the MathResults using System.Math and Accord.

System.Math MathResults: Only the final row of data has not been classified properly, but inspection of the 4 predictor inputs explains why (i.e. dirtfloor = 0, land = 0, householdsize = 0)



DevTreks –social budgeting that improves lives and livelihoods

label	locality_id	household	poverty_in	hhsz	dirtfloor	land	female_hh	label	probability
laplacian=	none	none	qcategory	none	none	none	none		
0	26	5	61	4	1	1	0	medium po	0.0158
1	26	5	61	4	1	1	0	medium po	0.0158
2	26	11	26	6	1	2	0	high pover	0.0255
3	26	11	26	6	1	2	0	high pover	0.0255
4	26	13	61	6	1	4	0	high pover	0.0342
5	26	13	61	6	1	4	0	high pover	0.0342
6	26	16	61	7	1	2	0	high pover	0.0158
7	26	16	61	7	1	2	0	high pover	0.0158
8	26	21	26	7	1	2	0	high pover	0.0158
9	26	21	26	7	1	2	0	high pover	0.0158
10	26	22	26	9	1	1	1	medium po	0.0164
11	26	22	26	9	1	1	1	medium po	0.0164
19784	88	15684	86	4	0	4	0	medium po	0.2341
19785	88	15685	86	3	0	16	0	low pover	0.5562
19786	88	15685	86	3	0	16	0	low pover	0.5562
19787	88	15686	61	1	1	0	1	medium po	0.0312
19788	88	15686	61	1	1	0	1	medium po	0.0312
19789	88	15691	86	9	0	0	0	medium po	0.113

[The Accord Nugget packages were removed from Version 2.1.6 due to their alpha status].

Accord MathResults: For testing purposes, the following table displays Accord’s integer transformation of the previous table’s data using Accord’s codebook. These results failed to classify the low poverty indexes properly. Accord’s documentation for codebooks explains about the need to preprocess some data and use more advanced processing options. The reason for these transformations suggest a learning curve (and closer attention to the R and Python examples).



DevTreks –social budgeting that improves lives and livelihoods

label	locality_id	household	poverty_in	hhsz	dirtfloor	land	female_hh	label	probability
laplacian=	none	none	qcategory	none	none	none	none	none	none
0	26	5	0	0	0	0	0	medium po	0.7534
1	26	5	0	0	0	0	0	medium po	0.7534
2	26	11	1	0	1	0	0	high pover	0.7534
3	26	11	1	0	1	0	0	high pover	0.7534
4	26	13	1	0	2	0	0	high pover	0.7534
5	26	13	1	0	2	0	0	high pover	0.7534
6	26	16	2	0	1	0	0	high pover	0.7534
7	26	16	2	0	1	0	0	high pover	0.7534
8	26	21	2	0	1	0	0	high pover	0.7534
9	26	21	2	0	1	0	0	high pover	0.7534
10	26	22	3	0	0	1	0	medium po	0.7534
11	26	22	3	0	0	1	0	medium po	0.7534
19784	88	15684	0	1	2	0	0	medium po	0.7534
19785	88	15685	4	1	3	0	0	medium po	0.7534
19786	88	15685	4	1	3	0	0	medium po	0.7534
19787	88	15686	5	0	4	1	0	medium po	0.7534
19788	88	15686	5	0	4	1	0	medium po	0.7534
19789	88	15691	3	1	4	0	0	medium po	0.7534

Example b. Deep Neural Network Predictions (algorithm1, subalgorithm_02)

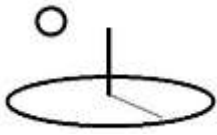
System.Math and MathNet URLs:

<https://www.devtreks.org/greentreks/preview/carbon/input/Machine Learning 8Ab/2147397565/none>

<http://localhost:5000/greentreks/preview/carbon/resourcepack/Resource Conservation Value Accounting Example 8A/549/none>

<http://localhost:5000/greentreks/preview/carbon/output/Machine Learning 8A/2141223503/none>

McCaffrey (MSDN, August and September, 2017) uses System.Math to demonstrate using this algorithm to make predictions for categories of dependent variables. The following image displays McCaffrey’s model. All of McCaffrey’s algorithms are available as open access.



Take a look at the DNN shown in **Figure 1**. The DNN has three hidden layers that have four, two and two nodes. The input node values are 3.80 and 5.00, which you can imagine as the normalized age (38 years old) and income (\$50,000 per year) of a person. Each of the eight hidden nodes has a value between -1.0 and +1.0 because the DNN uses tanh activation.

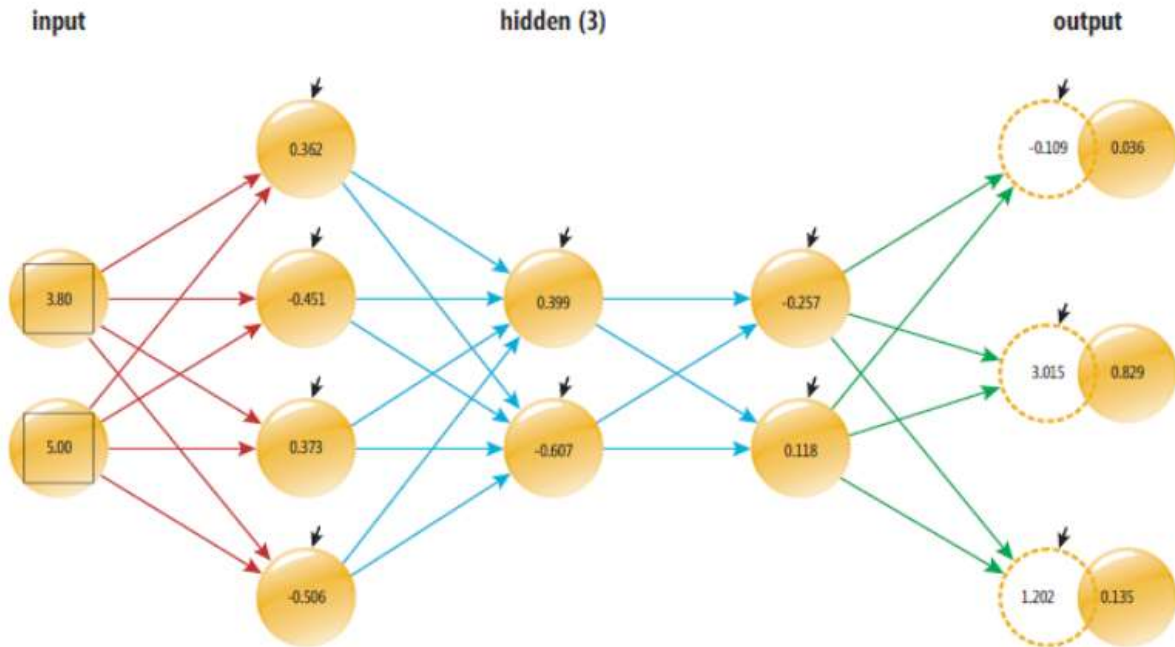
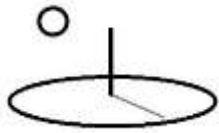


Figure 1 An Example 2-(4,2,2)-3 Deep Neural Network



```
Begin deep net training demo

Generating 2000 artificial training data items

Done. Training data is:
[ 0] 2.24 1.91 2.52 2.41 0.00 1.00 0.00
[ 1] -3.61 2.50 1.29 -1.57 0.00 0.00 1.00
[ 2] 3.18 -1.67 3.22 0.71 1.00 0.00 0.00
...
[1999] 1.30 -2.41 -3.18 0.11 1.00 0.00 0.00

Creating a 4-(10,10,10)-3 deep neural network (tanh &
softmax)

Setting maxEpochs = 2000
Setting learnRate = 0.001
Setting momentum = 0.010

Start training using back-prop with mean squared error

epoch = 200 MS error = 0.0986 accuracy = 0.8015
input-to-hidden [0][0] gradient = -0.002536

epoch = 400 MS error = 0.0779 accuracy = 0.8525
input-to-hidden [0][0] gradient = -0.000551

epoch = 600 MS error = 0.0699 accuracy = 0.8595
input-to-hidden [0][0] gradient = -0.000141

epoch = 800 MS error = 0.0597 accuracy = 0.8815
input-to-hidden [0][0] gradient = -0.159148

epoch = 1000 MS error = 0.0533 accuracy = 0.8980
input-to-hidden [0][0] gradient = -0.000009

epoch = 1200 MS error = 0.0473 accuracy = 0.9105
input-to-hidden [0][0] gradient = -0.000020

epoch = 1400 MS error = 0.0437 accuracy = 0.9160
input-to-hidden [0][0] gradient = 0.000003
```

The following image shows that the algorithm is trying to predict the health expenditure classification of households in the Gertler dataset. The 2nd row in the table specifies the types of data transformation to use in each column.

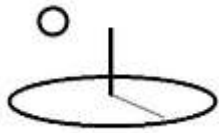


DevTreks –social budgeting that improves lives and livelihoods

label	locality	househ	health_	treatme	poverty	female_hh	hhsiz	dirtfloor
none	none	none	qcategory	none	zscore	none	minmax	none
1	26	5	19.5809	1	55.95054	0	4	1
3	26	11	2.398854	1	46.05873	0	6	1
4	26	13	0	1	54.09583	0	6	1
6	26	16	20.02691	1	56.9034	0	7	1
9	26	21	16.66469	1	46.90881	0	7	1
11	26	22	0.115546	1	48.18393	1	9	1
12	26	23	0	1	49.69088	0	7	0
14	26	25	17.2237	1	57.10974	0	9	0
17	26	27	1.284901	1	49.69088	0	7	1
19	26	28	6.417482	1	52.62751	0	9	1
20	26	29	2.529971	1	53.94127	0	4	0
23	26	31	1.427871	1	50.18547	0	9	0
24	26	37	11.09324	1	52.12519	0	5	1

The following Indicator.MathExpression specifies the subset of column data used in the analysis.

I1.Q1.treatment_locality + I1.Q2.hhsiz + I1.Q3.dirtfloor



Indicator 1

Neural Network 01

Indicator 1 Description

This machine learning example can be found in the Social Performance Analysis 3 reference.

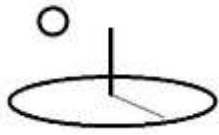
Indicator 1 URL

http://localhost:5000/resources/network_carbon/resourcepack_549/resource_1989/Example8A-b.csv;http://localhost:5000/resources/network_carbon/resourcepack_549/resource_1990/Example8A-btest.csv

Label 1	Rel Label 1
ML02	classify_systemmath
Date 1	Dist Type 1
05/03/2018	none
Q1 1	Q1 Unit 1
13.0000	low cost
Q2 1	Q2 Unit 1
21.0000	medium cost
Q3 1	Q3 Unit 1
150.0000	high cost
Q4 1	Q4 Unit 1
0.0000	none
Q5 1	Q5 Unit 1
0.0000	none
Math Operator 1	BaseIO 1
equalto	none

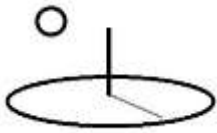
algorithm1	subalgorithm_02
QT D1 1	QT D1 Unit 1
0.0000	none
QT D2 1	QT D2 Unit 1
0.0000	none
QT Most 1	QT Most Unit 1
6.5000	medium cost
QT Low 1	QT Low Unit 1
6.4958	lower 80 % ci
QT High 1	QT High Unit 1
6.5042	upper 80 % ci
Math Expression 1	
I1.Q1.treatment_locality + I1.Q2.hhszize + I1.Q3.dirtflo	
Math Result 1	
ml results label,locality_identifier,household_identifier,health_expenditures,treatment_locality,hhszize,dirtfloor,accuracy,mse,probability,qtm,class,ctl,qtu none,none,none,qcategory,none,minmax,none 1,26,5,17.0000,1.0000,0.2857,1.0000,0.1031,0.7940,76.1228,6.5000,low cost,6.4958,6.5042 3,26,11,6.5000,1.0000,0.5714,1.0000,0.1031,0.7940,94.4234,6.5000,low cost,6.4958,6.5042 4,26,13,6.5000,1.0000,0.5714,1.0000,0.1031,0.7940,94.4234,6.5000,low cost,6.4958,6.5042 6,26,16,17.0000,1.0000,0.7143,1.0000,0.1031,0.7940,94.9390,6.5000,low cost,6.4958,6.5042 4373,51,6163,85.5000,0.0000,0.4286,1.0000,0.1031,0.7940,73.3428,6.5000,low cost,6.4958,6.5042 4376,51,6165,85.5000,0.0000,0.5714,1.0000,0.1031,0.7940,85.4483,6.5000,low cost,6.4958,6.5042	

The Score.Iterations in the following image specifies the subset of training dataset rows used in this analysis (i.e. 2000 out of 5600+). The complete dataset runs very slowly and is a candidate for remote data sciences machines. When the confidence interval is filled in, it will be used to



DevTreks –social budgeting that improves lives and livelihoods

calculate a confidence interval for the Indicator.QTM and Score.QTM. Alternatively, Example c demonstrates that a threshold value can be added to the first column of the ML instructions row. In the latter case, the value is subtracted from Indicator.QTM to derive Indicator.QTL or added to get Indicator.QU.



← → ↻ Secure | https://ww... 🔍 ☆ ⋮

Score	Score Unit
<input type="text" value="6.5000"/>	<input type="text" value="most likely cost"/>
Score D1	Score D1 Unit
<input type="text" value="7.0000"/>	<input type="text" value="mean"/>
Score D2	Score D2 Unit
<input type="text" value="4.0000"/>	<input type="text" value="sd"/>
Score Dist Type	Iterations
<input style="border: 1px solid #ccc; border-radius: 4px; padding: 2px 5px; background-color: #f0f0f0; display: inline-block; text-align: center; width: 100px; height: 20px; vertical-align: middle;" type="button" value="normal"/> ▼	<input type="text" value="2000"/>
Confidence Interval	Random Seed
<input type="text" value="80"/>	<input type="text" value="56"/>
Score BaseIO	
<input style="border: 1px solid #ccc; border-radius: 4px; padding: 2px 5px; background-color: #f0f0f0; display: inline-block; text-align: center; width: 100px; height: 20px; vertical-align: middle;" type="button" value="none"/> ▼	
Score Most Likely	Score Most Unit
<input type="text" value="6.8937"/>	<input type="text" value="health savings"/>
Score Low Estimate	Score Low Unit
<input type="text" value="6.7777"/>	<input type="text" value="lower 80 % ci"/>
Score High Estimate	Score High Unit
<input type="text" value="7.0097"/>	<input type="text" value="upper 80 % ci"/>
Score Math Type	Score Math Sub Type
<input style="border: 1px solid #ccc; border-radius: 4px; padding: 2px 5px; background-color: #f0f0f0; display: inline-block; text-align: center; width: 100px; height: 20px; vertical-align: middle;" type="button" value="algorithm1"/> ▼	<input type="text" value="subalgorithm1"/>
Score Math Result	
<div style="border: 1px solid #ccc; padding: 5px; font-family: monospace; font-size: 0.9em;">sampled descriptive statistics N,Total,Mean,Median,StdDev,Var,Min,Max 2000,10707,1000,0.0007,0.0011,1.0001,10.1001</div>	
Joint Data	

The following 2 images of the Indicator.MathResults compares the difference when the dataset is trained with 2000 rows versus the full 5600+ rows in the second table. Neural networks are



DevTrek –social budgeting that improves lives and livelihoods

important machine learning algorithms and will continue to be explored as more ML algorithms are developed and tested.

label	locality_id	household	health_exp	treatment	hsize	dirtfloor	mse	accuracy	probability	qtm	class	qtl	qtu
none	none	none	qcategory	none	minmax	none	0.1039	0.7994	none	none	none	none	none
1	26	5	17	1	0.2857	1	10.5	FALSE	77.4352	6.5	low cost	6.497	6.503
3	26	11	6.5	1	0.5714	1	0	TRUE	92.993	6.5	low cost	6.497	6.503
4	26	13	6.5	1	0.5714	1	0	TRUE	92.993	6.5	low cost	6.497	6.503
6	26	16	17	1	0.7143	1	10.5	FALSE	93.6949	6.5	low cost	6.497	6.503
4373	51	6163	85.5	0	0.4286	1	79	FALSE	56.4987	6.5	low cost	6.497	6.503
4376	51	6165	85.5	0	0.5714	1	79	FALSE	75.5134	6.5	low cost	6.497	6.503
11192	88	15716	85.5	0	0.1429	1	0	TRUE	38.4459	85.5	high cost	85.497	85.503
11194	35	15721	6.5	0	0.7143	1	0	TRUE	87.1488	6.5	low cost	6.497	6.503
11196	35	15723	85.5	0	0.2857	1	79	FALSE	40.4898	6.5	low cost	6.497	6.503
11197	35	15728	85.5	0	0.1429	1	0	TRUE	38.4459	85.5	high cost	85.497	85.503
11199	35	15729	85.5	0	0.1429	1	0	TRUE	38.4459	85.5	high cost	85.497	85.503
11201	35	15733	6.5	0	0.8571	0	0	TRUE	87.2422	6.5	low cost	6.497	6.503
11204	35	15735	6.5	0	1	1	0	TRUE	93.3427	6.5	low cost	6.497	6.503
11206	35	15736	17	0	0.2857	1	10.5	FALSE	40.4898	6.5	low cost	6.497	6.503
11207	35	15739	85.5	0	0.8571	1	79	FALSE	91.9055	6.5	low cost	6.497	6.503
11210	35	15740	85.5	0	0.4286	1	79	FALSE	56.4987	6.5	low cost	6.497	6.503
11211	35	15742	17	0	0.7143	1	10.5	FALSE	87.1488	6.5	low cost	6.497	6.503
11214	35	15743	6.5	0	0.8571	0	0	TRUE	87.2422	6.5	low cost	6.497	6.503
11216	60	15745	85.5	0	0.1429	1	0	TRUE	38.4459	85.5	high cost	85.497	85.503
11218	60	15746	85.5	0	0.4286	1	79	FALSE	56.4987	6.5	low cost	6.497	6.503
11219	60	15747	17	0	0.7143	1	10.5	FALSE	87.1488	6.5	low cost	6.497	6.503
11221	60	15748	6.5	0	0.8571	1	0	TRUE	91.9055	6.5	low cost	6.497	6.503
11223	60	15750	6.5	0	0.8571	1	0	TRUE	91.9055	6.5	low cost	6.497	6.503
11226	60	15754	6.5	0	0.4286	1	0	TRUE	56.4987	6.5	low cost	6.497	6.503
11228	60	15755	6.5	0	0.4286	1	0	TRUE	56.4987	6.5	low cost	6.497	6.503
11229	60	15757	85.5	0	0.4286	1	79	FALSE	56.4987	6.5	low cost	6.497	6.503
11231	60	15759	17	0	0.2857	1	10.5	FALSE	40.4898	6.5	low cost	6.497	6.503
11234	60	15760	17	0	0	1	68.5	FALSE	41.844	85.5	high cost	85.497	85.503
11235	60	15761	17	0	0.5714	1	10.5	FALSE	75.5134	6.5	low cost	6.497	6.503
11238	60	15762	17	0	0.4286	1	10.5	FALSE	56.4987	6.5	low cost	6.497	6.503
11240	60	15764	6.5	0	0.4286	1	0	TRUE	56.4987	6.5	low cost	6.497	6.503
11242	60	15765	6.5	0	0.7143	1	0	TRUE	87.1488	6.5	low cost	6.497	6.503



DevTreks –social budgeting that improves lives and livelihoods

label	locality_id	household	health_exp	treatment	hhsz	dirtfloor	mse	accuracy	probability	qtm	class	qtl	qtu
none	none	none	qcategory	none	minmax	none	0.1423	0.6767	none	none	none	none	none
1	26	5	17	1	0.2857	1	10.5	FALSE	80.2768	6.5	low cost	6.4976	6.5024
3	26	11	6.5	1	0.5714	1	0	TRUE	93.0707	6.5	low cost	6.4976	6.5024
4	26	13	6.5	1	0.5714	1	0	TRUE	93.0707	6.5	low cost	6.4976	6.5024
6	26	16	17	1	0.7143	1	10.5	FALSE	95.6589	6.5	low cost	6.4976	6.5024
4373	51	6163	85.5	0	0.4286	1	68.5	FALSE	54.0513	17	medium cc	16.9976	17.0024
4376	51	6165	85.5	0	0.5714	1	79	FALSE	48.3063	6.5	low cost	6.4976	6.5024
11192	88	15716	85.5	0	0.1429	1	0	TRUE	49.7445	85.5	high cost	85.4976	85.5024
11194	35	15721	6.5	0	0.7143	1	0	TRUE	48.3817	6.5	low cost	6.4976	6.5024
11196	35	15723	85.5	0	0.2857	1	68.5	FALSE	54.8363	17	medium cc	16.9976	17.0024
11197	35	15728	85.5	0	0.1429	1	0	TRUE	49.7445	85.5	high cost	85.4976	85.5024
11199	35	15729	85.5	0	0.1429	1	0	TRUE	49.7445	85.5	high cost	85.4976	85.5024
11201	35	15733	6.5	0	0.8571	0	0	TRUE	52.0865	6.5	low cost	6.4976	6.5024
11204	35	15735	6.5	0	1	1	0	TRUE	75.4539	6.5	low cost	6.4976	6.5024
11206	35	15736	17	0	0.2857	1	0	TRUE	54.8363	17	medium cc	16.9976	17.0024
11207	35	15739	85.5	0	0.8571	1	79	FALSE	51.5612	6.5	low cost	6.4976	6.5024
11210	35	15740	85.5	0	0.4286	1	68.5	FALSE	54.0513	17	medium cc	16.9976	17.0024
11211	35	15742	17	0	0.7143	1	10.5	FALSE	48.3817	6.5	low cost	6.4976	6.5024
11214	35	15743	6.5	0	0.8571	0	0	TRUE	52.0865	6.5	low cost	6.4976	6.5024
11216	60	15745	85.5	0	0.1429	1	0	TRUE	49.7445	85.5	high cost	85.4976	85.5024
11218	60	15746	85.5	0	0.4286	1	68.5	FALSE	54.0513	17	medium cc	16.9976	17.0024
11219	60	15747	17	0	0.7143	1	10.5	FALSE	48.3817	6.5	low cost	6.4976	6.5024
11221	60	15748	6.5	0	0.8571	1	0	TRUE	51.5612	6.5	low cost	6.4976	6.5024
11223	60	15750	6.5	0	0.8571	1	0	TRUE	51.5612	6.5	low cost	6.4976	6.5024
11226	60	15754	6.5	0	0.4286	1	10.5	FALSE	54.0513	17	medium cc	16.9976	17.0024
11228	60	15755	6.5	0	0.4286	1	10.5	FALSE	54.0513	17	medium cc	16.9976	17.0024
11229	60	15757	85.5	0	0.4286	1	68.5	FALSE	54.0513	17	medium cc	16.9976	17.0024
11231	60	15759	17	0	0.2857	1	0	TRUE	54.8363	17	medium cc	16.9976	17.0024
11234	60	15760	17	0	0	1	68.5	FALSE	72.1718	85.5	high cost	85.4976	85.5024
11235	60	15761	17	0	0.5714	1	10.5	FALSE	48.3063	6.5	low cost	6.4976	6.5024
11238	60	15762	17	0	0.4286	1	0	TRUE	54.0513	17	medium cc	16.9976	17.0024
11240	60	15764	6.5	0	0.4286	1	10.5	FALSE	54.0513	17	medium cc	16.9976	17.0024
11242	60	15765	6.5	0	0.7143	1	0	TRUE	48.3817	6.5	low cost	6.4976	6.5024

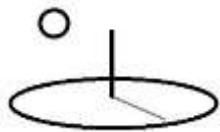
Cognitive Toolkit URL:

The following image displays the code for this algorithm. Note that the current version of this toolkit is only compatible with netframework, not the netstandard framework used by DevTreks. This example assumes the toolkit will eventually migrate over, but until it does, the libraries are not included with DevTreks. For now, this example is included as evidence that .Net libraries support machine learning.

Under planning for a future release.

Example c. Time Series Neural Network (algorithm1, subalgorithm_03)

System.Math and MathNet URL:



DevTreks –social budgeting that improves lives and livelihoods

<https://www.devtreks.org/greentreks/preview/carbon/input/Machine Learning 8Ac/2147397566/none>

<http://localhost:5000/greentreks/preview/carbon/input/Machine Learning 8Ac/2147409851/none>

McCaffrey (MSDN, October, 2017, April, 2018) uses System.Math to demonstrate using this algorithm to make time series predictions. Example b’s dataset was altered to time series data by adding 4 additional columns of health expenditures collected for 4 quarters (+-5% and +-10% of actual health cost). The threshold used to determine the accuracy of the dataset has been set to \$2.50 in the first column of the ML instructions row. Predictions that are within \$2.50 return TRUE in the accuracy column. The following Indicator.MathResults occur when the Score.Iterations is set to 2000, thereby truncating the data used to train the network (from the full 5600+ row datasets). Given that the dataset has been sorted from low to high health expenditures, it’s not surprising that the algorithm does a reasonable job of making time series predictions for lower health care costs (i.e. compare the health_expenditures column with the qtm column).

label	locality_id	household	health_exp	q1	q2	q3	q4	mse	accuracy	qtm	qtl	qtu
2.5	none	none	none	none	none	none	none	0.0001	1	none	none	none
4	26	13	0	0	0	0	0	0.0008	TRUE	-0.0008	-2.5008	2.4992
3	26	11	2.3989	2.5188	2.2789	2.6387	2.159	0.1177	TRUE	2.5166	0.0166	5.0166
11214	35	15743	9.9167	10.4125	9.4209	10.9084	8.925	0.3689	TRUE	10.2856	7.7856	12.7856
11226	60	15754	10.8225	11.3637	10.2814	11.9048	9.7403	0.274	TRUE	11.0965	8.5965	13.5965
11204	35	15735	11.3129	11.8786	10.7473	12.4442	10.1817	0.1916	TRUE	11.5045	9.0045	14.0045
11201	35	15733	11.3189	11.8849	10.753	12.4508	10.187	0.1904	TRUE	11.5093	9.0093	14.0093
11242	60	15765	11.5798	12.1588	11.0008	12.7378	10.4218	0.1367	TRUE	11.7165	9.2165	14.2165
11228	60	15755	11.6964	12.2812	11.1116	12.8661	10.5268	0.1105	TRUE	11.8069	9.3069	14.3069
11194	35	15721	11.9141	12.5098	11.3184	13.1055	10.7227	0.0578	TRUE	11.9719	9.4719	14.4719
11221	60	15748	12.0777	12.6816	11.4738	13.2855	10.87	0.015	TRUE	12.0927	9.5927	14.5927
11223	60	15750	12.7036	13.3388	12.0685	13.974	11.4333	0.1746	TRUE	12.529	10.029	15.029
11240	60	15764	12.9251	13.5713	12.2788	14.2176	11.6325	0.2516	TRUE	12.6735	10.1735	15.1735
11231	60	15759	14.4735	15.1972	13.7498	15.9209	13.0262	0.9277	TRUE	13.5458	11.0458	16.0458
11219	60	15747	16.1345	16.9413	15.3278	17.748	14.5211	1.8931	TRUE	14.2414	11.7414	16.7414
11206	35	15736	17.1026	17.9577	16.2475	18.8128	15.3923	2.5498	FALSE	14.5528	12.0528	17.0528
11238	60	15762	17.2092	18.0696	16.3487	18.9301	15.4883	2.6258	FALSE	14.5834	12.0834	17.0834
11234	60	15760	18.0169	18.9177	17.116	19.8185	16.2152	3.2217	FALSE	14.7952	12.2952	17.2952
1	26	5	19.5809	20.5599	18.6019	21.539	17.6228	4.4644	FALSE	15.1165	12.6165	17.6165

The following results take place when the full dataset, consisting of 5600+ rows, is used to train the network. Now the algorithm does a reasonable job only with high initial health care costs. Production machine learning algorithm development requires additional understanding and work.



DevTreks –social budgeting that improves lives and livelihoods

label	locality_id	household	health_exp	q1	q2	q3	q4	mse	accuracy	qtm	qtl	qtu
2.5	none	none	none	none	none	none	none	41.1285	0.2983	none	none	none
4	26	13	0	0	0	0	0	5.8209	FALSE	5.8209	3.3209	8.3209
3	26	11	2.3989	2.5188	2.2789	2.6387	2.159	3.4219	FALSE	5.8208	3.3208	8.3208
11214	35	15743	9.9167	10.4125	9.4209	10.9084	8.925	4.0959	FALSE	5.8208	3.3208	8.3208
11226	60	15754	10.8225	11.3637	10.2814	11.9048	9.7403	5.0017	FALSE	5.8208	3.3208	8.3208
11204	35	15735	11.3129	11.8786	10.7473	12.4442	10.1817	5.1829	FALSE	6.13	3.63	8.63
11201	35	15733	11.3189	11.8849	10.753	12.4508	10.187	5.086	FALSE	6.2329	3.7329	8.7329
11242	60	15765	11.5798	12.1588	11.0008	12.7378	10.4218	8.6492	FALSE	20.229	17.729	22.729
11228	60	15755	11.6964	12.2812	11.1116	12.8661	10.5268	8.5339	FALSE	20.2303	17.7303	22.7303
11194	35	15721	11.9141	12.5098	11.3184	13.1055	10.7227	8.3162	FALSE	20.2303	17.7303	22.7303
11221	60	15748	12.0777	12.6816	11.4738	13.2855	10.87	8.1526	FALSE	20.2303	17.7303	22.7303
11223	60	15750	12.7036	13.3388	12.0685	13.974	11.4333	7.5267	FALSE	20.2303	17.7303	22.7303
11240	60	15764	12.9251	13.5713	12.2788	14.2176	11.6325	7.3052	FALSE	20.2303	17.7303	22.7303
11231	60	15759	14.4735	15.1972	13.7498	15.9209	13.0262	5.7568	FALSE	20.2303	17.7303	22.7303
11219	60	15747	16.1345	16.9413	15.3278	17.748	14.5211	4.0958	FALSE	20.2303	17.7303	22.7303
11206	35	15736	17.1026	17.9577	16.2475	18.8128	15.3923	3.1277	FALSE	20.2303	17.7303	22.7303
11238	60	15762	17.2092	18.0696	16.3487	18.9301	15.4883	3.0211	FALSE	20.2303	17.7303	22.7303
11234	60	15760	18.0169	18.9177	17.116	19.8185	16.2152	2.2134	TRUE	20.2303	17.7303	22.7303
1	26	5	19.5809	20.5599	18.6019	21.539	17.6228	0.6494	TRUE	20.2303	17.7303	22.7303

Cognitive Toolkit URL:

The following image displays the code for this algorithm.

Under planning for a future release.

B. Conclusions

Algorithm1 uses open source statistical libraries that have been compiled and run within DevTreks to develop custom algorithms. This approach accesses the libraries directly through their exposed APIs. Unlike Examples 8B and 8C, they require no separate statistical packages with separate deployment, related batch files, and DLL syntax. Their primary advantages include fine-tuned control over algorithm development, increased performance, flexible deployment, and no reliance on “black box” algorithms. Their primary drawbacks involve the maturity and institutional backing of the platforms when compared to alternative statistical libraries (with allowances for System.Math and CNTK).

Footnotes

1. System.Math is the internal mathematical library found in the open source .NET Standard 2.0. MathNet and Accord are open source statistical libraries that use MIT licenses which are similar to DevTreks license (see the References section). This release runs the



DevTreks –social budgeting that improves lives and livelihoods

libraries from their latest Nuget packages (i.e. compiled and distributed with the DevTreks source code). Prior to this release, the MathNet source had been included as a separate project, but proved hard to keep updated. DevTreks may also employ a Nuget package approach for the data layer in an upcoming release. Version 2.1.6 removed the Accord Nuget package because the source is still in alpha testing status.

2. MathNet and Accord, like DevTreks, appear to originate with one person. These technologists may have been displeased with the conventional approaches used by most of their peers and decided to take an independent, automated, digital, approach in their field of specialty. Caution must be exercised because these efforts seldom have the institutional backing of large teams with sufficient budgets. They have their role, but the next generation needs to critically assess whether they can not only “do it right”, but also “do it better”.

References

McCaffrey, James. MSDN Magazine. Various issues referenced in the Example. Microsoft Corporation. Source code is available as open access by going to the referenced issues.

Version 2.1.6 removed the Accord Nuget package because the source is still in alpha testing status

Accord. Last accessed April 24, 2018 (currently being reviewed):

<http://accord-framework.net/>

<https://github.com/accord-net/framework>

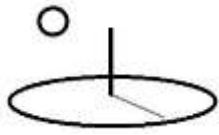
MathNet. Last accessed April 24, 2018:

<https://numerics.mathdotnet.com/>

<https://github.com/mathnet/mathnet-numeric>

System.Math. Last accessed April 24, 2018:

[https://msdn.microsoft.com/en-us/library/system.math\(v=vs.110\).aspx](https://msdn.microsoft.com/en-us/library/system.math(v=vs.110).aspx)



DevTreks –social budgeting that improves lives and livelihoods

<https://docs.microsoft.com/en-us/dotnet/api/?view=netstandard-2.0&term=math>

<https://github.com/dotnet/standard>

[Cognitive Toolkit. Last accessed April 24, 2018:

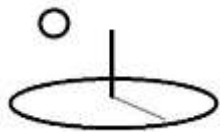
Not yet compatible with netstandard 2.0 and not distributed with DevTreks:

<https://docs.microsoft.com/en-us/cognitive-toolkit/>

<https://github.com/Microsoft/CNTK>]

review

<https://blogs.msdn.microsoft.com/dotnet/2018/05/07/introducing-ml-net-cross-platform-proven-and-open-source-machine-learning-framework/>



DevTreks –social budgeting that improves lives and livelihoods

Example 8B. SDG Stakeholder Resource Conservation Value Accounting using R

Algorithms: algorithm2, various subalgorithms;

URLs: under planning for 216

A. ML R Examples (algorithm2, subalgorithm_xx)

This example uses the R statistical library to demonstrate the following examples of basic ML algorithms. The R Library, as explained in the CTA02 reference, must be deployed as a stand-alone software library that is accessed using batch TEXT files.

Version 2.1.4 began distinguishing machine learning (ML) subalgorithms from standard statistical subalgorithms for algorithm2 using the following conventions:

- a. Subalgorithm Names:** Names must use a “_xx” suffix, such as “subalgorithm_01”.
- b. Indicator.URL and Score.URL Datasets:** ML datasets must use 2 separate TEXT files –the first is for an R statistical script TEXT file, the second for a ML TEXT dataset. The URLs to these datasets must be stored using the Indicator.URL property for Indicators and Score.URL for Scores.
- c. Training, Test, and Subset Datasets:** The R statistical script TEXT file must hard code how the Data TEXT file is broken into training, testing, and subset datasets.
- d. Data Transformation.** The R statistical script TEXT file must hard code any dynamic data transformation needed in the dataset.
- e. Manual and Automatic Indicator Meta Completion:** uses the same rules as Example 8A employed using R statistical script.

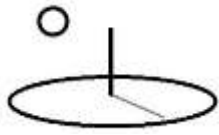
Example a. Naïve Bayes (algorithm2, subalgorithm_01)

The following image displays the code for this algorithm.

Under planning for a future release.

Example b. Logistic Regression (algorithm2, subalgorithm_02)

The following image displays the code for this algorithm.



DevTreks –social budgeting that improves lives and livelihoods

Under planning for a future release.

Example c. Time Series Neural Network (algorithm2, subalgorithm_03)

The following image displays the code for this algorithm.

Under planning for a future release.

B. Conclusions

R (algorithm2) fully supports standard and ML algorithm development.



DevTreks –social budgeting that improves lives and livelihoods

Example 8C. SDG Stakeholder Resource Conservation Value Accounting using Python (1*)

Algorithms: algorithm3, various subalgorithms;

URLs: : under planning for 216

A. ML Python Examples (algorithm3, subalgorithm_xx)

This example uses the Python statistical library to demonstrate the following examples of basic ML algorithms. The Python Library, as explained in the CTA03 reference, must be deployed as a stand-alone software library that is accessed using batch TEXT files.

Version 2.1.4 began distinguishing machine learning (ML) subalgorithms from standard statistical subalgorithms for algorithm3 using the following conventions:

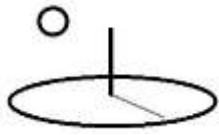
- a. Subalgorithm Names:** Names must use a “_xx” suffix, such as “subalgorithm_01”.
- b. Indicator.URL and Score.URL Datasets:** ML datasets must use 2 separate TEXT files –the first is for a Python statistical script TEXT file, the second for a ML TEXT dataset. The URLs to these datasets must be stored using the Indicator.URL property for Indicators and Score.URL for Scores.
- c. Training, Test, and Subset Datasets:** The Python statistical script TEXT file must hard code how the Data TEXT file is broken into training, testing, and subset datasets.
- d. Data Transformation.** The Python statistical script TEXT file must hard code any dynamic data transformation needed in the dataset.
- e. Manual and Automatic Indicator Meta Completion:** uses the same rules as Example 8A employed using the Python statistical script.

Example a. Naïve Bayes (algorithm3, subalgorithm_01)

The following image displays the code for this algorithm.

Under planning for a future release.

Example b. Logistic Regression (algorithm3, subalgorithm_02)



DevTreks –social budgeting that improves lives and livelihoods

The following image displays the code for this algorithm.

Under planning for a future release.

Example c. Time Series Neural Network (algorithm3, subalgorithm_03)

The following image displays the code for this algorithm.

B. Under planning for a future release.

C. Conclusions

Python (algorithm3) fully supports standard and ML algorithm development.



DevTreks –social budgeting that improves lives and livelihoods

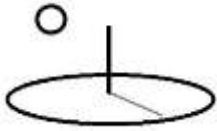
Appendix A. Standard Statistical Impact Evaluation Analysis

A. Introduction to Standard Impact Evaluation

The following examples demonstrate carrying out Impact Evaluation using some of the standard statistical analysis techniques introduced in the Gertler et al (2016) reference. Sections B and C modify the following statistical scripts found in Gertler et al’s Technical Companion (Version 1, September, 2016) to related statistical software scripts.

Example a. Randomized Assignment in a Regression Framework

The following Stata script concludes that health expenditures decreased by \$10.14 from the intervention (with an average expenditure without the treatment of \$17.98).



To implement this in Stata (see Stata Example 1), we regress our outcome of interest (health expenditures) on a binary treatment variable (`treatment_locality`), which is equal to 1 if the household is located in a treatment area. We run this regression on the sample of households that were eligible for HISP and based on data collected after the program has been administered (`round = 1`). The estimated regression coefficient for δ is -10.14, indicating that eligible households exposed to HISP spent \$10.14 less on health expenditures than eligible households in the comparison group. The standard error of the coefficient is 0.396.⁸ The t-statistic calculated in the regression, -25.63, shows that this coefficient is statistically significant at the 1 percent level.

Stata Example 1. Randomized Assignment in a Regression Framework (Linear Regression)

* In this context, the program is randomized at the village level, and you compare follow-up situation of eligible households in treatment and comparison villages.

```
*Select the relevant data
use "evaluation.dta", clear
keep if eligible==1
```

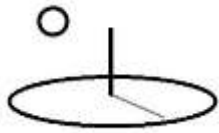
```
reg health_expenditures treatment_locality if round ==1, cl(locality_identifier)
```

```
Linear regression              Number of obs =    5629
                              F( 1, 196) = 656.77
                              Prob > F    = 0.0000
                              R-squared    = 0.3004
                              Root MSE  = 7.7283
```

(Std. Err. adjusted for 197 clusters in `locality_identifier`)

	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
<code>health_expenditu-s</code>						
<code>treatment_locality</code>	-10.14037	.3956824	-25.63	0.000	-10.92071	-9.36003
<code>_cons</code>	17.98055	.3066373	58.64	0.000	17.37582	18.5852

The following images of multivariate regression analysis explains why, with this randomized dataset, multivariate regression doesn't contribute much to the statistical conclusions.



Stata Example 4. Randomized Assignment in a Regression Framework (Multivariate Regression)

```
reg health_expenditures treatment_locality age_hh age_sp educ_hh educ_sp female_hh
indigenous_hhsize dirtfloor bathroom land hospital_distance if round ==1,
cl(locality_identifier)
```

Linear regression

Number of obs = 5629
F(12, 196) = 135.95
Prob > F = 0.0000
R-squared = 0.4297
Root MSE = 6.9844

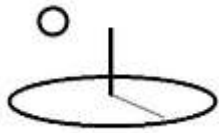
(Std. Err. adjusted for 197 clusters in locality_identifier)

health_expenditu-s	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
treatment_locality	-10.01032	.3412294	-29.34	0.000	-10.68327	-9.337363
age_hh	.0411975	.0146714	2.81	0.005	.0122635	.0701316
age_sp	.0031789	.0171833	0.18	0.853	-.0307089	.0370667
educ_hh	-.0389384	.0468306	-0.83	0.407	-.1312951	.0534182
educ_sp	-.0223849	.049144	-0.46	0.649	-.1193038	.0745339
female_hh	.6430682	.4442419	1.45	0.149	-.2330395	1.519176
indigenous	-1.905311	.3496098	-5.45	0.000	-2.594791	-1.215831
hhsize	-1.603432	.0655058	-24.48	0.000	-1.732619	-1.474245
dirtfloor	-1.849394	.2776092	-6.66	0.000	-2.396878	-1.301909
bathroom	.2850031	.2463569	1.16	0.249	-.2008474	.7708537
land	.0380483	.0376021	1.01	0.313	-.0361083	.112205
hospital_distance	-.0026034	.0042288	-0.62	0.539	-.0109431	.0057363
_cons	27.56544	.8635174	31.92	0.000	25.86246	29.26841

In the case of HISP, the coefficient estimated by multivariate regression in Stata Example 4 (-10.01) is very similar to the coefficient estimated by bivariate regression in Stata Example 1 (-10.14). This is because the HISP program was randomized in the first place, so that the additional covariates do not affect the consistency of the estimator.

Example b. Randomized Treatment Regression Analysis

The following Stata script concludes that health expenditures decreased by \$10.72 from the intervention (with an average expenditure without the treatment of \$20.06).



Stata Example 6: "Local Average Treatment Effect" (LATE) 2SLS IV Estimates¹³

* You can back out 'local average treatment effect' estimates on complier units that do take-up the program in treatment communities

```
ivreg health_expenditures (enrolled = treatment_locality) if round ==1, first
```

First-stage regressions

Source	SS	df	MS			
Model	885.675858	1	885.675858	Number of obs =	9914	
Residual	1192.5756	9912	.120316344	F(1, 9912) =	7361.23	
				Prob > F =	0.0000	
				R-squared =	0.4262	
				Adj R-squared =	0.4261	
				Root MSE =	.34687	
Total	2078.25146	9913	.209649093			

enrolled_ro	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
treatment_locality	.5977823	.0069674	85.80	0.000	.5841248	.6114397
_cons	2.12e-14	.0049282	0.00	1.000	-.0096602	.0096602

Instrumental variables (2SLS) regression

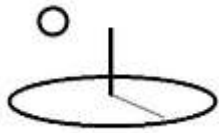
Source	SS	df	MS			
Model	334474.116	1	334474.116	Number of obs =	9914	
Residual	1067039.56	9912	107.651288	F(1, 9912) =	944.81	
				Prob > F =	0.0000	
				R-squared =	0.2387	
				Adj R-squared =	0.2386	
				Root MSE =	10.376	
Total	1401513.68	9913	141.381386			

health_exp~s	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
Enrolled	-10.71629	.348636	-30.74	0.000	-11.39969	-10.03289
_cons	20.06416	.1474116	136.11	0.000	19.7752	20.35312

Instrumented: enrolled
 Instruments: treatment_locality

Example c. Difference in Differences Regression Analysis

The following Stata script concludes that health expenditures decreased by \$8.16 from the intervention (with an average expenditure without the treatment of \$20.79).



Stata Example 8 illustrates how to estimate equation (3) using Stata for the HISP case. Let’s assume that we have data only from localities where the program has been offered. In these localities, we have data both for households that participate in the program, as well as households that do not to participate. Data are available for a baseline survey collected before the program, and a follow-up survey collected after the program. To obtain the difference-in-differences estimates, we first create a new variable (enrolled_round), which is the interaction between participation in the program and the time at which the data are measured. The outcome variable is then regressed on this new variable, along with dummy variables capturing whether or not the household participated in the program, and the time at which each data point is observed. The coefficient of the new interaction variable is our difference-in-differences impact estimate. The results below indicate that health expenditures for households that enrolled in the program were \$8.16 lower than among households that did not enroll.

Stata Example 8. Difference-in-Differences in a Regression Framework

* In this method, you compare the change in health expenditures over time
* between enrolled and nonenrolled households in the treatment localities.

```
*Select the relevant data
use "evaluation.dta", clear
keep if treatment_locality==1
```

```
gen eligible_round=eligible*round
```

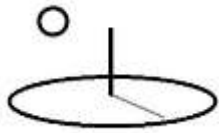
```
reg health_expenditures eligible_round round eligible, cl(locality_identifier)
```

```
Linear regression                               Number of obs =   9919
                                                F(   3,   99) =  813.98
                                                Prob > F      =  0.0000
                                                R-squared     =  0.3436
                                                Root MSE     =  7.9128

                                (Std. Err. adjusted for 100 clusters in locality_identifier)
```

health_expen~s	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
eligible_round	-8.162931	.3191368	-25.58	0.000	-8.796168	-7.529695
round	1.513416	.3564533	4.25	0.000	.8061355	2.220697
eligible	-6.3018	.193082	-32.64	0.000	-6.684917	-5.918684
_cons	20.79149	.1722887	120.68	0.000	20.44964	21.13335

The following Stata script concludes that health expenditures decreased by \$8.16 from the intervention (with an average expenditure without the treatment of \$17.02).



Stata Example 10. Calculating Difference-in-Difference Estimates by Taking the Difference between Before-After Differences in the Treatment and Comparison Groups

```

xtset household_identifier round
      panel variable:  household_identifier (unbalanced)
      time variable:   round, 0 to 1
                    delta: 1 unit

gen xtenrolled=0

replace xtenrolled=1 if enrolled==1 & round==1

xtreg health_expenditures xtenrolled round if treatment_locality==1, fe vce(cluster
locality_identifier)

Fixed-effects (within) regression              Number of obs   =       9919
Group variable: household_~r                  Number of groups =       4960

R-sq:  within = 0.2437                        Obs per group:  min =         1
      between = 0.3852                          avg =         2.0
      overall  = 0.2805                          max =         2

corr(u_i, Xb) = 0.2632                          F(2,99)         =       633.70
                                                Prob > F        =       0.0000

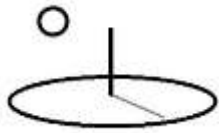
```

(Std. Err. adjusted for 100 clusters in locality_identifier)

health_exp~s	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
xtenrolled	-8.163337	.3190874	-25.58	0.000	-8.796476	-7.530198
round	1.513416	.3564353	4.25	0.000	.8061711	2.220661
_cons	17.02477	.1203165	141.50	0.000	16.78603	17.2635
sigma_u	7.1354916					
sigma_e	6.5169842					
rho	.54521089	(fraction of variance due to u_i)				

Example d. Discontinuity Regression Analysis

The following Stata script concludes that health expenditures decreased by \$11.19 from the intervention (with an average expenditure without the treatment of \$20.55).



Stata Example 12. Regression Discontinuity Design Estimates

```

* REGRESSION DISCONTINUITY DESIGN
* In this context, you compare health expenditures at follow-up between households
just above
* and just below the poverty index threshold, in the treatment localities.

*Select the relevant data
use "evaluation.dta", clear
keep if treatment_locality==1

*Normalize the poverty index
gen poverty_index_left=poverty_index-58 if poverty_index<=58
(8570 missing values generated)

replace poverty_index_left=0 if poverty_index>58
(8570 real changes made)

gen poverty_index_right=poverty_index-58 if poverty_index>58
(11257 missing values generated)

replace poverty_index_right=0 if poverty_index<=58
(11257 real changes made)

reg health_expenditures poverty_index_left poverty_index_right eligible if round ==1

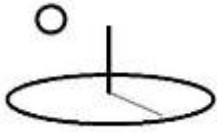
```

Source	SS	df	MS			
Model	257911.257	3	85970.4191	Number of obs =	4960	
Residual	505111.412	4956	101.919171	F(3, 4956) =	843.52	
Total	763022.67	4959	153.866237	Prob > F =	0.0000	
				R-squared =	0.3380	
				Adj R-squared =	0.3376	
				Root MSE =	10.096	

health_expenditures	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
poverty_index_left	.1755687	.0302796	5.80	0.000	.1162073	.23493
poverty_index_right	.2202764	.0315631	6.98	0.000	.1583987	.2821541
eligible	-11.19171	.4661828	-24.01	0.000	-12.10564	-10.27779
_cons	20.55449	.3376327	60.88	0.000	19.89258	21.2164

Example e. Propensity Score Matching Regression Analysis

The following Stata script concludes that savings in health care expenditures can be explained best by the dirt floor and household size variables.



Stata Example 13. Propensity Score Matching Estimates

```
* MATCHING
* In this context, you compare health expenditures at follow-up between enrolled
* households and a set of matched nonenrolled households from both treatment and
* comparison villages.

*Select the relevant data
```

29

```
use "evaluation.dta", clear

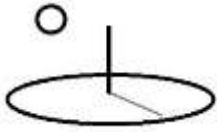
* reshape the database
reshape wide health_expenditures age_hh age_sp educ_hh educ_sp hospital,
i(household_identifier) j(round)

probit enrolled age_hh age_sp educ_hh educ_sp female_hh indigenous hhsize dirtfloor
bathroom land hospital_distance
```

```
Iteration 0: log likelihood = -6047.086
Iteration 1: log likelihood = -5510.3753
Iteration 2: log likelihood = -5506.5201
Iteration 3: log likelihood = -5506.5196
```

```
Probit regression                               Number of obs   =    9,913
                                                LR chi2(11)    =   1081.13
                                                Prob > chi2    =    0.0000
Log likelihood = -5506.5196                    Pseudo R2      =    0.0894
```

enrolled	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
age_hh	-.0131411	.0017648	-7.45	0.000	-.0166001	-.0096822
age_sp	-.0078841	.0020507	-3.84	0.000	-.0119035	-.0038648

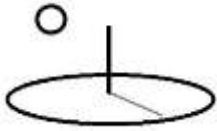


DevTreks –social budgeting that improves lives and livelihoods

enrolled	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
age_hh	-.0131411	.0017648	-7.45	0.000	-.0166001 -.0096822
age_sp	-.0078841	.0020507	-3.84	0.000	-.0119035 -.0038648
educ_hh	-.0215019	.0062476	-3.44	0.001	-.033747 -.0092568
educ_sp	-.0155054	.0067557	-2.30	0.022	-.0287462 -.0022645
female_hh	-.0204807	.0518766	-0.39	0.693	-.1221569 .0811955
indigenous	.1613552	.031199	5.17	0.000	.1002062 .2225041
hhsizes	.1188953	.0067088	17.72	0.000	.1057462 .1320443
dirtfloor	.3758706	.0308276	12.19	0.000	.3154496 .4362916
bathroom	-.1245256	.0289856	-4.30	0.000	-.1813364 -.0677149
land	-.0277659	.0049886	-5.57	0.000	-.0375435 -.0179884
hospital_distance	.0015885	.0003514	4.52	0.000	.0008998 .0022772
_cons	-.4974732	.0904964	-5.50	0.000	-.6748429 -.3201035

```
predict pscore
kdensity pscore if enrolled ==1, gen(take1 den1)
kdensity pscore if enrolled ==0, gen(take0 den0)
graph twoway (line den0 take0, lpattern(solid)) (line den1 take1, lpattern(dash))
```

The following Stata script concludes that health expenditures decreased by \$9.97 from the intervention.



```
set seed 100
generate u=runiform()
sort u
```

```
psmatch2 enrolled age_hh age_sp educ_hh educ_sp female_hh indigenous hhsiz dirtfloor:
bathroom land hospital_distance, out(health_expenditures1)
```

```
Probit regression                               Number of obs   =    9,913
                                                LR chi2(11)    =   1081.13
                                                Prob > chi2    =    0.0000
Log likelihood = -5506.5196                    Pseudo R2      =    0.0894
```

enrolled	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
age_hh	-.0131411	.0017648	-7.45	0.000	-.0166001	-.0096822
age_sp	-.0078841	.0020507	-3.84	0.000	-.0119035	-.0038648
educ_hh	-.0215019	.0062476	-3.44	0.001	-.033747	-.0092568
educ_sp	-.0155054	.0067557	-2.30	0.022	-.0287462	-.0022645
female_hh	-.0204807	.0518766	-0.39	0.693	-.1221569	.0811955
indigenous	.1613552	.031199	5.17	0.000	.1002062	.2225041
hhsiz	.1188953	.0067088	17.72	0.000	.1057462	.1320443
dirtfloor	.3758706	.0308276	12.19	0.000	.3154496	.4362916
bathroom	-.1245256	.0289856	-4.30	0.000	-.1813364	-.0677149
land	-.0277659	.0049886	-5.57	0.000	-.0375435	-.0179884
hospital_distance	.0015885	.0003514	4.52	0.000	.0008998	.0022772
_cons	-.4974732	.0904964	-5.50	0.000	-.6748429	-.3201035

There are observations with identical propensity score values.
The sort order of the data could affect your results.
Make sure that the sort order is random before calling psmatch2.

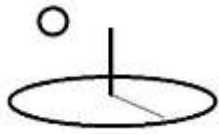
Variable	Sample	Treated	Controls	Difference	S.E.	T-stat
health_expendi-1	Unmatched	7.83977335	20.70746	-12.8676866	.226604141	-56.78
	ATT	7.83977335	17.8088716	-9.96909828	.263484213	-37.84

Note: S.E. does not take into account that the propensity score is estimated.

```
| psmatch2:
psmatch2: | Common
Treatment | support
```

assignment	On suppor	Total
Untreated	6,949	6,949
Treated	2,964	2,964

B. Introduction to Standard Impact Evaluation using .NetStandardLibraries (algorithm1)



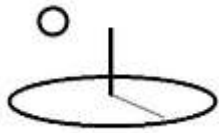
DevTreks –social budgeting that improves lives and livelihoods

Stock Input Calculator

[http://localhost:5000/greentreks/preview/carbon/input/Impact Evaluation 7, 1 and 2/2147409847/none](http://localhost:5000/greentreks/preview/carbon/input/Impact%20Evaluation%207,%201%20and%202/2147409847/none)

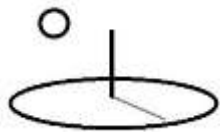
Example a. Randomized Assignment in a Regression Framework (algorithm1, subalgorithm6)

The following script and image display the equivalent results using MathNet.



Math Type 1		Math Sub Type 1		
algorithm1		subalgorithm6		
QT D1 1	QT D1 Unit 1			
0.0000				
QT D2 1	QT D2 Unit 1			
0.0000				
QT Most 1	QT Most Unit 1			
17.9851	actual most hh savings			
QT Low 1	QT Low Unit 1			
17.7931	lower 80 % ci			
QT High 1	QT High Unit 1			
18.1772	upper 80 % ci			
Math Expression 1				
l1.Q1.treatment_locality				
Math Result 1				
regression results				
dependent variable: health_expenditures				
source	df	SS	MS	
model	1	144334.8001	144334.8001	
error	5624	336016.5412	59.7469	
total	5625	480351.3413		
R-squared	0.3005	Adj R-squared	0.3004	
F value	2415.7707	prob > F	0.0001	
variable	coefficient	stand error	T-ratio	prob > T
intercept	17.98515	0.1498	120.0269	0.0001
treatment_locality	-10.14497	0.2064	-49.1505	0.0001
durbin-watson: 1.7118				
estimate	predicted	lower 80%	upper 80%	
QTM CI	17.9851	17.7931	18.1772	
QTM PI	17.9851	8.0762	27.8941	
row 5627				
CI	17.9851	17.7931	18.1772	

The following image shows the result when the MathExpression didn't follow the required algorithm1 dataset conventions: the 1st 3 columns are row identifiers, the 4th column is the dependent variable, and the remaining independent variables are identified by the



DevTreks –social budgeting that improves lives and livelihoods

Example a is the only regression algorithm completed using the .NetStandard libraries. It's very likely the remaining regressions can be developed using these libraries, but this reference's current focus is on developing Machine Learning algorithms. Demonstration of the remaining regression techniques is deferred to a possible future release.

C. Introduction to Standard Impact Evaluation using R

Cloud testing is planned for 216.

[https://www.devtreks.org/greentreks/preview/carbon/resourcepack/Resource Conservation Value Accounting, RCA Example 7B/1562/none](https://www.devtreks.org/greentreks/preview/carbon/resourcepack/Resource%20Conservation%20Value%20Accounting,%20RCA%20Example%207B/1562/none)

[http://localhost:5000/greentreks/preview/carbon/resourcepack/Impact Evaluation, RCA Example 7B/548/none](http://localhost:5000/greentreks/preview/carbon/resourcepack/Impact%20Evaluation,%20RCA%20Example%207B/548/none)

M&E Input Calculator

[http://localhost:5000/greentreks/preview/carbon/input/Impact Evaluation 7, 1 and 2/2147409847/none](http://localhost:5000/greentreks/preview/carbon/input/Impact%20Evaluation%207,%201%20and%202/2147409847/none)

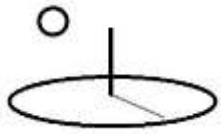
Stock Input Calculator

[http://localhost:5000/greentreks/preview/carbon/input/Impact Evaluation 7A-3/2147409848/none](http://localhost:5000/greentreks/preview/carbon/input/Impact%20Evaluation%207A-3/2147409848/none)

The following examples modify the Stata statistical scripts found in Gertler et al's Technical Companion (Version 1, September, 2016) to R statistical package scripts (4*). Given that related CTA references already introduce statistical libraries, including R and Python (i.e. algorithms 2 and 3), which support those techniques, this example does not provide additional documentation for the statistical techniques.

Example a. Randomized Assignment in a Regression Framework (algorithm2, subalgorithm1)

The following script and image demonstrate using R to run this regression. The image also demonstrates that Version 2.1.4 has been upgraded to more fully support the R and Python algorithms (algorithms 2, 3, 4, and 5 are being upgraded for use with machine learning algorithms).



DevTreks –social budgeting that improves lives and livelihoods

```
args <- commandArgs(TRUE)

url <- args[1]

print(url)

dataset1 <- read.table(url, header=TRUE, sep=",")

#dataset2 is dataset1 minus the last 3 lines of data

dataset2 <- head(dataset1, n=-3)

#dataset3 is the last 3 lines of data used for ci

dataset3 <- tail(dataset1, n=3)

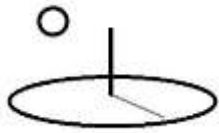
model <- lm(health_expenditures ~ treatment_locality, data=dataset2)

f1 <- summary(model)

print(f1)

ci <- predict(model, dataset3, interval='confidence')

print(ci)
```



Indicator 1

Impact Evaluation 7A1

Indicator 1 Description

This example can be found as Example 7A in the Social Performance Analysis tutorial

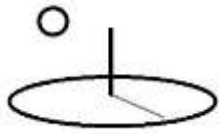
Indicator 1 URL

http://localhost:5000/resources/network_carbon/resourcepack_548/resource_1974/Example7A-1.txt;http://localhost:5000/resources/network_carbon/resourcepack_548/resource_1973/Example7A-1.csv

Label 1	Rel Label 1
<input type="text" value="1"/>	<input type="text"/>
Date 1	Dist Type 1
<input type="text" value="04/04/2018"/>	<input type="text" value="none"/>
Q1 1	Q1 Unit 1
<input type="text" value="17.9852"/>	<input type="text" value="benchmark most hh savings"/>
Q2 1	Q2 Unit 1
<input type="text" value="17.6914"/>	<input type="text" value="benchmark low hh savings"/>
Q3 1	Q3 Unit 1
<input type="text" value="18.2789"/>	<input type="text" value="benchmark high hh savings"/>
Q4 1	Q4 Unit 1
<input type="text" value="17.9852"/>	<input type="text" value="target most hh savings"/>
Q5 1	Q5 Unit 1
<input type="text" value="17.6914"/>	<input type="text" value="target low hh savings"/>
Math Operator 1	BaseIO 1
<input type="text" value="equalto"/>	<input type="text" value="none"/>
QT 1	QT Unit 1
<input type="text" value="18.2789"/>	<input type="text" value="target high hh savings"/>

<input type="text" value="18.2789"/>	<input type="text" value="target high hh savings"/>
Math Type 1	Math Sub Type 1
<input type="text" value="algorithm2"/>	<input type="text" value="subalgorithm1"/>
QT D1 1	QT D1 Unit 1
<input type="text" value="0.0000"/>	<input type="text" value="none"/>
QT D2 1	QT D2 Unit 1
<input type="text" value="0.0000"/>	<input type="text" value="none"/>
QT Most 1	QT Most Unit 1
<input type="text" value="17.9852"/>	<input type="text" value="actual most hh savings"/>
QT Low 1	QT Low Unit 1
<input type="text" value="17.6914"/>	<input type="text" value="actual low hh savings"/>
QT High 1	QT High Unit 1
<input type="text" value="18.2789"/>	<input type="text" value="actual high hh savings"/>
Math Expression 1	
<input type="text" value="I1.Q1.health_expenditures + I1.Q2.treatment_locality"/>	
Math Result 1	
<pre> ---r results [1] "http://localhost:5000/resources/network_carbon/resourcepack_548/resource_1973/Example7A-1.csv" Call: lm(formula = health_expenditures ~ treatment_locality, data = dataset2) Residuals: Min 1Q Median 3Q Max -8.719 -5.353 -1.987 2.994 79.540 Coefficients: Estimate Std. Error t value Pr(> t) (Intercept) 17.9851 0.1498 120.03 <2e-16 *** treatment_locality -10.1450 0.2064 -49.15 <2e-16 *** --- Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 Residual standard error: 7.73 on 5624 degrees of freedom Multiple R-squared: 0.3005, Adjusted R-squared: 0.3004 F-statistic: 2416 on 1 and 5624 DF, p-value: < 2.2e-16 </pre>	

The following image shows that a DevTreks convention with several algorithms is to use the last 3 rows of a dataset to generate confidence intervals for 3 predicted ranges of values. Version



2.1.4 upgraded the rules for several algorithms to use these 3 predicted ranges in the same manner as introduced in Example 5 -to fill in Indicator metadata properties for most likely, low, and high, predictions for 3 sets of Indicator.metadata properties. The analyst must fill in appropriate units for the 3 predictions (i.e. Example 5’s benchmarks, targets, and actuals).

In this image, the reason why Gertler’s last 3 rows of data generate the same prediction, even though each has different household expenditures has something to do with the R command’s predict() function, but is not explored here (i.e. F value = 2416?, binomial explanatory variables?) (3*).

```
Residual standard error: 7.73 on 5624 degrees of freedom
Multiple R-squared: 0.3005, Adjusted R-squared: 0.3004
F-statistic: 2416 on 1 and 5624 DF, p-value: < 2.2e-16
```

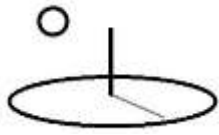
```
fit lwr upr
5627 17.98515 17.6914 18.27889
5628 17.98515 17.6914 18.27889
5629 17.98515 17.6914 18.27889
```

Gertler’s Stata regression scripts include a cluster command, or cl(locality_identifier) which can be replicated using the R package’s miceadds package with the following script. Given that this script only changes the standard errors, but not the coefficients or statistical validity of the results, these scripts will not cluster this dataset further.

```
> dataset1 <- read.table("http://localhost:8000/resources/network_carbon/resourcepack_043/Resource_19/3/example/A-1.csv", )
> #dataset2 is dataset1 minus the last 3 lines of data
> dataset2 <- head(dataset1, n=3)
> #dataset3 is the last 3 lines of data used for ci
> dataset3 <- tail(dataset1, n=3)
> model <- miceadds::lm.cluster(health_expenditures ~ treatment_locality, data=dataset2, cluster = "locality_identifier" )
> f1 <- summary(model)
R^2 = 0.30048
```

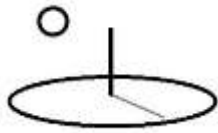
```
Estimate Std. Error t value Pr(>|t|)
(Intercept) 17.98515 0.3062078 58.73510 0.000000e+00
treatment_locality -10.14497 0.3953497 -25.66074 3.208171e-145
```

The following R Script and image confirm that the 10 variable limit for explanatory variables is only enforced with algorithm1. Even if not required, DevTreks still recommends the 10 explanatory variable limit.

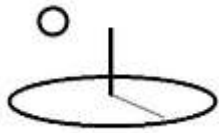


DevTreks –social budgeting that improves lives and livelihoods

```
model <- lm(health_expenditures ~ treatment_locality + age_hh + age_sp + educ_hh +  
educ_sp + female_hh + indigenous + hhsiz + dirtfloor + bathroom + land +  
hospital_distance, data=dataset2)
```



Indicator 2 Impact Evaluation 7A1 Indicator 2 Description This example can be found as Example 7A in the Social Performance Analysis tutorial Indicator 2 URL http://localhost:5000/resources/network_carbon/resourcepack_548/resource_1968/Example7A-1b.txt ; http://localhost:5000/resources/network_carbon/resourcepack_548/resource_1973/Example7A-1.csv		Math Type 2 algorithm2 Math Sub Type 2 subalgorithm1																																																		
Label 2 2	Rel Label 2	QT D1 2 0.0000 QT D2 2 0.0000 QT Most 2 16.7279 QT Low 2 16.1618 QT High 2 17.2939																																																		
Date 2 04/04/2018	Dist Type 2 none	QT D1 Unit 2 QT D2 Unit 2 QT Most Unit 2 actual most hh savings QT Low Unit 2 actual low hh savings QT High Unit 2 actual high hh savings																																																		
Q1 2 16.4047	Q1 Unit 2 benchmark most hh savings	Math Expression 2 $I2.Q1.treatment_locality + I2.Q2.age_hh + I2.Q3.age_sp + I2.Q4.edu$																																																		
Q2 2 15.7978	Q2 Unit 2 benchmark low hh savings	Math Result 2 <pre>--r results [1] "http://localhost:5000/resources/network_carbon/resourcepack_548/resource_1973/Example7A-1.csv"</pre>																																																		
Q3 2 17.0116	Q3 Unit 2 benchmark high hh savings	Call: $lm(formula = health_expenditures \sim treatment_locality + age_hh + age_sp + educ_hh + educ_sp + female_hh + indigenous + hhsz + dirtfloor + bathroom + land + hospital_distance, data = dataset2)$																																																		
Q4 2 13.8149	Q4 Unit 2 target most hh savings	Residuals: Min 1Q Median 3Q Max -16.737 -4.378 -1.341 2.861 77.933																																																		
Q5 2 13.2636	Q5 Unit 2 target low hh savings	Coefficients: <table border="1"> <thead> <tr> <th></th> <th>Estimate</th> <th>Std. Error</th> <th>t value</th> <th>Pr(> t)</th> </tr> </thead> <tbody> <tr> <td>(Intercept)</td> <td>27.570906</td> <td>0.615931</td> <td>44.763</td> <td>< 2e-16 ***</td> </tr> <tr> <td>treatment_locality</td> <td>-10.012304</td> <td>0.186888</td> <td>-53.574</td> <td>< 2e-16 ***</td> </tr> <tr> <td>age_hh</td> <td>0.041138</td> <td>0.012297</td> <td>3.345</td> <td>0.000827 ***</td> </tr> <tr> <td>age_sp</td> <td>0.003092</td> <td>0.014317</td> <td>0.216</td> <td>0.829005</td> </tr> <tr> <td>educ_hh</td> <td>-0.039319</td> <td>0.043528</td> <td>-0.903</td> <td>0.366396</td> </tr> <tr> <td>educ_sp</td> <td>-0.022416</td> <td>0.045231</td> <td>-0.496</td> <td>0.620206</td> </tr> <tr> <td>female_hh</td> <td>0.642431</td> <td>0.371010</td> <td>1.732</td> <td>0.083404 .</td> </tr> <tr> <td>indigenous</td> <td>-1.903504</td> <td>0.205928</td> <td>-9.244</td> <td>< 2e-16 ***</td> </tr> <tr> <td>hhsz</td> <td>-1.603272</td> <td>0.049035</td> <td>-32.697</td> <td>< 2e-16 ***</td> </tr> </tbody> </table>		Estimate	Std. Error	t value	Pr(> t)	(Intercept)	27.570906	0.615931	44.763	< 2e-16 ***	treatment_locality	-10.012304	0.186888	-53.574	< 2e-16 ***	age_hh	0.041138	0.012297	3.345	0.000827 ***	age_sp	0.003092	0.014317	0.216	0.829005	educ_hh	-0.039319	0.043528	-0.903	0.366396	educ_sp	-0.022416	0.045231	-0.496	0.620206	female_hh	0.642431	0.371010	1.732	0.083404 .	indigenous	-1.903504	0.205928	-9.244	< 2e-16 ***	hhsz	-1.603272	0.049035	-32.697	< 2e-16 ***
	Estimate	Std. Error	t value	Pr(> t)																																																
(Intercept)	27.570906	0.615931	44.763	< 2e-16 ***																																																
treatment_locality	-10.012304	0.186888	-53.574	< 2e-16 ***																																																
age_hh	0.041138	0.012297	3.345	0.000827 ***																																																
age_sp	0.003092	0.014317	0.216	0.829005																																																
educ_hh	-0.039319	0.043528	-0.903	0.366396																																																
educ_sp	-0.022416	0.045231	-0.496	0.620206																																																
female_hh	0.642431	0.371010	1.732	0.083404 .																																																
indigenous	-1.903504	0.205928	-9.244	< 2e-16 ***																																																
hhsz	-1.603272	0.049035	-32.697	< 2e-16 ***																																																
Math Operator 2 equalto	BaseIO 2 none																																																			
QT 2 14.3662	QT Unit 2 target high hh savings																																																			



The reason why the final 3 predicted ranges differ from the previous results has not been explored for this release.

```
Coefficients:
      Estimate Std. Error t value Pr(>|t|)
(Intercept)  27.570906  0.615931  44.763 < 2e-16 ***
treatment_locality -10.012304  0.186888 -53.574 < 2e-16 ***
age_hh       0.041138  0.012297   3.345 0.000827 ***
age_sp       0.003092  0.014317   0.216 0.829005
educ_hh     -0.039319  0.043528  -0.903 0.366396
educ_sp     -0.022416  0.045231  -0.496 0.620206
female_hh   0.642431  0.371010   1.732 0.083404 .
indigenous  -1.903504  0.205928  -9.244 < 2e-16 ***
hhszsize    -1.603272  0.049035 -32.697 < 2e-16 ***
dirtfloor   -1.848968  0.219555  -8.421 < 2e-16 ***
bathroom    0.282634  0.193330   1.462 0.143818
land        0.038350  0.036448   1.052 0.292752
hospital_distance -0.002581  0.002336  -1.105 0.269271
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

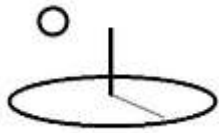
Residual standard error: 6.986 on 5613 degrees of freedom
Multiple R-squared:  0.4297, Adjusted R-squared:  0.4285
F-statistic: 352.4 on 12 and 5613 DF, p-value: < 2.2e-16

      fit  lwr  upr
5627 16.72786 16.16179 17.29393
5628 16.40474 15.79783 17.01164
5629 13.81492 13.26364 14.36619
```

+ Indicator 3

Example b. Randomized Treatment Regression Analysis (algorithm2, subalgorithm1)

The following script and image demonstrate using R to run this regression. This statistical analysis technique, known as two stage least squares regression analysis, requires loading a custom R package, AER. The model’s formula employs “subset” syntax because this dataset is the full 19,000+ row dataset but only 9,000+ rows are needed by this regression (5*). The R predict() function did not return valid confidence intervals for the 3 predicted ranges

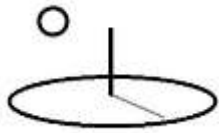


DevTreks –social budgeting that improves lives and livelihoods

demonstrated in the previous regression example and was replaced by the alternative `confint()` function.

```
install.packages("AER")
library(AER)
args <- commandArgs(TRUE)
url <- args[1]
print(url)
dataset1 <- read.table(url, header=TRUE, sep=",")
model <- ivreg(health_expenditures ~ enrolled | treatment_locality, data=dataset1, subset =
round == "1")
f1 <- summary(model)
print(f1)
ci <- confint(model, 'enrolled', level=0.80)
print(ci)
```

The following images show that the R results match the Stata results. The `confint()` function returned an 80% ci for the estimated housing savings coefficient (i.e. \$10.72 +- 0.44) which was used to manually fill in the QTM, QTL, and QTU properties. Similar techniques can be used for the benchmark and targets but are not demonstrated in this example (3*).



- Indicator 3

Description 3

Indicator 3 URL

Label 3

Rel Label 3

Date 3

Dist Type 3

Q1 3

Q1 Unit 3

Q2 3

Q2 Unit 3

Q3 3

Q3 Unit 3

Q4 3

Q4 Unit 3

Q5 3

Q5 Unit 3

Math Operator 3

BaseIO 3

QT 3

QT Unit 3

Math Type 3

Math Sub Type 3

QT Most 3

QT Most Unit 3

QT Low 3

QT Low Unit 3

QT High 3

QT High Unit 3

Math Expression 3

Math Result 3

```

---r results
package 'AER' successfully unpacked and MD5 sums checked

The downloaded binary packages are in
C:\Users\kpboyle1\AppData\Local\Temp\RtmpMIP1ht\downloaded_packages
[1] "http://localhost:5000/resources/network_carbon/resourcepack_548/resource_1975/Example7A.csv"

Call:
ivreg(formula = health_expenditures ~ enrolled |
treatment_locality,
data = dataset1, subset = round == "1")

Residuals:
Min 1Q Median 3Q Max
-11.857 -6.789 -2.785 3.545 96.766

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 20.0642   0.1474 136.11 <2e-16 ***
enrolled  -10.7163   0.3486 -30.74 <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

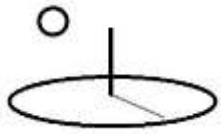
Residual standard error: 10.38 on 9912 degrees of freedom
Multiple R-Squared: 0.2387, Adjusted R-squared: 0.2386
Wald test: 944.8 on 1 and 9912 DF, p-value: < 2.2e-16

            10 %   90 %
enrolled -11.16309 -10.2695

```

+ Indicator 4

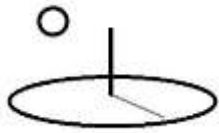
Example c. Difference in Differences Regression Analysis (algorithm2, subalgorithm1)



DevTreks –social budgeting that improves lives and livelihoods

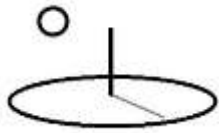
The following script and image show that the R results match the Stata results. In this example, the QTM, QTL, and QTU, values were manually entered from the R results and the calculations were run a second time to process the edits.

```
args <- commandArgs(TRUE)
url <- args[1]
print(url)
dataset1 <- read.table(url, header=TRUE, sep=",")
#did regression
model <- lm(health_expenditures ~ round * eligible, data=dataset1, subset = treatment_locality
== "1")
f1 <- summary(model)
print(f1)
ci <- confint(model, 'round:eligible', level=0.80)
print(ci)
```



Impact Evaluation, Example 3a		QT Most 1	QT Most Unit 1
Indicator 1 Description		8.1700	most likely hh savings
Example 3 in Social Performance Analysis 3 reference.		QT Low 1	QT Low Unit 1
Indicator 1 URL		7.7500	low 80 percent ci
http://localhost:5000/resources/network_carbon/resourc epack_548/resource_1976/Example7A- 3a.txt;http://localhost:5000/resources/network_carbon/re		QT High 1	QT High Unit 1
Label 1	Rel Label 1	8.5800	high 80 percent ci
RCA7A3	none	Math Expression 1	
Date 1	Dist Type 1	I1.Q1.health_expenditures + I1.Q2.round + I1.Q3.eligible	
04/12/2018	none	Math Result 1	
Q1 1	Q1 Unit 1	--r results [1] "http://localhost:5000/resources/network_carbon/resour cepack_548/resource_1975/Example7A.csv"	
0.0000	none	Call: lm(formula = health_expenditures ~ round * eligible, data = dataset1, subset = treatment_locality == "1")	
Q2 1	Q2 Unit 1	Residuals: Min 1Q Median 3Q Max -14.057 -4.910 -0.942 3.221 91.730	
0.0000	none	Coefficients: Estimate Std. Error t value Pr(> t) (Intercept) 20.7915 0.1772 117.362 < 2e-16 *** round 1.5134 0.2505 6.041 1.59e-09 *** eligible -6.3018 0.2291 -27.501 < 2e-16 *** round:eligible -8.1629 0.3241 -25.190 < 2e-16 ***	
Q3 1	Q3 Unit 1	--- Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1	
0.0000	none	Residual standard error: 7.913 on 9915 degrees of freedom Multiple R-squared: 0.3436, Adjusted R-squared: 0.3434 F-statistic: 1730 on 3 and 9915 DF, p-value: < 2.2e-16	
Q4 1	Q4 Unit 1	10 % 90 % round:eligible -8.578249 -7.747614	
0.0000	none		
Q5 1	Q5 Unit 1		
0.0000	none		
Math Operator 1	BaselO 1		
equalto	none		
QT 1	QT Unit 1		
0.0000	none		
Math Type 1	Math Sub Type 1		
algorithm2	subalgorithm1		

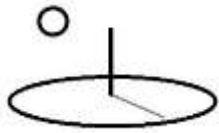
The following R script and image shows similar results. The confit() function for generating a confidence interval does not work with the plm package. In practice, the interval must be



DevTreks –social budgeting that improves lives and livelihoods

calculated from the `xtenrolled` t-statistic. For convenience, they were taken from the previous example.

```
args <- commandArgs(TRUE)
url <- args[1]
print(url)
dataset1 <- read.table(url, header=TRUE, sep=",")
#generate a new column
dataset1$xtenrolled <- '0'
#change column values based on condition
dataset2 <- within(dataset1, xtenrolled[enrolled == '1' & round == '1'] <- '1')
#fixed effects regression
install.packages("plm")
library(plm)
model <- plm(health_expenditures ~ xtenrolled + round, data=dataset2, subset =
treatment_locality == "1", index=c("household_idenfifier", "round"), model="within")
f1 <- summary(model)
print(f1)
```

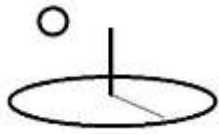


Math Type 2 <input type="text" value="algorithm2"/>	Math Sub Type 2 <input type="text" value="subalgorithm1"/>
QT D1 2 <input type="text" value="0.0000"/>	QT D1 Unit 2 <input type="text" value="low"/>
QT D2 2 <input type="text" value="0.0000"/>	QT D2 Unit 2 <input type="text" value="high"/>
QT Most 2 <input type="text" value="8.1700"/>	QT Most Unit 2 <input type="text" value="most likely hh savings"/>
QT Low 2 <input type="text" value="7.7500"/>	QT Low Unit 2 <input type="text" value="lower 80 % ci"/>
QT High 2 <input type="text" value="8.5800"/>	QT High Unit 2 <input type="text" value="upper 80 % ci"/>
Math Expression 2 <input type="text" value="I1.Q1.health_expenditures + I1.Q2.round + I1.Q3.eligible"/>	
Math Result 2 <pre>---r results package 'plm' successfully unpacked and MD5 sums checked The downloaded binary packages are in C:\Users\kpboyle1 \AppData\Local\Temp\RtmpeqiAOI\downloaded_packages [1] "http://localhost:5000/resources/network_carbon/resource pack_548/resource_1975/Example7A.csv" Oneway (individual) effect Within Model Call: plm(formula = health_expenditures ~ xtenrolled + round, data = dataset2, subset = treatment_locality == "1", model = "within", index = c("household_idenfifier", "round")) Unbalanced Panel: n=4960, T=1-2, N=9919 Residuals : Min. 1st Qu. Median 3rd Qu. Max. -40.90 -2.35 0.00 2.35 40.90 Coefficients : Estimate Std. Error t-value Pr(> t) xtenrolled1 -8.16334 0.26690 -30.5858 < 2.2e-16 *** round1 1.51342 0.20634 7.3345 2.587e-13 *** --- Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1</pre>	

+ Indicator 3

Example d. Discontinuity Regression Analysis (algorithm2, subalgorithm1)

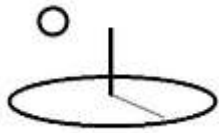
The following R script and image derive from the Baumer reference. Although not a perfect match to the Stata results, the statistical results are similar. A new Version 2.1.4 rule caused the



DevTreks –social budgeting that improves lives and livelihoods

quantitative Indicator.meta properties to be filled in automatically. When the final lines of script in algorithms 2, 3, 4, and 6, return a 3 row, 4 column, dataset with titles that include the term “predict”, or when a predict() function is used in the script, the Indicator.meta quantitative properties will be filled in automatically based on the final 3 rows of data. The analyst must still fill in their units of measurement manually.

```
args <- commandArgs(TRUE)
url <- args[1]
print(url)
dataset1 <- read.table(url, header=TRUE, sep=",")
dataset2 <- subset(dataset1, dataset1$round == '1' & dataset1$treatment_locality == '1')
model<-lm(dataset2$health_expenditures~I(dataset2$poverty_index-58)*dataset2$eligible)
f1 <- summary(model)
print(f1)
mean = coef(summary(model))[1]
mean2 = mean-2.5
mean3 = mean+2.5
savings <- coef(summary(model))[3] * -1
df1<-data.frame(predict=(c(mean,mean2,mean3)),low=(c(mean-savings,mean2-savings,mean3-
savings)),high=(c(mean+savings,mean2+savings,mean3+savings)))
print(df1)
```



Q1 1	Q1 Unit 1	18.0545	bm most
Q2 1	Q2 Unit 1	6.8628	bm low
Q3 1	Q3 Unit 1	29.2462	bm high
Q4 1	Q4 Unit 1	23.0545	target most
Q5 1	Q5 Unit 1	11.8628	target low
Math Operator 1	BaselO 1	equalto	none
QT 1	QT Unit 1	34.2462	target high
Math Type 1	Math Sub Type 1	algorithm2	subalgorithm1
QT D1 1	QT D1 Unit 1	0.0000	none
QT D2 1	QT D2 Unit 1	0.0000	none
QT Most 1	QT Most Unit 1	20.5545	most likely hh savings
QT Low 1	QT Low Unit 1	9.3628	low hh savings
QT High 1	QT High Unit 1	31.7462	high hh savings

Math Expression 1

I1.Q1.health_expenditures + I1.Q2.poverty_index + I1.Q3.eligible

Math Result 1

```

--r results
[1] "http://localhost:5000/resources/network_carbon/resourcepack_548/resource_1975/Example7A.csv"

Call:
lm(formula = dataset2$health_expenditures ~ I(dataset2$poverty_index - 58) * dataset2$eligible)

Residuals:
    Min     1Q   Median     3Q      Max
-16.816  -6.449  -2.293   3.440  89.489

Coefficients:
                Estimate Std. Error t value
(Intercept)      20.55449   0.33763  60.878
I(dataset2$poverty_index - 58)  0.22028   0.03156   6.979
dataset2$eligible  -11.19171   0.46618 -24.007
I(dataset2$poverty_index - 58).dataset2$eligible -0.04471   0.04374  -1.022

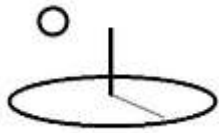
Pr(>|t|)
(Intercept) < 2e-16 ***
I(dataset2$poverty_index - 58) 3.37e-12 ***
dataset2$eligible < 2e-16 ***
I(dataset2$poverty_index - 58).dataset2$eligible 0.307
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 10.1 on 4956 degrees of freedom
Multiple R-squared: 0.338, Adjusted R-squared: 0.3376
F-statistic: 843.5 on 3 and 4956 DF, p-value: < 2.2e-16

predict low high
1 20.55449 9.362776 31.7462
2 18.05449 6.862776 29.2462
3 23.05449 11.862776 34.2462

```

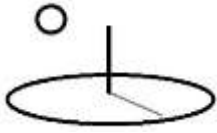
Example e. Propensity Score Matching Regression Analysis (algorithm2, subalgorithm1)



DevTreks –social budgeting that improves lives and livelihoods

The following script and image demonstrate using R to run this regression. This script does not reshape the data but returns similar results.

```
args <- commandArgs(TRUE)
url <- args[1]
print(url)
dataset1 <- read.table(url, header=TRUE, sep=",")
#data_wide <- dcast(dataset1, household_identifier ~ round,
#value.var = c("health_expenditures", "age_hh", "age_sp", "educ_hh", "educ_sp", "hospital"))
model <- glm(enrolled ~
age_hh+age_sp+educ_hh+educ_sp+female_hh+indigenous+hhsize+dirtfloor
+bathroom+land+hospital_distance, family = binomial(link = "probit"), data = dataset1)
f1 <- summary(model)
print(f1)
dataset2 <- tail(dataset1, n=3)
ci <- predict(model, dataset2, interval='confidence')
print(ci)
```



Math Result 2

```
---r results
[1] "http://localhost:5000/resources/network_carbon/resourcepack_54
8/resource_1975/Example7A.csv"

Call:
glm(formula = enrolled ~ age_hh + age_sp + educ_hh + educ_sp +
  female_hh + indigenous + hhsz + dirtfloor + bathroom +
  land + hospital_distance, family = binomial(link = "probit"),
  data = dataset1, subset = round == "1")

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.5405 -0.8723 -0.6088  1.1725  2.6206

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)  -0.4780164  0.0911052  -5.247 1.55e-07 ***
age_hh       -0.0127868  0.0017741  -7.208 5.70e-13 ***
age_sp       -0.0083686  0.0020634  -4.056 5.00e-05 ***
educ_hh      -0.0217746  0.0062244  -3.498 0.000468 ***
educ_sp      -0.0153541  0.0067221  -2.284 0.022365 *
female_hh    -0.0294526  0.0521053  -0.565 0.571903
indigenous    0.1624462  0.0313125   5.188 2.13e-07 ***
hhsz         0.1193654  0.0066731  17.888 < 2e-16 ***
dirtfloor    0.3749278  0.0308346  12.159 < 2e-16 ***
bathroom     -0.1243927  0.0290425  -4.283 1.84e-05 ***
land         -0.0275939  0.0050322  -5.483 4.17e-08 ***
hospital_distance 0.0016014  0.0003533   4.533 5.82e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 12097  on 9913  degrees of freedom
Residual deviance: 11014  on 9902  degrees of freedom
AIC: 11038
```

The following R script and image shows similar results. R contains several more advanced packages for conducting propensity score matching. The Indicator.meta displays 3 predicted household expenditures (i.e. benchmark, target, actual) based on the last 3 rows of data.

```
args <- commandArgs(TRUE)
url <- args[1]
print(url)
```



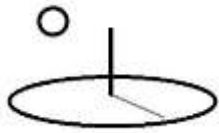

DevTreks –social budgeting that improves lives and livelihoods

```
dataset2 <- read.table(url, header=TRUE, sep=",")
dataset3 <- subset(dataset2, round == '1')
# generate propensity scores for all of the data
ps.model <- glm(enrolled ~
  age_hh+age_sp+educ_hh+educ_sp+female_hh+indigenous+hhsize+dirtfloor
  +bathroom+land+hospital_distance, data = dataset3, family=binomial(link="logit"),
  na.action=na.pass)
#summary(ps.model)
# add pscores to study data
dataset3$pscore <- predict(ps.model, newdata = dataset3, type = "response")
# distribution of ps
#summary(dataset3$pscore)
dim(dataset3)
# restrict data to ps range .10 <= ps <= .90=
dataset4 <- dataset3[dataset3$pscore >= .10 & dataset3$pscore <=.90,]
summary(dataset4$pscore)

# regression with controls on propensity score screened data set
model <-
lm(health_expenditures~enrolled+age_hh+age_sp+educ_hh+educ_sp+female_hh+indigenous+h
hsize+dirtfloor
  +bathroom+land+hospital_distance, data = dataset4)

f1 <- summary(model)
print(f1)
dataset2 <- tail(dataset4, n=3)
ci <- predict(model, dataset2, interval='confidence')

print(ci)
```



Math Type 2 **Math Sub Type 2**

algorithm2 subalgorithm1

QT D1 2 QT D1 Unit 2

0.0000 none

QT D2 2 QT D2 Unit 2

0.0000 none

QT Most 2 QT Most Unit 2

15.2267 none

QT Low 2 QT Low Unit 2

14.6403 none

QT High 2 QT High Unit 2

15.8130 none

Math Expression 2

I2.Q2.age_hh + I2.Q3.age_sp + I2.Q4.educ_hh + I2.Q5.educ_sp

Math Result 2

```

---r results
[1] "http://localhost:5000/resources/network_carbon/resourcepack_548/resource_1980/Example7NoHospital.csv"
[1] 9914 23
   Min. 1st Qu.  Median    Mean 3rd Qu.   Max.
 0.1000 0.2083 0.3081 0.3208 0.4224 0.7054

Call:
lm(formula = health_expenditures ~ enrolled + age_hh + age_sp + educ_hh + educ_sp + female_hh + indigenous + hhsize + dirtfloor + bathroom + land + hospital_distance, data = dataset4)

Residuals:
   Min     1Q   Median     3Q    Max
-18.321 -5.096 -1.455  3.261 86.998

```

Residuals:

	Min	1Q	Median	3Q	Max
	-18.321	-5.096	-1.455	3.261	86.998

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	28.020130	0.577436	48.525	< 2e-16 ***
enrolled	-9.975192	0.193555	-51.537	< 2e-16 ***
age_hh	0.048446	0.011157	4.342	1.43e-05 ***
age_sp	0.012922	0.013119	0.985	0.32464
educ_hh	-0.002834	0.039185	-0.072	0.94234
educ_sp	-0.016638	0.042397	-0.392	0.69475
female_hh	0.833563	0.323174	2.579	0.00992 **
indigenous	-2.149103	0.197686	-10.871	< 2e-16 ***
hhsize	-1.820919	0.043823	-41.552	< 2e-16 ***
dirtfloor	-1.560514	0.195206	-7.994	1.46e-15 ***
bathroom	0.365344	0.183002	1.996	0.04592 *
land	0.055423	0.032767	1.691	0.09078 .
hospital_distance	-0.004459	0.002230	-2.000	0.04558 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 8.275 on 9034 degrees of freedom

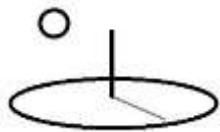
Multiple R-squared: 0.4006, Adjusted R-squared: 0.3998

F-statistic: 503.2 on 12 and 9034 DF, p-value: < 2.2e-16

	fit	lwr	upr
9912	15.22665	14.64026	15.81303
9913	14.58105	13.78674	15.37536
9914	15.93932	15.38537	16.49327

+
Indicator 3

D. Introduction to Standard Impact Evaluation using Python (algorithm3) (1*)



DevTreks –social budgeting that improves lives and livelihoods

The following examples demonstrate applying these principles. This example uses algorithm3 to modify Example 7's Stata statistical scripts (1*).

Example a. Randomized Assignment in a Regression Framework (algorithm3, subalgorithm1)

The following script and image display the equivalent results.

Example b. Randomized Treatment Regression Analysis (algorithm3, subalgorithm1)

The following script and image display the equivalent results.

Example c. Difference in Differences Regression Analysis (algorithm3, subalgorithm1)

The following script and image display the equivalent results.

Example d. Discontinuity Regression Analysis (algorithm3, subalgorithm1)

The following script and image display the equivalent results.

Example e. Propensity Score Matching Regression Analysis (algorithm3, subalgorithm1)

The following script and image display the equivalent results.

Case Study References

Example 5, plus the following.

<https://cran.r-project.org/doc/manuals/r-release/R-intro.html>



DevTreks –social budgeting that improves lives and livelihoods

<https://cran.r-project.org/web/views/>

<https://cran.r-project.org/doc/manuals/r-release/R-data.html>

Standard Stats

James, Gareth, Daniela Witten, Trevor Hastie, Robert Tibshirani. An Introduction to Statistical Learning with Applications in R. Springer Texts in Statistics. 2017

<https://www.rdocumentation.org/packages/AER/versions/1.2-5/topics/ivreg>

<https://dss.princeton.edu/training/>

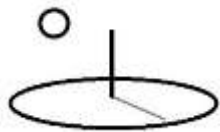
Baumer, Patricia. Regression Discontinuity. last accessed: April 13, 2018:

faculty.smu.edu/kyle/courses/7312/presentations/baumer/Baumer_RD.pdf

<https://sejdemyr.github.io/r-tutorials/statistics/tutorial8.html>

Footnotes

1. Although Python examples haven't been prepared yet, that doesn't imply any preference for statistical packages. DevTreks endorses any package that is open source or does not charge fees, can be automated by software developers, and carries out calculations that can be verified using reference datasets. DevTreks assumes additional statistical packages, such as Julia (algorithm6), will be supported in future releases.



DevTreks –social budgeting that improves lives and livelihoods

Appendix B. Machine Learning Statistical Impact Evaluation Analysis

A. Introduction to Machine Learning Impact Evaluation

The following examples demonstrate carrying out Impact Evaluation using Machine Learning techniques that correspond to Gertler et al’s (2016) standard statistical techniques.

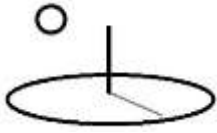
This section introduces several ML algorithms that address specific impact evaluation questions. Wikipedia URLs provide the definition for each algorithm. Bontempi (2017) provides the mathematical basis for each of these algorithms. Smola and Vishwanathan (2008) provide the actual algorithm. Only the first example demonstrates how these authors explain the algorithms. Examples 8A, 8B, and 8C demonstrate using alternative statistical software to run each algorithm.

In the following examples, the ML term “labels” refer to the equivalent statistical term “y = dependent variable”. The ML term “features” refer to the equivalent statistical term “x = independent variables”. ML training datasets teach ML algorithms how to make predictions for ML testing datasets.

Example a. Naïve Bayes. This classification algorithm uses training datasets to predict the classification of the members of a testing dataset. The following examples predict the poverty index classification of households and villages in the Gertler dataset.

Wikipedia definition: https://en.wikipedia.org/wiki/Naive_Bayes_classifier

The following image (Smola and Vishwanathan, 2008) displays this algorithm.



Algorithm 1.1 Naive Bayes

Train(\mathbf{X}, \mathbf{Y}) {reads documents \mathbf{X} and labels \mathbf{Y} } Compute dictionary D of \mathbf{X} with n words. Compute m, m_{ham} and m_{spam} . Initialize $b := \log c + \log m_{\text{ham}} - \log m_{\text{spam}}$ to offset the rejection threshold Initialize $p \in \mathbb{R}^{2 \times n}$ with $p_{ij} = 1, w_{\text{spam}} = n, w_{\text{ham}} = n$.

{Count occurrence of each word}

 {Here x_i^j denotes the number of times word j occurs in document x_i } for $i = 1$ to m do if $y_i = \text{spam}$ then for $j = 1$ to n do $p_{0,j} \leftarrow p_{0,j} + x_i^j$ $w_{\text{spam}} \leftarrow w_{\text{spam}} + x_i^j$

end for

else

 for $j = 1$ to n do $p_{1,j} \leftarrow p_{1,j} + x_i^j$ $w_{\text{ham}} \leftarrow w_{\text{ham}} + x_i^j$

end for

end if

end for

{Normalize counts to yield word probabilities}

 for $j = 1$ to n do $p_{0,j} \leftarrow p_{0,j} / w_{\text{spam}}$ $p_{1,j} \leftarrow p_{1,j} / w_{\text{ham}}$

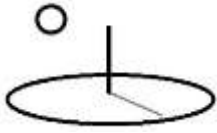
end for

Classify(x) {classifies document x } Initialize score threshold $t = -b$ for $j = 1$ to n do $t \leftarrow t + x^j (\log p_{0,j} - \log p_{1,j})$

end for

 if $t > 0$ return spam else return ham

The following image (Bontempi, 2017) displays the underlying mathematics.



8.2.1 Naive Bayes classifier

The Naive Bayes (NB) classifier has shown in some domains a performance comparable to that of neural networks and decision tree learning.

8.2. NONLINEAR CLASSIFICATION

215

Consider a classification problem with n inputs and a random output variable \mathbf{y} that takes values in the set $\{c_1, \dots, c_K\}$. The Bayes optimal classifier should return

$$c^*(x) = \arg \max_{j=1, \dots, K} \text{Prob} \{y = c_j | x\}$$

We can use the Bayes theorem to rewrite this expression as

$$\begin{aligned} c^*(x) &= \arg \max_{j=1, \dots, K} \frac{\text{Prob} \{x | y = c_j\} \text{Prob} \{y = c_j\}}{\text{Prob} \{x\}} \\ &= \arg \max_{j=1, \dots, K} \text{Prob} \{x | y = c_j\} \text{Prob} \{y = c_j\} \end{aligned}$$

How to estimate these two terms on the basis of a finite set of data? It is easy to estimate each of the a priori probabilities $\text{Prob} \{y = c_j\}$ simply by counting the frequency with which each target class occurs in the training set. The estimation of $\text{Prob} \{x | y = c_j\}$ is much harder. The NB classifier is based on the simplifying assumption that the input values are conditionally independent given the target value:

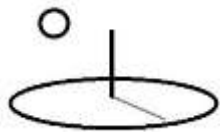
$$\text{Prob} \{x | y = c_j\} = \text{Prob} \{x_1, \dots, x_n | y = c_j\} = \prod_{h=1}^n \text{Prob} \{x_h | y = c_j\}$$

The NB classification is then

$$c_{NB}(x) = \arg \max_{j=1, \dots, K} \text{Prob} \{y = c_j\} \prod_{h=1}^n \text{Prob} \{x_h | y = c_j\}$$

If the inputs x_h are discrete variables the estimation of $\text{Prob} \{x_h | y = c_j\}$ boils down to the counting of the frequencies of the occurrences of the different values of x_h for a given class c_j .

The impact evaluation question addressed by this example is:



DevTreks –social budgeting that improves lives and livelihoods

How accurately can the poverty status of individual households and villages be predicted?

The following image shows that the dependent variable, `poverty_index`, for the dataset used in this example has not yet been classified into categories. The 2nd row passes the following instructions to the algorithm:

column1: `laplacian=true` (use Laplacian data transformation),

column4: `qcategory` (classify the `poverty_index` into categories that can be defined by `Indicator.Q1` to `Q5`).

This dataset used all of the rows in the Gertler dataset but only a subset of columns. Up to 5 categories can be classified this way by filling in properties for Q1 to Q5. Example 8A demonstrates that the category names will be taken from the corresponding `Q1Unit` to `Q5Unit` properties.

label	locality_id	household_id	poverty_index	hhsz	dirtfloor	land	female_hh
<code>laplacian=true</code>	none	none	<code>qcategory</code>	none	none	none	none
0	26	5	55.95054245	4	1	1	0
1	26	5	55.95054245	4	1	1	0
2	26	11	46.05873108	6	1	2	0
3	26	11	46.05873108	6	1	2	0
4	26	13	54.0958252	6	1	4	0
5	26	13	54.0958252	6	1	4	0
6	26	16	56.90340042	7	1	2	0
7	26	16	56.90340042	7	1	2	0
8	26	21	46.90880966	7	1	2	0
9	26	21	46.90880966	7	1	2	0
10	26	22	48.18392563	9	1	1	1

Examples

<https://web.stanford.edu/class/cs124/lec/naivebayes.pdf>

Example b. Deep Neural Network. This algorithm uses training datasets to predict the value of an explanatory variable for the members of a testing dataset. The following examples predict a



DevTreks –social budgeting that improves lives and livelihoods

household health care expenditure category (high costs, medium costs, low costs) for households and villages in the Gertler dataset.

Wikipedia definitions:

https://en.wikipedia.org/wiki/Deep_learning

https://en.wikipedia.org/wiki/Artificial_neural_network

The impact evaluation question addressed by this example is:

How accurately can the approximate size of household health care expenditures be predicted for individual households and villages?

In the following dataset, the label is the health_expenditures. The features are the hhsized, dirtfloor, land, and genderhh, parameters. By 2.1.4 conventions, the final 3 rows of the Test TEXT dataset are used to fill in Indicator.meta properties.

label	locality_id	household	health_ex	treatment	eligible	enrolled	poverty_in	round	age_hh	age_sp	educ_hh	educ_sp	female_hh	indigenous	hhsized	dirtfloor	bathroom	land
0	26	5	15.18546	1	1	1	55.95054	0	24	23	0	6	0	0	4	1	0	1
1	26	5	19.5809	1	1	1	55.95054	1	25	24	0	6	0	0	4	1	0	1
2	26	11	13.07626	1	1	1	46.05873	0	30	26	4	0	0	0	6	1	0	2
3	26	11	2.398854	1	1	1	46.05873	1	31	27	4	0	0	0	6	1	0	2
4	26	13	0	1	1	1	54.09583	1	59	57	0	0	0	0	6	1	0	4
5	26	13	15.28635	1	1	1	54.09583	0	58	56	0	0	0	0	6	1	0	4
6	26	16	20.02691	1	1	1	56.9034	1	36	25	3	0	0	0	7	1	0	2
7	26	16	11.31176	1	1	1	56.9034	0	35	24	3	0	0	0	7	1	0	2

Examples:

https://github.com/Microsoft/CNTK/blob/master/Tutorials/CNTK_101_LogisticRegression.ipynb

Example c. Time Series Neural Network. This deep learning algorithm uses training datasets to predict the health care expenditures for upcoming years for households and villages in the Gertler dataset.

Wikipedia definition https://en.wikipedia.org/wiki/Time_series

https://en.wikipedia.org/wiki/Artificial_neural_network

The impact evaluation question addressed by this example is:



DevTreks –social budgeting that improves lives and livelihoods

What are the trends in household expenditures for individual poor households and villages?

. The following table shows that Gertler’s dataset has been modified to include time series data. . Example b’s dataset was altered to time series data by adding 4 additional columns of health expenditures collected for 4 quarters (+-5% and +-10% of actual health cost).

label	locality	househ	health_expendit	q1	q2	q3	q4
2.5	none	none	none	none	none	none	none
4	26	13	0	0	0	0	0
3	26	11	2.398854494	2.518797	2.278912	2.63874	2.158969
11214	35	15743	9.916697502	10.41253	9.420863	10.90837	8.925028
11226	60	15754	10.82252502	11.36365	10.2814	11.90478	9.740273
11204	35	15735	11.31294537	11.87859	10.7473	12.44424	10.18165
11201	35	15733	11.31893444	11.88488	10.75299	12.45083	10.18704
11242	60	15765	11.57980633	12.1588	11.00082	12.73779	10.42183
11228	60	15755	11.6964159	12.28124	11.1116	12.86606	10.52677

Examples:

<https://www.bankofengland.co.uk/-/media/boe/files/working-paper/2017/machine-learning-at-central-banks.pdf?la=en&hash=EF5C4AC6E7D7BDC1D68A4BD865EEF3D7EE5D7806>