



DevTreks –social budgeting that improves lives and livelihoods

## CTA Algorithm 1 Examples

**Last Updated: September 13, 2018; First Released: January 09, 2015**

**Author: Kevin Boyle, President, DevTreks**

**Version: DevTreks 2.1.6**

### A. Algorithm 1 Introduction

The sibling reference, *Conservation Technology Assessment* (CTA), introduces the background numerical techniques for completing CTAs. This reference introduces examples of CTAs completed using Algorithm 1, MathNet and System.Math Algorithms (1\*).

Algorithm 1 is a front end to custom algorithms developed by DevTreks. All of these algorithms employ System.Math and MathNet mathematical libraries (refer to the references). The goal of most algorithms will be to produce confidence intervals for an Indicator's QTM, QTL, and QTU and a Score's ScoreM, ScoreL, and ScoreU, properties (most likely, low, and high estimates). The advantage to algorithms using this library include: 1) fine-tuned control by developers over how an algorithm works and how results get displayed, 2) compiled binary code which can be optimized for cloud computing performance, and 3) ease of stepping outside the boundaries of conventional statistical libraries.

The following subalgorithms are currently available with this algorithm. **Appendix A** gives examples for each subalgorithm.

**subalgorithm1: Monte Carlo:** Example 1a and the examples in the *Resource Stock Calculation* reference use this option to introduce basic risk analysis. This algorithm uses the distribution of QT (QT, QTD1, and QTD2) or TEXT datasets with a mathematical library to produce QTM, QTL, and QTU. Unlike a full probabilistic risk analysis, this algorithm does not account for correlations between indicators.

**subalgorithm2: Normal Copula:** Example 1b uses this option to carry out probabilistic risk analysis, accounting for correlations between indicators.



DevTreks –social budgeting that improves lives and livelihoods

**subalgorithm3: Eigen Copula with Normal Distribution:** Examples 1c, 1d, and 1e use this option to carry out probabilistic risk analysis, accounting for correlations between indicators.

**subalgorithm4: Eigen Copula with Uniform Distribution.** Example 1f uses this option to demonstrate a slight variation of subalgorithm3. The only difference is the use of a Normal distribution (sub3) or Uniform distribution (sub4) in the calculations. The resultant calculations have slight differences that might be significant for some circumstances (i.e. health and safety Indicators).

**subalgorithm5: Simulated Annealing.** Example 1g uses this option to introduce combinatorial optimization analysis.

**subalgorithm6: Regression.** Example 1h uses this option to introduce probabilistic statistics that employ regression analysis.

**subalgorithm7: Neural Network.** Example 1l uses this option to introduce prediction and classification analysis.

**subalgorithm8: Anova.** Example 1m uses this option to introduce probabilistic statistics that employ analysis of variance to analyze randomized experimental data.

**subalgorithm9: Disaster Risk Reduction (DRR):** The associated *Conservation Technology Assessment 2* tutorial demonstrates that this algorithm uses Disaster Risk Reduction algorithms to calculate confidence intervals for Benefit Cost Ratios and Cost Effectiveness Ratios. This algorithm focuses on measuring the direct monetary savings from disaster prevention interventions, especially those associated with climate change.

**subalgorithm10: Disaster Risk Index (DRI):** The associated *Conservation Technology Assessment 2* tutorial demonstrates that this algorithm uses Disaster Risk Index algorithms to calculate confidence intervals for Benefit Cost Ratios and Cost



DevTreks –social budgeting that improves lives and livelihoods

Effectiveness Ratios. These Indexes measure both the direct and indirect savings from disaster prevention interventions, especially those associated with climate change.

**subalgorithm11: Risk Management Index (RMI):** The associated *Conservation Technology Assessment 2* tutorial demonstrates that this algorithm uses Risk Management Index algorithms to calculate confidence intervals for Cost Effectiveness Ratios and Multi-Criteria Assessment Ratings. These Indexes measure a community's ability to manage disasters, especially those associated with climate change.

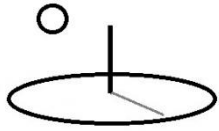
**subalgorithm12: Resiliency Index (RI):** The associated *Conservation Technology Assessment 2* tutorial demonstrates that this algorithm uses Resiliency Indexes algorithms to calculate confidence intervals for Cost Effectiveness Ratios and Multi-Criteria Assessment Ratings. These Indexes measure a community's ability to monitor and evaluate their disaster prevention goals.

**subalgorithm13, 14, 15, 16, and 17: Resource Conservation Accounting (RCA) Value Framework:** The Performance and Social Performance Analysis tutorials document that Version 2.1.0+ uses these algorithms in a RCA Value Framework to measure social performance.

## Calculator Patterns

Versions 2.1.4 and 2.1.6 upgraded the calculator patterns used by these algorithms. The primary pattern enforced in the upgrade is to place greater emphasis on using the Indicator.URL property to store data and/or scripts for the specific Indicator holding the URL. This promotes consistency with how calculations are run for the remaining algorithms (i.e. R and Python).

The legacy patterns of using a combination of the Score.DataURL and Score.JointDataURL for running joint calculations has been retained and simplified. These patterns allow all Indicator and Score data to be stored in 1 dataset and are useful when separate Indicator datasets are considered overkill or when the separate Indicators must be calculated together. The examples in



DevTreks –social budgeting that improves lives and livelihoods

Appendix A demonstrate that these patterns allow different combinations of Score and Indicator properties to be used to fill in Indicator and Score MathResults.

## Summary and Conclusion

This reference demonstrates how to use algorithms based on the System.Math and MathNet libraries to complete CTAs. CTAs may help people to reach decisions that improve their lives and livelihoods in sustainable ways.

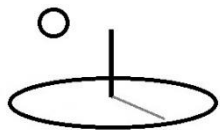
## Footnotes

1. Version 2.0.2 supports additional mathematical libraries discussed in the sibling CTA references. The CTA 2 reference, R algorithms, supports the use of R and Intel Math Kernel mathematical libraries. The CTA 3 reference, Python algorithms, supports the use of Python mathematical libraries. The CTA 4 reference, Azure Machine Learning algorithms, supports the use of AML mathematical libraries. Porting subalgorithms from 1 math library to another may not be particularly difficult. The Version 2.0.4 and 2.0.6 releases upgraded the Monitoring and Evaluation tools so that they can use these algorithms.
2. The confidence interval generated for this gamma distribution is too close to be useful for modeling the uncertainty of this Indicator. Networks should consider including experts in statistics, mathematics, or domain-specific fields such as disaster assessment, to provide uniform guidance to their clubs about how to use specific CTA algorithms (i.e. our role is demonstrate what you should be doing rather than what you are doing). For example, the gamma's shape and scale parameters can also be estimated using maximum likelihoods methods similar to the following:

$$\text{shape} = (1 / 4A) * (1 + (1 + (4A/3))^2)$$

$$\text{scale} = \text{mean}(x) / \text{shape}$$

where A, a transformed value for n observation, is calculated:



DevTreks –social budgeting that improves lives and livelihoods

$$A = \ln(\text{mean}(x) - (\text{sum of } \ln(x) / n))$$

3. The McCaffrey reference makes it clear that the algorithms employed are introductory examples. The neural network algorithm is a good example. The author points out ways that developers may want to take to improve the algorithm (i.e. stopping techniques). Several major information technology companies (i.e. Facebook, Google, and Microsoft) are starting to make their algorithms available as open source. These are logical replacements for several of the algorithms introduced in this reference.
4. Some M&E practitioners may legitimately disagree. The jobs and economic activity associated with private investments do have legitimate societal benefits –capitalism works. A fuller example is left to those practitioners or a future release.
5. These types of algorithms need the full time commitment of staff who specialize in thoroughly understanding their use and abuse. That’s not necessarily the role of software developers, but it is the responsibility of the “owners of the algorithms” who generally will be members and clubs or full time DevTreks, or DevTreks-like, staff. That is, staff who work in nonconventional institutions and work hard to “do it right”.

## References

Anderson, John; Harri, Ardian; Coble, Keith. Simulation from Mixed Marginal Distributions with Application to Whole-Farm Revenue Simulation. Journal of Agricultural Resource and Resource Economics. April, 2009.

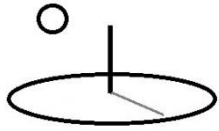
Brebbia, C.A. Risk Analysis VII. 2013 (last accessed on the web in December, 2014).

Azure Machine Learning (AML). Microsoft. 2015 (last accessed August, 2016 at <http://azure.microsoft.com/en-us/documentation/services/machine-learning/>)

MathNet. Last accessed April 24, 2018:

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

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



DevTreks –social budgeting that improves lives and livelihoods

McCaffrey, James. Microsoft Developers Network (MSDN) Magazine (various issues)

Mendenhall, William, Sincich, Terry. A Second Course in Business Statistics: Regression Analysis. Third Edition. Dellen Publishing Company. 1989

Piwcewicz, Bartosz. Assessment of Diversification Benefit in Insurance Portfolios. Institute of Actuaries of Australia, 2005

Studeman, A.H. Using Econometrics, a Practical Guide. 2<sup>nd</sup> edition. Harpers Collins Publishers. 1992

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)

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

<https://github.com/dotnet/standard>

## References Note

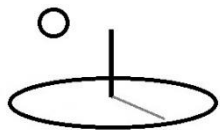
We try to use references that are open access or that do not charge fees.

## Improvements, Errors, and New Features

Please notify DevTreks (devtrekkers@gmail.com) if you find errors in these references. Also please let us know about suggested improvements or recommended new features.

## A video tutorial explaining this reference can be found at:

<https://www.devtreks.org/commontreks/preview/commons/resourcepack/Technology>  
Assessment 1/1526/none



DevTreks –social budgeting that improves lives and livelihoods

## **Appendix A. Algorithm 1 Examples.**

These datasets are owned by the Natural Resource Stock club in the GreenTreks network group (if testing on localhost, switch clubs). Some of these algorithms, such as the neural network algorithm, have been replaced by Machine Learning algorithms introduced in the Social Performance Analysis 3 reference. The latter reference also demonstrates more advanced algorithms than the regression and ANOVA algorithms introduced here. These algorithms will be retained as examples of “homegrown” algorithms that are built from scratch, as contrasted to relying on prebuilt algorithms in statistical packages, such as R and Python.

Version 2.1.6 upgraded security, with <http://localhost:5000> URLs being redirected, automatically, to <https://localhost:5001> URLs.

### **Example 1. Algorithm 1. Subalgorithm 1. Monte Carlo Simulation with Uncertain Net Benefits**

#### **URLs**

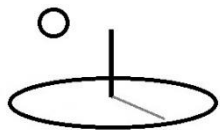
<https://www.devtreks.org/greentreks/preview/carbon/resourcepack/Conservation Technology Assessments Media/1534/none>

#### **Uncertain Output Net Benefits**

<https://www.devtreks.org/greentreks/preview/carbon/outputseries/NIST 451 Net Benefits/2141212685/none>

<http://localhost:5000/greentreks/preview/carbon/outputseries/NIST 5-4-1 Net Benefits/2141212696/none>

#### **Uncertain Total Benefits**



DevTreks –social budgeting that improves lives and livelihoods

[https://www.devtreks.org/greentreks/preview/carbon/outputseries/NIST 451 Total Benefits/2141212686/none](https://www.devtreks.org/greentreks/preview/carbon/outputseries/NIST%20451%20Total%20Benefits/2141212686/none)

[https://www.devtreks.org/greentreks/preview/carbon/input/NIST 451 Total Costs/2147397542/none](https://www.devtreks.org/greentreks/preview/carbon/input/NIST%20451%20Total%20Costs/2147397542/none)

[https://www.devtreks.org/greentreks/preview/carbon/component/NIST 451 Total Costs/2194/none](https://www.devtreks.org/greentreks/preview/carbon/component/NIST%20451%20Total%20Costs/2194/none)

[https://www.devtreks.org/greentreks/preview/carbon/outcome/NIST 451 Total Benefits/5766/none](https://www.devtreks.org/greentreks/preview/carbon/outcome/NIST%20451%20Total%20Benefits/5766/none)

[https://www.devtreks.org/greentreks/preview/carbon/investment/NIST 451 Net Benefits/429/none](https://www.devtreks.org/greentreks/preview/carbon/investment/NIST%20451%20Net%20Benefits/429/none)

[https://localhost:5000/greentreks/preview/carbon/investment/NIST 451 CTA/433/none](https://localhost:5000/greentreks/preview/carbon/investment/NIST%20451%20CTA/433/none)

[http://localhost:5000/greentreks/preview/carbon/outputseries/NIST 5-4-1 Total Benefits/2141212697/none](http://localhost:5000/greentreks/preview/carbon/outputseries/NIST%205-4-1%20Total%20Benefits/2141212697/none)

Uncertain Correlated Net Benefits

[https://www.devtreks.org/greentreks/preview/carbon/outputseries/NIST 451 Net Benefits, SubAlg3/2141212687/none](https://www.devtreks.org/greentreks/preview/carbon/outputseries/NIST%20451%20Net%20Benefits,%20SubAlg3/2141212687/none)

[http://localhost:5000/greentreks/preview/carbon/outputseries/NIST 5-4-1 Net Benefits Correlations/2141212698/none](http://localhost:5000/greentreks/preview/carbon/outputseries/NIST%205-4-1%20Net%20Benefits%20Correlations/2141212698/none)

Examples 1 to 5 in the introductory *Resource Stock Calculation* reference use algorithm1 and subalgorithm1 to demonstrate how to use Monte Carlo techniques to calculate uncertain emission and environmental performance indicators. Examples in the *M&E Calculation* reference





DevTreks –social budgeting that improves lives and livelihoods

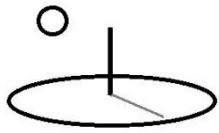
demonstrate these techniques for malnutrition project performance indicators. This example focuses on economic performance indicators.

The following example derives from Section 5.4.1, Storage Facility Simulation Example, found in the NIST (1988) reference. The reference introduces the example as follows:

A private investor wants to compute the Net Benefits (NB), Benefit Cost Ratio (BCR), and Adjusted Internal Rate of Return (AIRR) measures of worth to evaluate the economic merits of constructing small scale warehouse storage facilities for rent. ... Examples of uncertain inputs that might affect the profitability of a warehouse are rental receipts, operating costs, resale value of the facility at the end of its holding period, and construction costs.

Further explanations for these “measures of worth”, or Performance Measures, can be found in the *Performance Analysis* tutorial. In the context of CTA, this investment could be for any public goods purpose, such as carbon, energy, health, or water, conservation technologies. The difficulty of measuring the returns from investments in public goods will be addressed in related tutorials (i.e. the *CTA-Prevention and Social Performance Analysis* tutorials). This example demonstrates the following 3 techniques for conducting an economic evaluation of this capital investment:

1. **Uncertain Output Net Benefits:** This method calculates Net Benefits for 1 Output base element. The Net Benefit is a type of uncertain Output revenue. This technique is appropriate for quick, summary, economic evaluations. Investments that lose money should use Inputs and be treated as a type of uncertain Input cost.
2. **Uncertain Capital Budget Net Benefits:** This method calculates Net Benefits using 1 Input that calculates the uncertainty of the total costs and 1 Output that calculates the uncertainty of the total benefits. The Input and Output are added to a Capital Budget and a Resource Stock Totals analysis is used to calculate the uncertainty of the final Net



DevTreks –social budgeting that improves lives and livelihoods

Benefits of the investment. This technique is appropriate for formal, full, economic evaluations.

3. **Uncertain Output Net Benefits with Correlated Indicators:** This method is similar to Method 1 except that the Units Rented, Resale Value, and Operating Costs, Indicators in the analysis are correlated. Examples 1b to 1e should be completed before reviewing this method.

The following image compares the results of the 3 methods. The results for Method 2 may reflect slightly different Indicator properties –Method 1 and 2’s properties were fine-tuned after Method 2 was already run. Given the random samples that are used to generate these numbers, the results do not appear to be significantly different (which is an empirical question that can be further tested).

Warehouse Net Benefits (\$), 90% Confidence Interval				
	Method 1, subalgorithm1	Method 2, subalgorithm1	Method 3, subalgorithm3	Method 3, subalgorithm4
ScoreM	2,929,843.01	3,029,737	2,931,168.74	2,929,435.89
ScoreL	2,924,958.50	3,028,244	2,926,930.84	2,925,185.53
ScoreU	2,934,727.52	3,031,231	2,935,406.65	2,933,686.25
Net Benefits Q6M	2,930,127.99		2,930,327.08	2,926,844.90
Net Benefits Q6L	2,925,244.47		2,925,375.53	2,921,887.61
Net Benefits Q6U	2,935,011.51		2,935,278.63	2,931,802.20

### Method 1. Uncertain Output Net Benefits

The following Indicators have been added to 1 Output base element. The probability distributions for four of the Indicators can be found in the NIST reference. A fictitious distribution was used for the Net Benefit Indicator. Selected properties of each Indicator are highlighted. The confidence interval for these indicators (x%) is defined using the Score.ConfidenceLevel property.



DevTreks –social budgeting that improves lives and livelihoods

**Units Rented Indicator 1:** This Indicator uses a gamma distribution of the number of units rented to calculate uncertain revenues. The distribution has 564 (mean) and 12 (standard deviation). Selected properties include:

Q1 = 600 units

Q2 = 0.94 occupancy rate

Math Expression =  $I1.Q1 * I1.Q2$

QT = 564 units rented

Distribution Type: gamma. The following shape and scale parameters were derived from the method of moments estimations from the 1.3.6.6.11.Gamma Distribution in the US NIST Engineering Statistics Handbook. Footnote 2 discusses limitations with the resultant confidence interval.

QTD1: 2209 shape parameter =  $(\text{mean} / \text{sd})^2$

QTD2 = 3.917 the inverse scale parameter =  $1 / (\text{sd}^2 / \text{mean})$

QTM = 563.9099 units rented

QTL = 563.6759 lower x% ci

QTU = 564.1438 upper x% ci

**Unit Rental Income Indicator 2:** This Indicator is not uncertain. Selected properties include:

Q1 = 1200 rent per unit per year

Q2 = 10 years

Q3 = .0386 real discount rate (derived from  $3.86 / 100$ )

Math Expression =  $I2.Q1 * (((1 + I2.Q3)^{I2.Q2} - 1) / (I2.Q3 * ((1 + I2.Q3)^{I2.Q2})))$

QT and QTM = 9,801.2643 uniform present value of rent per unit over 10 years

QTM has to be manually added

**Resale Value Indicator 3:** This Indicator uses a normal distribution of the warehouse resale value to calculate uncertain revenues. Selected properties include:

Q1 = 1,980,000 resale value in 10 years

Q2 = 10 years

Q3 = .0386 real discount rate (derived from  $3.86 / 100$ )



DevTreks –social budgeting that improves lives and livelihoods

Math Expression =  $I3.Q1 * (1 / (1 + I3.Q3) ^ I3.Q2)$

QT = 1,350,547 present value of resale value

Distribution Type: normal

QTD1: 1,355,758 (mean)

QTD2 = 230,479 (standard deviation)

QTM = 1,356,235.5922 present resale value

QTL = 1,351,742.0740 lower x% ci

QTU = 1,360,729.1105 upper x% ci

**Operating Cost Indicator 4:** This Indicator uses a normal distribution of the operating costs to calculate uncertain costs. Selected properties include:

Q1 = 156,000 annual operating costs

Q2 = 10 years

Q3 = .0386 real discount rate (derived from  $3.86 / 100$ )

Math Expression =  $I4.Q1 * (((1 + I4.Q3)^{I4.Q2}) - 1) / (I4.Q3 * ((1 + I4.Q3)^{I4.Q2}))$

QT = 1,274,164.3621 uniform present value of operating costs

Distribution Type: normal

QTD1: 1,274,161 (mean)

QTD2 = 127,416 (standard deviation at 10% of mean)

QTM = 1,274,053.6717 operating costs

QTL = 1,271,559.6809 lower x% ci

QTU = 1,276,547.6625 upper x% ci

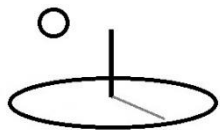
**Construction Costs Indicator 5:** This Indicator uses a lognormal distribution of the construction costs to calculate uncertain costs. The distribution has 1,800,000 (mean) and 180,000 (standard deviation). Selected properties include:

Q1 = 900,000 site preparation

Q2 = 900,000 construction costs

Math Expression =  $I5.Q1 + I5.Q2$

QT = 1,800,000.0000



DevTreks –social budgeting that improves lives and livelihoods

Distribution Type: lognormal

QTD1: 14.3983 = shape parameter =  $\text{LN}(\text{mean} / (1 + (\text{variance} / \text{mean}^2)^{0.5}))$  or  $\text{LN}(\text{mean}) - 0.5 * (\text{scale}^2)$  using the LN function of a calculator (not Excel LOG)

QTD2 = 0.00997 = scale parameter =  $\text{LN}((1 + (\text{variance} / \text{mean}^2))^{0.5})$

QTM = 1,799,956.0168 construction cost

QTL = 1,796,459.5882 lower x% ci

QTU = 1,803,452.4454 upper x% ci

**Net Benefit Indicator 6:** This indicator is used to subtract the cost Indicators from the benefit Indicators to calculate uncertain Net Benefits. It is also used to update the Output.Price to the calculated Net Benefits.

Q1 = 180,000 land purchase cost

Q2 = 700,000.0000 tax adjustment cost (for the TX variable in the NIST NB formula)

Math Expression =  $((I1.QTM * I2.QTM) + I3.QTM) - (I4.QTM + I5.QTM + I6.Q1 + I6.Q2)$

QT = 2,930,624.5261

Distribution Type: normal

QTD1: 2,925,000.0000 (mean)

QTD2 = 300,000 (standard deviation)

QTM = 2,931,071.7309 net benefits

QTL = 2,925,154.7017 lower x% ci

QTU = 2,936,988.7601 upper x% ci

BaseIO = benprice (updates the base element's Output.Price property with QTM)

The Score properties have been set to return the same results as the Net Benefits Indicator. The following image displays the Resource Stock calculated results for this Output.



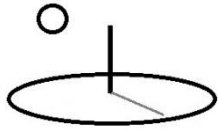
## Indicators

Math Expression: I6.QTM	Score Unit: net benefits
Score Amount: 2,928,702.0412	Score D1 Unit: mean
Score D1 Amount: 2,935,000.0000	Score D2 Unit: standard deviation
Score D2 Amount: 300,000.0000	Math Type: algorithm1
Distribution Type: normal	Score Most Unit: net benefits
Score Most Amount: 2,928,702.0412	Score Low Unit: lower 90% ci
Score Low Amount: 2,923,781.3996	Score High Unit: upper 90% ci
Score High Amount: 2,933,622.6828	Math Sub Type: subalgorithm1
Iterations: 10000	Random Seed: 5
Confid Int: 90	
Base IO: none	
Score Math Result: sampled descriptive statistics N,Total,Mean,Median,StdDev,Var,Min,Max 10000, 29287020411.8005, 2928702.0412, 2929586.9058, 298220.7041, 88935588383.0786, 1824176.0571, 4006434.3390, sampled cumulative density function 0.00,0.10,0.20,0.30,0.40,0.50,0.60,0.70,0.80,0.90,1.00 1824176.0571,2544896.2472,2676505.4723,2770676.2352,2854843.0512,2929603.8838,3004	
<b>Indic 1 Name:</b> Units Rented	Label: UR
Date: 05/22/2015	Rel Label:
Math Type: algorithm1	Dist Type: gamma
Q1 Amount: 600.0000	Q1 Unit: total units
Q2 Amount: 0.9400	Q2 Unit: occupancy rate
Q3 Amount: 0.0000	Q3 Unit: none
Q4 Amount: 0.0000	Q4 Unit: none
Q5 Amount: 0.0000	Q5 Unit: none
Math Express: I1.Q1 * I1.Q2	Math Operator: equalto
QT Amount: 564.0000	QT Unit: units rented
QT D1 Amount: 2,209.0000	QT D1 Unit: shape
QT D2 Amount: 3.9170	QT D2 Unit: scale
QT Most Amount: 563.9687	QT Most Unit: units rented
QT Low Amount: 563.7721	QT Low Unit: lower 90% ci
QT High Amount: 564.1653	QT High Unit: upper 90% ci
Math Sub Type: subalgorithm1	Base IO: none
Indic 1 Description: This indicator is used in a CTA tutorial.	
<b>Indic 2 Name:</b> Unit Rental Price	Label: NIST2
Date: 05/22/2015	Rel Label:
Math Type: none	Type: none
Q1 Amount: 1,200.0000	Q1 Unit: rent per unit per year
Q2 Amount: 10.0000	Q2 Unit: years
Q3 Amount: 0.0386	Q3 Unit: real discount rate
Q4 Amount: 0.0000	Q4 Unit: none
Q5 Amount: 0.0000	Q5 Unit: none
Math Express: I2.Q1 * (((1 + I2.Q3)^I2.Q2) - 1) / (I2.Q3 * ((1 + I2.Q3)^I2.Q2)))	Math Operator: equalto
QT Amount: 9,801.2643	QT Unit: upv rent
QT D1 Amount: 0.0000	QT D1 Unit: low
QT D2 Amount: 0.0000	QT D2 Unit: high
QT Most Amount: 9,801.2643	QT Most Unit: upv rent
QT Low Amount: 0.0000	QT Low Unit: lower 90% ci
QT High Amount: 0.0000	QT High Unit: upper 90% ci



## DevTreks –social budgeting that improves lives and livelihoods

<b>Indic 3 Name:</b> Resale Price	Label: RV
Date: 05/22/2015	Rel Label:
Math Type: algorithm1	Type: normal
Q1 Amount: 1,980,000.0000	Q1 Unit: resale in 10 years
Q2 Amount: 10.0000	Q2 Unit: years
Q3 Amount: 0.0386	Q3 Unit: real discount rate
Q4 Amount: 0.0000	Q4 Unit: none
Q5 Amount: 0.0000	Q5 Unit: none
Math Express: $I3.Q1 * (1 / (1 + I3.Q3) ^ I3.Q2)$	Math Operator: equalto
QT Amount: 1,355,757.4752	QT Unit: upv rent
QT D1 Amount: 1,355,758.0000	QT D1 Unit: mean
QT D2 Amount: 230,479.0000	QT D2 Unit: sd
QT Most Amount: 1,350,919.5092	QT Most Unit: pv resale value
QT Low Amount: 1,347,139.1607	QT Low Unit: lower 90% ci
QT High Amount: 1,354,699.8577	QT High Unit: upper 90% ci
Math Sub Type: subalgorithm1	Base IO: none
Indic 3 Description: This indicator is used in a CTA tutorial.	
<b>Indic 4 Name:</b> Operating Costs	Label: OC
Date: 05/22/2015	Rel Label:
Math Type: algorithm1	Type: normal
Q1 Amount: 156,000.0000	Q1 Unit: operating costs
Q2 Amount: 10.0000	Q2 Unit: years
Q3 Amount: 0.0386	Q3 Unit: discount rate
Q4 Amount: 0.0000	Q4 Unit: none
Q5 Amount: 0.0000	Q5 Unit: none
Math Express: $I4.Q1 * (((1 + I4.Q3)^{I4.Q2}) - 1) / (I4.Q3 * ((1 + I4.Q3)^{I4.Q2}))$	Math Operator: equalto
QT Amount: 1,274,164.3621	QT Unit: upv op costs
QT D1 Amount: 1,274,161.0000	QT D1 Unit: mean
QT D2 Amount: 127,416.0000	QT D2 Unit: sd
QT Most Amount: 1,271,486.1309	QT Most Unit: upv op costs
QT Low Amount: 1,269,396.2360	QT Low Unit: lower 90% ci
QT High Amount: 1,273,576.0258	QT High Unit: upper 90% ci
Math Sub Type: subalgorithm1	Base IO: none
Indic 4 Description: This indicator is used in a CTA tutorial.	
<b>Indic 5 Name:</b> Construction Costs	Label: NIST5
Date: 05/22/2015	Rel Label:
Math Type: algorithm1	Type: lognormal
Q1 Amount: 900,000.0000	Q1 Unit: site prep
Q2 Amount: 900,000.0000	Q2 Unit: construction cost
Q3 Amount: 0.0000	Q3 Unit: none
Q4 Amount: 0.0000	Q4 Unit: none
Q5 Amount: 0.0000	Q5 Unit: none
Math Express: $I5.Q1 + I5.Q2$	Math Operator: equalto
QT Amount: 1,800,000.0000	QT Unit: construction cost
QT D1 Amount: 14.3983	QT D1 Unit: shape
QT D2 Amount: 0.0100	QT D2 Unit: scale
QT Most Amount: 1,790,739.9397	QT Most Unit: mean construction cost
QT Low Amount: 1,790,446.2597	QT Low Unit: lower 90% ci
QT High Amount: 1,791,033.6197	QT High Unit: upper 90% ci
Math Sub Type: subalgorithm1	Base IO: none



DevTreks –social budgeting that improves lives and livelihoods

<b>Indic 6 Name:</b> Net Benefits	<b>Label:</b> NIST6
<b>Date:</b> 05/22/2015	<b>Rel Label:</b>
<b>Math Type:</b> algorithm1	<b>Type:</b> normal
<b>Q1 Amount:</b> 180,000.0000	<b>Q1 Unit:</b> land purchase cost
<b>Q2 Amount:</b> 700,000.0000	<b>Q2 Unit:</b> tax adjustment cost
<b>Q3 Amount:</b> 0.0000	<b>Q3 Unit:</b> none
<b>Q4 Amount:</b> 0.0000	<b>Q4 Unit:</b> none
<b>Q5 Amount:</b> 0.0000	<b>Q5 Unit:</b> none
<b>Math Express:</b> ((I1.QTM * I2.QTM) + I3.QTM) - Math Operator: equalto	
<b>(I4.QTM + I5.QTM + I6.Q1 + I6.Q2)</b>	
<b>QT Amount:</b> 2,936,299.7242	<b>QT Unit:</b> net benefits
<b>QT D1 Amount:</b> 2,935,000.0000	<b>QT D1 Unit:</b> mean
<b>QT D2 Amount:</b> 300,000.0000	<b>QT D2 Unit:</b> sd
<b>QT Most Amount:</b> 2,928,702.0412	<b>QT Most Unit:</b> net benefits
<b>QT Low Amount:</b> 2,923,781.3996	<b>QT Low Unit:</b> lower 90% ci
<b>QT High Amount:</b> 2,933,622.6828	<b>QT High Unit:</b> upper 90% ci
<b>Math Sub Type:</b> subalgorithm1	<b>Base IO:</b> benprice
<b>Indic 6 Description:</b> This indicator is used in a CTA tutorial.	

[Feedback About carbon/outputseries/NIST 451 Net Benefits/2141212685/none](#)

**Search IRIs:**

[https://www.devtreks.org/greentreks/linkedviews/carbon/outputseries/NIST 451 Net  
benefits/2141212685/none](https://www.devtreks.org/greentreks/linkedviews/carbon/outputseries/NIST%20451%20Net%20Benefits/2141212685/none)

The following partial image of the equivalent M&E calculated results shows that the Score is treated as just another Indicator located in the zero index, or Indicator 0, position of the collection of Indicators.





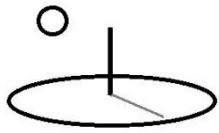
Output Group: CTA Output Examples

Output : NIST 5-4-1 Net Benefits

Output Series : NIST 5-4-1 Net Benefits

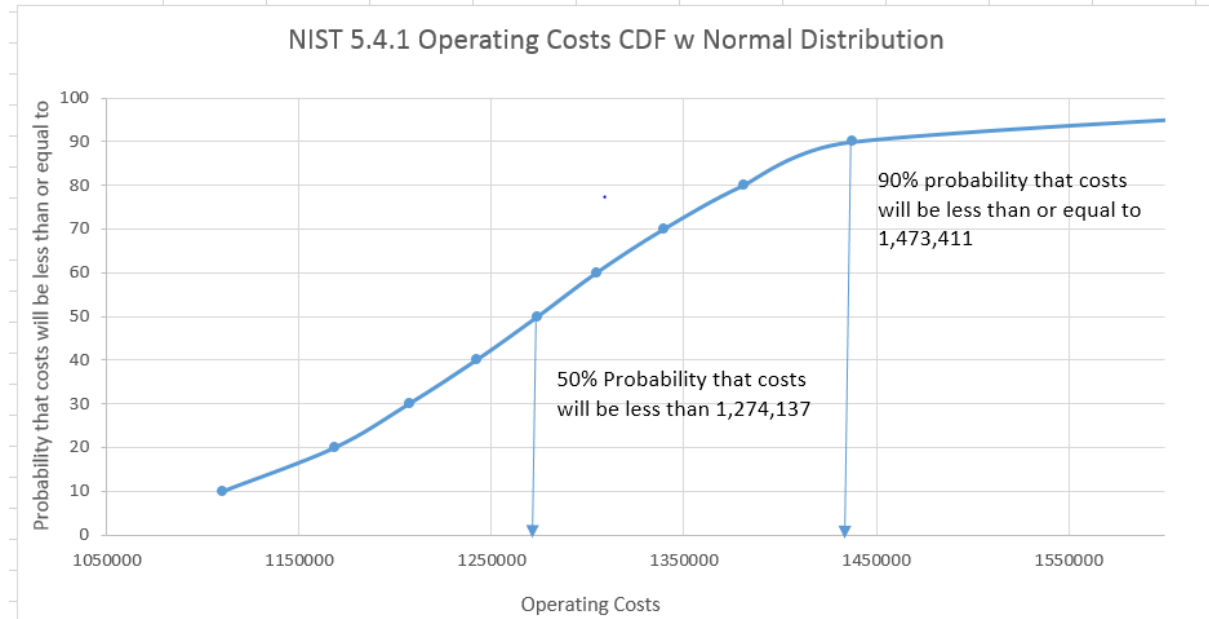
Indicators Details			
<b>Indic 0 Name:</b> Net Benefits Label: NIST0 Date: 11/15/2016 Rel Label: NIST6 Math Type: algorithm1 Dist Type: normal Math Sub Type: subalgorithm1 Base IO: none Math Express: I6.QTM Math Operator: equalto QT Amount: 2,924,717.6328 QT Unit: net benefits QT D1 Amount: 2,925,000.0000 QT D1 Unit: mean QT D2 Amount: 300,000.0000 QT D2 Unit: sd QT Most Amount: 2,924,717.6328 QT Most Unit: net benefits QT Low Amount: 2,908,939.1073 QT Low Unit: lower 90 % ci QT High Amount: 2,940,496.1583 QT High Unit: upper 90 % ci Score Math Result: sampled descriptive statistics N,Total,Mean,Median,StdDev,Var,Min,Max 1000, 2924717632.7501, 2924717.6328, 2916365.6379, 302400.4765, 91446048201.4029, 1886976.2069, 3783973.9671, sampled cumulative density function 0.00,0.10,0.20,0.30,0.40,0.50,0.60,0.70,0.80,0.90,1.00 1886976.2069,2549071.9518,2671646.9356,2758945.9445		I5.Q2 QT Amount: 1,800,000.0000 QT Unit: construction cost QT D1 Amount: 14.3983 QT D1 Unit: shape QT D2 Amount: 0.0100 QT D2 Unit: scale QT Most Amount: 1,791,101.4709 QT Most Unit: construction cost QT Low Amount: 1,790,159.5532 QT Low Unit: lower 90 % ci QT High Amount: 1,792,043.3886 QT High Unit: upper 90 % ci Indic 5 Description: This example supports the M and E tutorials.	
<b>Indic 1 Name:</b> Units Label: NIST1 Rented Date: 11/15/2016 Rel Label: none Math Type: algorithm1 Dist Type: gamma Math Sub Type: subalgorithm1 Base IO: none Q1 Amount: 600.0000 Q1 Unit: units Q2 Amount: 0.9400 Q2 Unit: occupancy rate Q3 Amount: 0.0000 Q3 Unit: none Q4 Amount: 0.0000 Q4 Unit: none Q5 Amount: 0.0000 Q5 Unit: none Math Express: I1.Q1 * I1.Q2 Math Operator: equalto QT Amount: 564.0000 QT Unit: units rented		<b>Indic 6 Name:</b> Net Benefits Label: NIST6 Date: 11/15/2016 Rel Label: none Math Type: algorithm1 Dist Type: normal Math Sub Type: subalgorithm1 Base IO: benprice Q1 Amount: 180,000.0000 Q1 Unit: land purchase cost Q2 Amount: 700,000.0000 Q2 Unit: tax adjustment Q3 Amount: 0.0000 Q3 Unit: none Q4 Amount: 0.0000 Q4 Unit: none Q5 Amount: 0.0000 Q5 Unit: none Math Express: ((I1.QTM * I2.QTM) + I3.QTM) - (I4.QTM + I5.QTM + I6.Q1 + I6.Q2) Math Operator: equalto QT Amount: 2,944,608.9012 QT Unit: net benefits QT D1 Amount: 2,925,000.0000 QT D1 Unit: mean QT D2 Amount: 300,000.0000 QT D2 Unit: sd QT Most Amount: 2,924,717.6328 QT Most Unit: net benefits QT Low Amount: 2,908,939.1073 QT Low Unit: lower 90 % ci QT High Amount: 2,940,496.1583 QT High Unit: upper 90 % ci Indic 6 Description: This example supports the M and E tutorials.	
		<a href="#">Feedback About carbon/outputseries/NIST 5-4-1 Net Benefits/2141212696/none</a>	

The following image is the cumulative density function (CDF) for the Operating Cost Indicator.

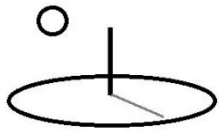


DevTreks –social budgeting that improves lives and livelihoods

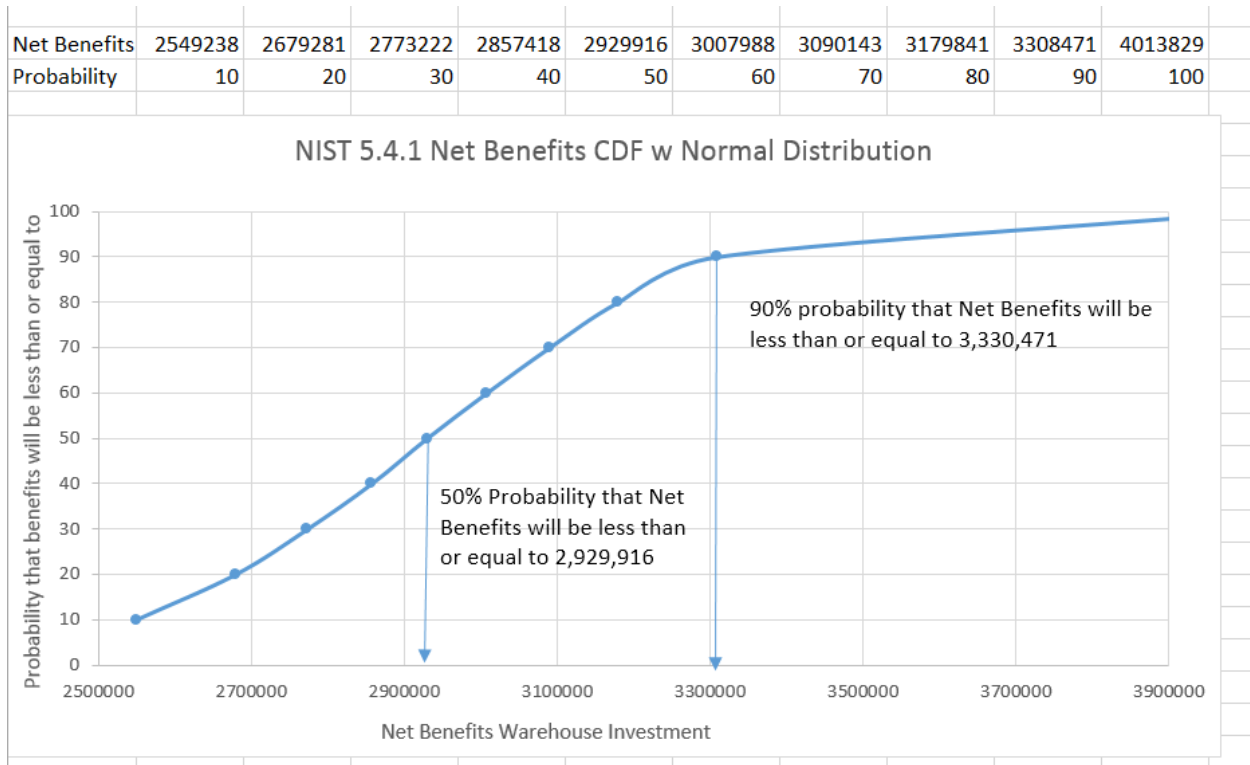
Operating Costs	1110455	1168574	1207137	1242136	1274137	1304590	1339595	1380912	1437411	1797417
Probability	10	20	30	40	50	60	70	80	90	100



The following image is the cumulative density function for the Net Benefits Indicator.



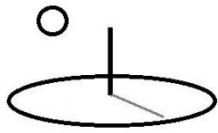
DevTreks –social budgeting that improves lives and livelihoods



The following image displays the updated Output.Price for this Output. The price was automatically updated by the calculator.

Input and Output Amounts can also be updated automatically by the calculator, but after being added to an Operation, Component, Outcome, Operating Budget, or Capital Budget, their Amounts are not automatically updated to the base Input or Output Amounts. Unlike prices, Input and Output quantities must be manually adjusted in budgets (or Operations, Outcomes, and Components) to define a specific technology. That’s why most DevTreks references recommend using *unit* Inputs and Outputs.

Although the Monitoring and Evaluation Indicators and Scores include the BaseIO property, this version does not automatically update the underlying base element properties. The current thinking is that the Resource Stock calculators should be used for that purpose because they



DevTreks –social budgeting that improves lives and livelihoods

aggregate their results in the same manner as base element costs and benefits –accumulating data from descendants to ancestors.

SubmitCancelClose

NIST 5-4-1 Net Benefits

Output Series +0

— NIST 5-4-1 Net Benefits

NIST 5-4-1 Net Benefits

D

U

Date Changed:

5/21/2015

Label:

BCA01

Date:

5/21/2015

Output Amount:

1.0000

Output Price:

2931071.7300

Output Unit:

each

Description

This example is used in a CTA tutorial. It comes from section 5.4.1 of the NIST publication Techniques for Treating Uncertainty and Risk in the Economic Evaluation of Building Investments.

Edit Linked Views

Views

## Method 2. Uncertain Capital Budget Net Benefits for Resource Stock calculations

The cost Indicators used in Method 1 have been added to an Input and the benefit Indicators to an Output. An additional Indicator was added to each base element to calculate Total Costs or Total Benefits. The Total Costs Indicator was used to update the Input.CAPPrice and Total



DevTreks –social budgeting that improves lives and livelihoods

Benefits Indicator updated the Output.Price. The Input and Output have been added to a Component and Outcome and those base elements have been added to a Capital Budget.

The following image displays the Resource Stock Totals Analysis of this capital investment. The Total Benefits Indicator and Total Costs Indicator can be used to communicate the probability of these uncertain benefits and costs.

The Score for the Input was given a negative number equal to Total Costs, while the Score for the Output had a positive number equal to Total Benefits, resulting in a Net Benefits Score as well. The uncertainty of that Performance Measure has been modeled independently and the results can also be communicated to decision makers. Example 1j demonstrates how to include emissions indicators in this type of analysis to analyze an uncertain Cost Effectiveness Performance Measure. The latter measure is often used in HTAs (substituting QALY or DALY Indicators for the emissions Indicators –see the Ireland HIQA reference and the Social Performance Analysis Examples reference).



## DevTreks –social budgeting that improves lives and livelihoods

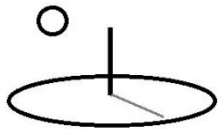
Investment : NIST 451 CTA									
Score	Score Unit	Score D1 Amount	Score D1 Unit	Score D2 Amount	Score D2 Unit	Iterations			
Score Most Amount	Score Most Unit	Score Low Amount	Score Low Unit	Score High Amount	Score High Unit	Distribution Type	Math Type	Math Sub Type	Observations
2,896,481.7095	total benefits	3,040,000.0000	low estimate	1,150,000.0000	high estimate	10000			
3,040,937.6198	total benefits	3,018,464.7297	lower 95% ci	3,063,410.5097	upper 95% ci	normal	algorithm1	subalgorithm1	2.0
I4.Q6M									
sampled descriptive statistics N,Total,Mean,Median,StdDev,Var,Min,Max 10000, 69984948640.5539, 6998494.8641, 6995601.4974, 749047.0208, 561071439336.6800, 4289285.2400, 9686392.5775, sampled cumulative density function 0.10,0.20,0.30,0.40,0.50,0.60,0.70,0.80,0.90,1.00 6041287.2274,6363527.1883,6590414.7756,6801062.4842,6995703.3931,7184066.4201,7384374.0315,7626573.5973,7970107.3090,9686392.5775									
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
Q6 Amount	Q6 Unit	Q6 D1 Amount	Q6 D2 Amount	Q6 Most Amount	Q6 Most Unit	Q6 Low Amount	Q6 Low Unit	Q6 High Amount	Q6 High Unit
<b>Operating Costs (1.0)</b>		NIST451A	05/22/2015		algorithm1	normal	1.0	equalto	subalgorithm1
156,000.0000	annual operating costs	10.0000	years	0.0386	real discount rate	0.0000	none	0.0000	none
1,274,164.3621	pv operating costs	1,274,161.0000	127,500.0000	1,272,333.0995	operating costs	1,269,811.3114	lower 95% ci	1,274,854.8876	upper 95% ci
This indicator is used in a CTA tutorial.									
I1.Q1 * (((1 + I1.Q3)^I1.Q2) - 1) / (I1.Q3 * ((1 + I1.Q3)^I1.Q2)))									
<b>Construction Costs (1.0)</b>		NIST451B	05/22/2015		algorithm1	lognormal	none	equalto	subalgorithm1
900,000.0000	site prep	900,000.0000	construction cost	0.0000	none	0.0000	none	0.0000	none
1,800,000.0000	construction costs	14.3983	0.0997	1,800,901.5597	construction costs	1,797,375.1620	lower 95% ci	1,804,427.9573	upper 95% ci
This indicator is used in a CTA tutorial.									
I2.Q1 + I2.Q2									
<b>Total Costs (1.0)</b>		NIST451C	05/22/2015		algorithm1	normal	caprice	equalto	subalgorithm1
180,000.0000	land purchase costs	700,000.0000	tax adjustment cost	0.0000	none	0.0000	none	0.0000	none
3,953,234.6592	total costs	3,960,000.0000	400,000.0000	3,960,831.6285	total costs	3,953,091.4465	lower 95% ci	3,968,571.8105	upper 95% ci
This indicator is used in a CTA tutorial.									
I1.Q6M + I2.Q6M + I3.Q1 + I3.Q2									



## DevTreks –social budgeting that improves lives and livelihoods

<b>Units Rented (1.0)</b>		NIST451D	05/22/2015		algorithm1	gamma	none	equalto	subalgorithm1
600.0000	total units	0.9400	percent units rented	0.0000	none	0.0000	none	0.0000	none
564.0000	units rented	2,209.0000	3.9170	563.9414	units rented	563.7060	lower 95% ci	564.1768	upper 95% ci
This indicator is used in a CTA tutorial.									
I1.Q1 * I1.Q2									
<b>Unit Rental Price (1.0)</b>		NIST451E	05/22/2015		none	none	none	equalto	none
1,200.0000	rent per unit per year	10.0000	years	0.0386	real discount rate	0.0000	none	0.0000	none
9,801.2643	present value unit rent	0.0000	0.0000	9,801.2643	present value unit rent	0.0000	none	0.0000	none
This indicator is used in a CTA tutorial.									
(I2.Q1 * (((1 + I2.Q3)*I2.Q2) - 1) / (I2.Q3 * ((1 + I2.Q3)*I2.Q2))))									
<b>Resale Price (1.0)</b>		NIST451F	05/22/2015		algorithm1	normal	none	equalto	subalgorithm1
1,980,000.0000	resale in 10 years	10.0000	years	0.0386	real discount rate	0.0000	none	0.0000	none
1,355,757.4752	present value resale price	1,355,758.0000	230,479.0000	1,358,370.1295	present value resale price	1,353,838.6961	lower 95% ci	1,362,901.5629	upper 95% ci
This indicator is used in a CTA tutorial.									
I3.Q1 * (1 / (1 + I3.Q3) ^ I3.Q2)									
<b>Total Benefits (1.0)</b>		NIST451G	05/22/2015		algorithm1	normal	benprice	equalto	subalgorithm1
0.0000	none	0.0000	none	0.0000	none	0.0000	none	0.0000	none
6,885,708.8406	total benefits	6,850,000.0000	700,000.0000	6,857,313.3380	total benefits	6,843,532.7469	lower 95% ci	6,871,093.9292	upper 95% ci
This indicator is used in a CTA tutorial.									
(I1.Q6M * I2.Q6M) + I3.Q6M									
Time Period : 2015 NIST 481 Example									

The following image displays the Net Present Value analysis of this budget. In this example, they don't equal the Net Benefits in the Score because the Inputs and Outputs in the NPV calculations have discounted interest added to them. The NPV totals and nets should also be reported to decision makers.



DevTreks –social budgeting that improves lives and livelihoods

Investment Group : CTA Tests

+ Investment Group Details

Investment : NIST 451 CTA

+ Investment Details

Total Ben : 6,940,801.66	Ann Ben : 6,940,801.66
Total OC Cost : 0.00	Ann OC Cost : 0.00
Net OC Returns : 6,940,801.66	Ann Net OC Returns : 6,940,801.66
Total AOH Cost : 0.00	Ann AOH Cost : 0.00
Net AOH Returns : 6,940,801.66	Ann Net AOH Returns : 6,940,801.66
Total CAP Cost : 4,009,055.06	Ann CAP Cost : 4,009,055.06
Net Returns : 2,931,746.60	Ann Net Returns : 2,931,746.60
Incent Ben : 6,940,801.66	Ann Incent Ben : 6,940,801.66
Incent Cost : 4,009,055.06	Ann Incent Cost : 4,009,055.06
Net Incent Cost : 2,931,746.60	Net Ann Incent Return : 2,931,746.60
Equiv Ann Ann : 2931746.60	

+ Time Period : 2015 NIST 451 Example

[Feedback About carbon/investment/NIST 451 CTA/433/none](#)

## Method 2. Uncertain Capital Budget Net Benefits for M &E calculations

This warehouse investment may not be a good example demonstrating full M&E analysis because, as the M&E references demonstrate, these analyses are typically carried out in the context of public, rather than private, investments (4\*). The following image, for a Health and Sanitation civil engineering project taken from the M&E Introduction reference, is a more typical example.





EXAMPLE 1: WATER AND SANITATION PROJECT			
OBJECTIVES (What you want to achieve)	INDICATORS (How to measure change)	MEANS OF VERIFICATION (Where & how to get information)	ASSUMPTIONS (What else to be aware of)
<b>Goal:</b> Reduce death and illness related to Water and Sanitation related diseases in the targeted communities	G1 % (percentage) reduction in water and sanitation related diseases among target population G2 % of children under 36 months with diarrhoea in the last two weeks	Ministry of Health / WHO statistics Records from village clinics	
<b>Outcome 1</b> Improved access to and use of sustainable sources of safe water in target communities	1a % of people in the target communities using minimum 25L of safe water per day 1b % of targeted households with access to an functional water source 1c % of water points managed by local WatSan committees 1d # hours spent by women in fetching water daily	1a,b,d Household survey 1c Key informant interviews with WatSan committee members	Civil war / hostilities do not return Improved access to clinical health facilities
<b>Outputs</b> 1.1 Community water points constructed or rehabilitated	1.1a # (number) of water points constructed to national standard (140) 1.1ab% of water handpumps rehabilitated to national standard (35)	"Community Facility Inspection" field report	Low rainfall does not limit overall water supply.
1.2 Community management of water points is improved	1.2a # of communities with a WatSan committee established 1.2b # of WatSan committees with technicians trained to perform basic maintenance on water points 1.2c % of WatSan committees collecting adequate charges to maintain the water points	1.2a Household survey Key informant interviews with WatSan committee members	No major disputes or conflicts within the community
<b>Outcome 2</b> Improved access to and use of sustainable sanitation facilities among targeted communities	2a % of people in the target communities using latrines on a daily basis 2b % of targeted households with access to functional latrines meeting national standard 2c % of latrines managed by local WatSan committees	2a,b Household survey 2c Key informant interviews with WatSan committee members	Civil war / hostilities do not return
<b>Outputs</b> 2.1 Sanitation facilities constructed	2.1a # of fully functioning household latrines constructed (3,500)	"Community Facility Inspection" field report	Flooding or other environmental problems do not affect sanitation facilities

EXAMPLE 1: WATER AND SANITATION PROJECT			
OBJECTIVES (What you want to achieve)	INDICATORS (How to measure change)	MEANS OF VERIFICATION (Where & how to get information)	ASSUMPTIONS (What else to be aware of)
2.2 Sanitation facility use is promoted	2.2a # of demonstration toilets constructed (25) 2.2b # of awareness session on use of latrines (25) 2.2c # of people reached by sanitation promotion activities (2000)	"Community Facility Inspection" field report	
2.3 Community management of sanitation facilities is improved	See also indicator 1.2a 2.3a # of community WatSan committees with technicians trained to perform basic maintenance	Key informant interviews with WatSan committee members	No major disputes or conflicts within the community

In order to replicate the results of the warehouse investment contained in a full Capital Budget Resource Stock Analysis, new M&E Indicators would have to be added to the Input, Output, Outcome, Component, Time Period, and Investment, base elements. Those Indicators would then replicate the cumulative totals contained in the Stock Analysis. Given that Resource Stock Totals Analysis generates those numbers automatically, that's overkill for most analyses. Instead, as the previous image shows, M&E should be carried out in the manner it's supposed to be carried out. For example, Time Period M&E Indicators can be calculated that account for the *impact* of the



DevTreks –social budgeting that improves lives and livelihoods

investment –has money been spent effectively in improving the lives and livelihoods of the intended beneficiaries? Both sets of Indicators complement one another and should be used together to improve decision making.

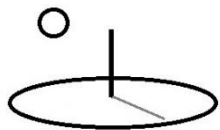
Given that the M&E Introduction reference already includes a complete example of a public goods investment in a malnutrition project, this reference will defer, at this time, from providing another complete example.

### **Method 3. Uncertain Output Net Benefits with Correlated Indicators**

Please review examples 1b to 1e before reviewing this method. Example 1e demonstrates how to calculate Scores using correlated Indicator sample observations. That technique will be used with correlated Units Rented, Resale Value, and Operating Costs, Indicators. This method is probably closest to the NIST example because the correlated Indicator sampled data is used to set the Score, or Net Benefits, Performance Measure. Method 2 can't be used because Indicator calculations can only be run for 1 base element at a time –Inputs and Outputs can't be calculated jointly yet.

The only properties different than Method 1 are the following Resource Stock Score properties:

- **Score Math Type and Math Sub Type:** algorithm 1 and subalgorithm3, Eigen Copula with Normal Distribution
- **Stock Joint Data URL or M&E Score URL:** The following fictitious Pearson correlation matrix has been uploaded to a Resource as a csv TEXT file and that URL has been added to this property. The NIST reference did not include this data. The only logic for these specific numbers is that the reference mentioned that these Indicators might have a positive correlation. The basic logic is that when the number of units rented increases, operating costs increase. Resale value increases when the units are maintained well, that is, when operating costs increase.



DevTreks –social budgeting that improves lives and livelihoods

pearson

UR,RV,OC

1,0.5,0.5

0.5,1,0.5

0.5,0.5,1

- **Score Distribution Type:** Set to none. The correlated indicator distributions are used to fill in the final Scores.
- **ScoreD1 and ScoreD2:** Both properties are set to zero. The correlated indicator distributions are used to fill in the final Scores.
- **Score Math Expression:** The following Score Math Expression is used to generate a Score for each row in the random sample data matrix. Only the 3 correlated Indicators with QT are in the sample matrix.

$$((I1.QT * I2.QTM) + I3.QT) - (I4.QT + I5.QTM + I6.Q1 + I6.Q2)$$

Example 1e mentions that the remaining non-correlated Indicators will be the same value for each row in the sample matrix (i.e. I2.QTM, I5.QTM, I6.Q1, I6.Q2). The QTM's derive from calculations that are run for the non-correlated Indicators prior to using them in this Math Expression.

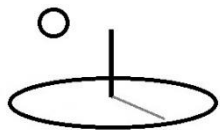
- **Confidence Interval:** 90 (this is new property in version 1.8.6)
- **ScoreM, ScoreL, and ScoreU:** Mean of Score with upper and lower 90% confidence intervals. The indicator distributions used 10,000 iterations, so 10,000 Scores were calculated and that vector generates these final results.

The following images for a Resource Stock calculation demonstrate that the Score properties for Method 3 follow the techniques explained in Example 1e for calculating correlated indicators. The results will be more meaningful when real correlation matrix numbers, or observational data sets, are used.



DevTreks –social budgeting that improves lives and livelihoods

<b>Score Math Expression</b> ((I1.QT * I2.QTM) + I3.QT) - (I4.QT + I5.QTM + I6.Q1 + I6.Q2)		<b>Score Math Result</b> original pearson DenseMatrix 3x3-Double 1 0.5 0.5 0.5 1 0.5 0.5 0.5 1 new pearson DenseMatrix 3x3-Double 1 0.506321 0.509766 0.506321 1 0.494718 0.509766 0.494718 1	
<b>Score</b> 2,742,368.1831	<b>Score Unit</b> net benefits		
<b>Score D1</b> 0.0000	<b>Score D1 Unit</b> low		
<b>Score D2</b> 0.0000	<b>Score D2 Unit</b> high		
<b>Score Dist Type</b> normal	<b>Iterations</b> 10000		
<b>Confidence Interval</b> 90	<b>Random Seed</b> 5		
<b>Score BaseIO</b> none			
<b>Score Most Likely</b> 2,936,299.5237	<b>Score Most Unit</b> net benefits		
<b>Score Low Estimate</b> 2,932,077.6499	<b>Score Low Unit</b> lower 90 % ci		
<b>Score High Estimate</b> 2,940,521.3975	<b>Score High Unit</b> upper 90 % ci		
<b>Score Math Type</b> algorithm1	<b>Score Math Sub Type</b> subalgorithm3		
<b>Score Math Result</b> original pearson DenseMatrix 3x3-Double		<b>Score Math Result</b> original pearson DenseMatrix 3x3-Double 1 0.506321 0.509766 0.506321 1 0.494718 0.509766 0.494718 1 sampled descriptive statistics N,Total,Mean,Median,StdDev,Var,Min,Max 10000, 29362995237.2461, 2936299.5237, 2938641.5510, 255871.1400, 65470040272.4238, 2056225.6331, 3964448.7210, sampled cumulative density function 0.00,0.10,0.20,0.30,0.40,0.50,0.60,0.70,0.80,0.90,1.00 2056225.6331,2607350.5397,2717632.5755,28028 85.0008,2872706.8011,2938666.1801,3001386.005 9,3068959.5536,3149636.9560,3262431.5356,3964 448.7210	
		<b>Joint Data</b> <a href="http://localhost:5000/resources/network_carbon/resourcepack_526/resource_1782/NIST451Pearson.cs">http://localhost:5000/resources/network_carbon/resourcepack_526/resource_1782/NIST451Pearson.cs</a>	
		<b>Calculations Description</b> This corresponds to Example 1 in the Conservation Technology Assessment 1 reference. v204d	
		<b>Media URL</b> <a href="http://localhost:5000/resources/network_carbon/resourcepack_166/resource_1739/Tradeoffs.png">http://localhost:5000/resources/network_carbon/resourcepack_166/resource_1739/Tradeoffs.png</a>	
		<b>Data URL</b> none	



DevTreks –social budgeting that improves lives and livelihoods

The following image for an equivalent M&E calculation shows similar results. In this example, the \$3,000 difference in the Most Likely Estimate is within acceptable bounds for numbers generated from random samples. Note the use of the Score URL instead of the Stock calculator's Joint Data URL.



DevTreks –social budgeting that improves lives and livelihoods

<b>Score</b> Net Benefits Score Math Expression $((I1.QT * I2.QTM) + I3.QT) - (I4.QT + I5.QTM + I6.Q1 + I6.Q2)$		High Estimate 2,946,832.5815 High Unit upper 90 % ci
Label NISTS1 Rel Label none		Math Type algorithm1 Math Sub Type subalgorithm3
Total Score 2,958,756.7618 Score Unit net benefits		Math Result original pearson DenseMatrix 3x3-Double 1 0.5 0.5 0.5 1 0.5 0.5 0.5 1 new pearson DenseMatrix 3x3-Double 1 0.523715 0.471227 0.523715 1 0.523573 0.471227 0.523573 1  sampled descriptive statistics N,Total,Mean,Median,StdDev,Var,Min,Max 1000, 2933693116.2779, 2933693.1163, 2940153.1977, 251822.0445, 63414342114.1770, 2058745.4657, 3715547.7931, sampled cumulative density function 0.00,0.10,0.20,0.30,0.40,0.50,0.60,0.70,0.80,0.90,1. 00 2058745.4657,2604583.3553,2718119.6337,28070 68.3780,2878098.1587,2940200.2312,3007056.073 7,3063616.3346,3152662.6247,3246026.8658,3715 547.7931
D1 0.0000 D1 Unit low		
D2 0.0000 D2 Unit high		
Date 11/15/2016 Dist Type normal		
Iterations 1000 Confidence Interval 90		
Random Seed 8 BaseIO none		
Math Operator equalto		
Most Likely 2,933,693.1163 Most Unit net benefits		Score URL <a href="http://localhost:5000/resources/network_carbon/resourcepack_526/resource_1782/NIST451Pearson.cs">http://localhost:5000/resources/network_carbon/resourcepack_526/resource_1782/NIST451Pearson.cs</a>
Low Estimate 2,920,553.6511 Low Unit lower 90 % ci		Calculations Description Example 1 in the Conservation Technology Assessment 1 reference. v204i

**Example 2. Algorithm 1. Subalgorithm 2. Probabilistic Risk: Normal Copulas**

**URLs:**



DevTreks –social budgeting that improves lives and livelihoods

[https://www.devtreks.org/greentreks/preview/carbon/outputseries/CTA Examples 1, Probabilistic Risk/2141212678/none](https://www.devtreks.org/greentreks/preview/carbon/outputseries/CTA%20Examples%201,%20Probabilistic%20Risk/2141212678/none)

[https://devtreks1.blob.core.windows.net/resources/network\\_carbon/resourcepack\\_1534/resource\\_7940/PCOR.csv](https://devtreks1.blob.core.windows.net/resources/network_carbon/resourcepack_1534/resource_7940/PCOR.csv)

Corresponding localhost URLs can be found by switching to the Carbon Emission club in the GreenTreks network group, searching for Outputs and following the data hierarchy: CTA Output Examples => Resource Stock Risk Examples => Output Stock Calculators => the 4 Output Series

The NIST (1988) reference used in the previous example mentions that simulation algorithms can also handle interdependencies between Indicators, such as when the Operating Cost Indicator is dependent on the Units Rented Indicator. A probabilistic risk assessment must account for these interdependencies, or correlations, among indicators when generating random samples of numbers to evaluate. **Appendix B** discusses this further. Otherwise, it's possible for an indicator to decrease, or not increase enough, when a positively correlated indicator increases. Example 1a demonstrates how Operating Costs may be expected to increase, and definitely not decrease, when the number of Units Rented increase.

Six correlated price and yield indicators are entered as follows:

- **Distribution Type:** The first simulation used normal distributions for prices and yields. A second simulation used the lognormal price and beta yield distributions shown in Table 2 of the Anderson 2009 reference. The second simulation was not replicated after version 1.8.7 because of Indicator property changes –time was better spent preparing the CTA-Prevention reference.
- **Q1 to Q5:** Standard variables used in Math Expressions. In this example, Q1 has been set equal to theoretical prices and yields.



DevTreks –social budgeting that improves lives and livelihoods

- **QT:** For correlated indicators, this is the correlated parameter that is represented in the correlation matrix and calculated as the mean of the random sample columns. For non-correlated indicators, it is the result of the Math Expression.
- **QTD1 and QTD2:** The normal simulation uses fictitious yield and price distributions. For tutorial purposes, the lognormal distribution uses the values in Table 2 of the Anderson 2009 reference (which wasn't carried out again after 1.8.7). Correlated indicators require correct distributions because they derive their initial random sample vectors from this distribution. Non correlated indicators are calculated in the regular manner.
- **Math Type; algorithm1, Sub Math Type: subalgorithm1 (Monte Carlo):** Each indicator starts with a random sample of numbers that are distributed according to the Distribution Type property.
- **Math Expression:** The Math Expression identifies which column of data to include in analyses and to calculate a QT for each row of data. Indicator 1 uses the following Expression. This example does not store data in a Data URL TEXT file, so the expression simply sets  $QT = Q1$ .

I1.Q1

- **QTM, QTU, QTL:** Mean of QT with upper and lower x% confidence intervals. Set from each indicator's correlated random samples explained shortly.

The Score properties are set to the following:

- **Math Type: algorithm1, Sub Math Type: subalgorithm2 (Normal Copula).** All correlated probabilistic risk analyses must use the Score Math Type, not each separate indicator's Math Type. Appendix B explains the basic steps used in the calculation.
- **Stock Joint Data URL or M&E Score URL:** The following correlation matrix is saved as a csv TEXT file, uploaded to a Base Resource element, and the URL is copied to this property. In this example, the Price and Yield indicator correlations are taken from a





DevTreks –social budgeting that improves lives and livelihoods

Pearson correlation matrix (derived from the Rank correlation matrix shown in Table 1 of the Anderson 2009 reference).

The first line must contain the name of the correlation matrix to use. The options are:

**pearson = Pearson correlation matrix**

**spearman = Spearman rho correlation matrix**

The second line must contain a comma-separated list of the correlated indicator labels:

**P1,Y1,P2,Y2,P3,Y3**

The remaining lines contain either a comma-separated real correlation matrix or be blank. If TEXT data referenced by the Data URL is being used and multiple observations of matched indicators is available, a blank matrix can be added and the software automatically generates the correlation matrix (see Example 1d). The Data URL data must contain at least 3 matched data elements for each observation. In this example, a known correlation matrix is derived from the techniques recommended by the IPCC 2006 and NASA 2011 references. The csv file appears as follows:

```
pearson
P1,Y1,P2,Y2,P3,Y3
1,-0.3645,0.5176,-0.1569,0.1047,-0.0524
-0.3645,1,-0.3129,0.7167,-0.0838,0.3129
0.5176,0.7167,1,-0.4363,0.2922,-0.1256
-0.1569,0.7167,-0.4363,1,-0.0733,0.2611
0.1047,-0.0838,0.2922,-0.0733,1,-0.2091
-0.0524,0.3129,-0.1256,0.2611,-0.2091,1
```

More than one Joint Data URL can be used (by using semicolon-delimited data urls), but if data files are being used to set the correlation matrixes, each Joint Data URL must have a corresponding Data URL file and they must be in the same order. The latter feature has not been tested with actual datasets.



DevTreks –social budgeting that improves lives and livelihoods

- **Score Math Expression:** The result of this Expression is the mean revenue for the three price and yield combinations.

$$((I1.QT * I2.QT) + (I3.QT * I4.QT) + (I5.QT * I6.QT)) / 3$$

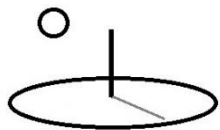
Refer to the third set of images to understand why the following expression doesn't work with this algorithm. The last version of this reference actually used this expression rather than the correct expression and showed the second set of images (5\*).

$$((I1.QTM * I2.QTM) + (I3.QTM * I4.QTM) + (I5.QTM * I6.QTM)) / 3$$

- **Score Distribution Type:** Both simulations used a normal distribution.
- **ScoreD1 and ScoreD2:** Set manually by some type of expert logic known about the Score (i.e. run the calculator to get the Score, set ScoreD1 and ScoreD2, and run the calculator again to get the final results). Example 1d shows how, by setting these to zero, the underlying indicator distributions can be used to fill in ScoreM, ScoreL, and ScoreU.
- **ScoreM, ScoreL, and ScoreU:** Mean of Score with upper and lower x% confidence intervals.
- **ScoreMUnit:** unit of measure for ScoreM.
- **Iterations:** Number of random samples to generate.
- **Confidence Interval:** 90

The calculator uses the following steps:

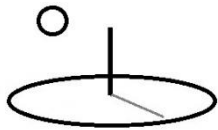
- **Step 1.** . Run an asynchronous loop that simultaneously iterates through each dataset in the Joint Data URL TEXT file. Parse the data and determine the type of correlation matrix to build, the labels of the correlated indicators, and optionally, an initial correlation matrix. Errors with datasets will be added to the Calculator.Description property.



DevTreks –social budgeting that improves lives and livelihoods

- **Step 2.** If the Data URL holds TEXT datasets, load the dataset corresponding to the Joint Data URL property. Use that matrix to automatically calculate a Pearson or Spearman correlation matrix.
- **Step 3.** Use the techniques explained in Appendix B to generate a matrix of correlated random samples.
- **Step 4.** Use the correlated random sample matrix to generate descriptive statistics for each vector of indicators. Set each correlation indicator's QTM, QTL, QTU, from the statistics and add a summary of the statistics to the Math Result property.
- **Step 5.** When all of the calculations are completed, Set ScoreM, ScoreL, and ScoreU using the regular properties of the Score.

The following images demonstrate that each indicator is calculated using random samples of correlated numbers and that the Pearson or Spearman coefficient matrix generated from those numbers is added to the Score Math Result. Note the slight difference between the starting Pearson matrix and the new matrix.



DevTreks –social budgeting that improves lives and livelihoods

Score Math Expression	
$((I1.QT*I2.QT) + (I3.QT*I4.QT) + (I5.QT*I6.QT)) / 3$	
Score	Score Unit
4,347.9297	mean cost
Score D1	Score D1 Unit
0.0000	mean
Score D2	Score D2 Unit
0.0000	sd
Score Dist Type	Iterations
normal	10000
Confidence Interval	Random Seed
90	5
Score BaseIO	
none	
Score Most Likely	Score Most Unit
4,140.7213	mean
Score Low Estimate	Score Low Unit
4,132.6825	lower 90 % ci
Score High Estimate	Score High Unit
4,148.7601	upper 90 % ci
Score Math Type	Score Math Sub Type
algorithm1	subalgorithm2
Score Math Result	
original pearson DenseMatrix 6x6-Double 1 -0.3645 0.5176 -0.1569 0.1047 -0.0524	



DevTreks –social budgeting that improves lives and livelihoods

original pearson  
DenseMatrix 6x6-Double  
1 -0.3645 0.5176 -0.1569 0.1047 -0.0524  
-0.3645 1 -0.3129 0.7167 -0.0838 0.3129  
0.5176 -0.3129 1 -0.4363 0.2922 -0.1256  
-0.1569 0.7167 -0.4363 1 -0.0733 0.2611  
0.1047 -0.0838 0.2922 -0.0733 1 -0.2091  
-0.0524 0.3129 -0.1256 0.2611 -0.2091 1  
new pearson  
DenseMatrix 6x6-Double  
1 -0.37182 0.456623 -0.115301 0.0996078 -0.040587  
-0.37182 1 -0.330845 0.574998 -0.0847187 0.273909  
0.456623 -0.330845 1 -0.289786 0.287141 -0.0715926  
-0.115301 0.574998 -0.289786 1 0.159931 0.0242184  
0.0996078 -0.0847187 0.287141 0.159931 1 -0.192733  
-0.040587 0.273909 -0.0715926 0.0242184 -0.192733 1  
  
sampled descriptive statistics  
N,Total,Mean,Median,StdDev,Var,Min,Max  
10000, 41407212.5890, 4140.7213, 4129.7881, 487.2009,  
237364.7585, 2667.1169, 6429.3110,  
sampled cumulative density function  
0.00,0.10,0.20,0.30,0.40,0.50,0.60,0.70,0.80,0.90,1.00  
2667.1169,3524.7891,3722.6354,3877.6401,4005.6031,4129.8554,  
4253.3262,4383.4175,4542.2669,4769.9264,6429.3110

Joint Data

[http://localhost:5000/resources/network\\_carbon/resourcepack\\_166/resource\\_1741/PCOR.csv](http://localhost:5000/resources/network_carbon/resourcepack_166/resource_1741/PCOR.csv)

Calculations Description

Example 1 in the CTA 1 reference. v204b

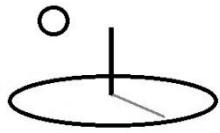
Media URL

[http://localhost:5000/resources/network\\_carbon/resourcepack\\_166/resource\\_1738/ResidualsinProduction.png](http://localhost:5000/resources/network_carbon/resourcepack_166/resource_1738/ResidualsinProduction.png)

Data URL

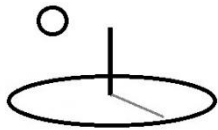
none

The following images shows that, for correlated indicators, the Indicator.QT property is calculated as the mean of the appropriate column of random samples. The initial random sample



DevTreks –social budgeting that improves lives and livelihoods

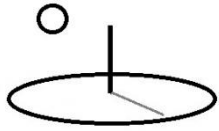
derive from the Indicator Distribution properties and those properties must be correct. That's also how the Indicator.QTM is set. For non-correlated indicators, QT is the result of the Math Expression.



DevTreks –social budgeting that improves lives and livelihoods

Indicator 1		Indicator 2	
Units Rented		563.9687	units rented
Indicator 1 Description		Math Type 1	Math Sub Type 1
This indicator is used in a CTA tutorial.		algorithm1	subalgorithm1
Indicator 1 URL		QT D1 1	QT D1 Unit 1
none		2,209.0000	shape
Label 1	Rel Label 1	QT D2 1	QT D2 Unit 1
UR	RV, OC	3.9170	scale
Date 1	Dist Type 1	QT Most 1	QT Most Unit 1
05/22/2015	gamma	563.9687	units rented
Q1 1	Q1 Unit 1	QT Low 1	QT Low Unit 1
600.0000	units	563.7721	lower 90 % ci
Q2 1	Q2 Unit 1	QT High 1	QT High Unit 1
0.9400	occupancy rate	564.1653	upper 90 % ci
Q3 1	Q3 Unit 1	Math Expression 1	
0.0000	none	I1.Q1 * I1.Q2	
Q4 1	Q4 Unit 1	Math Result 1	
0.0000	none	sampled descriptive statistics N,Total,Mean,Median,StdDev,Var,Min,Max 10000, 5639686.7936, 563.9687, 564.1348, 11.9137, 141.9371, 519.9388, 611.0904, sampled cumulative density function 0.00,0.10,0.20,0.30,0.40,0.50,0.60,0.70,0.80,0.90,1.00 519.9388,548.8212,553.7468,557.7319,561.03 38,564.1390,567.0013,570.1077,573.8649,579 .3309,611.0904	
Q5 1	Q5 Unit 1		
0.0000	none		
Math Operator 1	BaseIO 1		
equalto	none		
QT 1	QT Unit 1		
563.9687	units rented		

The following image shows a new MathExpression used to test Version 2.0.4. DevTreks hadn't run the correlated indicator examples in a while and assumed this expression was "better" than



DevTreks –social budgeting that improves lives and livelihoods

the original expression in its use of QTMs. Wrong assumption. The Score Math Expression uses the random sample matrix to calculate Scores. That matrix uses the QT, rather than QTM, properties to store numbers in the matrix. All this expression is doing is repeating the same numbers from the existing Indicator QTMs –it is not using the random sample data at all (5\*).

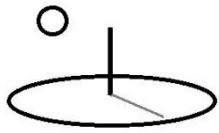
$$((I1.QTM * I2.QTM) + (I3.QTM * I4.QTM) + (I5.QTM * I6.QTM)) / 3$$





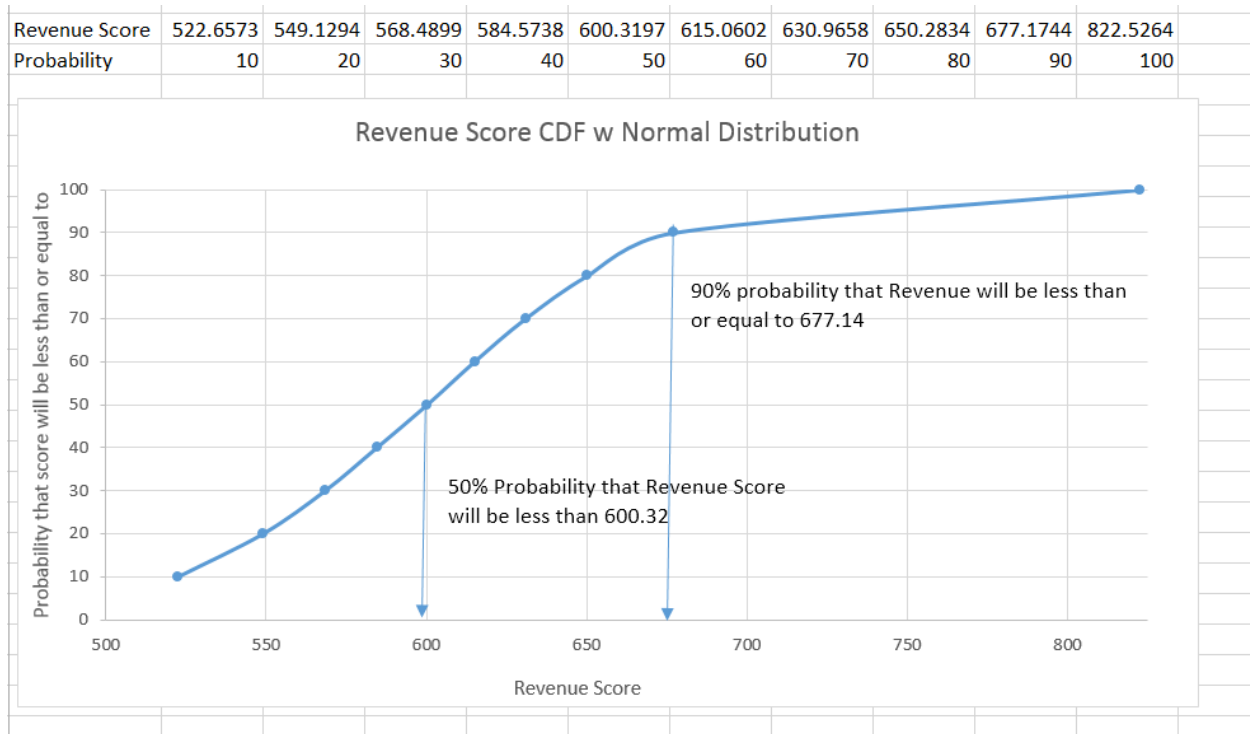
DevTreks –social budgeting that improves lives and livelihoods

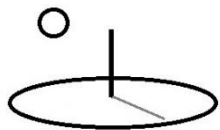
<b>Score Math Expression</b> ((I1.QTM * I2.QTM) + (I3.QTM * I4.QTM) + (I5.QTM * I6.QTM)) / 3		<b>Score Math Result</b> original pearson DenseMatrix 6x6-Double <pre> 1 -0.3645 0.5176 -0.1569 0.1047 -0.0524 -0.3645 1 -0.3129 0.7167 -0.0838 0.3129 0.5176 -0.3129 1 -0.4363 0.2922 -0.1256 -0.1569 0.7167 -0.4363 1 -0.0733 0.2611 0.1047 -0.0838 0.2922 -0.0733 1 -0.2091 -0.0524 0.3129 -0.1256 0.2611 -0.2091 1 </pre> new pearson DenseMatrix 6x6-Double <pre> 1 -0.37182 0.456623 -0.115301 0.0996078 -0.040587 -0.37182 1 -0.330845 0.574998 -0.0847187 0.273909 0.456623 -0.330845 1 -0.289786 0.287141 -0.0715926 -0.115301 0.574998 -0.289786 1 0.159931 0.0242184 0.0996078 -0.0847187 0.287141 0.159931 1 -0.192733 -0.040587 0.273909 -0.0715926 0.0242184 -0.192733 1 </pre> sampled descriptive statistics N,Total,Mean,Median,StdDev,Var,Min,Max 10000, 41491930.0000, 4149.1930, 4149.1930, 0.0000, 0.0000, 4149.1930, 4149.1930, sampled cumulative density function 1.00,1.00,1.00,1.00,1.00,1.00,1.00,1.00,1.00,1.00,1.00, 0.00 4149.1930,4149.1930,4149.1930,4149.1930,4149.1930,4149.1930, 4149.1930,4149.1930,4149.1930,4149.1930,4149.1930,4149.1930, 4149.1930,4149.1930	
<b>Score</b> 4,149.1930	<b>Score Unit</b> mean cost		
<b>Score D1</b> 0.0000	<b>Score D1 Unit</b> mean		
<b>Score D2</b> 0.0000	<b>Score D2 Unit</b> sd		
<b>Score Dist Type</b> normal	<b>Iterations</b> 10000		
<b>Confidence Interval</b> 90	<b>Random Seed</b> 5		
<b>Score BaseIO</b> none			
<b>Score Most Likely</b> 4,149.1930	<b>Score Most Unit</b> mean		
<b>Score Low Estimate</b> 4,149.1930	<b>Score Low Unit</b> lower 90 % ci		
<b>Score High Estimate</b> 4,149.1930	<b>Score High Unit</b> upper 90 % ci		
<b>Score Math Type</b> algorithm1	<b>Score Math Sub Type</b> subalgorithm2		
<b>Score Math Result</b> original pearson DenseMatrix 6x6-Double		<b>Joint Data</b> <a href="http://localhost:5000/resources/network_carbon/resourcepack_166/resource_1741/PCOR.csv">http://localhost:5000/resources/network_carbon/resourcepack_166/resource_1741/PCOR.csv</a>	
		<b>Calculations Description</b> Example 1 in the CTA 1 reference. v204a	



DevTreks –social budgeting that improves lives and livelihoods

The following image, from an earlier version, displays one potential way to communicate the results of this analysis to decision makers.





DevTreks –social budgeting that improves lives and livelihoods

### **Example 3. Algorithm 1. Subalgorithm 3. Probabilistic Risk: Eigen Copulas with Normal Distributions**

#### **URL:**

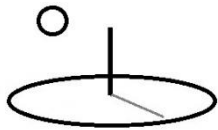
[https://www.devtreks.org/greentreks/preview/carbon/outputseries/CTA Examples 2, Probabilistic Risk/2141212679/none](https://www.devtreks.org/greentreks/preview/carbon/outputseries/CTA%20Examples%20Probabilistic%20Risk/2141212679/none)

[https://devtreks1.blob.core.windows.net/resources/network\\_carbon/resourcepack\\_1534/resource\\_7943/SCOR.csv](https://devtreks1.blob.core.windows.net/resources/network_carbon/resourcepack_1534/resource_7943/SCOR.csv)

Corresponding localhost URLs can be found by switching to the Carbon Emission club in the GreenTreks network group, searching for Outputs and following the data hierarchy: CTA Output Examples => Resource Stock Risk Examples => Output Stock Calculators => the 4 Output Series

This example uses algorithm 1 and subalgorithm3, Eigen Copula with Normal Distribution, for the Score Math Type and Score Math Sub Type properties. The changes from Example 1 include different Indicator distributions and the use of an Eigen decomposition function from the math library and a Spearman correlation. Appendix B shows the source code. The Anderson 2009 reference explains the steps used with this algorithm (i.e. using the square root of the eigenvalues). Note that a chapter in the Brebbia (2013) reference finds faults with this algorithm (but that reference is not open access and therefore of limited usefulness in this context).

The following images display a normal distribution test using the Spearman correlation matrix.



DevTreks –social budgeting that improves lives and livelihoods

Score Math Expression	
$((I1.QT*I2.QT) + (I3.QT*I4.QT) + (I5.QT*I6.QT)) / 3$	
Score	Score Unit
3,279.2814	mean cost
Score D1	Score D1 Unit
0.0000	low
Score D2	Score D2 Unit
0.0000	high
Score Dist Type	Iterations
normal	10000
Confidence Interval	Random Seed
90	0
Score BaseIO	
make selection	
Score Most Likely	Score Most Unit
4,072.2255	mean
Score Low Estimate	Score Low Unit
4,045.0633	lower 90 % ci
Score High Estimate	Score High Unit
4,099.3877	upper 90 % ci
Score Math Type	Score Math Sub Type
algorithm1	subalgorithm3
Score Math Result	
original spearman DenseMatrix 6x6-Double 1 -0.3645 0.5176 -0.1569 0.1047 -0.0524	



DevTreks –social budgeting that improves lives and livelihoods

original spearman
DenseMatrix 6x6-Double
1 -0.3645 0.5176 -0.1569 0.1047 -0.0524
-0.3645 1 -0.3129 0.7167 -0.0838 0.3129
0.5176 -0.3129 1 -0.4363 0.2922 -0.1256
-0.1569 0.7167 -0.4363 1 -0.0733 0.2611
0.1047 -0.0838 0.2922 -0.0733 1 -0.2091
-0.0524 0.3129 -0.1256 0.2611 -0.2091 1
new spearman
DenseMatrix 6x6-Double
1 -0.355953 0.495145 -0.162403 0.107786 -0.0515359
-0.355953 1 -0.295688 0.704287 -0.0815161 0.294089
0.495145 -0.295688 1 -0.414832 0.287288 -0.117438
-0.162403 0.704287 -0.414832 1 -0.0753068 0.244177
0.107786 -0.0815161 0.287288 -0.0753068 1 -0.214165
-0.0515359 0.294089 -0.117438 0.244177 -0.214165 1
sampled descriptive statistics
N,Total,Mean,Median,StdDev,Var,Min,Max
10000, 40722254.6824, 4072.2255, 3913.9444, 1646.1931,
2709951.6632, -2611.6235, 11207.0878,
sampled cumulative density function
0.00,0.10,0.20,0.30,0.40,0.50,0.60,0.70,0.80,0.90,1.00
-2611.6235,2098.1437,2655.3583,3113.6816,3520.2841,3914.1773,
4328.0869,4839.9266,5421.5167,6242.1747,11207.0878
Joint Data
<a href="http://localhost:5000/resources/network_carbon/resourcepack_166/resource_1742/SCOR.csv">http://localhost:5000/resources/network_carbon/resourcepack_166/resource_1742/SCOR.csv</a>
Calculations Description
Example 1 in the CTA 1 reference. v206a
Media URL
<a href="http://localhost:5000/resources/network_carbon/resourcepack_166/resource_1738/ResidualsinProduction.png">http://localhost:5000/resources/network_carbon/resourcepack_166/resource_1738/ResidualsinProduction.png</a>
Data URL
none



DevTreks –social budgeting that improves lives and livelihoods

### **Example 3. Algorithm 1. Subalgorithm 3. Probabilistic Risk: Eigen with Data URL and no Score Distribution**

#### **URL:**

[https://www.devtreks.org/greentreks/preview/carbon/outputseries/CTA Examples 3, Probabilistic Risk/2141212680/none](https://www.devtreks.org/greentreks/preview/carbon/outputseries/CTA%20Examples%203,%20Probabilistic%20Risk/2141212680/none)

Corresponding localhost URLs can be found by switching to the Carbon Emission club in the GreenTreks network group, searching for Outputs and following the data hierarchy: CTA Output Examples => Resource Stock Risk Examples => Output Stock Calculators => the 4 Output Series

Stock datasets (note the use of the Indicator Label)

[https://devtreks1.blob.core.windows.net/resources/network\\_carbon/resourcepack\\_1534/resource\\_7939/PCORPYs.csv](https://devtreks1.blob.core.windows.net/resources/network_carbon/resourcepack_1534/resource_7939/PCORPYs.csv)

[https://devtreks1.blob.core.windows.net/resources/network\\_carbon/resourcepack\\_1534/resource\\_7942/PYs.csv](https://devtreks1.blob.core.windows.net/resources/network_carbon/resourcepack_1534/resource_7942/PYs.csv)

[http://localhost:5000/resources/network\\_carbon/resourcepack\\_526/resource\\_1753/PYs.csv](http://localhost:5000/resources/network_carbon/resourcepack_526/resource_1753/PYs.csv)

[http://localhost:5000/resources/network\\_carbon/resourcepack\\_166/resource\\_1743/PCORPYs.csv](http://localhost:5000/resources/network_carbon/resourcepack_166/resource_1743/PCORPYs.csv)

M&E datasets (note the use of the Indicator index position)

[https://devtreks1.blob.core.windows.net/resources/network\\_carbon/resourcepack\\_1534/resource\\_9106/PCORPYs.csv](https://devtreks1.blob.core.windows.net/resources/network_carbon/resourcepack_1534/resource_9106/PCORPYs.csv)

[https://devtreks1.blob.core.windows.net/resources/network\\_carbon/resourcepack\\_1534/resource\\_9107/PYs.csv](https://devtreks1.blob.core.windows.net/resources/network_carbon/resourcepack_1534/resource_9107/PYs.csv)

[http://localhost:5000/resources/network\\_carbon/resourcepack\\_526/resource\\_1878/PYs.csv](http://localhost:5000/resources/network_carbon/resourcepack_526/resource_1878/PYs.csv)

[http://localhost:5000/resources/network\\_carbon/resourcepack\\_166/resource\\_1879/PCORPYs.csv](http://localhost:5000/resources/network_carbon/resourcepack_166/resource_1879/PCORPYs.csv)

This example changes the techniques employed in Example 1c by using a blank correlation matrix stored in a Joint Data URL property and a sample dataset of Prices and Yields stored in a



DevTreks –social budgeting that improves lives and livelihoods

Data URL property. In addition, Indicator distributions are used to set Scores, rather than Score Distributions. Normal price and yield distributions are used. The following properties differ from Example 1c:

- **Indicator.Math Expression.** The Expression must include terms that are associated with data column names using the Ix.Qx.ColName convention. This example uses only one price or yield variable, so the Math Expression simple sets  $QT = Q1$ :

I1.Q1.PriceorYield

- **Stock Joint Data URL or M&E Score URL:** This file holds 2 lines of csv data for the first simulation. The first line specifies the type of correlation, the second line specifies the labels of the correlated Indicators.

Stock data

pearson

P1, Y1, P2, Y2, P3, Y3

M&E data [must follow DataURL dataset conventions]

pearson

2, 1, 4, 3, 6, 5

- **Score Distribution Type:** Set to none. The indicator distributions are used to fill in the final Scores.
- **ScoreD1 and ScoreD2:** Both properties are set to zero. The indicator distributions are used to fill in the final Scores.
- **Score Math Expression:** The correlated random sample matrix (R) generated in Examples 1 and 2 contain each indicator's QT property only. Sample statistics generated from the indicator vectors in matrix R are used to set the final QTM, QTL, and QTU properties. A Score can be set for each row of R by using a Score Math Expression that includes each row of indicator QTs. The following Score Math Expression is used to generate a Score for each row in the R matrix:



DevTreks –social budgeting that improves lives and livelihoods

$$((I1.QT*I2.QT) + (I3.QT*I4.QT) + (I5.QT*I6.QT)) / 3$$

Additional indicator properties can be included in the expression, but they won't come from data in the R matrix –they'll come directly from each indicator (i.e. they'll be the same for each row). Non-correlated Indicator calculations are run before this Math Expression so that QTM terms can be used in the Expression (see Example 1a).

- **Data URL:** This csv file holds 11 fictitious data observation for each of the six price and yield indicators. This data is used to automatically generate the appropriate correlation matrix. Once the correlation matrix is built, this data is not used again because each indicator's distribution properties are used to run calculations. The following is the first row used in this example.

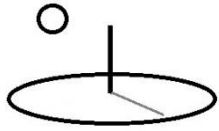
label, date, output, none, PriceorYield

The actual data starts on the second line. The QT value will be calculated for each row of data using the Indicator.MathExpression and whatever Qx properties are in the equation. The calculated QT columns are used to build a new correlation matrix. The first 4 lines as follows:

Stock dataset

P1,12/30/2014,barley,0,22  
P1,12/31/2014,barley,0,21  
P1,1/1/2015,barley,0,21.25  
P1,1/2/2015,barley,0,21.5  
P1,1/3/2015,barley,0,21.75  
P1,1/4/2015,barley,0,23  
P1,1/5/2015,barley,0,22.75  
P1,1/6/2015,barley,0,22.5  
P1,1/7/2015,barley,0,22.25  
P1,1/8/2015,barley,0,21.9  
P1,1/9/2015,barley,0,22.1





DevTreks –social budgeting that improves lives and livelihoods

Y1,12/30/2014,barley,0,11.5

Y1,12/31/2014,barley,0,10.5

Y1,1/1/2015,barley,0,10.75

M&E dataset

index, date, output, none, PriceorYield

2,12/30/2014,barley,0,22

2,12/31/2014,barley,0,21

...

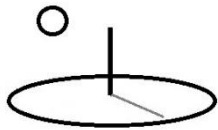
1,12/30/2014,barley,0,11.5

1,12/31/2014,barley,0,10.5

More than one Data URL can be used (by using semicolon-delimited data urls), but each Data URL must have a corresponding Joint Data URL file and they must be in the same order. All algorithms with multiple datasets are run asynchronously and simultaneously.

The following images shows these properties.

Stock Score



DevTreks –social budgeting that improves lives and livelihoods

Score Math Expression													
((I1.QT*I2.QT) + (I3.QT*I4.QT) + (I5.QT*I6.QT)) / 3													
Score	Score Unit												
185.6579	mean revenue												
Score D1	Score D1 Unit												
0.0000	mean												
Score D2	Score D2 Unit												
0.0000	sd												
Score Dist Type	Iterations												
normal	10000												
Confidence Interval	Random Seed												
90	6												
Score BaseIO													
none													
Score Most Likely	Score Most Unit												
218.8626	mean												
Score Low Estimate	Score Low Unit												
218.2301	lower 90 % ci												
Score High Estimate	Score High Unit												
219.4951	upper 90 % ci												
Score Math Type	Score Math Sub Type												
algorithm1	subalgorithm3												
Score Math Result													
original pearson DenseMatrix 6x6-Double <table border="1"> <tr> <td>1</td> <td>0.98939</td> <td>0.299951</td> <td>0.287029</td> <td>-0.0441807</td> <td>0.545457</td> </tr> <tr> <td>0.98939</td> <td>1</td> <td>0.181769</td> <td>0.267319</td> <td>0.000866288</td> <td>0.642968</td> </tr> </table>		1	0.98939	0.299951	0.287029	-0.0441807	0.545457	0.98939	1	0.181769	0.267319	0.000866288	0.642968
1	0.98939	0.299951	0.287029	-0.0441807	0.545457								
0.98939	1	0.181769	0.267319	0.000866288	0.642968								



DevTreks –social budgeting that improves lives and livelihoods

original pearson	
DenseMatrix 6x6-Double	
1	0.98939 0.299951 0.287029 -0.0441807 0.545457
0.98939	1 0.181769 0.267319 0.000866288 0.642968
0.299951	0.181769 1 0.29485 -0.36692 -0.32406
0.287029	0.267319 0.29485 1 0.106469 -0.224588
-0.0441807	0.000866288 -0.36692 0.106469 1 0.1684
0.545457	0.642968 -0.32406 -0.224588 0.1684 1
new pearson	
DenseMatrix 6x6-Double	
1	0.98938 0.298957 0.279974 -0.0364252 0.549544
0.98938	1 0.180308 0.258363 0.00800573 0.6472
0.298957	0.180308 1 0.302114 -0.359737 -0.328657
0.279974	0.258363 0.302114 1 0.107365 -0.233689
-0.0364252	0.00800573 -0.359737 0.107365 1 0.171247
0.549544	0.6472 -0.328657 -0.233689 0.171247 1
sampled descriptive statistics	
N,Total,Mean,Median,StdDev,Var,Min,Max	
10000, 2188625.7531, 218.8626, 217.3981, 38.3345, 1469.5315,	
89.3540, 365.1794,	
sampled cumulative density function	
0.00,0.10,0.20,0.30,0.40,0.50,0.60,0.70,0.80,0.90,1.00	
89.3540,170.7992,186.1245,197.4281,208.2039,217.4030,227.0331,	
237.2656,249.9978,269.6391,365.1794	
Joint Data	
<a href="http://localhost:5000/resources/network_carbon/resourcepack_166/resource_1743/PCORPYs.csv">http://localhost:5000/resources/network_carbon/resourcepack_166/resource_1743/PCORPYs.csv</a>	
Calculations Description	
Example 1 in the CTA 1 reference. v206a	
Media URL	
<a href="http://localhost:5000/resources/network_carbon/resourcepack_166/resource_1738/ResidualsinProduction.png">http://localhost:5000/resources/network_carbon/resourcepack_166/resource_1738/ResidualsinProduction.png</a>	
Data URL	
<a href="http://localhost:5000/resources/network_carbon/resourcepack_526/resource_1753/PYs.csv">http://localhost:5000/resources/network_carbon/resourcepack_526/resource_1753/PYs.csv</a>	

M and E Score



DevTreks –social budgeting that improves lives and livelihoods

original pearson

DenseMatrix 6x6-Double

```
1 0.98939 0.299951 0.287029 -0.0441807 0.545457
0.98939 1 0.181769 0.267319 0.000866288 0.642968
0.299951 0.181769 1 0.29485 -0.36692 -0.32406
0.287029 0.267319 0.29485 1 0.106469 -0.224588
-0.0441807 0.000866288 -0.36692 0.106469 1 0.1684
0.545457 0.642968 -0.32406 -0.224588 0.1684 1
```

new pearson

DenseMatrix 6x6-Double

```
1 0.989068 0.300714 0.273382 0.00611149 0.529796
0.989068 1 0.181144 0.251071 0.056225 0.630938
0.300714 0.181144 1 0.307627 -0.391699 -0.336507
0.273382 0.251071 0.307627 1 0.100929 -0.242916
0.00611149 0.056225 -0.391699 0.100929 1 0.228016
0.529796 0.630938 -0.336507 -0.242916 0.228016 1
```

sampled descriptive statistics

N,Total,Mean,Median,StdDev,Var,Min,Max

1000, 221303.0035, 221.3030, 220.0376, 38.4309, 1476.9320, 114.9264, 363.5916,

sampled cumulative density function

0.00,0.10,0.20,0.30,0.40,0.50,0.60,0.70,0.80,0.90,1.00

114.9264,173.1169,187.6765,199.1440,210.5364,220.0549,231.0023,240.5294,251.8735,  
270.4640,363.5916

Score URL

[http://localhost:5000/resources/network\\_carbon/resourcepack\\_166/resource\\_1879/PCORPYs.csv](http://localhost:5000/resources/network_carbon/resourcepack_166/resource_1879/PCORPYs.csv)

Calculations Description

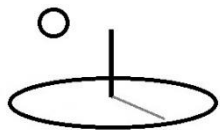
This Monitoring and Evaluation tool tracks up to 15 generic indicators that support the monitoring and evaluation of projects, programs, and technologies.v206c

Media URL

[http://localhost:5000/resources/network\\_carbon/resourcepack\\_166/resource\\_314/EPALCA1.PNG](http://localhost:5000/resources/network_carbon/resourcepack_166/resource_314/EPALCA1.PNG)

Data URL

[http://localhost:5000/resources/network\\_carbon/resourcepack\\_526/resource\\_1878/PYs.csv](http://localhost:5000/resources/network_carbon/resourcepack_526/resource_1878/PYs.csv)



DevTreks –social budgeting that improves lives and livelihoods

#### **Example 4. Algorithm 1. Subalgorithm 4. Probabilistic Risk: Eigen Copulas with Uniform Distributions**

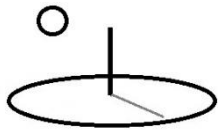
##### **URL:**

[https://www.devtreks.org/greentreks/preview/carbon/outputseries/NIST 451 Net Benefits, SubAlg4/2141212688/none](https://www.devtreks.org/greentreks/preview/carbon/outputseries/NIST%20451%20Net%20Benefits,SubAlg4/2141212688/none)

[https://devtreks1.blob.core.windows.net/resources/network\\_carbon/resourcepack\\_1534/resource \\_7978/NIST451Pearson.csv](https://devtreks1.blob.core.windows.net/resources/network_carbon/resourcepack_1534/resource_7978/NIST451Pearson.csv)

This algorithm uses a slight variation of subalgorithm3 that is explained in Appendix B. Since the same exact techniques are used, only the results of an example calculation will be used to explain this algorithm.

The following image shows the results of running this algorithm for the same data used with Example 1a, Method 3. That example used subalgorithm3. The resultant differences are slight, may not be statistically significant, but may be useful for some Indicators. The reason for including it is that some of Appendix B's references recommend this technique.



DevTreks –social budgeting that improves lives and livelihoods

Score Dist Type		Iterations	
normal		10000	
Confidence Interval			
90			
Score Most Likely		Score Most Unit	
2,929,435.8913		net benefits	
Score Low Estimate		Score Low Unit	
2,925,185.5318		lower 90% ci	
Score High Estimate		Score High Unit	
2,933,686.2508		upper 90% ci	
Score Math Type		Score Math Sub Type	
algorithm1		subalgorithm4	
Score Math Result			
<pre> original pearson DenseMatrix 3x3-Double   1 0.5 0.5 0.5 1 0.5 0.5 0.5 1 new pearson DenseMatrix 3x3-Double   1 0.49237 0.497996 0.49237 1 0.488183 0.497996 0.488183 1  sampled descriptive statistics N,Total,Mean,Median,StdDev,Var,Min,Max 10000, 29294358912.7858, 2929435.8913, 2927859.9384, 257597.5439, 66356494630.0903, 2255492.9536, 3609223.2549, sampled cumulative density function 0.10,0.20,0.30,0.40,0.50,0.60,0.70,0.80,0.90,1.00 2583527.0252,2700486.2241,2790393.9995,2862248.7955,2927893.2776,2998043.3369,3074408.7131,3159216.4352,3 271260.0443,3609223.2549 </pre>			
Joint Data			
http://localhost/resources/network_carbon/resourcepack_526/resource_1782/NIST451Pearson.csv			



DevTreks –social budgeting that improves lives and livelihoods

### **Example 5. Algorithm 1. Subalgorithm 5. Combinatorial Optimization: Simulated Annealing (3\*)**

#### **URL:**

<https://www.devtreks.org/greentreks/preview/carbon/output/CTA Examples 5, Combinatorial Optimization/2141223457/none>

[https://devtreks1.blob.core.windows.net/resources/network\\_carbon/resourcepack\\_1534/resource\\_7944/SimAnn1.csv](https://devtreks1.blob.core.windows.net/resources/network_carbon/resourcepack_1534/resource_7944/SimAnn1.csv)

<http://localhost:5000/greentreks/preview/carbon/output/Resource Stock Combo Optimization Examples/2141223462/none>

#### **Stock dataset**

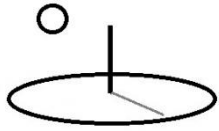
[http://localhost:5000/resources/network\\_carbon/resourcepack\\_166/resource\\_1746/SimAnn1.csv](http://localhost:5000/resources/network_carbon/resourcepack_166/resource_1746/SimAnn1.csv)

#### **M&E dataset**

[http://localhost:5000/resources/network\\_carbon/resourcepack\\_166/resource\\_1874/SimAnn2.csv](http://localhost:5000/resources/network_carbon/resourcepack_166/resource_1874/SimAnn2.csv)

This algorithm derives from the McCaffrey reference (January, 2012 issue). That reference contains examples of dozens of numeric algorithms, often used by software developers, to solve numeric problems involving cost minimization, best scheduling, machine learning, and artificial intelligence. This reference will include additional examples of these algorithms in future upgrades.

The simulated annealing algorithm solves combinatorial optimization problems. It was developed by engineers to calculate the best way to cool down specific materials, such as molten metals. They used the algorithm to figure out the best, or optimum, amount of energy that could be used for the cool down.



DevTreks –social budgeting that improves lives and livelihoods

In this example, it tries to minimize the total amount of time spent by 5 workers to complete 6 tasks. Each task must be assigned to one worker and each task can be completed by one of three workers. The total potential combinations of worker/tasks is  $3^6$ , or 729. The usefulness of the algorithm increases as the potential combinations increase –the McCaffrey reference uses the example of 20 tasks that can be completed by 12 workers, or  $12^{20}$  combinations. The algorithm loops through a sample size defined by an Iterations property to keep selecting better and better (lower number of hours) combinations of workers and tasks.

The algorithm can be applied in creative ways. The author found a recent reference (not cited) where the algorithm has been proposed as a better way to generate random samples of correlated numbers than conventional copula methods (by minimizing differences in Pearson coefficients).

The following properties are the initial parameters used by the indicators being calculated. This example uses the terms “workers, tasks, and hours” for illustrative purposes. The terms “rows, columns, and best result” are also appropriate.

- **Label:** For Stock calculators, must correspond to the Label used in the Joint Data URL property. For M&E calculators does not need to correspond –but the Indicator must be in the Index position identified in the Score URL dataset.
- **Distribution Type:** none. Results will not be expressed in terms of confidence intervals.
- **Q1:** current temperature
- **Q2:** alpha, the cooling rate
- **Q3:** in this example, a penalty when a worker has more than 1 task to complete
- **Math Type algorithm1, Sub Math Type: subalgorithm5 (simulated\_annealing):** Run a custom DevTreks algorithm to carry out simulated annealing analysis.
- **Math Expression:** blank. Q1 to Q3 will be passed to the algorithm and the algo will generate QT and QTM.
- **QT and QTM:** QT and QTM are equal and reflect the total number of hours (i.e. energy) required in the final solution.





DevTreks –social budgeting that improves lives and livelihoods

- **Math Result:** this algorithm requires deleting previous Math Results or the new Math Results get added to the previous results

The Stock Score properties are set to the following:

- **Score.DataURL:** The following table shows that this style of algorithm employs a standard TEXT dataset structure for jointly calculated Indicators. Unlike a Data URL, the numeric data in these datasets do not usually coincide with QT to Q1 properties. In addition, the Math Expression does not need to identify which columns of data to analyze. The data is saved as a csv TEXT file, uploaded to a Base Resource element, and the URL is copied to this property.

The first line, or header row, substitutes a Row Name column for a Data URL's Date column and allows up to 11 columns of numbers to be analyzed. Columns and rows throughout algorithms are restricted until testing reveals more about their consequences.

Indicator	Row	none	Col 1	Col 2	Col 3	...	Col 11
Label	Name		Amount	Amount	Amount		Amount
CO2	W1	0	6.543	7.000	3.26	...	1.500

The actual data for the 2 indicators being analyzed appears as follows;

label,rowname,none,T1, T2, T3, T4, T5, T6

TW1,W1,,7.5,3.5,2.5,0,0,0

TW1,W2,,0,1.5,4.5,3.5,0,0

TW1,W3,,0,0,3.5,5.5,3.5,0

TW1,W4,,0,0,0,6.5,1.5,4.5

TW1,W5,,2.5,0,0,0,2.5,2.5

TW2,W1,,7,2.5,3.5,0,0,0

TW2,W2,,0,3.5,1.5,3.5,0,0



DevTreks –social budgeting that improves lives and livelihoods

TW2,W3,,0,0,3.5,5.5,2.5,0

TW2,W4,,0,0,0,6.5,1.5,4

TW2,W5,,2.5,0,0,0,2.5,2.5

- **Math Type:** algorithm1 and subalgorithm1. In this example, the score is running a standard probabilistic risk function.
- **Score Math Expression:** I1.QTM + I2.QTM
- **Score Distribution Type:** normal
- **Score and ScoreM, ScoreLow, ScoreHigh:** The result of the Math Expression and Score algorithm
- **ScoreMUnit:** unit of Measure for ScoreM.
- **Score.MathResult:** The result of running the Score algorithm or none.

The following M&E Score properties differ from the Stock Score properties:

- **Score.DataURL:** The following table shows this dataset differs from the Stock dataset by using the Indicators' Index position, rather than Label, to identify the Indicator where the calculations get run.

Indicator	Row	none	Col 1	Col 2	Col 3	...	Col 11
Index	Name		Amount	Amount	Amount		Amount
1	W1	0	6.543	7.000	3.26	...	1.500

The actual data for the 2 indicators being analyzed appears as follows;

index,rowname,none,T1, T2, T3, T4, T5, T6

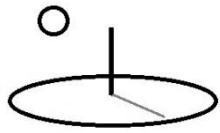
1,W1,,7.5,3.5,2.5,0,0,0

1,W2,,0,1.5,4.5,3.5,0,0

1,W3,,0,0,3.5,5.5,3.5,0

1,W4,,0,0,0,6.5,1.5,4.5

1,W5,,2.5,0,0,0,2.5,2.5



DevTreks –social budgeting that improves lives and livelihoods

2,W1,,7,2.5,3.5,0,0,0

2,W2,,0,3.5,1.5,3.5,0,0

2,W3,,0,0,3.5,5.5,2.5,0

2,W4,,0,0,0,6.5,1.5,4

2,W5,,2.5,0,0,0,2.5,2.5

The calculator uses the following steps:

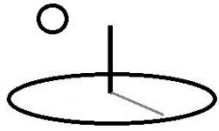
- **Step 1.** Run an asynchronous loop that simultaneously iterates through each dataset in the Joint Data URL TEXT file. Parse the data into a double[,] array. Errors with datasets will be added to the Calculator.Description property.
- **Step 2.** Use the indicator Q1 to Q3 properties to initiate a simulated annealing object. Pass in a double[,] array holding the initial worker/task hours. Run simultaneous simulations.
- **Step 3.** Set Indicator 1's QT and QTM properties to the solution's best energy amount. Set the MathResult property to the final optimum worker/task matrix.
- **Step 4.** Loop to the next indicator and carry out Steps 1 to 3 for that indicator.
- **Step 5.** Set the Score properties from the indicators properties when all the simulations have been completed.

The following image shows the results. This example is not particularly important in itself. The importance lies in the context in which the algorithm can be used.



DevTreks –social budgeting that improves lives and livelihoods

<a href="https://www.devtreks.org/greentreks/search/watershed/output/none/0/none?Keywords">DevTreks [US]</a> <a href="https://www.devtreks.org/greentreks/search/watershed/output/none/0/none?Keywords">https://www.devtreks.org/greentreks/search/watershed/output/none/0/none?Keywords</a>									
<b>Output Series: CTA Examples 5, Combinatorial Optimization</b>									
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
40.0000	hours	0.0000	none	0.0000	none	10000	90	2	none
40.0000	hours labor	0.0000	none	0.0000	none	none	algorithm1	subalgorithm5	1
I1.QTM+H2.QTM									
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
<b>Work Performance Indicator 1</b>		TW1	01/01/2015		algorithm1	none	none	none	subalgorithm5
1,000.0000	temperature	0.9950	alpha	3.5000	hours penalty	0.0000	none	0.0000	none
20.5000	hours labor	0.0000	0.0000	20.5000	hours labor	0.0000	none	0.0000	none
Example 6 in CTA reference.									
none									
<b>Work Performance Indicator 2</b>		TW2	01/01/2015		algorithm1	none	none	equalto	subalgorithm5
1,000.0000	temperature	0.9950	alpha	3.5000	hours penalty	0.0000	none	0.0000	none
19.5000	hours labor	0.0000	0.0000	19.5000	hours labor	0.0000	none	0.0000	none
Example 6 in CTA reference.									
none									
<a href="#">Feedback About carbon/output/CTA Examples 5, Combinatorial Optimization/2141223457/none</a>									



DevTreks –social budgeting that improves lives and livelihoods

### **Example 7. Algorithm 1. Subalgorithm 7. Classification and Prediction: Neural Network (3\*)**

#### **URLs:**

Simulation 1. 4 input variables with 3 possible output values

<https://www.devtreks.org/greentreks/preview/carbon/output/CTA Examples 7a, Neural Network/2141223458/none>

[https://devtreks1.blob.core.windows.net/resources/network\\_carbon/resourcepack\\_1534/resource\\_7955/NeuralEx1.csv](https://devtreks1.blob.core.windows.net/resources/network_carbon/resourcepack_1534/resource_7955/NeuralEx1.csv)

[https://devtreks1.blob.core.windows.net/resources/network\\_carbon/resourcepack\\_1534/resource\\_7956/NeuralEx2.csv](https://devtreks1.blob.core.windows.net/resources/network_carbon/resourcepack_1534/resource_7956/NeuralEx2.csv)

<http://localhost:5000/greentreks/preview/carbon/output/Resource Stock Neural Network Example 1/2141223463/none>

Stock dataset

[http://localhost:5000/resources/network\\_carbon/resourcepack\\_526/resource\\_1755/NeuralEx1.csv](http://localhost:5000/resources/network_carbon/resourcepack_526/resource_1755/NeuralEx1.csv)

M&E dataset

[http://localhost:5000/resources/network\\_carbon/resourcepack\\_526/resource\\_1876/NeuralEx1.csv](http://localhost:5000/resources/network_carbon/resourcepack_526/resource_1876/NeuralEx1.csv)

Version 2.1.4, as documented in the Social Performance Analysis 3 reference, demonstrates the latest techniques employed by DevTreks for running machine learning algorithms. This algorithm is being retained as another example of how to run joint calculations.



DevTreks –social budgeting that improves lives and livelihoods

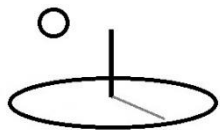
This algorithm derives from the McCaffrey reference (July, 2012 issue). This algorithm classifies and predicts the value of an output variable, given a set of predictive input variables. The neural network term refers to the techniques employed by the algorithm to learn how to classify and predict an output value based on an initial dataset of observed, or expertly surmised, input variables.

Botanists first developed this algorithm to predict the classification of a plant's species based on a species' input characteristics (i.e. sepal length, sepal width, petal length, petal width). The McCaffrey reference cites examples where the algorithm has been used to classify a loan applicant's credit score based on their income and expenses, and a hospital patient's cancer status based on blood test variables.

Although the algorithm supports classifications that are either numeric or textual, the nature of these Resource Stock indicators means that only numeric outcome values are supported in this version. Analysts should provide multimedia support (graphics, tables) that communicates the raw data numeric results in a manner that decision makers will understand.

McCaffrey uses data containing 4 fictitious input variables that are used to classify or predict 3 possible flower colors. The predicted output can take the values: 0=red, 1=green, 2=blue, and the code turns each value into an N-1 vector (i.e. red = {1, 0, 0}). The referenced code has been modified to accept up to 9 input variables and 10 potential values for the output being predicted (i.e. but was not debugged in version 1.9.2 because it will be either upgraded or replaced at an appropriate time). The size of the neural network used to make predictions directly relates to the number of input variables and output options (hence the restrictions). This version of the algorithm uses 500 iterations to train the neural network.

The algorithm employs a Particle Swarming optimization algorithm to minimize the error associated with training the neural network. The algorithm uses random numbers for what it calls "cognitive and social randomizations". The result is that running the algorithm consecutive times returns different results. To test the sensitivity of the different results, one test ran the calculator

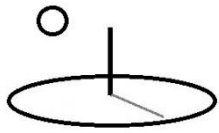


DevTreks –social budgeting that improves lives and livelihoods

10 consecutive times for one output and output series with this example's data. The test generated the following results: 0 predicted 10 out of 20 times with 90% average accuracy, 2 predicted 5 out of 20 times with 90% average accuracy, and 1 predicted 5 out of 20 times with 86 percent average accuracy. The reference points out that alternative ways for training the neural network can be used.

The author is not an expert using this type of algorithm, but assumes these types of results may be acceptable for some types of data (i.e. predicting advertisement use). He further assumes those types of results support a predicted output value of 0 with 90% accuracy (rather than 50%). As pointed out in Footnote 3, these algorithms will evolve.

A simulation that uses 1 indicator is displayed in the following image.



Indicator 1

Plant Color

Indicator 1 Description

This indicator uses a neural network algorithm to predict flower color.

Indicator 1 URL

none

Label 1

C1

Rel Label 1

none

Date 1

01/15/2015

Dist Type 1

none

Q1 1

2.0000

Q1 Unit 1

green

Q2 1

5.0000

Q2 Unit 1

sepal length

Q3 1

3.0000

Q3 Unit 1

sepal width

Q4 1

2.0000

Q4 Unit 1

petal length

Q5 1

3.0000

Q5 Unit 1

petal width

Math Operator 1

equalto

BaselO 1

none

QT 1

0.0000

QT Unit 1

plant color

Math Type 1

algorithm1

Math Sub Type 1

subalgorithm7

The following Indicator properties are set using a similar pattern to Example 5:

- **Label:** Must correspond to the Labels used in the Joint Data URL property.

64





DevTreks –social budgeting that improves lives and livelihoods

- **Distribution Type:** none. Results will not be expressed in terms of confidence intervals.
- **Q1 to Q5 Amounts:** Not required because 20% of the data rows are test data –the first row of that test data can correspond to Qx variables and the calculated result is displayed in the Math Result. Rather than leaving these set to zero, the values for the first row of test data can be used to improve reporting. When 2 indicators are used, the second indicator will contain the final calculations.
- **Q1 to Q5 Units:** Not required, but rather than leaving empty, they can correspond to the first row of test data.
- **Math Type algorithm1, Sub Math Type: subalgorithm7 (neural\_network):** Run a custom DevTreks algorithm to carry out neural network analysis.
- **Math Expression:** blank. The Qx properties will be passed to the algorithm and the algo will generate QT and QTM.
- **QT and QTM Amounts:** The QT and QTM Amounts will be equal and reflect the classified/predicted output variable. In this example, Q1 is the guessed output variable.
- **QTL:** If the Distribution Type is set to none, records the percent accuracy of the network in making predications.

The following Score properties are set using a similar pattern to Example 5:

- **Score.DataURL:** The following table shows that this style of algorithm employs the same data structure as algorithm 5. In addition, the first 80% of rows are used to train the neural network while the remaining 20% are used to test the network. The prediction results for the first 3 rows of test data are included in the Math Result and can correspond to specific Qxs used to make predictions. The data is saved as a csv TEXT file, uploaded to a Base Resource element, and the URL is copied to this property. *In this algorithm, the Col 0 Amount must be an output value. The input values go into Col 1 to Col 10.*

Indicator Label	Species	Row Name	Col 1 Amount	Col 2 Amount	Col 3 Amount	...	Col 11 Amount
C1	rose	red	1	6.543	7.000	...	1.500



**Indicator 1.** 4 input variables and 3 output value options.

This data contains 100 rows of artificial data taken directly from the McCaffrey reference. The following tables displays the first few lines of data. The output value, with 3 possible numeric values, is stored in the 3<sup>rd</sup> column with the header “color”. The input values are in the remaining 4 columns.

label,species,color2,color,length,width,slength,swidth

C1,rose,green,2,8,5,9,5

C1,rose,blue,1,9,5,2,2

C1,rose,red,0,6,9,4,6

C1,rose,blue,1,9,2,3,3

C1,rose,red,0,7,6,9,8

**Indicator 2.** 8 input variables and 5 output value options.

This data contains 100 rows of artificial data. The following tables displays the first few lines of data. The output value, with 5 possible numeric values, is stored in the 3<sup>rd</sup> column. The input values are in the remaining 8 columns.

label,species,color2,color,length,width,slength,swidth,var5,var6,var7,var8

C2,rose,green,2,8,5,9,5,8,5,9,5

C2,rose,blue,1,9,5,2,2,9,5,2,2

C2,rose,red,0,6,9,4,6,6,9,4,6

C2,rose,white,4,9,2,3,3,9,2,3,3

C2,rose,red,0,7,6,9,8,7,6,9,8

C2,rose,red,0,9,8,8,6,9,8,8,6



DevTreks –social budgeting that improves lives and livelihoods

C2,rose,blue,1,3,8,6,2,3,8,6,2

- **Iterations:** This property is not used by the algorithm (the algorithm uses 500 iterations). This property can be used if confidence intervals are being generated.

The M&E Score properties are as follows:

- **Score.DataURL:** M&E calculators use the same property to store these types of datasets, but they place an Indicator's Index position, rather than Label, in the first column. The Score's Index position, 0, can be included in these datasets.

index,species,color2,color,length,width,length,width,var5,var6,var7,var8

1,rose,green,2,8,5,9,5,8,5,9,5

1,rose,blue,1,9,5,2,2,9,5,2,2

The calculator uses the following steps:

- **Step 1.** Run an asynchronous loop that simultaneously iterates through each dataset in the Joint Data URL TEXT file. Parse the data stored in the Joint Data URL file (which must be uploaded to a base Resource element).
- **Step 2.** Use one or more indicator Q1 to Q5 properties as an observed set of output and input variables. Parse the data into a matrix and pass the output/input matrix to the algorithm. Run the simulation.
- **Step 3.** Use the first 80% of the data observations to train the neural networks. This data is normalized to values between -1.0 and 1.0 using the intercept and slope of the min and max values in the data. Use the final 20% of observations to see how well the new neural network matrix predicts and classifies the output variable. Add the Qx variables to the last row of the test matrix. Use this last row to set the indicator's QT and QTM properties. Do not use that row when determining the percent accuracy of the neural network. Add summary data from the simulation to the MathResult property.



DevTreks –social budgeting that improves lives and livelihoods

- **Step 4.** Loop to the next unique indicator and carry out Steps 1 to 3 for that indicator.
- **Step 5.** When all of the calculations have been completed, set the Score properties from the indicators properties.

The following images shows the Math Result property of the first simulation. The first row of numbers, Input, in the *Examples of the network accuracy* are the indicator Qx results (using data that has been normalized to values between -1 and 1). The second row, Output, uses an indexed value of 1 to show the actual Output value found in the dataset. The third row, Predicted, uses fractions to show the network's predicted probabilities for each output value. The fraction with the highest value is the predicted output (i.e. a winner takes all approach). The predicted color is 0, or red. In general, having 0 as an option is discouraged because the results appear as if nothing happened. The network's prediction accuracy is around 85%. Decision makers can decide if that level of accuracy for making predictions is good or bad (i.e. it might be bad for a health and safety-related indicator, but acceptable for an advertising sales prediction). Note that the exact same calculator properties were run at the same time for the child output series and produced the same predicted color with a 76% accuracy.



DevTreks –social budgeting that improves lives and livelihoods

Best weights found:

```
5.00 5.00 -5.00 5.00 5.00 -3.75 -2.18 -3.62 -2.03 0.95 1.28 0.96
-1.35 -5.00 1.28 2.84 1.57 -0.19 -5.00 -2.20 -0.66 1.51 0.54 2.80
3.37 -5.00 5.00 1.85 -5.00 1.90 2.12 5.00 -1.54 1.65 0.43 -5.00
0.48 -5.00 5.00 4.43 3.53 0.39 -1.95
```

Examples of the neural network accuracy:

```
-----
Input:  -1.00 -0.78 0.56 -0.33
Output:  0.0 1.0 0.0 (output 2)
Predicted: 1.0 0.0 0.0 (output 1)
wrong
```

```
-----
Input:  -0.56 0.11 0.11 1.00
Output:  0.0 1.0 0.0 (output 2)
Predicted: 0.5 0.3 0.2 (output 1)
wrong
```

```
-----
Input:  -1.00 0.33 0.78 -0.11
Output:  1.0 0.0 0.0 (output 1)
Predicted: 1.0 0.0 0.0 (output 1)
correct
```

```
-----
Input:  -0.78 0.33 0.11 -0.11
Output:  0.0 0.0 1.0 (output 3)
Predicted: 1.0 0.0 0.0 (output 1)
wrong
```

...

Correct = 14

Wrong = 7

Prediction accuracy = 0.6667

Predicted QT = 0.0000



DevTreks –social budgeting that improves lives and livelihoods

### **Example 6. Algorithm 1. Subalgorithm 6. Probabilistic Statistics: Regression Analysis**

<https://www.devtreks.org/greentreks/preview/carbon/input/DevTreks OLS 1/2147397534/none>

<http://localhost:5000/greentreks/preview/carbon/inputseries/Example 6, Regression/2147380293/none>

Stock datasets

[https://devtreks1.blob.core.windows.net/resources/network\\_carbon/resourcepack\\_1534/resource\\_7951/Ex6.csv](https://devtreks1.blob.core.windows.net/resources/network_carbon/resourcepack_1534/resource_7951/Ex6.csv)

[http://localhost:5000/resources/network\\_carbon/resourcepack\\_526/resource\\_1749/Ex6.csv](http://localhost:5000/resources/network_carbon/resourcepack_526/resource_1749/Ex6.csv)

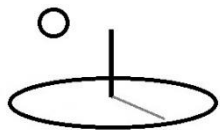
M&E datasets

[https://devtreks1.blob.core.windows.net/resources/network\\_carbon/resourcepack\\_1534/resource\\_9108/Ex6.csv](https://devtreks1.blob.core.windows.net/resources/network_carbon/resourcepack_1534/resource_9108/Ex6.csv)

[http://localhost:5000/resources/network\\_carbon/resourcepack\\_526/resource\\_1875/Ex6.csv](http://localhost:5000/resources/network_carbon/resourcepack_526/resource_1875/Ex6.csv)

This algorithm uses regression analysis to estimate probable values for an indicator's QTM property (i.e. dependent variable), given its Qx values (i.e. independent variables). It also makes predictions about that value. At the Score level, the algorithm uses whatever indicator properties are used in the Score.MathExpression and Score.MathType to generate Score, ScoreM, ScoreL, and ScoreU properties. Up to 10 explanatory variables can be analyzed. Use algorithm 2 or 3 when more explanatory variables need to be analyzed. Multiple base elements can be calculated at one time using standard techniques.

An indicator's QT Amount is the dependent variable, y, of a typical regression expressed by the following equation. The x variables that will be estimated and predicted are added to the last 3 rows of datasets and serve as a scoring dataset. All of the data being analyzed must be added to a comma-separated-value file found using the Data URL property.



DevTreks –social budgeting that improves lives and livelihoods

$$y = b_0 + b_1x + b_nx$$

In the context of regression analysis and CTA, the *probabilistic* model for any particular observed value of y is:

$$y = (\text{mean value of } y \text{ for a given value of } x_i) + (\text{random error})$$

$$y = b_0 + b_1 + b_nx + e$$

This version requires that all data being analyzed to be transformed (log, negative numbers, polynomials, weights) prior to being used in the algorithm.

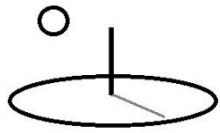
More than one indicator can be used in the regression analysis by using appropriate Math Expressions and dataset columns.

The Indicator.MathExpression terms must correspond to column names used in the dataset. The math terms must use the standard Ix.Qx with a required “.commoncolname” suffix. The “commoncolname” suffix must exactly match a dataset column name, without the period delimiter. The following two Math Expressions show that this convention allows columns in the dataset, and Ix.Qx variables, to be ignored by not including their name in the Math Expression. Math Expression with 2 variables in 1 indicator:

$$I1.Q1.housesize1 + I1.Q2.housesize2$$

Math Expression with 10 variables in 1 indicator (calculators don’t display Q6 to Q10 properties, but datasets can contain up to 10 independent variables).

$$I1.Q1.housesize1 + I1.Q2.housesize2 + I1.Q3.locationamenity1 + I1.Q4. \\ locationamenity2 + I1.Q5. locationamenity3 + I1.Q6.lotsize1 + I1.Q7.lotsize2 + \\ I1.Q8.meansalesprice + I1.Q9.constructionqualityrating + I1.Q10.locationrating$$



DevTreks –social budgeting that improves lives and livelihoods

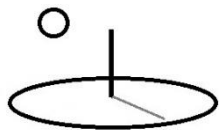
The first example derives from a college text book (Mendenhall and Sincich, 1989). The text book example provided datasets, regression examples, mathematical matrix examples, and SAS reports, needed to develop and test every major feature in this version of the algorithm. The dependent variable,  $y$ , is household monthly energy use. The first independent variable,  $x_1$ , is size of house. The second independent variable,  $x_2$ , is size of house squared.

Abstract equation:  $y = B_0 + B_1x_1 + B_2x_1^2 + e$

Math Expression:  $I1.Q1.housesize1 + I1.Q2.housesize2$

The following images display the completed properties.





Indicator 1

Regression Example 1

Indicator 1 Description

Example 6 in the CTA01 reference.

Indicator 1 URL

http://localhost:5000/resources/network\_carbon/resourcepack\_526/resource\_1749/Ex6.csv

Label 1

B2

Rel Label 1

none

Date 1

01/22/2015

Dist Type 1

none

Q1 1

1,500.0000

Q1 Unit 1

sq ft housesize

Q2 1

2,250,000.0000

Q2 Unit 1

sq ft housesizesquare

Q3 1

0.0000

Q3 Unit 1

none

Q4 1

0.0000

Q4 Unit 1

none

Q5 1

0.0000

Q5 Unit 1

none

Math Operator 1

equalto

BaseIO 1

none

QT 1

1,400.0000

QT Unit 1

guessed energy use

Math Type 1

algorithm1

Math Sub Type 1

subalgorithm6

QT D1 1

0.0000

QT D1 Unit 1

none

QT D2 1

0.0000

QT D2 Unit 1

none

QT Most 1

1,369.6609

QT Most Unit 1

estimated energy use

QT Low 1

1,324.9886

QT Low Unit 1

lower 95 % ci

QT High 1

1,414.3331

QT High Unit 1

upper 95 % ci

Math Expression 1

I1.Q1.housesize1 + I1.Q2.housesize2

Math Result 1

regression results

dependent variable: energyuse

source	df	SS	MS
model	2	831069.5464	415534.7732
error	7	15332.5536	2190.3648
total	9	846402.1000	

R-squared

0.9819

Adj R-squared

0.9767

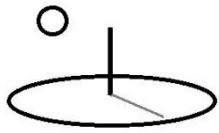
F value

189.7103

prob > F

0.0001

These properties are set as follows:



DevTreks –social budgeting that improves lives and livelihoods

- **Label:** For Stock calculators, must correspond to the Label used in the Joint Data URL property. For M&E calculators does not need to correspond –but the Indicator must be in the Index position identified in the Score URL dataset.
- **Description:** Explanation of the regression equation being analyzed ( $y = B_0 + B_1x_1 + B_2x_1^2 + e$ ) and the results.
- **Distribution Type:** none. This algorithm automatically generates confidence intervals for the estimated QTM. The Math Result includes additional confidence intervals for both estimated and predicted QTM amounts.
- **Q1 to Q5 Amounts:** Set of independent, or explanatory, variables used to estimate and predict the dependent variable, QTM. These are not actually used in any calculation –the last 3 rows of data are used to “score” the model. The last row of data is used to complete the QTM, QTL, and QTM properties. Because Q1 to Q5 are displayed in Stock Total Analyses, set them equal to the five most significant independent variables in the last dataset row. In this example, the calculation is trying to estimate and predict the energy use for a 1500 square foot house.
- **Q1 to Q5 Units:** Units of measurement for the independent variables. They do not need to match their corresponding dataset names.
- **QT Amount and Unit:** Prior guess about the amount of energy use associated with the independent variables. Every dataset must include data for the QT Amount property as the observed values of the dependent variable.
- **Math Type algorithm1, Sub Math Type: subalgorithm6 (Regression):** Each indicator is using regression analysis to generate its QTM, QTL, and QTM properties.
- **URL:** Regressions are run for indicators found in datasets, not for individual indicators. The regression examples demonstrate the required data conventions. The *Resource Stock Calculation* reference explains the following conventions more thoroughly.

Version 2.1.4 and 2.1.6 refactored examples 1 and 5 to 8 by placing greater emphasis on using the Indicator.URL property to store datasets. The 1<sup>st</sup> column of data in the following dataset uses the standard R and Python data convention of including row identifiers (i.e. 1, 2, 3, ...) in the label column.



DevTreks –social budgeting that improves lives and livelihoods

The actual data starts on the second line. The third column, which supports a custom data column, has been left blank because it is not being used. The last 3 rows of data is not used in the regression analysis but is used to score the model –in this case, used in a sensitivity analysis of the confidence interval. The last row of data is used to complete the QTM, QTL, and QTU properties.

```
label,date,none,energyuse,housesize1,housesize2
```

```
1,1/30/2015,,1182,1290,1664100
```

```
2,2/30/2015,,1172,1350,1822500
```

```
3,3/30/2015,,1264,1470,2160900
```

The previous pattern of using the Score.DataURL property can still be used. In that case, the label column must correspond to an Indicator.Label.

```
label,date,none,energyuse,housesize1,housesize2
```

```
A1,1/30/2015,,1182,1290,1664100
```

```
A1,2/30/2015,,1172,1350,1822500
```

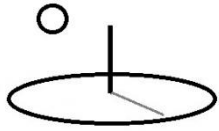
```
A1,3/30/2015,,1264,1470,216090
```

If the Score.DataURL property is being used with the M&E calculators, the following dataset shows that the label column must store an integer identifying the index position of an indicator.

```
index, date, none, energyuse, housesize1, housesize2
```

```
1,1/30/2015,,1182,1290,1664100
```

```
1,2/30/2015,,1172,1350,1822500
```



**Math Expression:** In the case of regression, this property is used to identify which columns of data in the TEXT file, and which independent variables, to include in the analysis. It's not actually parsed and run independently. The algorithm only uses the generic  $y = b_0 + b_1x + b_nx$  expression to run the regression. If the column name is not found in the Math Expression terms, the data column is ignored in the regression analysis. Each column of data in the TEXT file being analyzed must be transformed appropriately in the data set.

In the following Math Expression, the I1.Q2.housesize2 term is represented by a column of independent variable data that has already been raised to the second power.

$$I1.Q1.housesize1 + I1.Q2.housesize2$$

The terms used in the Math Expression must end in a corresponding column name used in the TEXT data file. They must start with the conventional Ix.Qx syntax. The following column names are used in the TEXT file:

housesize1, housesize2

- **QTM Amounts:** The QTM Amount will be calculated from the algorithm as the estimated value of the dependent variable. The last row of data in the TEXT file is used to set QTM, QTL, and QTU properties.
- **QTL and QTU Amounts:** These properties will be calculated from the algorithm as the lower and upper x% confidence interval for the estimated QTM Amount. The calculation uses a 2 sided T statistic test (t025). The Math Results display confidence intervals for both the estimated and predicted amounts.
- **Math Result:** This property includes standard descriptive statistics for regression analysis, including T statistics for each coefficient and R squared and F statistics for the estimated QTM. It also includes confidence intervals for estimated and predicted amounts for the Qx variables for the last 3 rows of data in the Data URL Text file. The results matched the referenced text book. The CTA 2 reference demonstrates that the results also match R and Python regression results.



DevTreks –social budgeting that improves lives and livelihoods

The following images demonstrate the Score properties.

Score Math Expression		Score Math Result	
<input type="text" value="I1.QTM"/>		sampled descriptive statistics N,Total,Mean,Median,StdDev,Var,Min,Max 1000, 1367617.9660, 1367.6180, 1368.9109, 77.0945, 5943.5666, 1113.4111, 1697.3510, sampled cumulative density function 0.00,0.10,0.20,0.30,0.40,0.50,0.60,0.70,0.80,0.90,1.0 0 1113.4111,1265.2150,1303.7926,1329.0152,1348.49 84,1369.0818,1387.8354,1408.6883,1431.1662,1468 .2713,1697.3510	
Score	Score Unit		
<input type="text" value="1,369.6609"/>	<input type="text" value="estimated energy use"/>		
Score D1	Score D1 Unit		
<input type="text" value="1,370.0000"/>	<input type="text" value="mean"/>		
Score D2	Score D2 Unit		
<input type="text" value="75.0000"/>	<input type="text" value="sd"/>		
Score Dist Type	Iterations		
<input type="button" value="normal"/> ▼	<input type="text" value="1000"/>		
Confidence Interval	Random Seed		
<input type="text" value="95"/>	<input type="text" value="0"/>		
Score BaseLO			
<input type="button" value="none"/> ▼			
Score Most Likely	Score Most Unit		
<input type="text" value="1,367.6180"/>	<input type="text" value="estimated energy use"/>		
Score Low Estimate	Score Low Unit		
<input type="text" value="1,362.8396"/>	<input type="text" value="lower 95 % ci"/>		
Score High Estimate	Score High Unit		
<input type="text" value="1,372.3964"/>	<input type="text" value="upper 95 % ci"/>		
Score Math Type	Score Math Sub Type		
<input type="button" value="algorithm1"/> ▼	<input type="text" value="subalgorithm1"/>		
Score Math Result		<b>Joint Data</b> <input type="text" value="none"/>	
		<b>Calculations Description</b> <input type="text" value="Example 1h in the CTA 1 reference. v216a"/>	
		<b>Media URL</b> <input type="text" value="http://localhost:5000/resources/network_carbon/resourcepack_166/resource_1739/Tradeoffs.png"/>	
		<b>Data URL</b> <input type="text" value="none"/>	
<b>Input Group: CTA Examples</b> <b>Input : Example 1h, Regression</b> <b>Input Series : Example 6, Regression</b>			



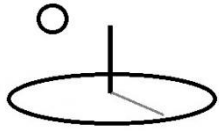
DevTreks –social budgeting that improves lives and livelihoods

The following properties show how Stock Score properties are set.

- **Score.MathExpression:** The following expression is just the result of the Indicator 1 QTM:  
I1.QTM
- **Score Math Type Properties:** none. The Score can be expressed as a confidence interval by using the properties explained in the previous examples, including this regression example (i.e. use the Score.JointDataURL to store the dataset).

The calculator uses the following steps:

- **Step 1.** Run an asynchronous loop that simultaneously iterates through each Indicator. Parse the data and build a vector of dependent variables (y) and a matrix of independent variables (xn). Errors with datasets will be added to the Calculator.Description property. The memory requirements for this technique must be considered when deciding which algorithm to use in a CTA.
- **Step 2.** Pass the data to the regression algorithm and run a regression for the indicator corresponding to the iteration loop. Errors with regression calculations will be added to the Math Result property of each indicator. To the extent possible, algorithms with multiple datasets run their calculations asynchronously and simultaneously.
- **Step 2a.** Use one or more indicator Q1 to Q5 properties as a set of independent variables from which a dependent variable can be estimated and predicted. Generate the dependent variable (QTM) and x% confidence intervals (QTL and QTU) and Score.ConfidenceInterval % prediction intervals for these specific values.
- **Step 3.** Add the results of the regression each Indicator's Math Result, QTM, QTL, and QTU properties.
- **Step 4.** Set the Score properties from the indicators properties. Score results may not always be meaningful if the regressions are primarily being run for their own sake, rather than for generating a Score. The Score is always reported, so we recommend using some meaningful combinations of Indicators.



DevTreks –social budgeting that improves lives and livelihoods

### **Example 6a. Algorithm 1. Subalgorithm 6. Probabilistic Statistics: Multiple Regression with Uncertain Costs**

#### **URLs**

The following Input contains 2 Input Series. One sets OC and AOH Prices. The other sets CAP and AOH Prices. Run the calculation at the Parent Input level and insert the calculator into the 2 children Input Series, then go into 1 of the Series, adjust the “wrong” price, rerun and save the calculation. Remember not to overwrite that child Series when running the Input calculation in the future. If calculations need to be rerun in the future and Input or Output prices must be updated in the base element, calculations must be run at the Series level –just updating them from the parent won’t work unless they are completely overwritten.

<https://www.devtreks.org/greentreks/preview/carbon/input/DevTreks OLS 3/2147397536/none>

The following Operation uses the Input with the OC and AOH price adjustments.

<https://www.devtreks.org/greentreks/preview/carbon/operation/Example 1I, Regression/2091557278/none>

The following Component uses the Input with the CAP and AOH price adjustments.

<https://www.devtreks.org/greentreks/preview/carbon/component/CTA Regress 1/1194/none>

The following dataset is used with the regression analysis.

[https://devtreks1.blob.core.windows.net/resources/network\\_carbon/resourcepack\\_1534/resource\\_7973/Ex6eshipping.csv](https://devtreks1.blob.core.windows.net/resources/network_carbon/resourcepack_1534/resource_7973/Ex6eshipping.csv)



DevTreks –social budgeting that improves lives and livelihoods

[http://localhost:5000/greentreks/preview/carbon/input/Example 1hb,  
Regression/2147409823/none](http://localhost:5000/greentreks/preview/carbon/input/Example%201hb,Regression/2147409823/none)

#### Score datasets

[http://localhost:5000/resources/network\\_carbon/resourcepack\\_526/resource\\_1779/Ex6eshipping.csv](http://localhost:5000/resources/network_carbon/resourcepack_526/resource_1779/Ex6eshipping.csv);  
[http://localhost:5000/resources/network\\_carbon/resourcepack\\_526/resource\\_1777/Ex6dshipping.csv](http://localhost:5000/resources/network_carbon/resourcepack_526/resource_1777/Ex6dshipping.csv)

#### M and E datasets

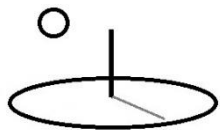
[http://localhost:5000/resources/network\\_carbon/resourcepack\\_526/resource\\_1881/Ex6eshipping.csv](http://localhost:5000/resources/network_carbon/resourcepack_526/resource_1881/Ex6eshipping.csv);  
[http://localhost:5000/resources/network\\_carbon/resourcepack\\_526/resource\\_1882/Ex6dshipping.csv](http://localhost:5000/resources/network_carbon/resourcepack_526/resource_1882/Ex6dshipping.csv)

This example extends Example 1i by calculating uncertain Input costs. The shipping cost dataset in that example will be used to demonstrate how to update an Input's quantitative properties. Two Indicators will be added to an Input to demonstrate how to update an Input's allocated overhead costs and operating costs. In the context of CTA, a logical way to use this technique is to calculate carbon prices and damages associated with climate change.

Only the Stock calculators can be used to update base Input and Output element properties. Although the M&E calculators include the BaseIO property, the current version does not use that property to update the underlying base element properties.

The following image demonstrates that the BaseIO property for the first Indicator is updating the Input.OCPrice. The amount of the Input will be changed once the Input is added to an Operation or Component and operating costs will be calculated by multiplying the fixed Input.OCPrice by the adjusted Input.Amount. The *Resource Stock and M&E Calculation* references recommend using "Unit Inputs and Outputs" so that these elements can be reused in any budget. In order to update the base Input with a "Unit OCPrice", the independent variables' amounts have been changed from the amounts used with Indicator 1 so that the Input.OCPrice corresponds to an

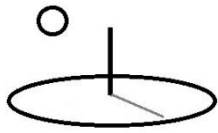




DevTreks –social budgeting that improves lives and livelihoods

Input.OCUnit of “shipping price per pound per mile”. That required changing the independent variables in the scoring dataset (the last 3 rows of data). For consistency, the five most important independent variables were added to the Qx properties, but these properties are only used in Resource Stock Totals Analyses.

Although not shown, the second Indicator had the exact same properties as the first Indicator except the Label matches its dataset and the BaseIO property is set to aohprice so that the Input.AOHPrice can be updated to calculate uncertain allocated overhead costs.



DevTreks –social budgeting that improves lives and livelihoods

← → ↺ DevTreks [US] <https://www.devtreks.org/greentreks/preview/carbon/i> ☆

Indicator 1

OC Price

Indicator 1 Description

Shipping cost is a function of package weight and distance shipped.

Indicator 1 URL

none

Label 1 Rel Label 1

P1 none

Date 1 Dist Type 1

02/13/2015 none

Q1 1 Q1 Unit 1

1.0000 pounds packagewt

Q2 1 Q2 Unit 1

1.0000 miles distance

Q3 1 Q3 Unit 1

1.0000 ptimesd

Q4 1 Q4 Unit 1

1.0000 psquared

Q5 1 Q5 Unit 1

1.0000 dsquared

Math Operator 1 BaseIO 1

equalto ocprice

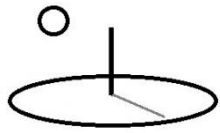
QT 1 QT Unit 1

0.2500 shipping cost



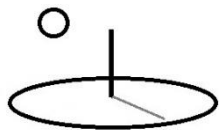
DevTreks –social budgeting that improves lives and livelihoods

Math Type 1		Math Sub Type 1		
algorithm1		subalgorithm6		
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		
0.3190		shipping cost		
QT Low 1		QT Low Unit 1		
-0.7274		lower 90% ci		
QT High 1		QT High Unit 1		
1.3653		upper 90% ci		
Math Expression 1				
I1.Q1.packagewt + I1.Q2.distancemiles + I1.Q3.ptimesd + I1.Q4.psquared + I1.Q5.dsqu				
Math Result 1				
regression results				
dependent variable: shippingcost				
source	df	SS	MS	
model	5	449.3408	89.8682	
error	14	2.7447	0.1961	
total	19	452.0855		
R-squared	0.9939	Adj R-squared	0.9918	
F value	458.3876	prob > F	0.0001	
variable	coefficient	stand error	T-ratio	prob > T
intercept	0.82702	0.7023	1.1776	0.2586
packagewt	-0.60914	0.1799	-3.3859	0.0044
distancemiles	0.00402	0.0080	0.5027	0.6230
ptimesd	0.00733	0.0006	11.4947	0.0001
psquared	0.08975	0.0202	4.4419	0.0006
dsquared	0.00002	0.0000	0.6718	0.5127
durbin-watson: 2.2811				
estimate	predicted	lower 90%	upper 90%	
QTM CI	0.3190	-0.7274	1.3653	
QTM PI	0.3190	-0.9860	1.6240	
row 21				



DevTreks –social budgeting that improves lives and livelihoods

The following image displays the updated \$0.32 Input OCPrice and \$0.32 Input.AOHPrice. The calculator automatically filled in these properties. No unit in the OC Unit list matches the actual “shipping price per pound per mile” unit so a default unit of “each” has been set.



DevTreks –social budgeting that improves lives and livelihoods

Dev... (US)https://www.devtrSearch

GreenTreks

Search

Preview

Select

Edit

Pack

Views

Club

Select

PackIt

Submit

Cancel

Close

DevTreks OLS 3

Input Series

0

+

Ex- 1J, AOH and CAP Costs

-

Ex- 1J, AOH and OC Costs

Ex- 1J, AOH and OC Costs

D

U

Date Changed:

2/13/2015

Label:

C1

Date:

2/13/2015

OC Amount:

1.000

OC Price:

0.3200

OC Unit:

each

AOH Price:

0.3200

AOH Unit:

shipping price per pound per mile

CAP Price:

0.0000

CAP Unit:

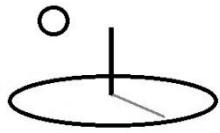
none or n/a

Description

Sample CTA using regression analysis.The actual OC Unit is shipping price per pound per mile.

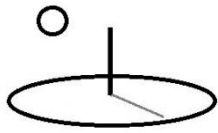
Edit Linked Views

Views



DevTreks –social budgeting that improves lives and livelihoods

The following image demonstrates that after this Input has been added to an Operation, the Input's OCAmount and AOHAmount are changed from 0 to 500 (i.e. 5 pounds \* 100 miles) to calculate the full Input costs. This is an important step in any technology assessment because it defines the exact nature of the technology.



DevTreks –social budgeting that improves lives and livelihoods

Dev... (US)https://www.devtrSearch

SubmitCancelClose

Example 1I, Regression

Select Input Series

Ex- 1J, AOH and OC Costs

Ex- 1J, AOH and OC Costs

D

U

Date:

2/13/2015

Label:

C1

Times:

1.000

InputUseAOHOnly?:

T

F

Incentive Amount:

0.000

Incentive Rate:

0.000

OC Amount:

500.000

OC Unit:

each

OC Price:

0.3200

Total:

160.000

AOH Amount:

500.000

AOH Unit:

shipping price per pound per mile

AOH Price:

0.3200

Total:

160.000

CAP Amount:

0.000

CAP Unit:

none or n/a

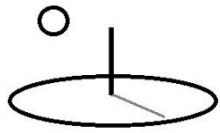
CAP Price:

0.0000

Total:

0.000

Description



DevTreks –social budgeting that improves lives and livelihoods

The following image displays a Resource Stock Total Analysis for an Operation containing this Input. The total operating costs for this input is the QT Most Amount shown for the first Indicator ( $\$159.50 = 500 * .32$ ). The uncertainty for this cost is defined using the QT Low (-477.65) and QT High (\$796.65) Amounts. The uncertainty of the Allocated Overhead costs is displayed in the second indicator.

The image also demonstrates that the Input.OCAmount is used as a stock multiplier when conducting Operation and Operating Budget Stock Analysis ( $Q_{xs} = 500$ ). For simplicity, all of the final calculated quantities are multiplied by 500. The Input.AOHAmount is never used as a stock multiplier. Certain types of allocated overhead costs may need to be dealt with differently in future upgrades.

Most persons using this analysis to make decisions about how to ship packages, or reduce climate change damages, will not be looking at this raw data, they'll be looking at the summarized tables and graphics of the analysis that have been referenced using the Media URL property. Those communication aids will include an explanation for all uncertain cost and benefits.





<a href="https://www.devtreks.org/greentreks/preview/carbon/operation/Ex">DevTreks [US] https://www.devtreks.org/greentreks/preview/carbon/operation/Ex</a>	
Indicators	
Math Expression: (I1.QTM + I2.QTM) Score Amount: 319.0000 Score D1 Amount: 320.0000 Score D2 Amount: 25.0000 Distribution Type: normal Score Most Amount: 319.5000 Score Low Amount: 318.5500 Score High Amount: 320.4500 Iterations: 2000 Score Math Result: sampled descriptive statistics N,Total,Mean,Median,StdDev,Var,Min,Max 2000, 1277.0.0503, 0.0025, 0.4549, 0.8186, sampled cumulative density function 0.00,0.10,0.20,0.30,0.40,0.50,0.60.0.4549,0.5763,0.5969,0.6122,0.6259,0.6385,0.6519,0.6654,0.6806,0.7040,0.8186 <b>Indic 1 Name:</b> OC Price Date: 02/13/2015 Math Type: algorithm1 Q1 Amount: 500.0000 Q2 Amount: 500.0000 Q3 Amount: 500.0000 Q4 Amount: 500.0000 Q5 Amount: 500.0000 Math Express: I1.Q1.packagewt + I1.Q2.distanceinmiles + I1.Q3.ptimesd + I1.Q4.psquared + I1.Q5.dsquared QT Amount: 125.0000 QT D1 Amount: 0.0000 QT D2 Amount: 0.0000 QT Most Amount: 159.4961 QT Low Amount: -363.6795 QT High Amount: 682.6718 Math Sub Type: subalgorithm6 Observations: 1.0 Indic 1 Description: Shipping cost is a function of package weight and distance shipped. <b>Indic 2 Name:</b> AOH Price Date: 02/05/2015 Math Type: algorithm1 Q1 Amount: 500.0000 Q2 Amount: 500.0000 Q3 Amount: 500.0000 Q4 Amount: 500.0000 Q5 Amount: 500.0000 Math Express: I2.Q1.packagewt + I2.Q2.distanceinmiles + I2.Q3.ptimesd + I2.Q4.psquared + I2.Q5.dsquared QT Amount: 125.0000 QT D1 Amount: 0.0000 QT D2 Amount: 0.0000 QT Most Amount: 159.4961 QT Low Amount: -363.6795 QT High Amount: 682.6718 Math Sub Type: subalgorithm6	Observations: 1.0 Score Unit: unit shipping cost Score D1 Unit: mean Score D2 Unit: standard deviation Math Type: algorithm1 Score Most Unit: unit shipping cost Score Low Unit: lower 90% ci Score High Unit: upper 90% ci Score Math Sub Type: subalgorithm1 Label: P1 Rel Label: none Dist Type: none Q1 Unit: pounds packagewt Q2 Unit: miles distance Q3 Unit: ptimesd Q4 Unit: psquared Q5 Unit: dsquared Math Operator: equalto QT Unit: shipping cost QT D1 Unit: none QT D2 Unit: none QT Most Unit: shipping cost QT Low Unit: lower 90% ci QT High Unit: upper 90% ci Base IO: ocprice Label: C1 Rel Label: none Type: none Q1 Unit: pounds packagewt Q2 Unit: miles distance Q3 Unit: ptimesd Q4 Unit: psquared Q5 Unit: dsquared Math Operator: equalto QT Unit: shipping cost QT D1 Unit: none QT D2 Unit: none QT Most Unit: shipping cost QT Low Unit: lower 90% ci QT High Unit: upper 90% ci Base IO: aohprice

The following image displays the associated Net Present Value calculation for this Operation.



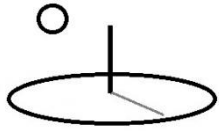
DevTreks –social budgeting that improves lives and livelihoods

Operation							
Date Applied	Label 1	Label 2	Amount	Eff. Life	Salv. Value	Incent. Amount	Incent. Rate
Example 1I, Regression (5/15/2015 12:00:00 AM)							
12/31/2015	none	none	1	1	0.0000	0.0000	0
Operation Unit:each			ResourceWeight	0	Rates (R and N)		0.0050 0.0300
Description	This operation is used in a CTA tutorial.						
Total Costs - Operation		Total Cost		Annual Cost		Interest Portion	
Total Operating Costs		164.27		164.27		4.27	
Total Allocated Overhead Costs		164.27		164.27		4.27	
Total Capital Costs		0.00		0.00		0.00	
Total Costs - Operation		328.54		328.54		8.54	
Total Costs - Operation w. Incentives		328.55		328.55			
Inputs							
Input Name	Date Applied	Times	Incent. Amount	Incent. Rate			
Ex- 1J, AOH and OC Costs							
	02/13/2015	1	0.0000	0			
Total Costs - Input		Amount	Unit	Price	Total	Interest	Total Cost
Total Operating Costs		500	each	0.3200	160.00	4.27	164.27
Total Allocated Overhead Costs		500	shipping price per pound per mile	0.3200	160.00	4.27	164.27
Total Capital Costs		0	none or n/a	0.0000	0.00	0.00	0.00
Total Costs with Incentives							328.55
Description	Sample CTA using regression analysis.The actual OC Unit is shipping price per pound per mile.						
<a href="#">Feedback About carbon/operation/Example 1I, Regression/2091557278/none</a>							

Search IRIs:

<https://www.devtreks.org/greentreks/linkedviews/carbon/operation/Example 1I, Regression/2091557278/none>

Two Indicators were added to a sibling Input Series to demonstrate how to update an Input's allocated overhead costs and capital costs. The only change made to the previous example is to



DevTreks –social budgeting that improves lives and livelihoods

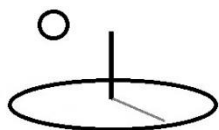
change Indicator 1's BaseIO property to caprice, rather than ocprice. The following Resource Stock Total Analysis and Net Present Value Analysis for a Component that uses this Input displays the resultant allocated overhead, capital costs, and stock totals. The Resource Stock Totals analysis for this Component is used to communicate the uncertainty of these costs.

The Resource Stock Totals Analysis also demonstrates that the Input.CAPAmount is used as a stock multiplier when conducting Component and Capital Budget Stock Analysis ( $Q_{xs} = 500$ ). The Input.AOHAmount is never used as a stock multiplier.



DevTreks –social budgeting that improves lives and livelihoods

<a href="https://www.devtreks.org/greentreks/preview/carbon/component/CTA%20Re">DevTreks [US]</a> <a href="https://www.devtreks.org/greentreks/preview/carbon/component/CTA%20Re">https://www.devtreks.org/greentreks/preview/carbon/component/CTA%20Re</a>	
Indicators	
Math Expression: (I1.QTM + I2.QTM) Score Amount: 319.0000 Score D1 Amount: 320.0000 Score D2 Amount: 25.0000 Distribution Type: normal Score Most Amount: 319.5000 Score Low Amount: 318.5500 Score High Amount: 320.4500 Iterations: 2000 Score Math Result: sampled descriptive statistics N,Total,Mean,Median,StdDev,Var,Min,Max 2000, 1277.9363, 0.630.0503, 0.0025, 0.4549, 0.8186, sampled cumulative density function 0.00,0.10,0.20,0.30,0.40,0.50,0.60,0.70,0.80, 0.4549,0.5763,0.5969,0.6122,0.6259,0.6385,0.6519,0.6654,0.6806,0.7040,0.8186 <b>Indic 1 Name:</b> CAP Price Date: 02/13/2015 Math Type: algorithm1 Q1 Amount: 500.0000 Q2 Amount: 500.0000 Q3 Amount: 500.0000 Q4 Amount: 500.0000 Q5 Amount: 500.0000 Math Express: I1.Q1.packagewt + I1.Q2.distancemiles + I1.Q3.ptimesd + I1.Q4.psquared + I1.Q5.dsquared QT Amount: 125.0000 QT D1 Amount: 0.0000 QT D2 Amount: 0.0000 QT Most Amount: 159.4961 QT Low Amount: -363.6795 QT High Amount: 682.6718 Math Sub Type: subalgorithm6 Observations: 1.0 Indic 1 Description: Shipping cost is a function of package weight and distance shipped.	Observations: 1.0 Score Unit: unit shipping cost Score D1 Unit: mean Score D2 Unit: standard deviation Math Type: algorithm1 Score Most Unit: unit shipping cost Score Low Unit: lower 90% ci Score High Unit: upper 90% ci Score Math Sub Type: subalgorithm1 Label: P1 Rel Label: none Dist Type: none Q1 Unit: pounds packagewt Q2 Unit: miles distance Q3 Unit: ptimesd Q4 Unit: psquared Q5 Unit: dsquared Math Operator: equalto QT Unit: shipping cost QT D1 Unit: none QT D2 Unit: none QT Most Unit: shipping cost QT Low Unit: lower 90% ci QT High Unit: upper 90% ci Base IO: caprice
<b>Indic 2 Name:</b> AOH Price Date: 02/05/2015 Math Type: algorithm1 Q1 Amount: 500.0000 Q2 Amount: 500.0000 Q3 Amount: 500.0000 Q4 Amount: 500.0000 Q5 Amount: 500.0000 Math Express: I2.Q1.packagewt + I2.Q2.distancemiles + I2.Q3.ptimesd + I2.Q4.psquared + I2.Q5.dsquared QT Amount: 125.0000 QT D1 Amount: 0.0000 QT D2 Amount: 0.0000 QT Most Amount: 159.4961 QT Low Amount: -363.6795 QT High Amount: 682.6718	Label: C1 Rel Label: none Type: none Q1 Unit: pounds packagewt Q2 Unit: miles distance Q3 Unit: ptimesd Q4 Unit: psquared Q5 Unit: dsquared Math Operator: equalto QT Unit: shipping cost QT D1 Unit: none QT D2 Unit: none QT Most Unit: shipping cost QT Low Unit: lower 90% ci QT High Unit: upper 90% ci



DevTreks –social budgeting that improves lives and livelihoods

Component							
Date Applied	Label 1	Label 2	Amount	Eff. Life	Salv. Value	Incent. Amount	Incent. Rate
CTA Regress 1 (5/15/2015 12:00:00 AM)							
12/31/2015	A1001	none	1	1	0.0000	0.0000	0
Component Unit:each			ResourceWeight	0	Rates (R and N)		0.0050 0.0300
Description	This component is used in a CTA tutorial.						
Total Costs - Component		Total Cost		Annual Cost		Interest Portion	
Total Operating Costs		0.00		0.00		0.00	
Total Allocated Overhead Costs		164.27		164.27		4.27	
Total Capital Costs		164.27		164.27		4.27	
Total Costs - Component		328.54		328.54		8.54	
Total Costs - Component w. Incentives		328.55		328.55			
Inputs							
Input Name	Date Applied	Times	Incent. Amount	Incent. Rate			
Example 1J, Regression							
	02/13/2015	1	0.0000	0			
Total Costs - Input		Amount	Unit	Price	Total	Interest	Total Cost
Total Operating Costs		0	none or n/a	0.0000	0.00	0.00	0.00
Total Allocated Overhead Costs		500	shipping price per pound per mile	0.3200	160.00	4.27	164.27
Total Capital Costs		500	none or n/a	0.3200	160.00	4.27	164.27
Total Costs with Incentives							328.55
Description	Sample CTA using regression analysis.The actual CAP Unit is shipping price per pound per mile.						
<a href="#">Feedback About carbon/component/CTA Regress 1/1194/none</a>							



DevTreks –social budgeting that improves lives and livelihoods

### **Example 6b. Algorithm 1. Subalgorithm 6. Probabilistic Statistics: Multiple Regression with 10 Explanatory Variables and Uncertain Performance Measures**

#### **URLs**

[https://www.devtreks.org/greentreks/preview/carbon/input/DevTreks OLS 4/2147397541/none](https://www.devtreks.org/greentreks/preview/carbon/input/DevTreks%20OLS%204/2147397541/none)

[https://devtreks1.blob.core.windows.net/resources/network\\_carbon/resourcepack\\_1534/resource\\_7974/ozone1.csv](https://devtreks1.blob.core.windows.net/resources/network_carbon/resourcepack_1534/resource_7974/ozone1.csv)

[https://devtreks1.blob.core.windows.net/resources/network\\_carbon/resourcepack\\_1534/resource\\_7975/Cost-Per-Unit-Pollution-cdf.PNG](https://devtreks1.blob.core.windows.net/resources/network_carbon/resourcepack_1534/resource_7975/Cost-Per-Unit-Pollution-cdf.PNG)

[http://localhost:5000/greentreks/preview/carbon/input/Example 1hc,  
Regression/2147409830/none](http://localhost:5000/greentreks/preview/carbon/input/Example%201hc,Regression/2147409830/none)

#### **Datasets**

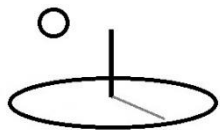
[http://localhost:5000/resources/network\\_carbon/resourcepack\\_526/resource\\_1778/ozone1.csv](http://localhost:5000/resources/network_carbon/resourcepack_526/resource_1778/ozone1.csv)

[http://localhost:5000/resources/network\\_carbon/resourcepack\\_526/resource\\_1880/ozone1.csv](http://localhost:5000/resources/network_carbon/resourcepack_526/resource_1880/ozone1.csv)

This example demonstrates how to include up to 10 explanatory variables in a regression and to calculate the uncertainty of Performance Measures. The 10 variable limit is arbitrary –the R and Python algorithms don’t impose this restriction. The following technique can be used when using more than 5 explanatory variables in an analysis:

- 1. Q6 to Q10 terms in Math Expressions:** Although calculators don’t contain Q6 to Q10 properties, these terms can still be used in Math Expressions to identify columns of data to include in analyses.

The following dataset contains 10 explanatory variables and 330 observations. The dependent variable is the amount of ozone measured in the atmosphere. The explanatory variables are



DevTreks –social budgeting that improves lives and livelihoods

weather and other factors influencing ozone concentrations. This algorithm uses the last 3 rows of data in the TEXT Data URL dataset to score the statistical model. The remaining rows are used to train the statistical model. The last row of data is used to fill in the QTM, QTL, and QTU properties. This is the same way that algorithms 2 and 3 use datasets.

Indicator.URL datasets:

```
label,date,location,O3,vh,wind,humidity,temp,ibh,dpg,ibt,vis,doy,ampm
1,5/15/2015,N45'37.75W121'46.25,3,5710,4,28,40,2693,-25,87,250,33,0
2,5/16/2015,N45'37.75W121'46.26,5,5700,3,37,45,590,-24,128,100,34,0
3,5/17/2015,N45'37.75W121'46.27,5,5760,3,51,54,1450,25,139,60,35,0
4,5/18/2015,N45'37.75W121'46.28,6,5720,4,69,35,1568,15,121,60,36,0
```

Score.DataURL datasets (label column must correspond to Indicator.Label for Stocks or Indicator index position for M&E):

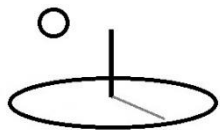
```
label,date,location,O3,vh,wind,humidity,temp,ibh,dpg,ibt,vis,doy,ampm
O3,5/15/2015,N45'37.75W121'46.25,3,5710,4,28,40,2693,-25,87,250,33,0
O3,5/16/2015,N45'37.75W121'46.26,5,5700,3,37,45,590,-24,128,100,34,0
O3,5/17/2015,N45'37.75W121'46.27,5,5760,3,51,54,1450,25,139,60,35,0
O3,5/18/2015,N45'37.75W121'46.28,6,5720,4,69,35,1568,15,121,60,36,0
```

The Math Expression for Indicator1 identifies the 10 columns of data to include in the analysis.

.

$$I1.Q1.vh + I1.Q2.wind + I1.Q3.humidity + I1.Q4.temp + I1.Q5.ibh + I1.Q6.dpg + I1.Q7.ibt + I1.Q8.vis + I1.Q9.doy + I1.Q10.ampm$$

The following image displays the results of this regression. In the context of CTA, a logical way to use this technique is to calculate the uncertainty of a Pollution Index (Score) that tracks several types of emission Indicators.



DevTreks –social budgeting that improves lives and livelihoods

To demonstrate using this technique to communicate the uncertainty of Performance Measures, a cost Indicator (I2) value from the previous example has been added to the Score.MathExpression to calculate uncertain operating costs. In practice, the cost Indicator is included as a separate Indicator and the Score.MathExpression includes the calculated cost. The following image demonstrates that the Score has been used to calculate the uncertainty of a Performance Measure –in this case, Cost per Unit Pollution Index. The Math Expression shows that the Cost Indicator (I2) is being divided by the Emissions Indicator (I1).





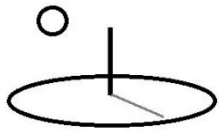
DevTreks –social budgeting that improves lives and livelihoods

QT 1	QT Unit 1																																																																																																										
<input type="text" value="5.0000"/>	<input type="text" value="shipping cost"/>																																																																																																										
Math Type 1	Math Sub Type 1																																																																																																										
<input type="text" value="algorithm1"/>	<input type="text" value="subalgorithm6"/>																																																																																																										
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="3.6848"/>	<input type="text" value="shipping cost"/>																																																																																																										
QT Low 1	QT Low Unit 1																																																																																																										
<input type="text" value="2.3418"/>	<input type="text" value="lower 80 % ci"/>																																																																																																										
QT High 1	QT High Unit 1																																																																																																										
<input type="text" value="5.0279"/>	<input type="text" value="upper 80 % ci"/>																																																																																																										
Math Expression 1																																																																																																											
<input type="text" value="I1.Q1.vh + I1.Q2.wind + I1.Q3.humidity + I1.Q4.temperature"/>																																																																																																											
Math Result 1																																																																																																											
<p>regression results</p> <p>dependent variable: O3</p> <table border="1"> <thead> <tr> <th>source</th> <th>df</th> <th>SS</th> <th>MS</th> </tr> </thead> <tbody> <tr> <td>model</td> <td>10</td> <td>14640.7379</td> <td>1464.0738</td> </tr> <tr> <td>error</td> <td>316</td> <td>6233.5068</td> <td>19.7263</td> </tr> <tr> <td>total</td> <td>326</td> <td>20874.2446</td> <td></td> </tr> </tbody> </table> <p>R-squared 0.7014      Adj R-squared 0.6919</p> <p>F value 74.2194      prob &gt; F 0.0001</p> <table border="1"> <thead> <tr> <th>variable</th> <th>coefficient</th> <th>stand error</th> <th>T-ratio</th> <th>prob &gt; T</th> </tr> </thead> <tbody> <tr> <td>intercept</td> <td>10.79053</td> <td>30.3552</td> <td>0.3555</td> <td>0.7225</td> </tr> <tr> <td>vh</td> <td>-0.00407</td> <td>0.0055</td> <td>-0.7378</td> <td>0.4612</td> </tr> <tr> <td>wind</td> <td>-0.03334</td> <td>0.1357</td> <td>-0.2457</td> <td>0.8061</td> </tr> <tr> <td>humidity</td> <td>0.08781</td> <td>0.0193</td> <td>4.5537</td> <td>0.0001</td> </tr> <tr> <td>temp</td> <td>0.27403</td> <td>0.0498</td> <td>5.4975</td> <td>0.0001</td> </tr> <tr> <td>ibh</td> <td>-0.00014</td> <td>0.0003</td> <td>-0.4539</td> <td>0.6502</td> </tr> <tr> <td>dpg</td> <td>-0.00589</td> <td>0.0114</td> <td>-0.5185</td> <td>0.6044</td> </tr> <tr> <td>ibt</td> <td>0.03201</td> <td>0.0137</td> <td>2.3393</td> <td>0.0199</td> </tr> <tr> <td>vis</td> <td>-0.00725</td> <td>0.0038</td> <td>-1.9060</td> <td>0.0576</td> </tr> <tr> <td>doy</td> <td>-0.00108</td> <td>0.0050</td> <td>-0.2154</td> <td>0.8296</td> </tr> <tr> <td>ampm</td> <td>-2.08092</td> <td>1.1051</td> <td>-1.8830</td> <td>0.0606</td> </tr> </tbody> </table> <p>durbin-watson: 1.4775</p> <table border="1"> <thead> <tr> <th>estimate</th> <th>predicted</th> <th>lower 80%</th> <th>upper 80%</th> </tr> </thead> <tbody> <tr> <td>QTM CI</td> <td>3.6848</td> <td>2.3418</td> <td>5.0279</td> </tr> <tr> <td>QTM PI</td> <td>3.6848</td> <td>-2.1750</td> <td>9.5446</td> </tr> </tbody> </table> <p>row 328</p> <table border="1"> <thead> <tr> <th></th> <th>CI</th> <th>PI</th> </tr> </thead> <tbody> <tr> <td>CI</td> <td>6.0204</td> <td>5.0205</td> </tr> <tr> <td>PI</td> <td>6.0204</td> <td>0.2296</td> </tr> </tbody> </table> <p>row 329</p> <table border="1"> <thead> <tr> <th></th> <th>CI</th> <th>PI</th> </tr> </thead> <tbody> <tr> <td>CI</td> <td>3.4885</td> <td>2.3364</td> </tr> <tr> <td>PI</td> <td>3.4885</td> <td>-2.3305</td> </tr> </tbody> </table>		source	df	SS	MS	model	10	14640.7379	1464.0738	error	316	6233.5068	19.7263	total	326	20874.2446		variable	coefficient	stand error	T-ratio	prob > T	intercept	10.79053	30.3552	0.3555	0.7225	vh	-0.00407	0.0055	-0.7378	0.4612	wind	-0.03334	0.1357	-0.2457	0.8061	humidity	0.08781	0.0193	4.5537	0.0001	temp	0.27403	0.0498	5.4975	0.0001	ibh	-0.00014	0.0003	-0.4539	0.6502	dpg	-0.00589	0.0114	-0.5185	0.6044	ibt	0.03201	0.0137	2.3393	0.0199	vis	-0.00725	0.0038	-1.9060	0.0576	doy	-0.00108	0.0050	-0.2154	0.8296	ampm	-2.08092	1.1051	-1.8830	0.0606	estimate	predicted	lower 80%	upper 80%	QTM CI	3.6848	2.3418	5.0279	QTM PI	3.6848	-2.1750	9.5446		CI	PI	CI	6.0204	5.0205	PI	6.0204	0.2296		CI	PI	CI	3.4885	2.3364	PI	3.4885	-2.3305
source	df	SS	MS																																																																																																								
model	10	14640.7379	1464.0738																																																																																																								
error	316	6233.5068	19.7263																																																																																																								
total	326	20874.2446																																																																																																									
variable	coefficient	stand error	T-ratio	prob > T																																																																																																							
intercept	10.79053	30.3552	0.3555	0.7225																																																																																																							
vh	-0.00407	0.0055	-0.7378	0.4612																																																																																																							
wind	-0.03334	0.1357	-0.2457	0.8061																																																																																																							
humidity	0.08781	0.0193	4.5537	0.0001																																																																																																							
temp	0.27403	0.0498	5.4975	0.0001																																																																																																							
ibh	-0.00014	0.0003	-0.4539	0.6502																																																																																																							
dpg	-0.00589	0.0114	-0.5185	0.6044																																																																																																							
ibt	0.03201	0.0137	2.3393	0.0199																																																																																																							
vis	-0.00725	0.0038	-1.9060	0.0576																																																																																																							
doy	-0.00108	0.0050	-0.2154	0.8296																																																																																																							
ampm	-2.08092	1.1051	-1.8830	0.0606																																																																																																							
estimate	predicted	lower 80%	upper 80%																																																																																																								
QTM CI	3.6848	2.3418	5.0279																																																																																																								
QTM PI	3.6848	-2.1750	9.5446																																																																																																								
	CI	PI																																																																																																									
CI	6.0204	5.0205																																																																																																									
PI	6.0204	0.2296																																																																																																									
	CI	PI																																																																																																									
CI	3.4885	2.3364																																																																																																									
PI	3.4885	-2.3305																																																																																																									



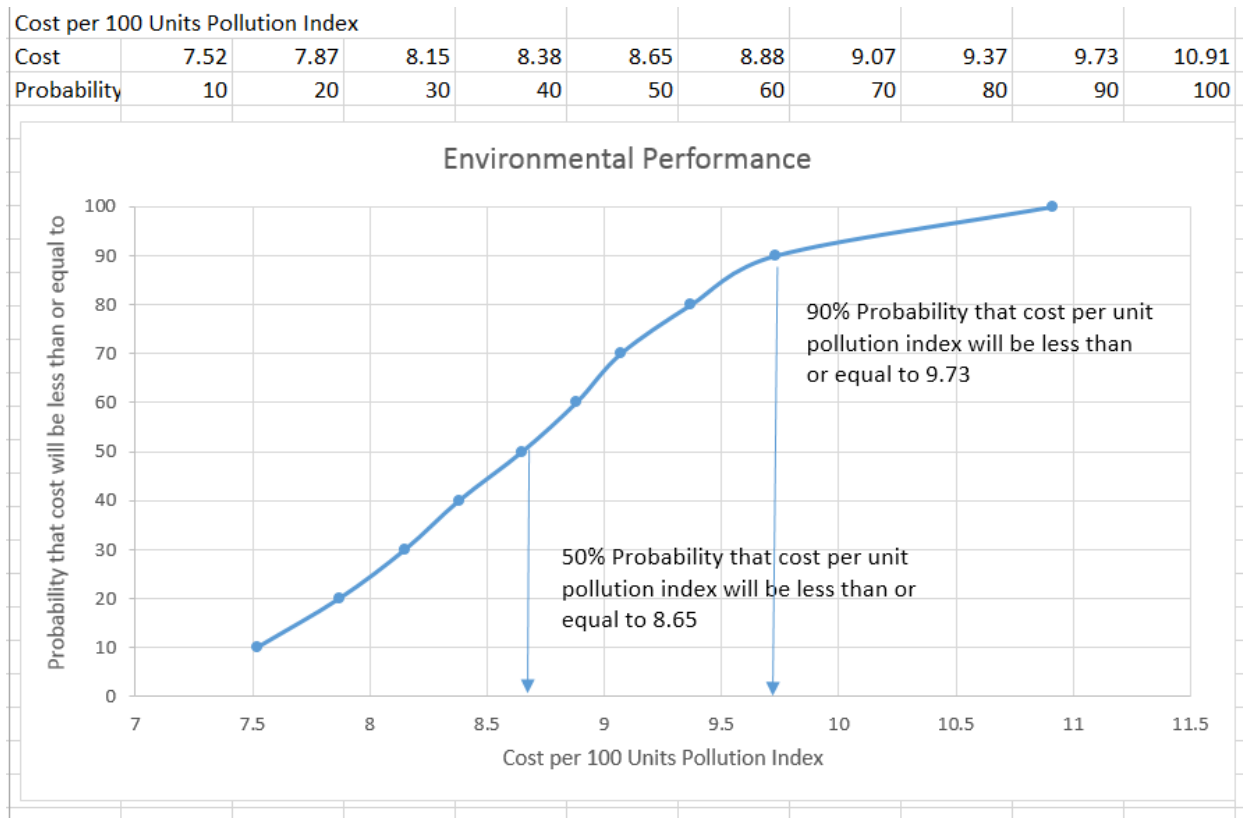
DevTreks –social budgeting that improves lives and livelihoods

<b>Score</b>			
Shipping Score		Most Likely	Most Unit
Score Math Expression		1.4292	estimated unit cost
I1.QT / I1.QTM		Low Estimate	Low Unit
Label		1.3989	lower 80 % ci
Rel Label		High Estimate	High Unit
S1	O1	1.4595	upper 80 % ci
Total Score	Score Unit	Math Type	Math Sub Type
1.3569	estimated unit cost per	algorithm1	subalgorithm1
D1	D1 Unit	Math Result	
1.4000	mean	sampled descriptive statistics N, Total, Mean, Median, StdDev, Var, Min, Max	
D2	D2 Unit	Score URL	
0.7500	sd	none	
Date	Dist Type	Calculations Description	
11/19/2016	normal	.v216d	
Iterations	Confidence Interval	Media URL	
1000	80	http://localhost:5000/resources/network_carbon/resourcepack_166/resource_314/EPALCA1.PNG	
Random Seed	BaseIO	Data URL	
6	none	none	
Math Operator			
equalto			
Most Likely	Most Unit		
1.4292	estimated unit cost		



DevTreks –social budgeting that improves lives and livelihoods

The following image demonstrates one way communicate the results of this type of Performance Measure (the GAO 2009 and NASA 2011 references explain using this type of cumulative density function to communicate the uncertainty of costs to decision makers).





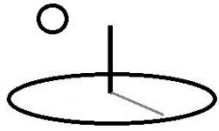
### **Example 8. Algorithm 1. Subalgorithm 8. Differences among Means: Analysis of Variance (ANOVA)**

This algorithm analyzes whether different subgroups of experimental data have statistically significant differences among their means. It also generates confidence intervals showing the exact difference between the means of specific subgroups of data. More explanation is offered for this algorithm because it introduces the analysis of randomized experimental data –an important characteristic of many CTAs. Three different experimental data designs can be used with the algorithm:

1. **Example 1ma. Completely randomized data:** The explanatory variables, or factors, are categorized into different levels, or treatments. The null hypothesis that the treatment means are equal is tested against the hypothesis that at least two of the means differ.
2. **Example 1mb. Randomized block data:** Besides treatments, blocks are used to further subdivide the data being analyzed. The null hypothesis that the treatment and/or block means are equal is tested against the hypothesis that at least two of the means differ.
3. **Example 1mc. Randomized factorial data:** Besides treatments and blocks, usually referred to as factors and levels, interactions between all factors are used to further analyze the data. The null hypothesis that the factor, level, and interaction effects means are equal is tested against the hypothesis that at least two of the means differ.
4. **Other random experimental data:** Examples of additional ANOVA techniques such as split-plot and covariance, are not included yet.

Each of the following 3 examples are taken from a college text book (Mendenhall and Sincich, 1989) demonstrating ANOVA. Because the Score and Indicator properties are similar to the regression algorithm (subalgorithm6), only selected properties for the 3 examples will be presented. All of the examples demonstrate running the first 3 ANOVAs:

1. **Method 1. Standard ANOVA:** This method carries out the analysis without using matrix mathematics and requires uniform sizes for treatments (and blocks, factors, levels). The primary test statistics for discerning differences among means are F statistics.



DevTreks –social budgeting that improves lives and livelihoods

2. **Method 2. Regression ANOVA:** This method carries out the analysis using regression analysis. The test statistics for discerning differences among means are F statistics and coefficient t-ratios.

As of 1.9.2, the following 2 features are being retained, but not debugged, until more advanced RCT algorithms are developed that can either replace, or enhance, this feature.

3. **[Method 3. Resource Stock Analysis ANOVA:** Unlike Method 1 and Method 2, the treatments and factors being analyzed are contained in different base elements. The data from all of the base elements being analyzed is combined into 1 data file and that file is used to carry out the analysis. Unlike standard Resource Stock Analysis, which analyzes differences in totals, these analyses examine the statistical differences among the treatment and factor means.
4. **Method 4. DevPacks Resource Stock Analysis ANOVA:** Version 1.9.2 deemed this less important than other algorithms at this time (i.e. CTA Prevention), so debugging is put off until the next release. Appendix B, DevPacks Stock Analysis, in the *Resource Stock Analysis* reference demonstrates how to run these same analyses using DevPacks.]



DevTreks –social budgeting that improves lives and livelihoods

### Example 8a. Completely Randomized Data

#### URLs:

<https://www.devtreks.org/greentreks/preview/carbon/input/Example 1ma, Anova, Complete Randomized/2147397543/none>

[https://devtreks1.blob.core.windows.net/resources/network\\_carbon/resourcepack\\_1534/resource\\_7983/Anova1.csv](https://devtreks1.blob.core.windows.net/resources/network_carbon/resourcepack_1534/resource_7983/Anova1.csv)

<https://www.devtreks.org/greentreks/preview/carbon/input/Example 1ma2, Anova, Complete Randomized/2147397544/none>

<https://www.devtreks.org/greentreks/preview/carbon/input/Example 1ma3, Anova, Complete Randomized/2147397545/none>

<http://localhost:5000/greentreks/preview/carbon/input/Example 1ma, Anova, Complete Randomized/2147409832/none>

Datasets (Version 2.1.6 no longer requires Indicator.Labels in the 1<sup>st</sup> column of data when the Indicator.URLs hold the datasets)

[http://localhost:5000/resources/network\\_carbon/resourcepack\\_526/resource\\_1784/Anova1.csv](http://localhost:5000/resources/network_carbon/resourcepack_526/resource_1784/Anova1.csv)

[http://localhost:5000/resources/network\\_carbon/resourcepack\\_526/resource\\_1877/Anova1.csv](http://localhost:5000/resources/network_carbon/resourcepack_526/resource_1877/Anova1.csv)

#### DOCUMENT CHANGE TO INDICATOR.URL

**Method 1: Standard ANOVA:** The following dataset shows a subset of the amount of debt owed (the y column) by delinquent credit card customers for 3 different income groups: A = under \$12,000, B = \$12,000 to \$25,000, and C = over \$25,000. The debts have been randomly selected from the 3 different groups. This algorithm tests whether the 3 different groups owe



DevTreks –social budgeting that improves lives and livelihoods

significantly different mean amounts of debts. The 3 income groups are referred to as treatments or levels. This algorithm runs a standard ANOVA when the following conventions are followed:

1. **Treatments:** The column name for the fifth column of data must be “treatment”. The data in the column must be doubles (or integers) that distinguish each treatment.
2. **Training Data:** The column name for the fourth column of data can follow standard conventions for dependent, or output, variable data. The data in the column must be doubles that contain the treatment observed data.
3. **Scoring Data:** In order to stay consistent with other algorithms, the final 3 rows of data must be scoring data. These 3 rows are not currently used by this algorithm.
4. **Indicator Math Expression:** The expression should be identical to regression analysis expressions (I1.Q1.treatment)

Stock dataset (The label column can be row identifiers when Indicator.URLs are used)

label,date,income,y,treatment

D1,12/3/2015,12,148,1

D1,12/4/2015,12,76,1

D1,12/5/2015,12,393,1

D1,12/6/2015,12,520,1

D1,12/7/2015,12,236,1

D1,12/8/2015,12,134,1

D1,12/9/2015,12,55,1

D1,12/10/2015,12,166,1

D1,12/11/2015,12,415,1

D1,12/12/2015,12,153,1

D1,12/3/2015,12to25,513,2

D1,12/4/2015,12to25,264,2

M&E dataset (the same convention is followed in the remaining ANOVA datasets)

index,date,income,y,treatment



DevTreks –social budgeting that improves lives and livelihoods

1,12/3/2015,12,148,1

1,12/4/2015,12,76,1

The following image shows the Math Result for the Indicator associated with this dataset. The results for the statistical tests and confidence intervals are explained in the following list. These results matched the reference text book.

1. **F Test:** The F Test Statistic, 3.48, is greater than the critical value for F at  $.05 = 3.35$ . The means are significantly different (at Score.ConfidenceLevel = 95). The p-value for the F statistic in the regression analysis also verifies this statistical significance.
2. **Treatment 1 Mean (< 12,000 debt):** The first row of the ANOVA confidence interval shows the mean for treatment 1, 229.6, lies somewhere in the interval 119.98 and 339.2. The terms used in the confidence interval, “base” and “xminus1” derive from the Change by Resource Stock analyzers and Method 3, below, demonstrates how they get displayed during a Change by Id Resource Stock Analysis.
3. **Treatment 2 Mean (12,000 to 25,000 debt) - Treatment 1 Mean (< 12,000 debt):** The second and third row of the ANOVA confidence interval shows the mean for the second treatment minus the mean for the first treatment, 80.3, lies somewhere in the interval - 74.73 and 235.33.
4. **Treatment 3 Mean (> 25,000 debt) - Treatment 2 Mean (12,000 to 25,000 debt):** The third row of the ANOVA confidence interval shows the mean for the third treatment minus the mean for the second treatment, 117.90, lies somewhere in the interval -37.13 and 272.93.
5. **Treatment 3 Mean (> 25,000 debt) - Treatment 1 Mean (< 12,000 debt):** The fifth row of the ANOVA confidence interval shows the mean for the third treatment minus the mean for the second treatment, 198.20, lies somewhere in the interval 43.17 and 353.23.





Indicator 1

Credit Card Debt

Indicator 1 Description

This indicator is used in a CTA Tutorial

Indicator 1 URL

http://localhost:5000/resources/network\_carbon/resourcepack\_526/resource\_1784/Anova1.csv

Label 1

DEBT1

Rel Label 1

none

Date 1

06/18/2015

Dist Type 1

none

Q1 1

1.0000

Q1 Unit 1

income category

Q2 1

1.0000

Q2 Unit 1

income category

Q3 1

0.0000

Q3 Unit 1

none

Q4 1

0.0000

Q4 Unit 1

none

Q5 1

0.0000

Q5 Unit 1

none

Math Operator 1

equalto

BaseIO 1

none

200.0000

guessed debt

Math Type 1

algorithm1

Math Sub Type 1

subalgorithm8

QT D1 1

0.0000

QT D1 Unit 1

none

QT D2 1

0.0000

QT D2 Unit 1

none

QT Most 1

198.2000

QT Most Unit 1

mean debt

QT Low 1

43.1724

QT Low Unit 1

lower 95 % ci

QT High 1

353.2276

QT High Unit 1

upper 95 % ci

Math Expression 1

I1.Q1.treatment

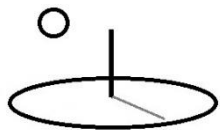
Math Result 1

anova results

source	df	SS	MS	F
treats	2	198772.4667	99386.2333	3.4819
error	27	770670.9000	28543.3667	
total	29	969443.3667		

estimate	mean diff	lower 95%	upper 95%
Treat 1 Mean	229.6000	119.9789	339.2211
xminus1 1	80.3000	-74.7276	235.3276
base 1	80.3000	-74.7276	235.3276
xminus1 2	117.9000	-37.1276	272.9276
base 2	198.2000	43.1724	353.2276

**Method 2: Regression ANOVA:** The following dataset changes the column names from “treatment” to any other acceptable name (such as x2 and x3). This tells subalgorithm8 to run the



DevTreks –social budgeting that improves lives and livelihoods

model as a regression analysis. The column x2 is coded 1 when the treatment, or income group, is 12,000 to 25,000, otherwise its coded 0. The column x3 is coded 1 when the treatment, or income group, is over 25,000, otherwise its coded 0.

Stock dataset (M&E datasets use the conventions explained for DataURL datasets)

label,date,income,y,x2,x3

D1,12/3/2015,12,148,0,0

D1,12/4/2015,12,76,0,0

D1,12/5/2015,12,393,0,0

D1,12/6/2015,12,520,0,0

D1,12/7/2015,12,236,0,0

D1,12/8/2015,12,134,0,0

D1,12/9/2015,12,55,0,0

D1,12/10/2015,12,166,0,0

D1,12/11/2015,12,415,0,0

D1,12/12/2015,12,153,0,0

D1,12/3/2015,12to25,513,1,0

D1,12/4/2015,12to25,264,1,0

D1,12/5/2015,12to25,433,1,0

D1,12/6/2015,12to25,94,1,0

D1,12/7/2015,12to25,535,1,0

D1,12/8/2015,12to25,327,1,0

D1,12/9/2015,12to25,214,1,0

D1,12/10/2015,12to25,135,1,0

D1,12/11/2015,12to25,280,1,0

D1,12/12/2015,12to25,304,1,0

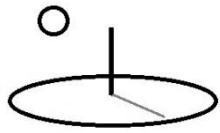
D1,12/13/2015,25,335,0,1

D1,12/14/2015,25,643,0,1

D1,12/15/2015,25,216,0,1

D1,12/16/2015,25,536,0,1

D1,12/17/2015,25,128,0,1



DevTreks –social budgeting that improves lives and livelihoods

D1,12/18/2015,25,723,0,1

D1,12/19/2015,25,258,0,1

D1,12/20/2015,25,380,0,1

D1,12/21/2015,25,594,0,1

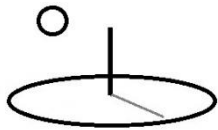
D1,12/22/2015,25,465,0,1

D1,12/23/2015,12,500,0,0

D1,12/24/2015,12to25,250,1,0

D1,12/25/2015,25,375,0,1

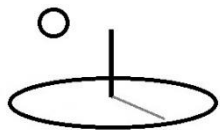
The following results demonstrate that the regression ANOVA returns the same results as the standard ANOVA.



DevTreks –social budgeting that improves lives and livelihoods

Math Type 1		Math Sub Type 1		
algorithm1		subalgorithm8		
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		
229.6000		mean debt		
QT Low 1		QT Low Unit 1		
119.9789		lower 95 % ci		
QT High 1		QT High Unit 1		
339.2211		upper 95 % ci		
Math Expression 1				
I1.Q1.x2 + I1.Q2.x3				
Math Result 1				
regression results				
dependent variable: y				
source	df	SS	MS	
model	2	198772.4667	99386.2333	
error	27	770670.9000	28543.3667	
total	29	969443.3667		
R-squared	0.2050	Adj R-squared	0.1462	
F value	3.4819	prob > F	0.0452	
variable	coefficient	stand error	T-ratio	prob > T
intercept	229.60000	53.4260	4.2975	0.0002
x2	80.30000	75.5558	1.0628	0.2973
x3	198.20000	75.5558	2.6232	0.0141

**Method 3: Resource Stock Analysis ANOVA:** As of 1.9.2, the following feature is being retained, but not debugged, until more advanced RCT algorithms are developed that can either



DevTreks –social budgeting that improves lives and livelihoods

replace, or enhance, this feature. The following 3 datasets are subsets of the data used in the standard ANOVA. Each dataset has been added to 3 sibling Input Series. Each Input Series and dataset represent a separate experimental treatment. The 3 rows of scoring data must be added to the last dataset (Series 3) in these types of analyses.

Stock dataset (M&E datasets use the conventions explained for DataURL datasets)

#### Input Series 1

label,date,income,y,treatment

D1,12/3/2015,12,148,1

D1,12/4/2015,12,76,1

D1,12/5/2015,12,393,1

D1,12/6/2015,12,520,1

D1,12/7/2015,12,236,1

D1,12/8/2015,12,134,1

D1,12/9/2015,12,55,1

D1,12/10/2015,12,166,1

D1,12/11/2015,12,415,1

D1,12/12/2015,12,153,1

#### Input Series 2

label,date,income,y,treatment

D1,12/3/2015,12to25,513,2

D1,12/4/2015,12to25,264,2

...

Input Series 3 (the 3 rows of scoring data must be added to the last dataset but are not currently used)

label,date,income,y,treatment

D1,12/13/2015,25,335,3

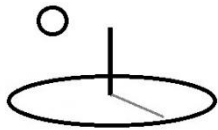


DevTreks –social budgeting that improves lives and livelihoods

D1,12/14/2015,25,643,3

...

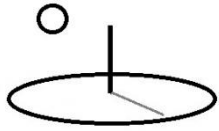
The following image shows that Input Series 1, Indicator 1's, properties have been set in a manner to a) cancel out the non-mathematical treatment column ( $I1.Q1.treatment - 1$ ) or ( $I1.Q1.treatment * 0$ ), and b) use the dataset's y (+ I1.QT), or dependent variable, column, to generate descriptive statistics. The properties of Input Series 2 and 3 were set in a similar manner (but subtracting 2 and 3 respectively in the expression or multiplying by 0).



DevTreks –social budgeting that improves lives and livelihoods

Math Operator 1	BaseIO 1
<input type="text" value="equalto"/>	<input type="text" value="none"/>
QT 1	QT Unit 1
<input type="text" value="229.6000"/>	<input type="text" value="mean debt"/>
Math Type 1	Math Sub Type 1
<input type="text" value="none"/>	<input type="text" value="none"/>
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="229.6000"/>	<input type="text" value="mean debt"/>
QT Low 1	QT Low Unit 1
<input type="text" value="147.0570"/>	<input type="text" value="lower 90% ci"/>
QT High 1	QT High Unit 1
<input type="text" value="312.1430"/>	<input type="text" value="upper 90% ci"/>
Math Expression 1	
<input type="text" value="(I1.Q1.treatment - 1) + I1.QT"/>	
Math Result 1	
<p>observed cumulative density function</p> <p>0.20,0.30,0.40,0.50,0.60,0.70,0.80,0.90,1.00,1.00</p> <p>76.0000,134.0000,148.0000,153.0000,166.0000,236.0000,393.0000,415.0000,520.0000,520.0000</p> <p>observed descriptive statistics</p> <p>N,Total,Mean,Median,StdDev,Var,Min,Max</p> <p>10,2296.0000,229.6000,159.5000,158.1962,25026.0444,55.0000,520.0000,</p> <p>observed means</p> <p>QT mean = 229.6, Q1 mean = 1,</p>	

The following image demonstrates that running a Change by Id Resource Stock Analysis required setting the following analyzer properties (new in 1.8.8):



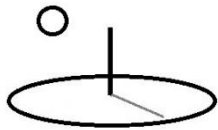
DevTreks –social budgeting that improves lives and livelihoods

- **Math Type** = algorithm1,
- **Math Sub Type** – subalgorithm8,
- **Math Expression** = I1.QTM.treatment.
- **Confidence Interval** = 95

The analysis fills in the following properties:

- **Data** = the data from each dataset is combined behind the scenes into 1 data file. All datasets are assumed to have the same column names and data content context.
- **Data Column Names** = the column names for the combined dataset is taken from the first dataset.
- **Math Result:** depending on the column names, displays the same results as the standard or regression ANOVA





DevTreks –social budgeting that improves lives and livelihoods

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

**Step 2 of 3. Analyze**

**+ Relations**

**What If Tag**

Base Resource (Input) Calculations To Analyze Type:  

Resource Stock Calculator 1

**Compare Using:**

**Math Type**

**Math Sub Type**

**Confidence Interval**

**Math Expression**

**Math Result**

anova results
 

source	df	SS	MS	F
treats	2	198772.4667	99386.2333	3.4819
error	27	770670.9000	28543.3667	
total	29	969443.3667		

estimate	mean diff	lower 95%	upper 95%
Treat 1 Mean	229.6000	119.9789	339.2211
xminus1 1	80.3000	-74.7276	235.3276
base 1	80.3000	-74.7276	235.3276
xminus1 2	117.9000	-37.1276	272.9276
base 2	198.2000	43.1724	353.2276



DevTreks –social budgeting that improves lives and livelihoods

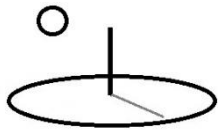
The following image demonstrates that running a Change by Id Resource Stock Analysis returns the identical Score properties as standard Resource Stock Analyses, but the Indicators compare the differences between the statistical means for each treatment. They use the same calculations as the ones used to set the confidence intervals for the standard ANOVA and regression ANOVA, but use the standard comparators used by the Resource Stock analyzers. They also use the same stylesheet as the standard analyses, which is interpreted as follows:

#### 1<sup>st</sup> Column

- Indicator.Total = Indicator.Mean
- Indicator.AmountChange = F Statistic for all Treatments
- Indicator.PercentChange = F Critical Value for all Treatments
- Indicator.BaseChange = Indicator.Mean
- Indicator.BasePercentChange = plus or minus confidence interval for the base mean alone

#### Remaining Columns

- Indicator.Total = Indicator.Mean
- Indicator.AmountChange = current Indicator Mean – xminus1 Indicator Mean
- Indicator.PercentChange = plus or minus confidence interval for the difference between the xminus1 means
- Indicator.BaseChange = current Indicator Mean – base Indicator Mean
- Indicator.BasePercentChange = plus or minus confidence interval for the difference between the base means



DevTreks –social budgeting that improves lives and livelihoods

<b>Score High Amount</b>	229.8853	309.8393	429.7934
<b>Score High Unit</b>	upper 90% ci	upper 90% ci	upper 90% ci
<b>Score High Amount Change</b>	0.0000	0.0000	119.9541
<b>Score High Percent Change</b>	0.0000	0.0000	38.7149
<b>Score High Base Change</b>	0.0000	79.9540	199.9081
<b>Score High Base Percent Change</b>	0.0000	34.7800	86.9599
<b>Indicator Observations</b>	1.0	1.0	1.0
<b>Name</b>	Anova randomized data	Anova randomized data	Anova randomized data
<b>Label</b>	D1	D1	D1
<b>Total</b>	229.6000	309.9000	427.8000
<b>Unit</b>	mean debt	mean debt	mean debt
<b>Amount Change</b>	3.4819	0.0000	117.9000
<b>Percent Change</b>	3.3541	0.0000	155.0276
<b>Base Change</b>	229.6000	80.3000	198.2000
<b>Base Percent Change</b>	109.6211	155.0276	155.0276
<a href="#">Feedback About carbon/input/Example 1ma3, Anova, Complete Randomized/2147409839/none</a>			



DevTreks –social budgeting that improves lives and livelihoods

### Example 8b. Randomized Block Data

#### URLs:

<https://www.devtreks.org/greentreks/preview/carbon/input/Example 1mb, Anova, Randomized Block/2147397546/none>

[https://devtreks1.blob.core.windows.net/resources/network\\_carbon/resourcepack\\_1534/resource\\_7988/Anova2.csv](https://devtreks1.blob.core.windows.net/resources/network_carbon/resourcepack_1534/resource_7988/Anova2.csv)

<https://www.devtreks.org/greentreks/preview/carbon/input/Example 1mb2, Anova, Randomized Block/2147397547/none>

<https://www.devtreks.org/greentreks/preview/carbon/input/Example 1mb3, Anova, Randomized Block/2147397548/none>

**Method 1: Standard ANOVA:** The object of the following dataset is to compare cost estimates for 3 cost estimators. Each cost estimator estimated the costs for the same 4 jobs. Randomized block data requires structuring data by treatments, the 3 cost estimators, and blocks, the 4 job cost estimates. This algorithm runs a standard ANOVA when the following conventions are followed:

5. **Treatments:** The column name for the fifth column of data must be “treatment”. The data in the column must be doubles (or integers) that distinguish each treatment.
6. **Blocks:** The column name for the sixth column of data must be “block”. The data in the column must be doubles (or integers) that distinguish each block.
7. **Training Data:** The column name for the fourth column of data can follow standard conventions for dependent, or output, variable data. The data in the column must be doubles that contain the treatment-block observed data.
8. **Scoring Data:** In order to stay consistent with other algorithms, the final 3 rows of data must be scoring data. These 3 rows are not currently used by this algorithm.



DevTreks –social budgeting that improves lives and livelihoods

9. **Indicator Math Expression:** The expression should be identical to regression analysis expressions (I1.Q1.treatment + I1.Q2.block)

Stock dataset (M&E datasets use the conventions explained for DataURL datasets)

label,estimator,job,y,treatment,block

B1,1,1,4.6,1,1

B1,2,2,6.3,2,2

B1,3,3,5.4,3,3

B1,1,4,6.6,1,4

B1,2,1,4.9,2,1

B1,3,2,5.9,3,2

B1,1,3,5,1,3

B1,2,4,6.8,2,4

B1,3,1,4.4,3,1

B1,1,2,6.2,1,2

B1,2,3,5.4,2,3

B1,3,4,6.3,3,4

B1,1,1,7,1,1

B1,2,2,4,2,2

B1,3,3,5,3,3

The following image shows the Math Result for the Indicator associated with this dataset. The results for the statistical tests and confidence intervals are explained in the following list. These results matched the reference text book.

1. **F Test Treatments:** The F test statistic for the treatments, 4.176, is greater than the critical value for F at .10 = 3.463. The treatment means are significantly different (at Score.ConfidenceLevel = 90).



DevTreks –social budgeting that improves lives and livelihoods

2. **F Test Blocks:** The F test statistic for the blocks, 72.464, is greater than the critical value for F at .10 = 3.463. The block means are significantly different (at Score.ConfidenceLevel = 90).
3. **Treatment Confidence Intervals:** The same confidence intervals are generated as the completely randomized example, with one exception. Instead of calculating the difference between the 2<sup>nd</sup> treatment and the 1<sup>st</sup> treatment means, these confidence intervals subtract the 1<sup>st</sup> from the 2<sup>nd</sup> and the 3<sup>rd</sup> from the 2<sup>nd</sup>. Differences in blocks are not displayed because of display issues with the Resource Stock Analysis (they all share the same code).

Blocks (job)					
		1	2	3	4
Treatments (ratio of raw material allocation)	1	4.6	6.2	5	6.6
	2	4.9	6.3	5.4	6.8
	3	4.4	5.9	5.4	6.3



DevTreks –social budgeting that improves lives and livelihoods

Math Type 1		Math Sub Type 1		
algorithm1		subalgorithm8		
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		
0.1000		mean diff 1 and 2		
QT Low 1		QT Low Unit 1		
-0.1424		lower 90% ci		
QT High 1		QT High Unit 1		
0.3424		upper 90% ci		
Math Expression 1				
I1.Q1.treatment + I1.Q2.block				
Math Result 1				
anova results				
source	df	SS	MS	F
treats	2	0.2600	0.1300	4.1786
blocks	3	6.7633	2.2544	72.4643
error	6	0.1867	0.0311	
total	11	7.2100		
F Crit treats	3.46330	F > F Critical	true	
F Crit blocks	3.28876	F > F Critical	true	
estimate	mean diff	lower 90%	upper 90%	
Treat 1 Mean	5.6000	5.4286	5.7714	
xminus1 1	-0.2500	-0.4924	-0.0076	
base 1	-0.2500	-0.4924	-0.0076	
xminus1 2	0.3500	0.1076	0.5924	
base 2	0.1000	-0.1424	0.3424	

**Method 2: Regression ANOVA:** The following dataset changes the column names from “treatment” and “block” to any other acceptable name (such as t1 and b1). This tells subalgorithm8 to run the model as a regression analysis. The column t1 is coded 0 when the



DevTreks –social budgeting that improves lives and livelihoods

treatment, or cost estimator, is number 1, otherwise its coded 1. The column b1 is coded 0 when the block, or job estimate, is number 1, otherwise its coded 1.

Stock dataset (M&E datasets use the conventions explained for DataURL datasets)

label,estimator,job,y,t1,b1

B1,1,1,4.6,0,0

B1,2,2,6.3,1,1

B1,3,3,5.4,1,1

B1,1,4,6.6,0,1

B1,2,1,4.9,1,0

B1,3,2,5.9,1,1

B1,1,3,5,0,1

B1,2,4,6.8,1,1

B1,3,1,4.4,1,0

B1,1,2,6.2,0,1

B1,2,3,5.4,1,1

B1,3,4,6.3,1,1

B1,1,1,7,0,0

B1,2,2,4,1,1

B1,3,3,5,1,1

The following image shows that running this dataset using subalgorithm8 will produce a standard regression analysis. In this analysis, the t-ratios for t1 (treatments or job estimators) and b1 (blocks or cost estimates) can be used to assess the differences among treatment and block means.





DevTreks –social budgeting that improves lives and livelihoods

Math Type 1		Math Sub Type 1		
algorithm1		subalgorithm8		
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		
4.5833		mean diff 1 and 2		
QT Low 1		QT Low Unit 1		
3.8275		lower 90% ci		
QT High 1		QT High Unit 1		
5.3392		upper 90% ci		
Math Expression 1				
l1.Q1.t1 + l1.Q2.b1				
Math Result 1				
regression results				
dependent variable: y				
source	df	SS	MS	
model	2	4.1494	2.0747	
error	9	3.0606	0.3401	
total	11	7.2100		
R-squared	0.5755	Adj R-squared	0.4812	
F value	6.1010	prob > F	0.0212	
variable	coefficient	stand error	T-ratio	prob > T
intercept	4.58333	0.4123	11.1152	0.0001
t1	0.07500	0.3571	0.2100	0.8383
b1	1.35556	0.3888	3.4868	0.0069
durbin-watson: 3.0724				
F Critical Value	3.00645	F > F Critical	true	
estimate	predicted	lower 90%	upper 90%	
Col 0 Mean CI	4.5833	3.8275	5.3392	
Col 1 - 0 Mean CI	0.0750	-0.5796	0.7296	
Col 2 - 0 Mean CI	1.3556	0.6429	2.0682	



DevTreks –social budgeting that improves lives and livelihoods

**Method 3: Resource Stock Analysis ANOVA:** As of 1.9.2, the following feature is being retained, but not debugged, until more advanced RCT algorithms are developed that can either replace, or enhance, this feature. The following 3 datasets are subsets of the data used in the standard ANOVA. Each dataset has been added to 3 sibling Input Series. Each Input Series and dataset represent a separate experimental treatment. An example of a Math Expression used to calculate Series 1 is  $(I1.Q1.treatment * 0) + (I1.Q2.block * 0) + I1.QT$ .

Stock dataset (M&E datasets use the conventions explained for DataURL datasets)

Input Series 1

label,estimator,job,y,treatment,block

B1,1,1,4.6,1,1

B1,1,4,6.6,1,4

B1,1,3,5,1,3

B1,1,2,6.2,1,2

Input Series 2

label,estimator,job,y,treatment,block

B1,2,2,6.3,2,2

B1,2,1,4.9,2,1

B1,2,4,6.8,2,4

B1,2,3,5.4,2,3

Input Series 3 (the 3 rows of scoring data must be added to the last dataset but are not currently used)

label,estimator,job,y,treatment,block

B1,3,3,5.4,3,3

B1,3,2,5.9,3,2

B1,3,1,4.4,3,1



DevTreks –social budgeting that improves lives and livelihoods

B1,3,4,6,3,3,4

B1,1,1,7,1,1

B1,2,2,4,2,2

B1,3,3,5,3,3

The following image demonstrates that running a Change by Id Resource Stock Analysis returns the identical Score properties as standard Resource Stock Analyses, but the Indicators compare the differences between the statistical means for each treatment. They use the same calculations as the ones used to set the confidence intervals for the standard ANOVA and regression ANOVA, but use the standard comparators used by the Resource Stock analyzers. They currently use the same stylesheet as the standard analyses, which cannot display the additional differences among blocks. Use the Score.MathResult to view the complete statistical results.

<b>Score High Base Percent Change</b>	0.0000	4.8322	-3.6441
<b>Indicator Observations</b>	1.0	1.0	1.0
<b>Name</b>	Mean Estimate	Mean Estimate	Mean Estimate
<b>Label</b>	B1	B1	B1
<b>Total</b>	5.6000	5.8500	5.5000
<b>Unit</b>	mean estimate	mean estimate	mean estimate
<b>Amount Change</b>	4.1786	0.0000	0.3500
<b>Percent Change</b>	3.4633	0.0000	0.2424
<b>Base Change</b>	5.6000	-0.2500	0.1000
<b>Base Percent Change</b>	0.1714	0.2424	0.2424



DevTreks –social budgeting that improves lives and livelihoods

### Example 8c. Randomized Factorial Data

#### URLs:

<https://www.devtreks.org/greentreks/preview/carbon/input/Example 1mc, Anova, Randomized Factorial/2147397549/none>

[https://devtreks1.blob.core.windows.net/resources/network\\_carbon/resourcepack\\_1534/resource\\_7993/Anova3.csv](https://devtreks1.blob.core.windows.net/resources/network_carbon/resourcepack_1534/resource_7993/Anova3.csv)

<https://www.devtreks.org/greentreks/preview/carbon/input/Example 1mc2, Anova, Randomized Factorial/2147397550/none>

<https://www.devtreks.org/greentreks/preview/carbon/input/Example 1mc3, Anova, Randomized Factorial/2147397551/none>

**Method 1: Standard ANOVA:** The object of the following subset of a full dataset is to compare the mean profit per unit raw material for 3 amounts of raw material amounts and 3 ratios of raw materials allocated to manufacturing lines. Randomized factorial data requires structuring data by factors, the ratios and raw materials, and levels, the 3 different amounts for each factor. The dataset represents 3 replications of a complete 3 x 3 factorial experiment. This algorithm runs a standard ANOVA when the following conventions are followed:

1. **Factors:** The column name for the fifth column of data must be “factor1”. The data in the column must be doubles (or integers) that distinguish each level of factor1. The column name for the sixth column of data must be “factor2”. The data in the column must be doubles (or integers) that distinguish each level of factor2. Additional factors will be supported in future releases.
2. **Training Data:** The column name for the fourth column of data can follow standard conventions for dependent, or output, variable data. The data in the column must be doubles that contain the factor-level observed data.



DevTreks –social budgeting that improves lives and livelihoods

3. **Scoring Data:** In order to stay consistent with other algorithms, the final 3 rows of data must be scoring data. These 3 rows are not currently used by this algorithm.
4. **Indicator Math Expression:** The expression should be identical to regression analysis expressions (I1.Q1.factor1 + I1.Q2.factor2)

Stock dataset (M&E datasets use the conventions explained for DataURL datasets)

label,ratioraw,rawmat,y,factor1,factor2

F1,0.50,15,23,0.50,15

F1,0.50,15,20,0.50,15

F1,0.50,15,21,0.50,15

F1,1.00,15,22,1.00,15

F1,1.00,15,20,1.00,15

F1,1.00,15,19,1.00,15

F1,2.00,15,18,2.00,15

F1,2.00,15,18,2.00,15

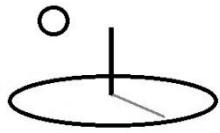
F1,2.00,15,16,2.00,15

F1,0.50,18,22,0.50,18

...

The following image shows the Math Result for the Indicator associated with this dataset. The results for the statistical tests and confidence intervals are explained in the following list. These results matched the reference text book.

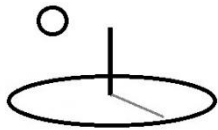
1. **F Test Factor1:** The F test statistic for factor1, 1.71, is less than the critical value for F at .05 = 3.55. The factor1 means are not significantly different (at Score.ConfidenceLevel = 95).
2. **F Test Factor2:** The F test statistic for factor2, 4.20, is greater than the critical value for F at .05 = 3.55. The factor2 means are significantly different (at Score.ConfidenceLevel = 95).



DevTreks –social budgeting that improves lives and livelihoods

3. **F Test Interactions:** The F test statistic for the interactive effects, 4.80, is greater than the critical value for F at  $.05 = 2.93$ . The interactive means are significantly different (at  $\text{Score.ConfidenceLevel} = 95$ ).
4. **Factor Confidence Intervals:** The confidence intervals reflect differences among factor-level data cells in the following data table. Specifically, the differences in the means for cells in position (0,0), (1,1), and (2,2) are compared.

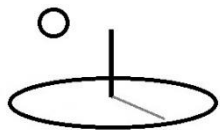
	A	D	C	D	E
	<b>Factor 1 (raw material supply in tons)</b>				
			15	18	21
<b>Factor 2 (ratio of raw material allocation)</b>	0.5	23,20,21	22,19,20	19,18,21	
	1	22,29,19	24,25,22	20,19,22	
	2	18,18,16	21,23,20	20,22,24	



DevTreks –social budgeting that improves lives and livelihoods

Math Operator 1		BaselO 1																																																																		
equalto		none																																																																		
QT 1		QT Unit 1																																																																		
0.0000		mean profit diff 1 and 2																																																																		
Math Type 1		Math Sub Type 1																																																																		
algorithm1		subalgorithm8																																																																		
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																																																																		
-0.6667		mean profit diff 1 and 2																																																																		
QT Low 1		QT Low Unit 1																																																																		
-3.3282		lower 95% ci																																																																		
QT High 1		QT High Unit 1																																																																		
1.9949		upper 95% ci																																																																		
Math Expression 1																																																																				
I1.Q1.factor1 + I1.Q2.factor2																																																																				
Math Result 1																																																																				
<table border="1"> <thead> <tr> <th colspan="5">anova results</th> </tr> <tr> <th>source</th> <th>df</th> <th>SS</th> <th>MS</th> <th>F</th> </tr> </thead> <tbody> <tr> <td>factor1</td> <td>2</td> <td>8.2222</td> <td>4.1111</td> <td>1.7077</td> </tr> <tr> <td>factor2</td> <td>2</td> <td>20.2222</td> <td>10.1111</td> <td>4.2000</td> </tr> <tr> <td>interacts</td> <td>4</td> <td>46.2222</td> <td>11.5556</td> <td>4.8000</td> </tr> <tr> <td>error</td> <td>18</td> <td>43.3333</td> <td>2.4074</td> <td></td> </tr> <tr> <td>total</td> <td>26</td> <td>118.0000</td> <td></td> <td></td> </tr> <tr> <td>F Crit factor1</td> <td>3.55456</td> <td>F &gt; F Critical</td> <td>false</td> <td></td> </tr> <tr> <td>F Crit factor2</td> <td>3.55456</td> <td>F &gt; F Critical</td> <td>true</td> <td></td> </tr> <tr> <td>F Crit Interacts</td> <td>2.92774</td> <td>F &gt; F Critical</td> <td>true</td> <td></td> </tr> <tr> <td>estimate</td> <td>mean diff</td> <td>lower 95%</td> <td>upper 95%</td> <td></td> </tr> <tr> <td>Treat 1 Mean</td> <td>21.3333</td> <td>19.4513</td> <td>23.2154</td> <td></td> </tr> <tr> <td>xminus1 1</td> <td>-2.3333</td> <td>-4.9949</td> <td>0.3282</td> <td></td> </tr> </tbody> </table>				anova results					source	df	SS	MS	F	factor1	2	8.2222	4.1111	1.7077	factor2	2	20.2222	10.1111	4.2000	interacts	4	46.2222	11.5556	4.8000	error	18	43.3333	2.4074		total	26	118.0000			F Crit factor1	3.55456	F > F Critical	false		F Crit factor2	3.55456	F > F Critical	true		F Crit Interacts	2.92774	F > F Critical	true		estimate	mean diff	lower 95%	upper 95%		Treat 1 Mean	21.3333	19.4513	23.2154		xminus1 1	-2.3333	-4.9949	0.3282	
anova results																																																																				
source	df	SS	MS	F																																																																
factor1	2	8.2222	4.1111	1.7077																																																																
factor2	2	20.2222	10.1111	4.2000																																																																
interacts	4	46.2222	11.5556	4.8000																																																																
error	18	43.3333	2.4074																																																																	
total	26	118.0000																																																																		
F Crit factor1	3.55456	F > F Critical	false																																																																	
F Crit factor2	3.55456	F > F Critical	true																																																																	
F Crit Interacts	2.92774	F > F Critical	true																																																																	
estimate	mean diff	lower 95%	upper 95%																																																																	
Treat 1 Mean	21.3333	19.4513	23.2154																																																																	
xminus1 1	-2.3333	-4.9949	0.3282																																																																	

**Method 2: Regression ANOVA:** The following dataset changes the column names from “factor1” and “factor2” to any other acceptable name (such as s1 and r1). This tells



DevTreks –social budgeting that improves lives and livelihoods

subalgorithm8 to run the model as a regression analysis. The column s1 is coded 0 when factor 1, or raw material ratio, is level 1, 0.5, otherwise its coded 1. The column r1 is coded 0 when factor 2, or raw material amount, is level 1, 15, otherwise its coded 1. The last column, sr, models the interactive effects between s1 and s2 and is coded as  $s1 * r1$ .

Note that adding additional interactive terms such as  $x1^2$  and  $x2^2$ , while using 0 and 1 codes for independent variables, is not supported by the matrix mathematical techniques used by the regression algorithm –they return an error message stating that the matrix is not positive definite.

Stock dataset (M&E datasets use the conventions explained for DataURL datasets)

label,ratoraw,rawmat,y,s1,r1,sr

F1,0.50,15,23,0.00,0,0

F1,0.50,15,20,0.00,0,0

F1,0.50,15,21,0.00,0,0

F1,1.00,15,22,1.00,0,0

F1,1.00,15,20,1.00,0,0

F1,1.00,15,19,1.00,0,0

F1,2.00,15,18,1.00,0,0

F1,2.00,15,18,1.00,0,0

F1,2.00,15,16,1.00,0,0

F1,0.50,18,22,0.00,1,0

F1,0.50,18,19,0.00,1,0

F1,0.50,18,20,0.00,1,0

F1,1.00,18,24,1.00,1,1

F1,1.00,18,25,1.00,1,1

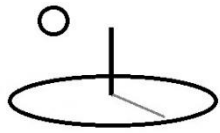
F1,1.00,18,22,1.00,1,1

F1,2.00,18,21,1.00,1,1

F1,2.00,18,23,1.00,1,1

F1,2.00,18,20,1.00,1,1





DevTreks –social budgeting that improves lives and livelihoods

F1,0.50,21,19,0.00,1,0

F1,0.50,21,18,0.00,1,0

F1,0.50,21,21,0.00,1,0

F1,1.00,21,20,1.00,1,1

F1,1.00,21,19,1.00,1,1

F1,1.00,21,22,1.00,1,1

F1,2.00,21,20,1.00,1,1

F1,2.00,21,22,1.00,1,1

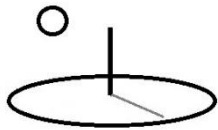
F1,2.00,21,24,1.00,1,1

F1,0.50,21,24,1.00,1,0

F1,1.00,21,24,1.00,1,0

F1,2.00,21,24,1.00,1,1

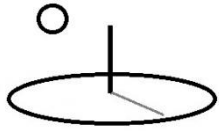
The following image shows that running this dataset using subalgorithm8 will produce a standard regression analysis. In this analysis, the t-ratios for s1 (factor1 or raw material ratios), r1 (factore2 or raw material amounts), and sr (interactive effects between s1 and r1), can be used to assess the differences among factors and interactive effect means.



DevTreks –social budgeting that improves lives and livelihoods

Matn Type 1		Matn Sub Type 1		
algorithm1		subalgorithm8		
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			
21.3333	mean profit diff 1 and 2			
QT Low 1	QT Low Unit 1			
19.5346	lower 90% ci			
QT High 1	QT High Unit 1			
23.1320	upper 90% ci			
Math Expression 1				
I1.Q1.s1 + I1.Q2.r1 + I1.Q3.sr				
Math Result 1				
regression results				
dependent variable: y				
source	df	SS	MS	
model	3	42.0000	14.0000	
error	23	76.0000	3.3043	
total	26	118.0000		
R-squared	0.3559	Adj R-squared	0.2719	
F value	4.2368	prob > F	0.0160	
variable	coefficient	stand error	T-ratio	prob > T
intercept	21.33333	1.0495	20.3271	0.0001
s1	-2.50000	1.2854	-1.9450	0.0641
r1	-1.50000	1.2854	-1.1670	0.2552
sr	4.50000	1.5742	2.8585	0.0089
durbin-watson: 1.8059				
F Critical Value 2.33873		F > F Critical true		
estimate	predicted	lower 90%	upper 90%	
Col 0 Mean CI	21.3333	19.5346	23.1320	
Col 1 - 0 Mean CI	-2.5000	-4.7030	-0.2970	
Col 2 - 0 Mean CI	-1.5000	-3.7030	0.7030	

**Method 3: Resource Stock Analysis ANOVA:** As of 1.9.2, the following feature is being retained, but not debugged, until more advanced RCT algorithms are developed that can either



DevTreks –social budgeting that improves lives and livelihoods

replace, or enhance, this feature. The following 3 datasets are subsets of the data used in the standard ANOVA. Each dataset has been added to 3 sibling Input Series. Each Input Series and dataset represent a separate experimental treatment. An example of a Math Expression used to calculate Series 1 is  $(I1.Q1.factor1 * 0) + (I1.Q2.factor2 * 0) + I1.QT$ .

Stock dataset (M&E datasets use the conventions explained for DataURL datasets)

#### Input Series 1

label,ratoraw,rawmat,y,factor1,factor2

F1,0.50,15,23,0.50,15

F1,0.50,15,20,0.50,15

F1,0.50,15,21,0.50,15

F1,0.50,18,22,0.50,18

F1,0.50,18,19,0.50,18

F1,0.50,18,20,0.50,18

F1,0.50,21,19,0.50,21

F1,0.50,21,18,0.50,21

F1,0.50,21,21,0.50,21

#### Input Series 2

label,ratoraw,rawmat,y,factor1,factor2

F1,1.00,15,22,1.00,15

F1,1.00,15,20,1.00,15

...

Input Series 3 (the 3 rows of scoring data must be added to the last dataset but are not currently used)

label,ratoraw,rawmat,y,factor1,factor2

F1,2.00,15,18,2.00,15

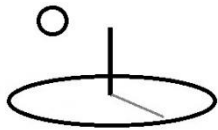
F1,2.00,15,18,2.00,15



DevTreks –social budgeting that improves lives and livelihoods

...

The following images demonstrates that running a Change by Id Resource Stock Analysis returns the identical Score properties as standard Resource Stock Analyses,, but the Indicator confidence intervals compare the differences between the statistical means for specific factor-level cells shown in the data table above. They use the same calculations as the ones used to set the confidence intervals for the standard ANOVA and regression ANOVA, but use the standard comparators used by the Resource Stock analyzers. They currently use the same stylesheet as the standard analyses, which cannot display the additional differences among all factor-level cells. Use the first image, the Score.MathResult, to view the complete statistical results.



DevTreks –social budgeting that improves lives and livelihoods

Base Resource (Input) Calculations To Analyze Type:

Resource Stock Calculator 1

Compare Using: **None** **Compare Only**

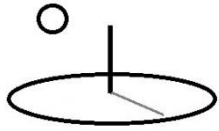
Math Type: **algorithm1** Math Sub Type: **subalgorithm8**

Confidence Interval: **95**

Math Expression: **I1.Q1.factor1 + I1.Q2.factor2**

Math Result

anova results				
source	df	SS	MS	F
factor1	2	8.2222	4.1111	1.7077
factor2	2	20.2222	10.1111	4.2000
interacts	4	46.2222	11.5556	4.8000
error	18	43.3333	2.4074	
total	26	118.0000		
F Crit factor1	3.55456	F > F Critical	false	
F Crit factor2	3.55456	F > F Critical	true	
F Crit Interacts	2.92774	F > F Critical	true	
estimate	mean diff	lower 95%	upper 95%	
Treat 1 Mean	21.3333	19.4513	23.2154	
xminus1 1	-2.3333	-4.9949	0.3282	
base 1	-2.3333	-4.9949	0.3282	
xminus1 2	1.6667	-0.9949	4.3282	
base 2	-0.6667	-3.3282	1.9949	



DevTreks –social budgeting that improves lives and livelihoods

<b>Score High Base Percent Change</b>	0.0000	2.6877	10.7508
<b>Indicator Observations</b>	1.0	1.0	1.0
<b>Name</b>	Mean Profit	Mean Profit	Mean Profit
<b>Label</b>	F1	F1	F1
<b>Total</b>	21.3333	23.6667	22.0000
<b>Unit</b>	mean profit	mean profit	mean profit
<b>Amount Change</b>	1.7077	0.0000	1.6667
<b>Percent Change</b>	3.5546	0.0000	2.6616
<b>Base Change</b>	21.3333	-2.3333	-0.6667
<b>Base Percent Change</b>	1.8820	2.6616	2.6616



## Appendix B. Correlated Uncertain Numbers

All of the main probabilistic-risk references (GAO, IPCC, NASA) explain the importance of accounting for correlated indicators in PRA. Failure to do so results in random samples that don't retain the correct correlations between indicators and therefore incorrect descriptive statistics. The following references, in particular, provide guidance about potential mathematical techniques that account for correlated indicators:

1. Piwcewicz of the Australian Actuaries Society (2005) provides an introduction to correlated multivariate analysis for practitioners. They introduce common algorithms, including pair-wise rank correlations (or Inman and Conover) and copula (mathematical formulas that manipulate matrixes) that analysts commonly use. Their approach is practical, recognizing that assumptions may have to be made about data, distributions, and correlations. They also explain common dangers posed by simulation techniques. The Brebbia (2013) reference provides more recent risk analysis techniques, such as using Kernel copulas (but it isn't an open access publication and therefore of limited usefulness).
2. Anderson, Harri, and Cable (2009) use an agricultural economics example relevant to CTA analysis to explain the difference between the conventional pair-wise rank correlations (Inman and Conover) and a matrix manipulation technique (eigenvalue decomposition). Advantages to the latter approach include better simulated numbers and ease of computation (i.e. using modern mathematical libraries). References to this latter approach can be found in the finance and engineering literature, but are not cited here.
3. Although not cited, several references were found that used techniques such as genetic algorithms, simulated annealing algorithms, and several sequencing methods (i.e. Hammersley) for carrying out these simulations. Potential advantages with some of the techniques involve speed and large data manipulation. References to these approaches can be found in the computer science literature, but are not cited here.

Searches on the web reveal that statisticians, software developers, and the mathematically-inclined, frequently answer this question in their forums. Many present statistical scripts, such as

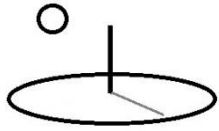


R package or Matlab, with concrete datasets demonstrating the answer. Judging from their online blogs, it appears that most practitioners recommend using mathematical matrix manipulation techniques based on copulas.

The open source mathematical library (Math.Net) used in this version supports a wide assortment of mathematical matrix manipulation. Algorithm1 with subalgorithm2, subalgorithm3, and subalgorithm4, use the following steps to generate statistics for correlated indicators:

- a) Use Monte Carlo simulation to generate a sample of random numbers using each indicator's Distribution Type, QT, QTD1, and QTD2. Combine each indicator's vector of random samples ( $F_n$ ) into a random sample matrix  $F$ . This matrix is also called marginal distributions.
- b) Generate a Pearson or Spearman correlation matrix  $R$  from the data or from expert knowledge. The matrix is based on the calculated QT variable.
- c) Use Monte Carlo simulation to generate a matrix  $Z$  of random numbers using a normal distribution ( $N(0,1)$ ) with the same rank as  $F$ .
- d) Select an  $n$ -copula  $C$  that is consistent with the data. A copula is multivariate uniform (or normal) distribution that uses correlation matrixes to define dependence among the uniform (or normal) variables. Subalgorithm2, or `pra_copulaN`, generates matrix  $X$  by multiplying a Cholesky decomposition matrix of  $R$  times  $Z$ . Subalgorithm3, or `pra_eigen`, and subalgorithm4, `pra_eigenU`, generate matrix  $X$  by multiplying an Eigen decomposition matrix of  $R$  times  $Z$ .
- e) Subalgorithm3 uses a Normal distribution and subalgorithm4 uses a Uniform distribution in the previous step. That's their only difference. The author found examples of both distributions being used and assumes it may matter. Tests confirm slightly different results.
- f) Generate a correlated random sample matrix  $R$  by using the probability of each variable in  $X$  to determine the inverse cumulative density function of each corresponding variable in  $F$ .



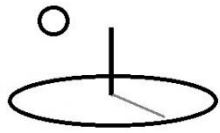


DevTreks –social budgeting that improves lives and livelihoods

- g) Verify the accuracy of the random correlated sample matrix by generating another Pearson or Spearman correlation matrix from R. Add the result to the Score Math Result.
- h) Generate sample statistics for each vector in R and add descriptive statistics to the associated indicator.
- i) Set the Score using standard Score properties.

The following source code displays the calculations used by subalgorithm3. Subalgorithm4 uses a Uniform distribution for the matrix u. The source code demonstrates why these types of algorithms should not be considered daunting. As mentioned, the analysis of large data sets should consider using algorithms that don't employ these exact techniques.

```
else if (HasMathType(MATH_TYPES.algorithm1, MATH_SUBTYPES.subalgorithm3))
{
    //eigenvalue decomposition
    //jointData is the correlation matrix
    var evd = jointData.Evd();
    if (evd == null)
    {
        this.SB1ScoreMathResult += string.Concat(" ",
            Errors.MakeStandardErrorMsg("MATRIX_BADEIGEN"));
        return null;
    }
    //take the square root of the diagonal eigenvalue matrix
    Matrix<double> squarerootEigenValues = evd.D.PointwisePower(0.5);
    //multiply the eigenvalues square root matrix by the eigenvectors diagonal matrix
    var v = evd.EigenVectors.Multiply(squarerootEigenValues);
    //random normal standards
    var u = Matrix<double>.Build.Random(this.SB1Iterations, cols);
    //generate correlated normal randoms by multiplying both matrixes together
    var X = u.TransposeAndMultiply(v);
}
```



DevTreks –social budgeting that improves lives and livelihoods

//random sample vector n = inverse of cumulative density function for  $X_n$

SetCorrelatedRandomSamples(X, sampleData, randomSampleData);

}