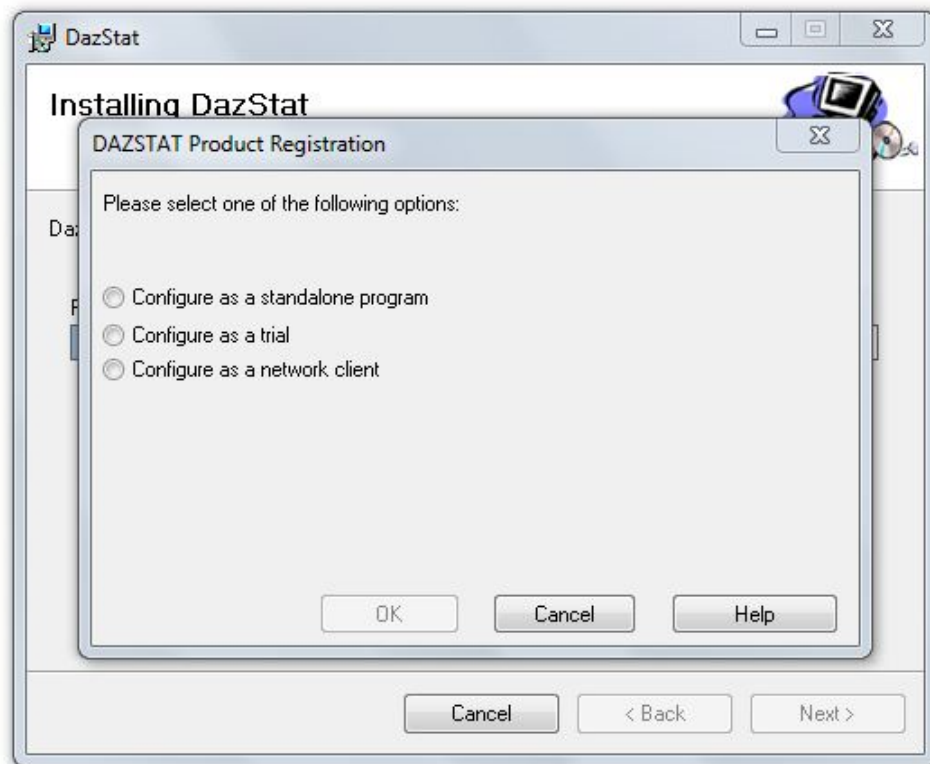# DazStat

Introduction

DazStat is an Excel add-in for Excel 2003 and Excel 2007.

DazStat is one of a series of Daz add-ins that are planned to provide increasingly sophisticated analytical functions particularly suited to investment analysts. Existing and planned modules include:

- DazRatio– analytical estimates of performance and risk
- DazBeta – analytical and functional estimates of correlation, covariance, and Beta
- DazAlpha – analytical estimates of Alpha, functional Alpha and Alpha stability
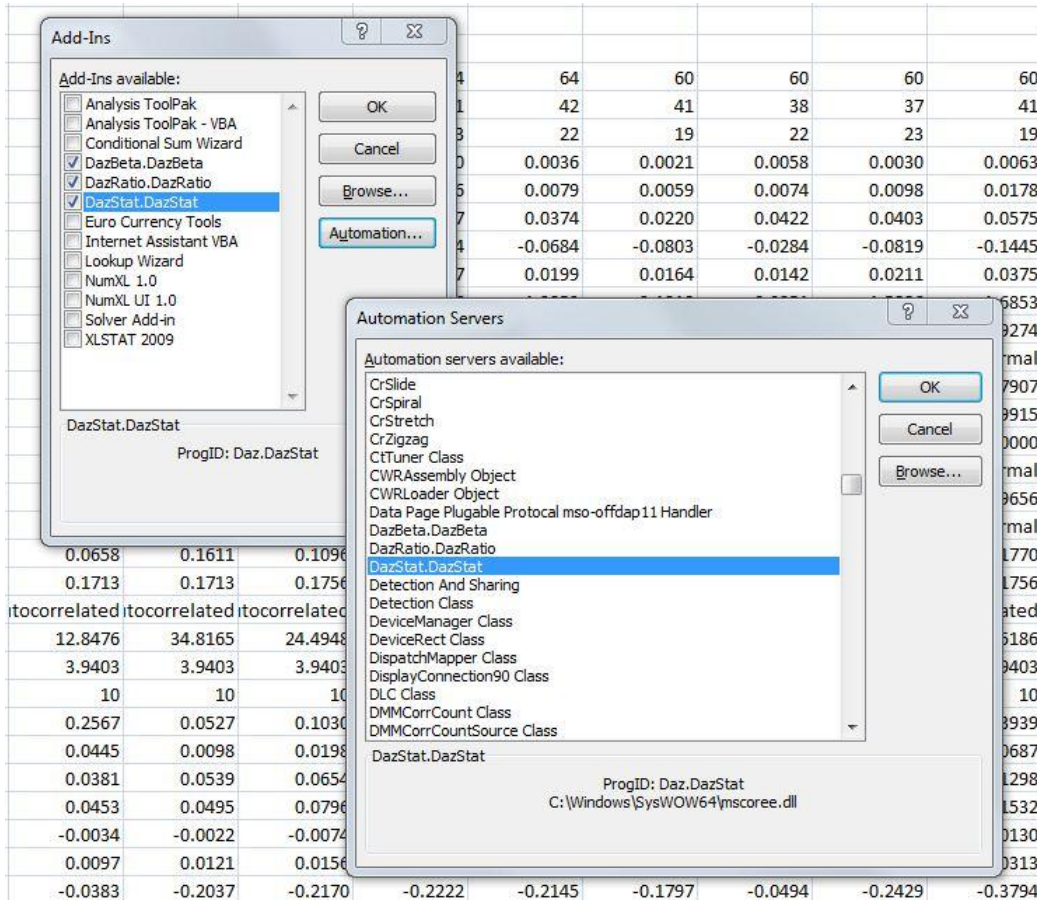- DazUtility – other functions that are useful and productive

Installation

DazStat is provided as an add-in application provided with an installer application. Double-click on the DazStatSetup installer, acknowledge the License Agreement, select whether to install for just the current user or all users on your computer and click through to completion.

DazStat uses a license manager application. This requires you to select the type of license you are registering. The license will be verified automatically.

After registration the License Manager will check the registration periodically when activated in Excel, but this takes just a moment. The License Manager is quite sophisticated and should not cause any problems if you need to re-install on a new drive, etc. If you have any problems please contact Information@DazRisk.com.

Add-Ins

Add-Ins available:
- Analysis ToolPak
- Analysis ToolPak - VBA
- Conditional Sum Wizard
- ☑ DazBeta.DazBeta
- ☑ DazRatio.DazRatio
- ☑ DazStat.DazStat
- Euro Currency Tools
- Internet Assistant VBA
- Lookup Wizard
- NumXL 1.0
- NumXL UI 1.0
- Solver Add-in
- XLSTAT 2009

OK Cancel Browse... Automation...

DazStat.DazStat
ProgID: Daz.DazStat

Automation Servers

Automation servers available:
- CrSlide
- CrSpiral
- CrStretch
- CrZigzag
- CtTuner Class
- CWRAssembly Object
- CWRLoader Object
- Data Page Plugable Protocal mso-offdap11 Handler
- DazBeta.DazBeta
- DazRatio.DazRatio
- DazStat.DazStat
- Detection And Sharing
- Detection Class
- DeviceManager Class
- DeviceRect Class
- DispatchMapper Class
- DisplayConnection90 Class
- DLC Class
- DMMCorrCount Class
- DMMCorrCountSource Class

OK Cancel Browse...

DazStat.DazStat
ProgID: Daz.DazStat
C:\Windows\SysWOW64\mscoree.dll

64	60	60	60	60
42	41	38	37	41
22	19	22	23	19
0.0036	0.0021	0.0058	0.0030	0.0063
0.0079	0.0059	0.0074	0.0098	0.0178
0.0374	0.0220	0.0422	0.0403	0.0575
-0.0684	-0.0803	-0.0284	-0.0819	-0.1445
0.0199	0.0164	0.0142	0.0211	0.0375

0.0658	0.1611	0.1096					1770
0.1713	0.1713	0.1756					1756
itocorrelated	itocorrelated	itocorrelated					ated
12.8476	34.8165	24.4948					6186
3.9403	3.9403	3.9403					9403
10	10	10					10
0.2567	0.0527	0.1030					3939
0.0445	0.0098	0.0198					0687
0.0381	0.0539	0.0654					1298
0.0453	0.0495	0.0796					1532
-0.0034	-0.0022	-0.0074					0130
0.0097	0.0121	0.0156					0313
-0.0383	-0.2037	-0.2170	-0.2222	-0.2145	-0.1797	-0.0494	-0.2429	-0.3794

In Excel 2003, go to Menu Tools/Add-ins. Select the Automation option, scroll down to DazStat (press D to get there fast) and click OK. DazStat will now show in the add-in list with a check mark to show that it is active.

In Excel 2007, go to the Microsoft Office Button and select Excel Options at the bottom. Select Add-in from the options on the left and click Go at the bottom. Select the Automation option, scroll down to DazStat (press D to get there fast) and click OK. DazStat will now show in the add-in list with a check mark to show that it is active.

DazStat has been tested with Excel 2003, 2007 and 2010 Beta on Windows XP, Vista and Windows 7.

The DazStat Examples file is a standard Excel worksheet that includes examples of the DazStat functions and how they may be used. These examples include the use with vertical return series, horizontal return series and return series in the percent not decimals. Note that the Stat switch can be referenced outside the function making it easy to select and change functions within an existing worksheet.

General

Function Variables

The function structure of DazStat is as follows:

>   =DazStat(Stat, Data, Freq, Pct, Label)

>   Stat (required)
>   The specific function result to be returned. The switch choices are listed below.

>   Data (required)
>   The data range for values to be used in the function.

>   Freq (optional)
>   The number of periods in a year to be used for annualized return values. The default value is 12.

>   Pct (optional)
>   A flag that forces DazStat to convert the data series into decimals (i.e. 4.32% or 0.0432) from percentages (i.e. 4.3200). The default value is True.

>   Label (optional)
>   A flag that forces DazStat to return the label (name) of the function rather than the result. The default value is False.

Note that in most cases you will only need to add the first two variables as the others are optional with reasonable defaults.

Rows and Columns

The data range specified in DazStat can be for either a row or column orientation.

Empty Values

As you can see in the Examples worksheet, DazStat allows you to run the analytics against data that may have initial and final empty values. Certain Excel function will automatically exclude empty values in the results while others treat an empty value as a zero value. DazStat creates its own internal array of valid values from the data range you specify to give you consistency and also ease in replicating the function across a range of data series.

Help

Unfortunately, no help structure is provided for so-called User-Defined Functions (UDF) in Excel. Some assistance is available in Excel if you enter =DazStat() in a cell and select the function builder *fx* next to the formula bar.

Statistics

Arrays are created by entering the DazStat function, selecting the number of cells for the array and pressing Ctrl-Shift-Enter.

0 Count

Cell: Returns the number of values included in the data range.

1 Positive

Cell: Returns the number of positive values included in the data range (includes zero).

2 Negative

Cell: Returns the number of negative values included in the data range.

3 Mean

Cell: Returns the mean (average) of the values included in the data range.

Note: First moment of the Normal distribution.

4 Median

Cell: Returns the median of the values included in the data range.

Note: Should be the same as the Mean if the distribution is Normal. One test of Normal-ity.

5 Max

Cell: Returns the maximum of the values included in the data range.

6 Min

Cell: Returns the minimum of the values included in the data range.

7 Standard Deviation

Cell: Returns the standard deviation of the values included in the data range.

Note: Second moment of the Normal distribution measures the dispersion of expected returns: 68.27% within +/- one Standard Deviation of the Mean. The basis for most investment risk measures – Annualized Volatility.

8 Skew

Cell: Returns the skew of the distribution of the values included in the data range.

Note: Third moment of the Normal distribution. Should be zero. One test of Normal-ity. Comparison with the Skew of the underlying asset returns would indicate active management, illiquidity or non-market pricing.

9 Excess Kurtosis

Cell: Returns the excess kurtosis of the distribution of the values included in the data range. This is equivalent to Excel's Kurt() function and is not the same as the actual Kurtosis of the distribution.

Note: Fourth moment of the Normal distribution. The expected Kurtosis is 3, so the actual value would be 3 + Excess Kurtosis. One test of Normal-ity.

Comparison with the Kurtosis of the underlying asset returns would indicate active management, illiquidity or non-market pricing.

10 Jarque-Bera

Cell: Returns a label describing the distribution of the values included in the data range as Normal or Non-Normal. Using a P-value of 0.05 as an acceptance level.

Array: Returns the label, the Jarque-Bera statistic (JB) and the Chi value for the 0.05 confidence.

Note: If the JB statistic is greater than the Chi value, then the Jarque-Bera test implies that the distribution of values departs from the Normal null-hypothesis with a greater than 0.05 confidence. Note that this implementation of the Jarque-Bera test is not robust with fewer observations than 200.

11 PPCC

Cell: Returns a label describing the distribution of the values included in the data range as Normal or Non-Normal using 0.99 as an acceptance level.

Array: Returns the label and the Probability Plot Correlation Coefficient for the data relative to a Normal distribution with the same mean and standard deviation.

Note: PPCC correlates the expected Normal distribution and the actual distribution, which produces a simple and easily understood result. However, the correlation coefficient should be above 0.9900 to have confidence that the actual distribution is Normal in a statistical sense.

12 Kolmogorov–Smirnov

Cell: Returns a label describing the distribution of the values included in the data range as Normal or Non-Normal using a P-value of 0.05 as an acceptance level.

Array: Returns the label, the Kolmogorov–Smirnov Z statistic for the differences between the expected Normal distribution and the actual data distribution and the approximate acceptance level of 1.36/Sqrt(n).

Note: Kolmogorov–Smirnov is a non-parametric statistic similar to PPCC, but measuring only a single point – the greatest distance between the expected and actual data.

13 Anderson-Darling

Cell: Returns a label describing the distribution of the values included in the data range as Normal or Non-Normal using a P-value of 0.05 as an acceptance level.

Array: Returns the label, the Anderson-Darling A^{*2} statistic for the differences between the expected Normal distribution and the actual data distribution and the P-value.

Note: Anderson-Darling is a powerful parametric test for goodness-of-fit to a parameterized distribution – Normal in this case.

14 Ljung-Box

Cell: Returns a label describing the distribution of the values included in the data range as Independent or Autocorrelated.

Array: Returns the label, the Ljung-Box Q statistic, the Chi value for the 0.05 acceptance level and the number of lag steps (k).

Note: If the Q statistic is greater than the Chi value, then the Ljung-Box test implies that the distribution

of returns is autocorrelated.

Note that the Ljung-Box test is considered to be very robust but depends on the number of lag steps used. DazStat sets the number of lag steps to 50 if the number of values is greater than 200, 3 if the number of values is less than 50 and otherwise uses 10 lag steps.

15 Wald-Wolfowitz

Cell: Returns a label describing the distribution of the values included in the data range as Random or Non-random.

Array: Returns the label, the Wald-Wolfowitz Z statistic and the P-value.

Note: Wald-Wolfowitz is a non-parametric test for randomness based on the number of observed runs in the data. Randomness is one of the properties that leads to a Normal distribution. Non-random may imply momentum, in-efficient market behavior and technical factors.

16 Cumulative Return

Cell: Returns the cumulative return of the values in the data range.

17 CAGR

Cell: Returns the Compound Annual Growth Rate of the values in the data range based on the Freq value entered into the function.

18 Annualized Volatility

Cell: Returns the annualized volatility of the values in the data range based on the Freq value entered into the function.

Note: Calculated as the Standard Deviation of the returns times the square root of the Freq. Not reliable as an estimate of 'risk' if the distribution of returns is not Normal. See Non-parametric Standard Deviation.

19 Non-parametric Standard Deviation - Proprietary

Cell: Returns the Non-parametric Standard Deviation.

Array: Returns the Non-parametric Standard Deviation, the Non-parametric Downside Dispersion and the Non-parametric Upside Dispersion.

Note: A Normal distribution has an *expected* 68.27% of the sample within +/- one Standard Deviation. The Non-parametric Standard Deviation measures the *actual* 68.27% range of the sample around the Mean return.

20 Non-parametric Annualized Volatility - Proprietary

Cell: Returns the Non-parametric Annualized Volatility derived from the Non-parametric Standard Deviation times the square root of the Freq.

Note: This is closer to the real volatility experienced by the return series. The Standard Deviation measured used for conventional Annualized Volatility assumes that the returns have a Normal distribution: symmetrical gains and losses, no excess Kurtosis or Skew. The Non-parametric Annualized

Volatility measures the actual range of returns (volatility or 'risk') rather than the expected range of returns.

21 Inter-quartile Volatility

Cell: Returns the annualized value of the inter-quartile volatility.

Array: returns the annualized value of the inter-quartile volatility, and the first and third quartile values.

Note: Measured as the difference between the 75 and 25 percentile values. This has been used as a non-parametric alternative to Standard Deviation, but it is measuring a different distribution (50%) rather than 69.27% - use Non-Parametric Standard Deviation for 68.27%. It is also measured without removing the Mean return. Inter-quartile Volatility is the 50% range of returns around the Median, which is assumed to be zero.

22 Max Drawdown

Cell: Returns the maximum cumulative drawdown of the returns.

Array: Returns the maximum, second and third non-overlapping cumulative drawdowns of the returns.

23 Max Run-up

Cell: Returns the maximum cumulative run-up of the returns.

Array: Returns the maximum, second and third non-overlapping cumulative run-ups of the values in the data range.

Ratios 98 and 99 are designed to provide a faster and more efficient method of producing all the DazRatio results by using an array rather than individual cell calculations.

98 All Results

Returns all 45 results in a single horizontal array.

99 All Results

Returns all 45 results in a single vertical array.

Tips

Take a look at the DazStat Examples spreadsheet for ideas on how to use DazStat.

Note that the DazStat structure is the same for all functions. This allows you to easily switch from one function to another or to use a cell value to determine the Stat result you want. That is one benefit of the DazStat module – you do not have to learn the eccentricities of Excel's implementation of functions.

Use the array results, even for individual functions. The array results provide information on the underlying or associated values to help you understand what is driving the value of the specific function result.

Use the Index function if you want to have a single cell that accesses a result from inside an array. For example, INDEX(DazStat(15,B$2:B$65,12),2) will return the Wald-Wolfowitz runs test Z score directly for the returns in B2:B65.

Look at the alternate functions for the same concept. For example, there are four different functions that measure the data's fit to a Normal distribution: Jarque-Bera, PPCC, Kolmogorov-Smirnov and Anderson-Darling. The differences tell you something about what happens when you change the assumptions.