



@RISK Getting Started Guide

Welcome to @RISK, the world's leading Excel-based Monte Carlo simulation package!

This short guide is designed to get you started using @RISK quickly. It is appropriate if you are either completely new to risk analysis using Monte Carlo simulation, or if you know those concepts well but are new to @RISK. If you are already familiar with a previous version of @RISK and want to learn about how @RISK 8 differs from previous versions, please consult the @RISK 8 Transition Guide.

Version 2 - Last Updated 5/15/2020



Minimize risk. Maximize potential.

Contents

Section 1: Introduction to Risk Analysis and @RISK	3
Beyond the point estimate	3
What is Monte Carlo simulation?	3
How does @RISK present the simulation results?.....	3
How does @RISK help me make informed decisions?	4
Do I have to be a programmer to use @RISK?	5
Section 2: @RISK Software Starter Kit	6
How do I get @RISK onto my computer?	6
How do I learn @RISK?	6
How do I get help?	6
Section 3: How to Create and Run Your Model	7
How do I know what to do in the @RISK ribbon?.....	7
What is an output? How do I add one to my model?.....	7
What is an input? What is a probability distribution? How do I add one to my model?	8
Which distribution should I choose?	9
Am I limited to a set of predefined distributions?.....	10
Can inputs be correlated? What is correlation in @RISK?.....	11
What are time series?	12
What happens during an @RISK simulation?	13
What's the difference between an iteration and a simulation?.....	13
How do I decide on a number of iterations?	13
What kinds of graphs can I get for an input or output?	14
What kinds of statistics can I get for an input or output?	15
What reports does @RISK make for senior management?	16
Can @RISK give me a summary of my whole model?	18
How does @RISK tell me which inputs are driving an output?	19

Section 1: Introduction to Risk Analysis and @RISK

Beyond the point estimate

Excel models are frequently used to assist in decision-making at all levels of an organization. Traditionally, these models contain single-point estimates of variables that are actually uncertain quantities, and in turn return single-point results that then don't include any measure of the inherent variability of the situation being modeled.

The future is uncertain. To truly provide a decision-maker with useful information, you need to generate a comprehensive range of potential outcomes and their relative likelihoods so the best possible decisions can be made.

@RISK helps you accomplish this task in Excel using a technique known as Monte Carlo simulation.

What is Monte Carlo simulation?

Monte Carlo simulation is a technique that produces information and insights from a model and its assumptions.

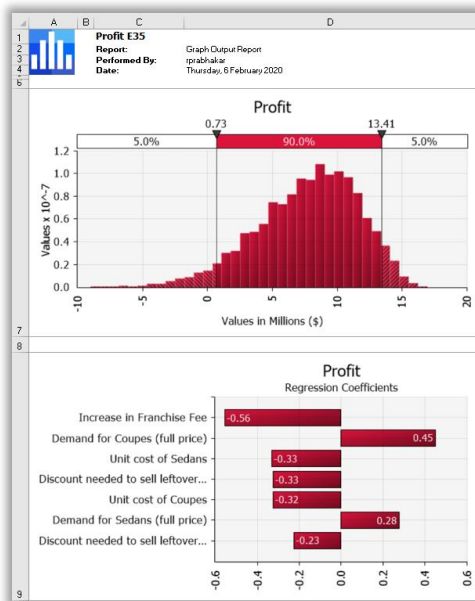
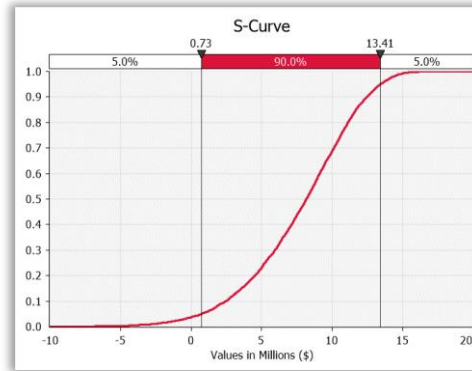
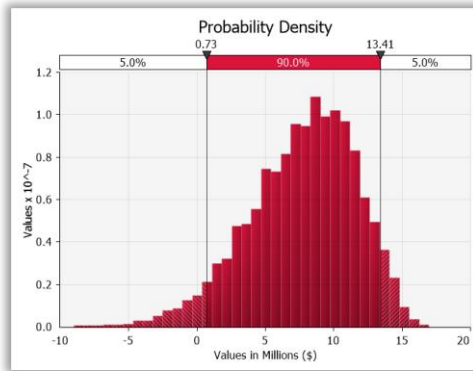
You start by replacing the fixed assumptions in your Excel model with probability distributions (also called "inputs") that describe the uncertainty around each of those assumptions. You then mark one or more other cells, called the "outputs", as being things you want to analyze.

During a simulation, @RISK randomly "samples" all the input distributions and recalculates the spreadsheet repeatedly, keeping track of the resulting output values. Each separate recalculation in this process is referred to as an "iteration". A single iteration represents one possible future set of circumstances in the model.

The entire simulation is made up of many such iterations, from tens of thousands to millions or more, depending on your specific application. As the sampling is random, commonly occurring input ranges and combinations of inputs appear more frequently in the simulation data. Similarly, rarer scenarios will be less likely. The simulation data for each output will then reasonably approximate samples from the output distribution. Statistical and graphical inferences can be made with reasonable confidence.

How does @RISK present the simulation results?

Presenting and reporting results effectively is key to maximizing the usefulness of @RISK. Visually appealing and informative charts, tables of key statistics, standard and editable reports, and even the simulation data itself are all available after running a simulation. Below are some examples. (From the top left: probability density and cumulative S-curve of an output, an output report, and a simulation data set.)



Name	Total Cost (incl... Output E31	Total Revenue Output E34	Profit Output E35
Description Cell Function	RiskOutput(D31)	RiskOutput(D34)	RiskOutput(D35)
1	\$104,485,490.48	\$102,178,508.80	-\$2,306,981.67
2	\$100,227,566.78	\$109,355,342.26	\$9,127,775.48
3	\$100,523,794.03	\$110,834,499.44	\$10,310,705.41
4	\$105,445,048.43	\$111,402,180.13	\$5,957,131.70
5	\$99,212,504.63	\$109,186,678.43	\$9,974,173.80
6	\$97,937,780.13	\$110,831,809.05	\$12,894,028.92
7	\$106,927,785.51	\$109,522,401.05	\$2,594,615.54
8	\$103,011,316.60	\$105,750,026.82	\$2,738,710.23
9	\$107,398,384.93	\$112,724,993.51	\$5,326,608.58
10	\$99,831,890.70	\$112,365,206.86	\$12,533,316.17
11	\$100,697,672.78	\$111,369,358.73	\$10,671,685.95
12	\$99,873,521.42	\$108,331,856.05	\$8,458,334.63
13	\$97,806,396.15	\$111,975,474.10	\$14,169,077.95
14	\$98,736,301.99	\$109,316,177.82	\$10,579,875.83
15	\$103,287,257.40	\$113,392,297.81	\$10,105,040.40
16	\$108,755,876.22	\$104,310,585.18	-\$4,445,291.04
17	\$96,140,244.31	\$108,816,138.79	\$12,675,894.47
18	\$97,689,573.09	\$110,178,353.09	\$12,488,780.00

How does @RISK help me make informed decisions?

A well-informed decision is based on comparing the decision-maker’s risk appetite (or tolerance) to the outcomes to which they are exposed. The reporting in @RISK shows information appropriate to this goal. Examples include the likelihood of making or losing money, or of reaching a specific target (e.g. the probability of at least \$1m revenue in the first year). Statistics are also calculated that are used to describe the various future outcomes a model shows, such as the mean or standard deviation of an output (e.g. NPV, project cost etc.). @RISK can identify your key inputs using sensitivity analyses such as regression and correlation tornado charts, prioritizing mitigation decisions including insurance policies, controls and the like.

All results generated by @RISK are calculated directly from the simulation data set. For example, if you run a simulation with 100,000 iterations and 62,400 of those have Year 1 Revenue above \$1m, @RISK reports a 62.4% of achieving at least that amount. Likewise, the mean revenue in Year 1 would be the

average of the values in that set of data, and the P90 of the Project Cost is the 90th percentile of the data.

The industrial edition of @RISK includes the ability to optimize your @RISK models, allowing you to determine which set of decision variable achieves the best outcome for your organization.

Do I have to be a programmer to use @RISK?

Absolutely not. @RISK is an Excel add-in, which includes its own custom ribbon to control the application. There is no unique programming language that you need to master in order to use @RISK. It is integrated into Excel, and can be operated in a near-identical way to other Excel features. @RISK functions can be inserted into the spreadsheet using the @RISK interface, the Excel Insert Function dialog, or simply by typing functions directly into spreadsheet cells. Simulations, results, and all of their associated options are controlled easily using your mouse and buttons on the ribbon.

Generally, the ribbon is used from left to right to build @RISK functions and features into your model and work processes. Broadly these flow across the groups Define → Simulation → Results, with additional features (Tools, Application) containing more advanced applications.

All of the above taken into consideration, if you are familiar with Excel's VBA programming language, @RISK is designed to be automatable using the @RISK for Excel Developer's Kit (@RISK XDK).

Section 2: @RISK Software Starter Kit

How do I get @RISK onto my computer?

Please refer to @RISK license and installation guide: <https://help.palisade.com/v8/en/Guides/@RISK-License-Installation-Guide.pdf>.

How do I learn @RISK?

There are several stages of learning all things @RISK. If you are new to the concepts of risk analysis with Monte Carlo simulation, the introductory topics in this document are a great start. If you understand quantitative risk analysis but need help with @RISK, in addition to this guide there are files available in Examples menu on the ribbon. These example files cover the basics of modelling with @RISK as well as specific topics including correlation, distribution fitting, and running multiple simulations. Where more intense and tailored training is required, Palisade offers both public and onsite training in all regions globally.

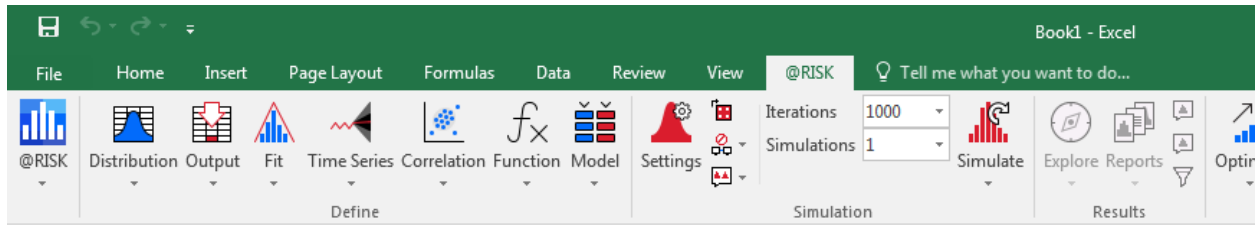
How do I get help?

There are several ways of getting help with @RISK; most are accessible directly from the Resources menu on the @RISK ribbon. From there you can access the @RISK help system, find links to several online guides (such as this one), and learn about Palisade's online webinars, training events, video library, and more.

Section 3: How to Create and Run Your Model

How do I know what to do in the @RISK ribbon?

The @RISK ribbon is organized in “groups” and you can think of the first three of these groups as running from left to right in terms of workflow: Define → Simulation → Results. The image below shows this part of the ribbon:



In the Define group, you will create both input distributions and outputs, fit distributions and time series to existing data, create correlations between inputs, etc. With the model created, the Simulation group allows you to adjust settings such as number of iterations and runtime graphing options, and start the simulation itself. After running the simulation, you move to the Results group to create graphs, tables, and reports.

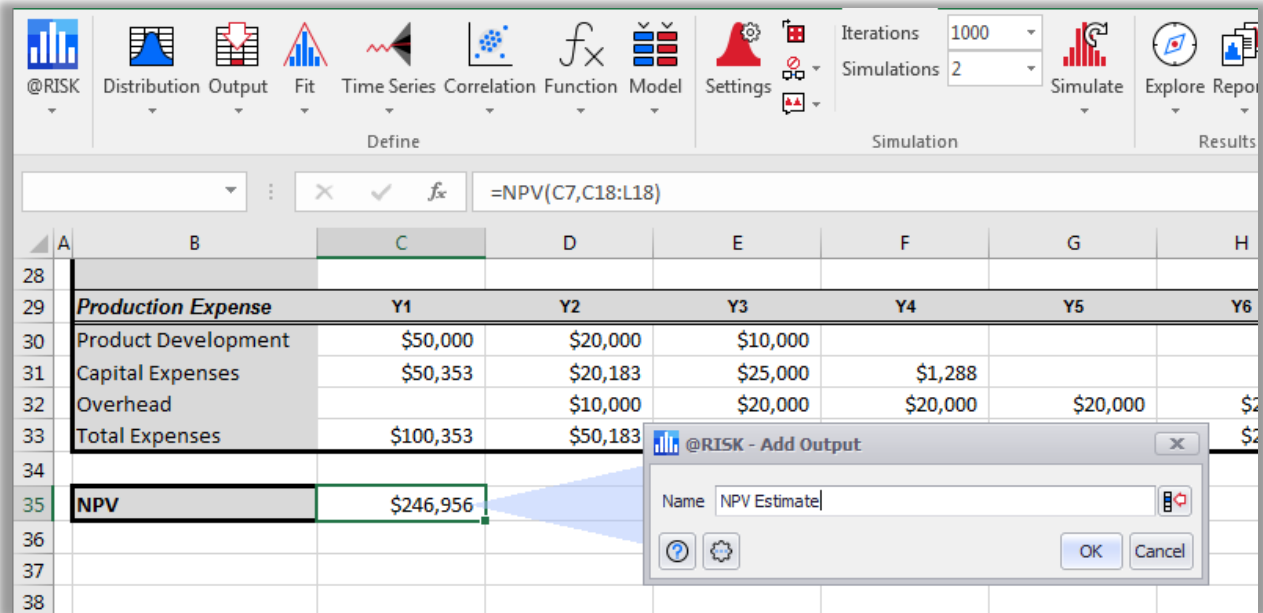
What is an output? How do I add one to my model?

The output is the reason your model exists! It is usually the “final” calculated cell in your model, dependent on all the inputs via formulas defined in your model logic. Examples include NPV (Net Portfolio Value), Total Project Cost and Portfolio Return. By specifying a cell as an output, you are telling @RISK that you are interested in collecting information about this cell and how it is impacted by the uncertainty of its related inputs. There may be multiple outputs (e.g. NPV and Payback Period), or interim outputs (annual revenue in each of the first 10 years of an investment). During a simulation @RISK will store the calculated value for the output for each iteration. Then when the simulation is completed, those results can be analyzed using @RISK.

For example, let’s assume you have the following function in cell C35 of a model:

```
=NPV(C7,C18:L18)
```

and you want to make it an @RISK output. The easiest way to accomplish this is by clicking on the Output button on the ribbon.



This displays a small dialog where you can optionally specify a name for the output. Clicking the OK button on the dialog will then add a RiskOutput function to your cell's formula:

`=RiskOutput("NPV Estimate") + NPV(C7,C18:L18)`

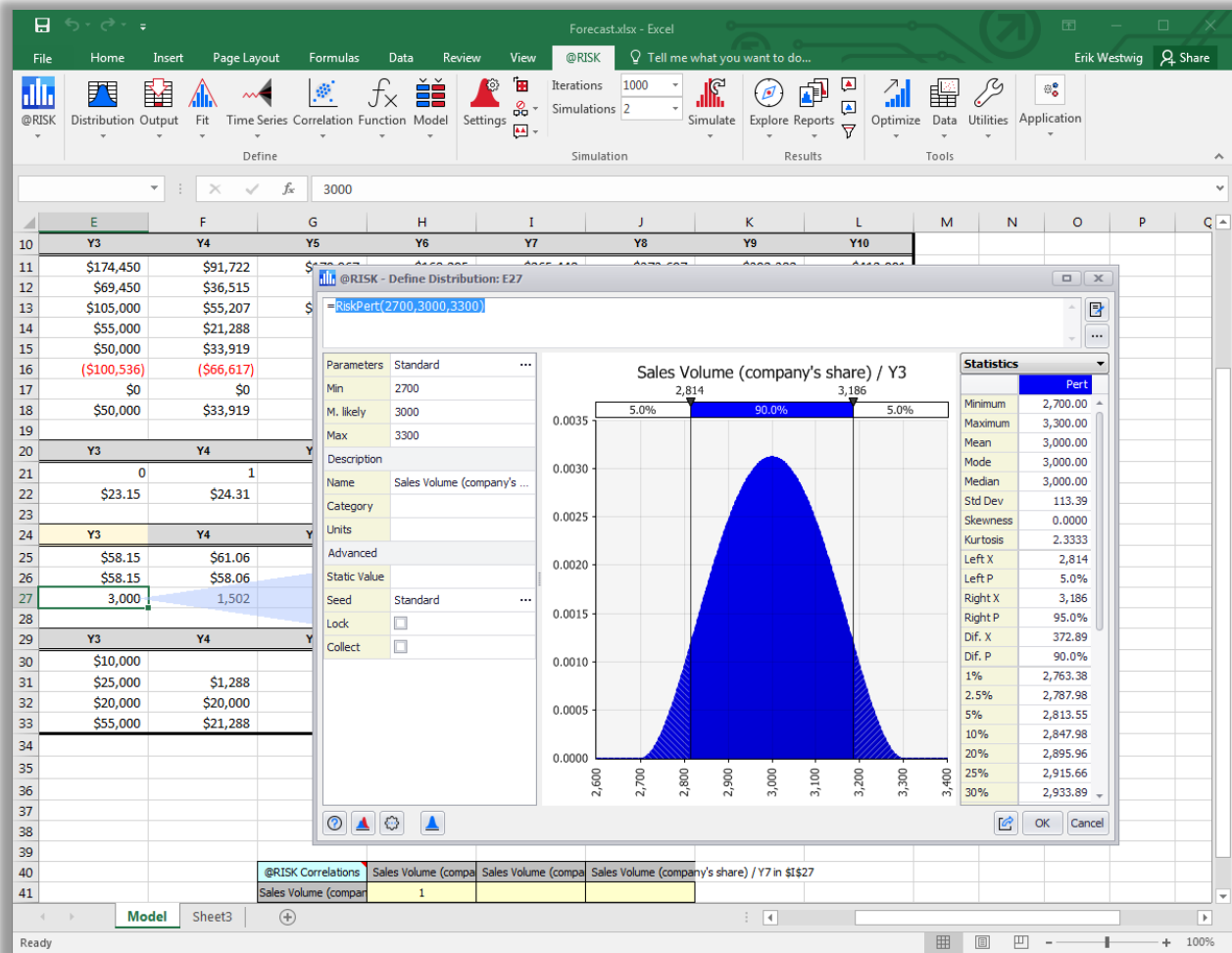
The RiskOutput function tells @RISK that this cell needs to be tracked during a simulation.

What is an input? What is a probability distribution? How do I add one to my model?

@RISK Inputs represent the uncertain values in your model. When you first made your model you may have used single-point estimates for these values. With @RISK, however, you will replace these single-point estimates with special @RISK functions that specify the amount of uncertainty around those estimates.

The most common type of uncertainty is due to randomness, and the most fundamental feature of @RISK is the ability to model randomness with probability distributions. In simple terms, probability distributions operate like dice – they are the ‘rules’ for which values are possible and how likely certain values or ranges of values are to occur.

For example, in the image below, the cell E27 had the fixed value of 3000, a point estimate for the sales volume of a new product. Clicking the Distribution button in @RISK brings up the Define Distribution Window from where you can replace that fixed value with a distribution.



In this example, a Pert distribution was used (a common choice for specifying a three-point estimate); in this case the estimate is configured with a minimum possible value of 2700, a most-likely value of 3000, and a maximum value of 3300. Once this dialog is closed, the cell will be updated with an @RISK distribution function:

`=RiskPert(2700,3000,3300)`

instead of the original fixed value. During the @RISK simulation process, random numbers will be drawn from the distribution.

As you get more comfortable with @RISK, you may choose to enter your distribution functions by typing them directly into the spreadsheet instead of using the graphical interface.

Which distribution should I choose?

To maximize confidence in your reported results, input distributions must be as realistic as practically possible. Each distribution has unique ranges of possible sampled values and associated likelihoods, so the 'wrong' distribution will create the 'wrong' simulation data. What is the 'right' distribution? In

general, this is a complicated question with many subtleties, a full discussion of which is beyond the scope of this document. There are guidelines to follow, however, that will enable you to create reasonable models. You might use the following hierarchy as a starting point.

Theory - Under certain circumstances, there will be a mathematical theory or logic that determines the technically correct distribution. These would usually form part of the formal training and education in a particular field. In common practice, some but definitely not all of your inputs could have a theoretical basis for selection. Caution must still be applied to these cases, as there may be assumptions in the theory (e.g. independence) that are not valid in your situation. Examples include the Central Limit Theorem, which indicates when the Normal distribution might be suitable for sums and Lognormal for products, Poisson processes that control the frequency and times between events, or using a Binomial to model the sum of several identical and independent Bernoulli processes.

Historical Data - @RISK allows for the fitting of distributions to historical data sets. This data is often a reasonable indicator of the distribution of future outcomes for the input, both in terms of the general shape as well as parameter estimates. You may also wish to include some personal judgement about the characteristics of the distribution, such as tail boundedness, and fit only to an appropriate subset of the possible distributions. @RISK's Fit feature enables you to perform these tasks in a simple GUI, returning a ranked list of fitted distributions for you to choose between. But historical data has many potential flaws that need to be considered. Is it truly representative of the potential future, i.e. how similar will the future conditions be to those in the past? Does the data only go back over a short window, and so are certain observations over or under represented?

Estimation - An absence of theory and reasonable (or any) data is not uncommon. In this situation, a suitable process must be followed to derive reasonable distributions and parameters. Most distributions are parameterized with Greek letters, the definition or practical understanding of which precludes robust estimation. There are very few distributions with common words for parameters such as "Minimum" and "Maximum", restricting your choices too severely and generally to bounded functions only. @RISK provides access to many distributions, both common and exotic, that can be parameterized using standard 3-point estimate techniques (as well as some 2- and 4-point estimates). These Alternate Parameter versions of the distributions are directly available in the Define Distribution window. A proper understanding of how to choose between these distributions is beyond the scope of this document, however it is reasonable to consider that an input that has a lot of variability or uncertainty would be reasonably modelled with a distribution that is unbounded in the direction of that 'riskiness'. Similarly, an input with relatively low uncertainty is probably best modelled with a bounded function.

Am I limited to a set of predefined distributions?

@RISK has over 90 functions in the Define Distribution window, although some are variations of one distribution (e.g. Normal, NormalAlt, NormalAltD). In addition to defining the classic distributions (Normal, Gamma, Binomial etc.) you can also create your own distributions in several ways. For

example, RiskCumul is a continuous function that requires user-specified Min, Max and multiple values and their associated cumulative percentiles. This allows you to completely specify the distribution shape to suit your needs. RiskGeneral is very similar but works with probability weights rather than percentiles. You can bootstrap from a data set using RiskResample, which has multiple sampling options such as “random with replacement”. The Distribution Artist window lets you draw a distribution curve in the window using your mouse, making it easy to visualize the distribution before using it in the model.

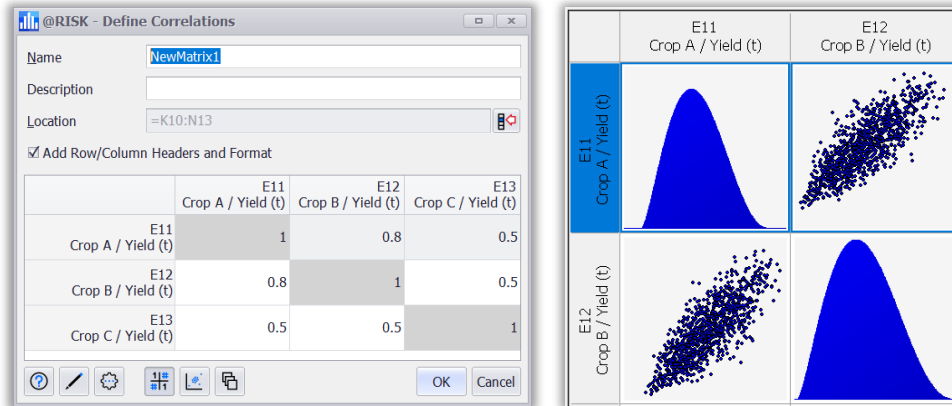
If you wish to build a distribution that isn't on the list, and can't be created using the other methods listed above, you just need its inverse Cumulative Distribution Function (CDF). Into this you place a RiskUniform(0,1) distribution, such that the resulting values after sampling from the Uniform will look like samples from your desired distribution. Finally, wrap a RiskMakeInput() function around your inverse CDF expression so that @RISK treats the formula as a regular input for purposes like sensitivity analysis and graphing.

Can inputs be correlated? What is correlation in @RISK?

Correlating inputs is a very important aspect of modelling. Many inputs are dependent on factors not present in the model, and as such must be correlated to generate realistic simulation data. Simply put, correlation controls the values that can be sampled *simultaneously* by two or more input distributions during a simulation without affecting their shape. This is often viewed as a scatterplot that shows the relationship. @RISK enables you to build correlation matrixes and define or fit copulas to represent the required dependency structure.

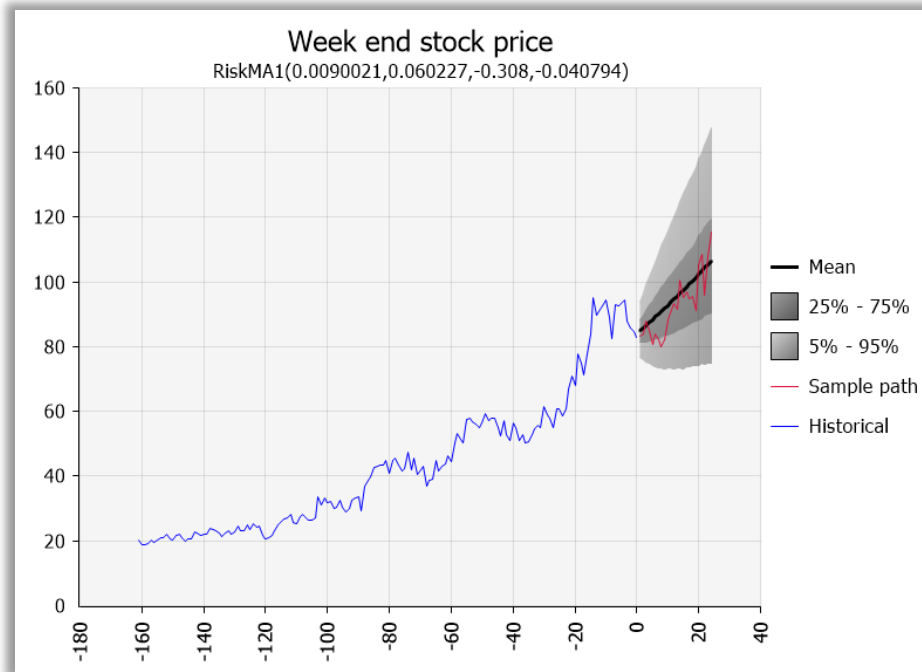
Copulas are generally a more sophisticated and complex concept. The matrix approach is more commonly used, involving the allocation of correlation coefficients to pairs of inputs. These coefficients can be estimated or calculated from historical data using @RISK's Data Viewer feature.

On the ribbon, click **Correlation** → **Define Correlation Matrix**, then choose the inputs to correlate. This leads to the Define Correlation Matrix window. Here you can import coefficients from an existing matrix in Excel, or define them directly in the window. After this, the matrix is placed in the spreadsheet where it is a 'live' parameter table, and can be edited or updated directly in Excel. In practice, correlating with @RISK means pre-ordering the sampled values to approximate the desired correlations in the simulation data. Below are views of the Define Correlation Matrix and a scatterplot of two positively correlated inputs.



What are time series?

A time series is a sequence of observations, typically measured at regularly spaced times, such as daily, weekly, or annually. Examples of time series include weekly currency exchange rates, quarterly product demand and annual commodity prices. Importantly, each observation is not completely independent – what happened in the past, especially more recently, influences the future. These relationships can be seen in such characteristics as long-term trending, seasonality and autocorrelation. @RISK includes three families of time series models – ARMA, GBM and ARCH models. Time series can be defined directly on the spreadsheet as an array function in a similar way to probability distributions. More commonly they will be fitted to historical observations of the series. These options, and the Batch Fit command, are found in **Define > Time Series**. Below is an example of a weekly stock price time series, both the historical data and forecast information.



What happens during an @RISK simulation?

When you click the Simulate button, @RISK recalculates your model over and over again, randomly sampling all of the probability distributions at the same time. The entire workbook is then recalculated by Excel to update all spreadsheet cells, and in particular all output cells, whose values are then recorded.

What's the difference between an iteration and a simulation?

A simulation is a process in which a spreadsheet model gets recalculated many times, each time with the input distributions returning randomly sampled values prior to the recalculation. The input data and corresponding output data from each recalculation is referred to as an "Iteration", and there are usually 10's or 100's of thousands of iterations, sometimes more. A "Simulation" is the collection of all iterations for a given set of model assumptions. Multiple simulations can be run, usually with a new set of assumptions for each simulation. The most common use of @RISK will have a single simulation run consisting of a reasonable number of iterations.

How do I decide on a number of iterations?

The number of iterations is precisely the amount of data you have with which to create graphs and statistics, and make your inferences about the future. In general, more data is better in that you will have more confidence when quoting an estimated statistic. However, additional data can come at a

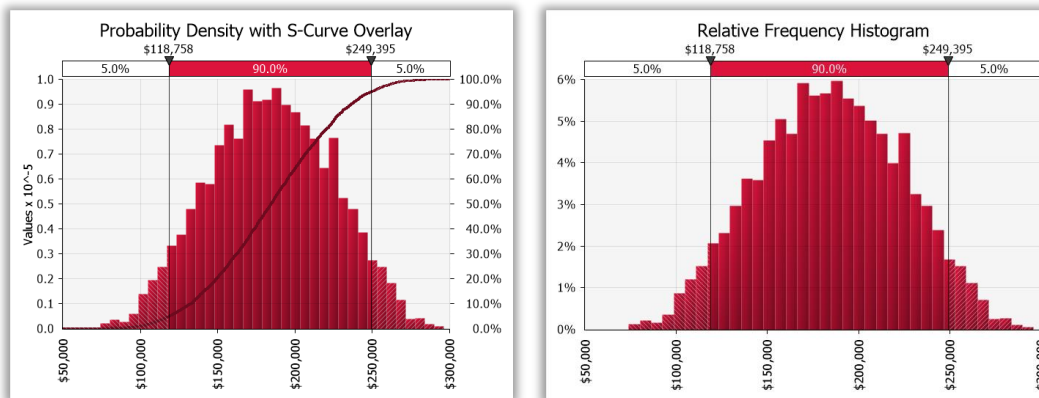
cost. There is the increase in simulation runtime, the time taken to create graphs and reports, and the space required to store the data.

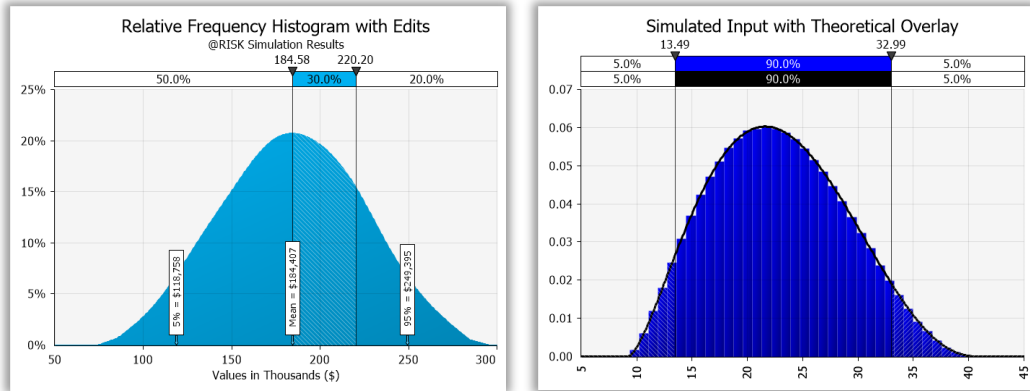
@RISK is capable of running millions of iterations for a single simulation! There is also the subtle consideration that improvements in estimated statistics taper off as more data is collected – there are diminishing benefits for the additional effort of running a larger simulation. One approach to determining the best compromise (in terms of iterations) is to use @RISK’s built-in convergence testing. Found in **Simulation Settings**, you can choose the statistic of an output to converge (mean, standard deviation, or any percentile), the acceptable tolerance, and statistical confidence of the test.

What kinds of graphs can I get for an input or output?

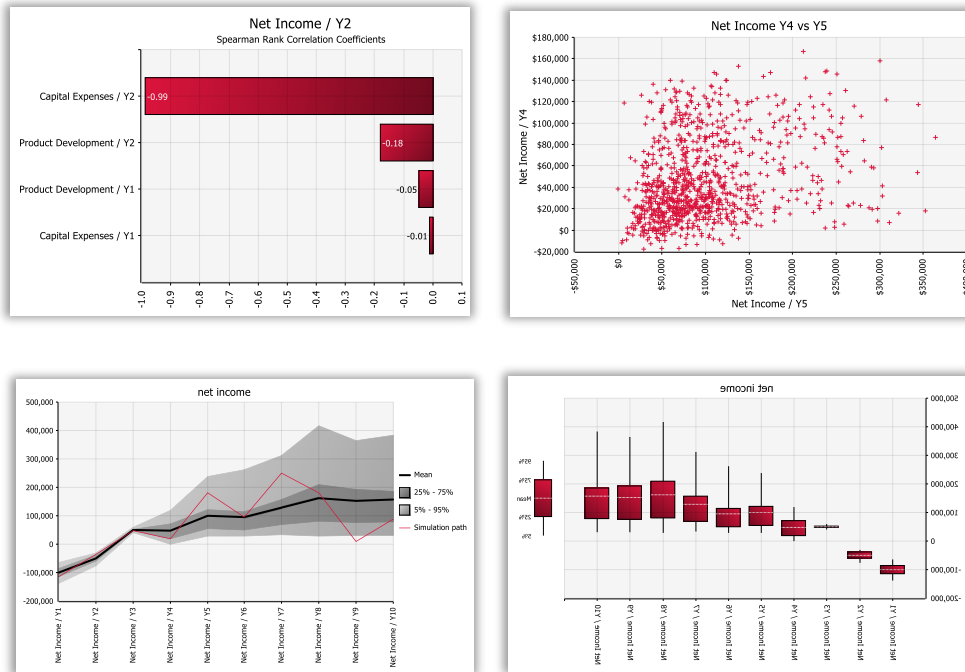
Distribution charts are the easiest way to view input and output distributions, and quickly understand their key characteristics. Both inputs and outputs can be represented in various ways, the most common being the Probability Density. Continuous inputs will look like smooth curves, and outputs are histograms.

The area under the curve is always one, as it represents the sum total of all possibilities. For this reason, the Y-axis can look strange if the variable is in large units (e.g. hundreds of millions) as it must maintain the area sum of 1. To achieve this, scientific notation is used (e.g. “Values x 10⁻⁸”). If this is an issue visually, the Relative Frequency Histogram can be used for Inputs or Outputs. The height of each bar is now the probability of sampling that region of the X-axis, and as such the heights of all bars sum to 100%. Conveniently this changes the values of the Y-axis to easy-to-read percentages. An alternative chart is the S-curve (and descending S-curve), where the Y-axis always goes from 0 to 1 and represents cumulative probability. These can be overlaid onto either histogram. Theoretical Input distributions can also be overlaid with their simulation data. All of the charts have many options for editing colors, statistics and other features.





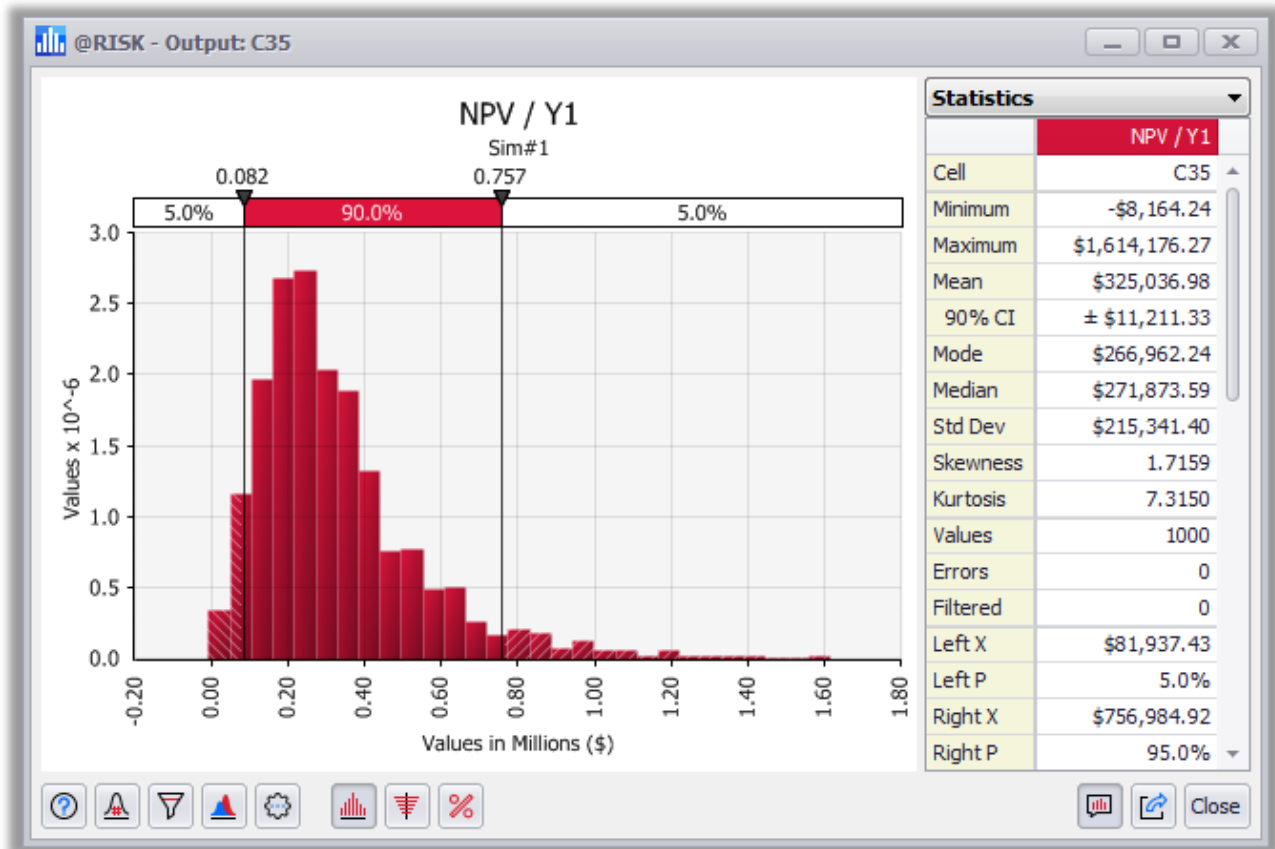
In addition, @RISK also includes many other graphs, including tornados, scatters, trends, and boxplots:



A list of all the possible output graphs is available on the Explore menu of the @RISK ribbon.

What kinds of statistics can I get for an input or output?

At the end of a simulation, the easiest way to get the most basic statistics for an input or output by using the delimiters and the statistics panel on the right side of an output results window:



The two delimiter lines allow you to answer basic questions about the distribution of values. In the graph above, for example, the lowest 5% of the values fall under \$82,000 while the top 5% are above \$757,000. By clicking on the delimiter lines and moving them with the mouse, or alternatively by typing in new values on the delimiter labels, other percentiles and targets can be explored.

The statistics panel to the right of the graph shows the most important measure of the distributed result (mean, standard deviation, median, etc.).

In addition, @RISK contains a wide array of statistic functions that can be typed directly into the spreadsheet. For example, the spreadsheet function

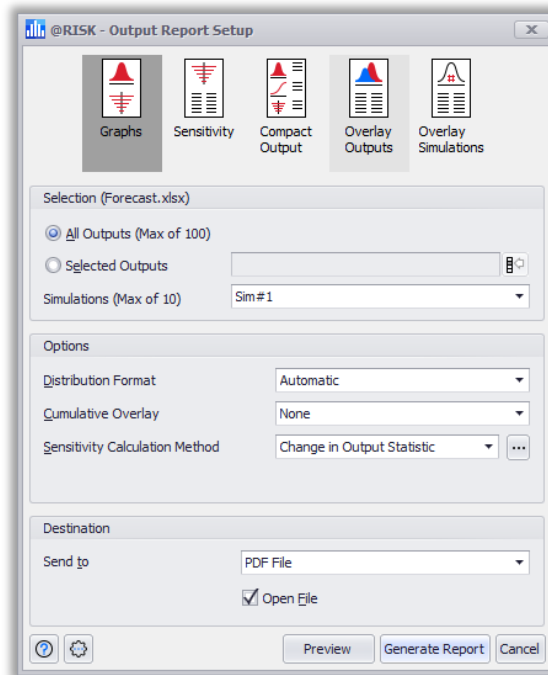
`=RiskMean(A1)`

Will return the simulated mean of the input our output in cell A1.

What reports does @RISK make for senior management?

Senior decision makers are interested in seeing the most important information quickly and easily, without clutter. This might include a distribution graph of one or more outputs, a handful of important statistics, and information on the key drivers (inputs) in the model. @RISK has an extensive range of standard reports, all of which are customizable.

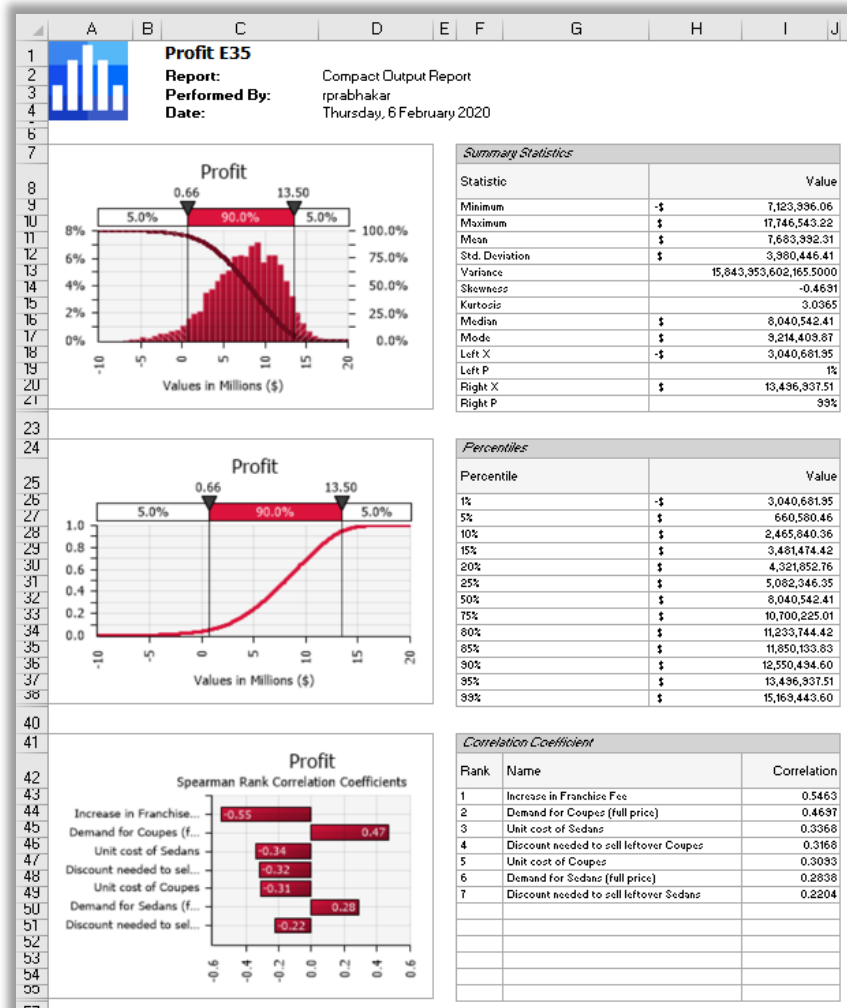
The Reports menu on the @RISK ribbon contains several categories of reports including Outputs, Detailed Statistics, and Sensitivities. Each category is associated with a dialog. For example, here's the dialog for the output reports:



At the top is a “gallery” of report formats. A good format to start your exploration of the reporting options is the first one: called “Graphs”. This generates a distribution graph of one or more outputs, as well as a tornado chart for each.

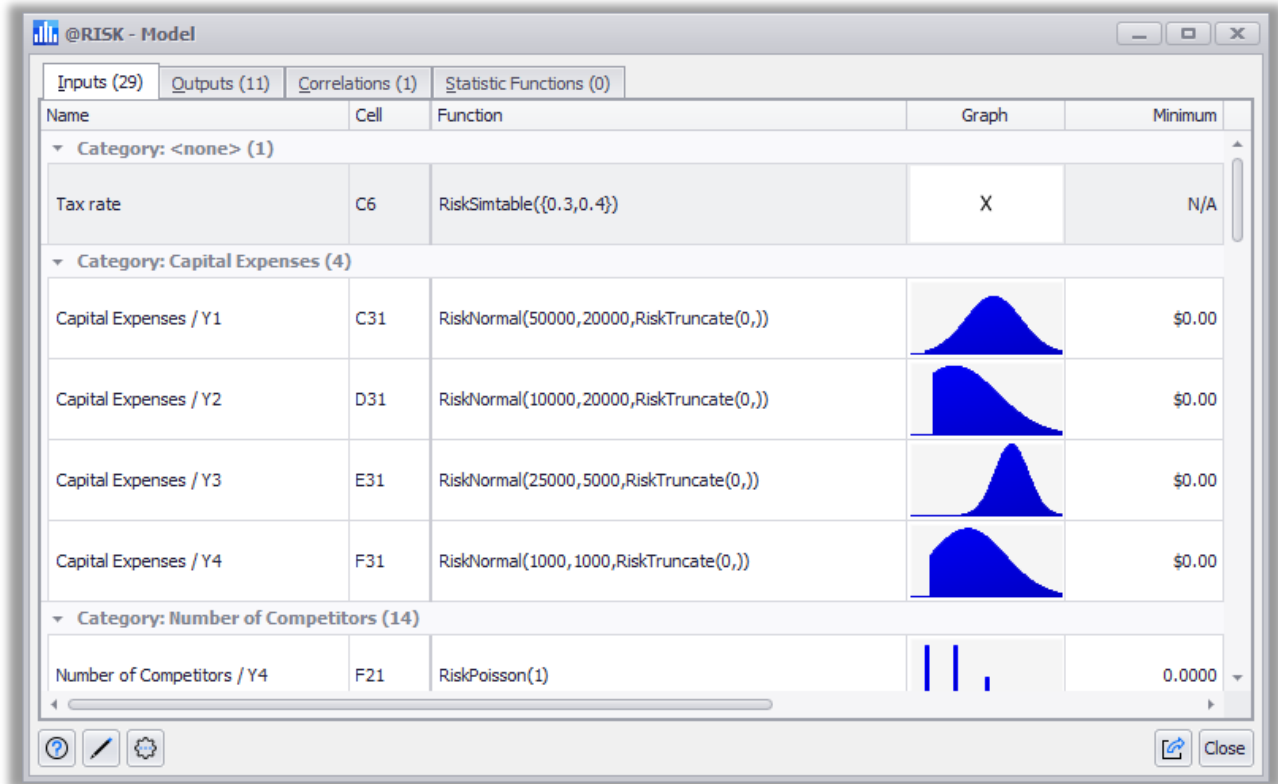
In the dialog, you can control the outputs to include in the report, the desired distribution format (density, S-curve etc.), the sensitivity calculation method (regression, change in statistic etc.). At the bottom the dialog, you control the destination for the report (e.g. new workbook, PDF file, etc).

An extra level of statistical detail is available from the “Compact Output” report, shown below. This includes smaller versions of the charts in the Graphs report, adds tables of statistics for the output (mean, standard deviation, various percentiles etc.) and the numerical values for the tornado chart.



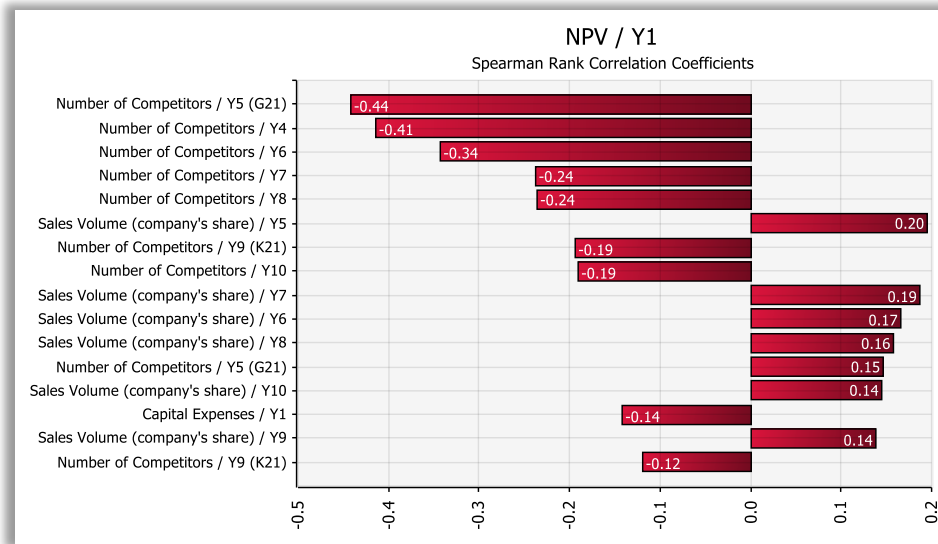
Can @RISK give me a summary of my whole model?

The @RISK Model Window gives you an overview of all the inputs, outputs, correlations, and statistic functions in your model each on a different tab of the window. To access this window, from the Define group of the @RISK ribbon, click the **Model** button.



How does @RISK tell me which inputs are driving an output?

@RISK provides several methods for examining the key inputs that affect each output. The most common way of showing these results is with a tornado chart, where the inputs are ranked by the magnitude of some form of sensitivity calculation. The length of each bar then represents the relative impact of its input. For example, in the tornado graph below, the inputs are ranked by their correlation coefficients.



There are sensitivity calculation methods, including regression, correlation and the “change in output statistic” method (a form of conditional sensitivity analysis). Many other analyses are also available. The reason for so many different sensitivity options is that there are many ways of defining and calculating what is “important” and the priorities of each decision maker could be different.

For example, the inputs with the “strongest” effect on an output is reasonably shown with a regression tornado, while the inputs with the most “consistent” effect on an output (i.e. they usually make a difference) are shown in a correlation tornado. These two analyses don’t always have the same rankings. This is especially noticeable when a model includes shocks or risk events; the binary nature of these inputs can give them a strong but inconsistent effect on the output.

The Change in Output Statistic tornado is a popular and flexible chart. Choose an important statistic to track (the default is the mean, but the mode or any percentile can be selected) and the tornado ranks inputs by their influence on the estimate of that statistic. This provides a very focused and specific version of how to rank inputs for importance.