

FasterAnalytics for Financial Services – A Case Study

Introduction

DecisionQ has developed FasterAnalytics, a unique analytics package that enables analysts and managers to use sophisticated predictive analytics from their desktops. FasterAnalytics is fast and creates high quality, predictive models from data that enable day-to-day review of financial and trading data, real-time hypothesis testing, and rapid decisions.

FasterAnalytics uses a modeling approach called Bayesian Networks to provide a mapping of the complex relationships in data, which can then be used to make high quality predictions. Users can:

- Get an instant global view of their data.
- Understand the driving factors in the data.
- Test hypotheses in real time in our model Explorer.
- Produce reports that can be exported to other applications.
- Make determinations that can help prioritize the use of scarce research resources.

Market Overview

The Financial Services industry spends billions of dollars on research globally. It is served by an array of software vendors selling tools that enable data analysis and modeling. The pace of financial markets requires that financial professionals constantly revise their models to minimize risk while maximizing profit.

Value to the Customer

FasterAnalytics enables both experts and non-experts in statistics to discover and leverage knowledge from large quantities of data quickly. Examples include:

- Automatically mapping data where targets are unknown to reveal correlations.
- Identifying both positively and negatively correlated relationships to a target variable.
- Discovering new relationships between variables and identifying new profit opportunities.
- Predicting the behavior of any factor or combination of factors in the model.
- Allowing analysts to develop new models in minutes, keeping pace with shifting markets.

FasterAnalytics is designed for real-time environments. Bayesian models are highly effective at identifying emerging trends that can be used to either to identify emerging profit opportunities or alert financial analysts to potential risk exposure.

Product and Technology

DecisionQ Corporation has produced a range of modules that perform data analysis, modeling, visualization, reporting, and decision optimization. FasterAnalytics modules include:

- *Discretizer.* Automatically configures the data for modeling.
- *Modeler.* Quickly creates a visual model of the data.
- *Explorer.* Allows real-time generation and testing of hypotheses.
- *Reporter.* Extracts insights and key points for inclusion in reports and presentations.

Using the System: An Arbitrage Example

The following is an example application of our software using publicly available market data. We have used a set comprised of 5 years of closing data across 42 indices, derivatives, and commodities. FasterAnalytics built the model in this example, from start to finish, in less than 15 minutes.

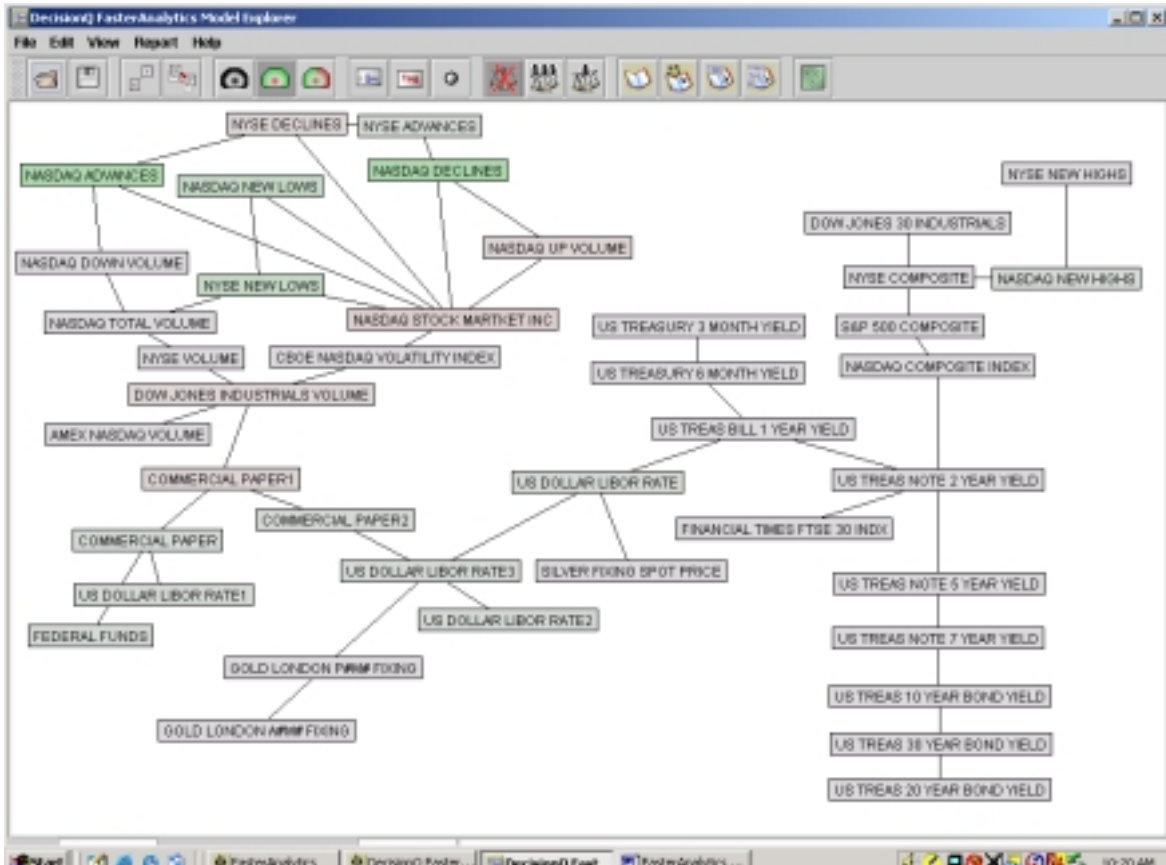
To build predictive models, our learning engine requires the data to be in a flat tabular format. (Figure 1) The data can be numerical, or variable character strings. Our software also handles missing values automatically and will either impute a value or treat missing values as a special category, at the user's discretion.

Figure 1: This example uses a financial data set held in an Excel spreadsheet as shown below (Partial).

Year	DOW JONES INDUSTRIALS	DOW JONES INDUSTRIALS VOLUME	AMEX NASDAQ VOLUME	CBOE NASDAQ VOLATILITY INDEX	NASDAQ COMPOSITE INDEX	NASDAQ ADVANCES	NASDAQ DECLINES	NASDAQ NEW HIGHS	NASDAQ NEW LOWS
1998	8343.728	53146.233	39.648	26.930	1770.533	2226.000	2021.000	2.000	0.240
1998	8350.458	53600.533	40.387	28.100	1758.540	1963.000	2294.000	2.040	0.290
1998	8384.830	57196.033	41.175	28.240	1757.140	2015.000	2229.000	3.500	0.290
1998	8339.240	51376.633	49.127	28.130	1759.700	1982.000	2176.000	3.630	0.380
1998	8444.330	57111.633	49.127	31.170	1711.930	1405.000	2822.000	0.990	0.440
1998	8369.390	56179.033	45.088	28.360	1753.490	2894.000	1313.000	3.720	0.280
1998	8567.140	60733.533	47.325	30.120	1725.160	1831.000	2403.000	3.930	0.380
1998	8643.130	53660.633	52.149	28.850	1748.530	2409.000	1903.000	3.930	0.370
1998	8675.730	48615.733	58.285	29.370	1756.850	2300.000	1971.000	2.200	0.270
1998	8659.560	46110.633	62.015	28.560	1764.060	2158.000	2025.000	3.930	0.250
1998	8802.530	48866.733	42.737	28.830	1771.680	2205.000	1920.000	3.930	0.370
1998	8718.850	49880.333	44.808	27.530	1788.180	2369.000	1869.000	2.230	0.330
1998	8749.990	51724.133	48.045	27.630	1779.380	1935.000	2310.000	3.940	0.360
1998	8775.400	48739.633	50.032	27.990	1788.280	2131.000	2023.000	3.930	0.340
1998	8803.050	54707.533	77.532	27.850	1769.980	2190.000	1963.000	2.530	0.310
1998	8906.430	78600.080	52.037	28.300	1789.160	3084.000	2111.000	3.630	0.240
1998	8816.230	52575.933	64.888	28.280	1792.530	1948.000	2343.000	3.780	0.310
1998	8904.440	46280.733	48.099	27.940	1813.440	2396.000	1903.000	2.030	0.280
1998	8872.800	38390.633	53.930	30.030	1824.530	2014.000	2209.000	2.680	0.270
1998	8846.890	52239.833	64.283	30.640	1828.540	2188.000	1983.000	3.590	0.410
1998	8796.080	57390.033	67.976	30.290	1823.630	2134.000	2038.000	3.930	0.380
1998	8782.120	44033.533	42.943	30.400	1818.700	1938.000	2318.000	3.990	0.270
1998	8799.810	63769.833	50.966	30.870	1825.680	2570.000	1762.000	2.300	0.280
1998	8868.530	79537.333	46.832	29.830	1847.680	2057.000	1998.000	2.680	0.380
1998	8886.640	74627.190	52.346	29.830	1852.960	2164.000	2133.000	2.890	0.340
1998	8883.410	65522.233	53.929	30.050	1855.480	2116.000	2119.000	2.880	0.290
1998	9033.230	76816.380	56.963	31.590	1829.140	1900.000	2521.000	2.300	0.340
1998	8956.500	74579.833	49.575	33.560	1798.730	1269.000	3034.000	3.040	0.340
1998	8891.480	71626.080	38.182	32.000	1807.030	2268.000	1926.000	3.190	0.330
1998	8994.860	58624.833	41.958	31.280	1820.240	2510.000	1701.000	3.630	0.380
1998	9012.300	62324.633	35.075	31.070	1824.950	1943.000	2296.000	2.220	0.370
1998	9110.200	61980.733	38.834	31.370	1843.030	2474.000	1839.000	2.680	0.380
1998	9162.270	67260.833	42.509	29.730	1863.260	2385.000	1936.000	2.790	0.330
1998	9076.570	64976.633	44.097	28.890	1858.240	1891.000	2438.000	2.110	0.410
1998	9167.500	62812.733	34.440	28.370	1866.600	2305.000	1864.000	3.830	0.380
1998	9141.840	60313.933	49.325	28.300	1887.140	2312.000	1981.000	2.370	0.410
1998	9184.940	61520.690	72.939	27.290	1903.870	2392.000	1930.000	2.790	0.230
1998	9176.720	53022.433	77.311	27.360	1917.630	2127.000	2170.000	2.500	0.230

Having selected the data, a fully automated process will continue until a full model is presented, or the user can stop each part of the process to manually change parameters in order to leverage particular domain expertise. The software begins by categorizing the data and 'binning' in accordance with the default settings; the data is then passed seamlessly to the Modeler for automated model development. Once the software has mapped all the complex correlations and causality in the data a graphical model is presented in the Explorer (Figure 2). This whole process takes only minutes.

Figure 2: Base case model of the data presented in Explorer



The display illustrates conditional dependence between variables and the pathways existing in the trading model. Notice that the network has several branches, one consisting of interest rates and commodity prices, one consisting of trading volume and directional data, and one consisting of index data.

In the example below, we examine how index performance affects volatility in bull vs. bear markets. We begin by selecting our target variable, CBOE NASDAQ VOLATILITY INDEX. The thick border indicates that this is the target selected, and its color red indicates that we are interested in analyzing how other variables behave when the target is 'high'. The coloring of the remaining nodes is red if the corresponding variable values are also 'high', and green if the corresponding variable values are 'low,' with tint of color being a relative measure of expression level.

Figure 3: CBOE NASDAQ VOLATILITY INDEX set to "high"

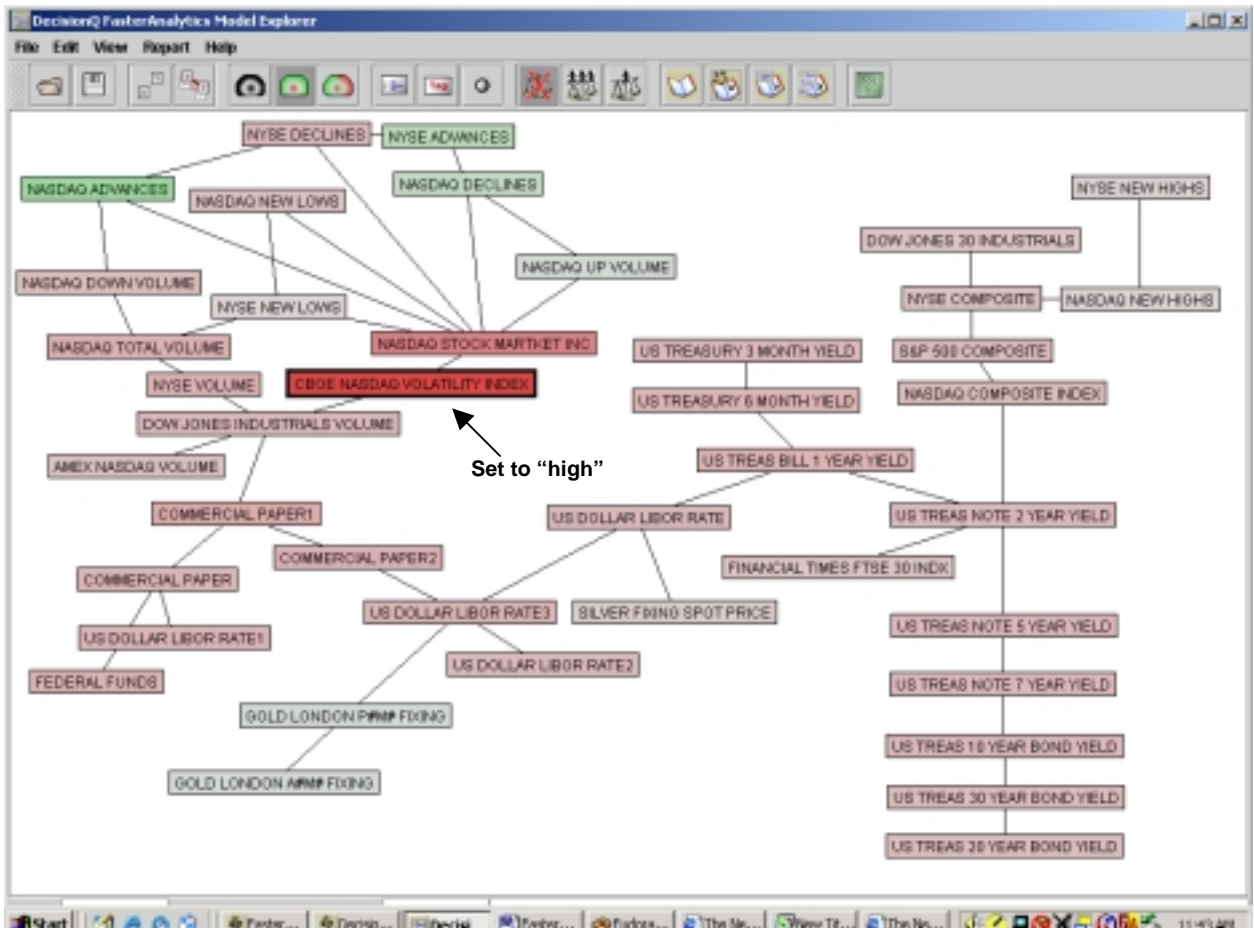
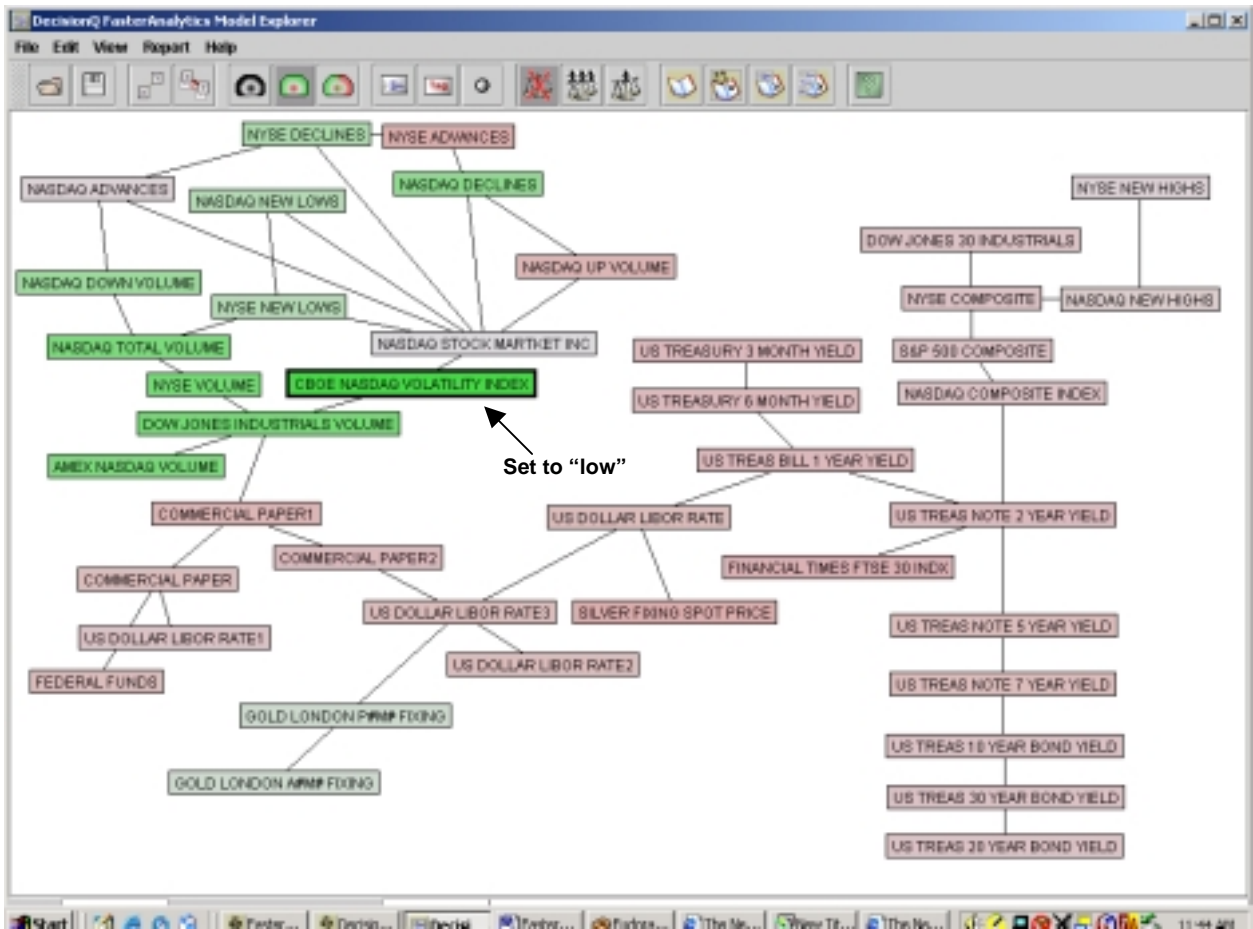


Figure 4: The complementary case analyzing CBOE NASDAQ VOLATILITY INDEX set to “low”

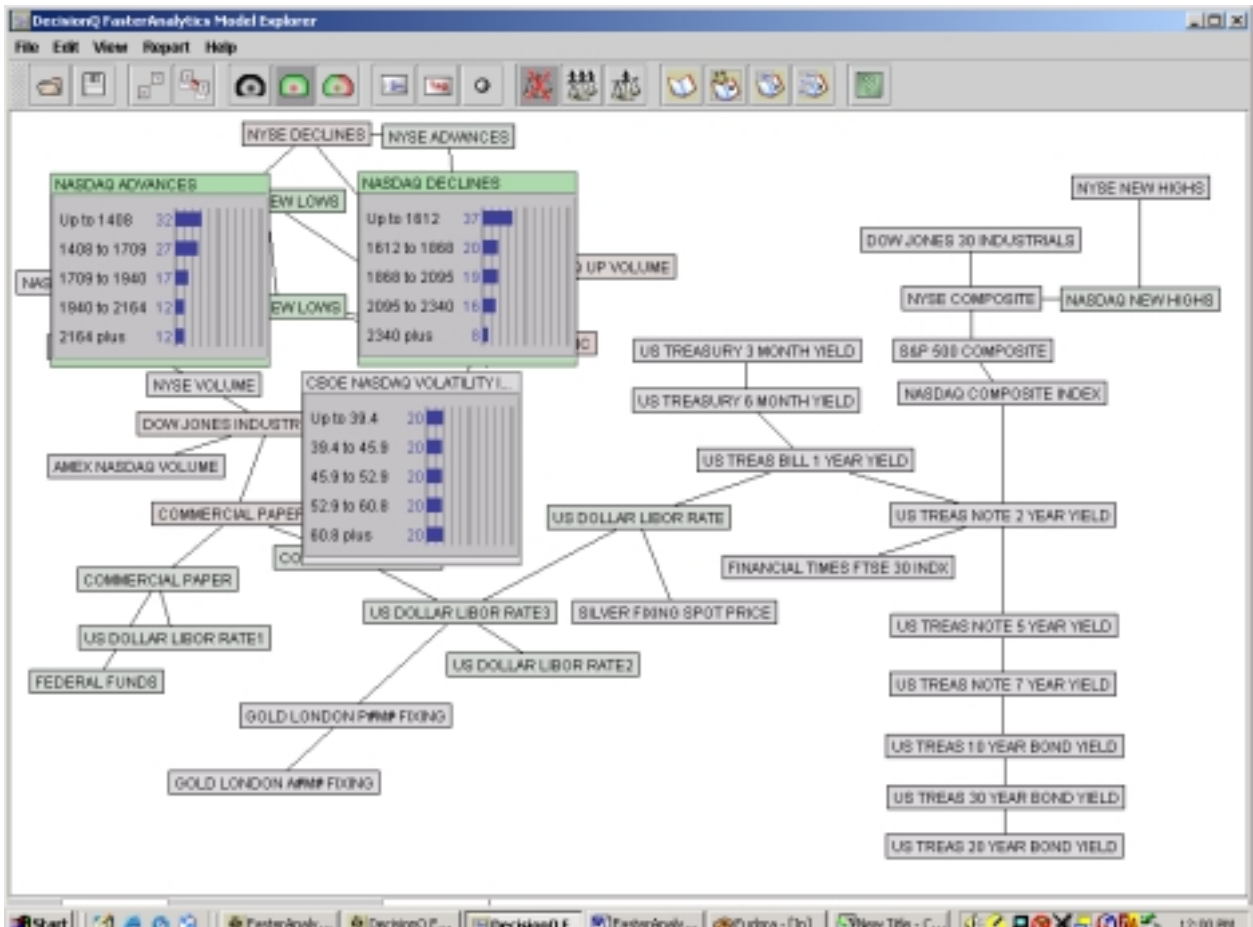


Compare the two models in Figure 3 and 4 above with the base level in Figure 1. It is also possible to select two or more variables simultaneously. The extent to which CBOE NASDAQ VOLATILITY INDEX is related to other variables in its neighborhood is intuitive and clear. This can be used to search for drivers of NASDAQ volatility.

Each variable can be expanded using the 'View' menu or icons to show quantitative information about the relationships. The population data is displayed as "cases" with bars that represent the marginal probability distribution of each case.

Suppose that we are interested in examining how NASDAQ DECLINES and NASDAQ ADVANCES affect CBOE NASDAQ VOLATILITY. We first select these nodes and click "Graph" to display the cases within these nodes. (Figure 5) This can be done for as many variables as we may choose.

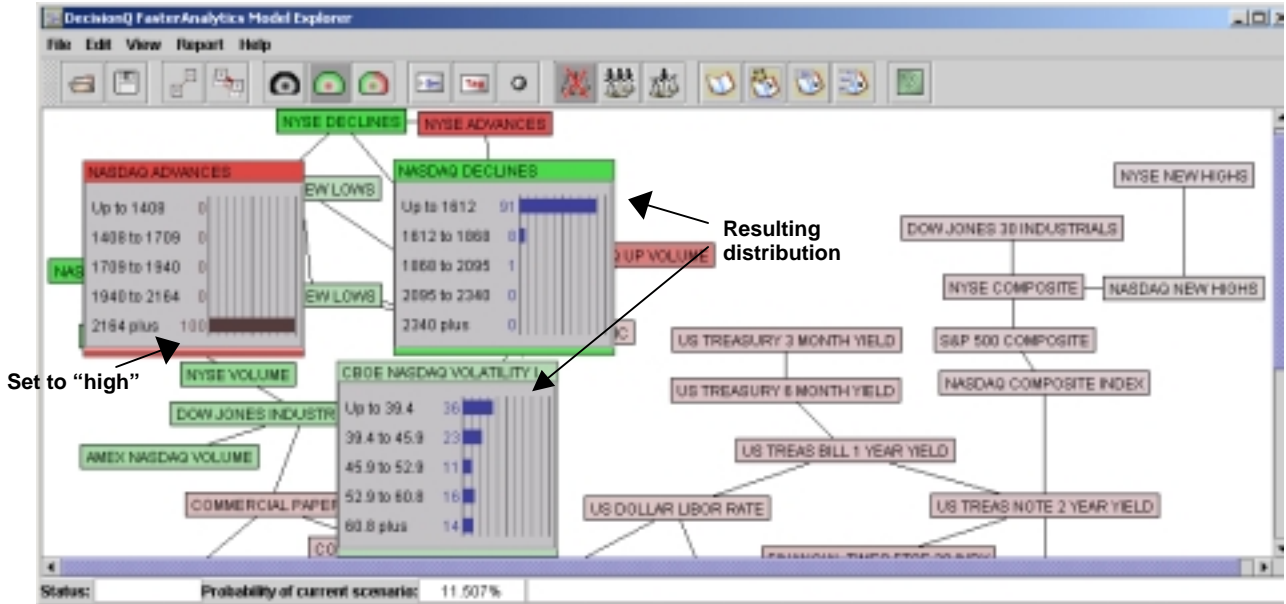
Figure 5: Quantitative information about NASDAQ declines and advances relative to volatility.



If we wish to test hypotheses, we can modify any node and see how our hypothesis affects the model. Notice how information flows through the network.

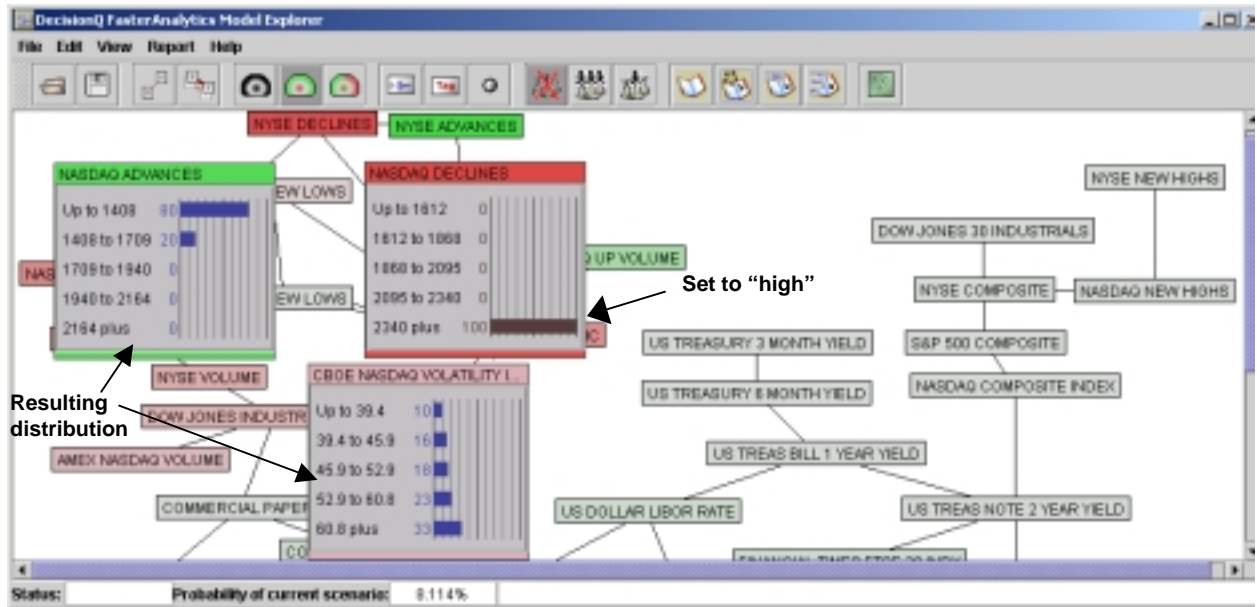
When we change the state of declines or advances, we see very clearly that a declining market has a higher level of volatility than an advancing market. We can see this graphically in Figures 6a and 6b below.

Figure 6a: NASDAQ advances and their effects on volatility.



S

Figure 6b: NASDAQ declines and their effects on volatility.



The Reporter module can be used to create a report that will show the conditional probabilities (or predicted likelihood) of any target variables, given the expression of any independent variable(s). Any part of the model visualization can be pasted into the Reporter module and then transferred into other applications. Figure 7 shows a sample report.

Figure 7: A sample report listing the probabilities of volatility given advances and declines.

Probability of case	Drivers		Target
	NASDAQ DECLINES	NASDAQ ADVANCES	CBOE NASDAQ VOLATILITY INDEX
0.406%	Up to 1612	Up to 1408	Up to 39.4 8.5 39.4 to 45.9 17.4 45.9 to 52.9 21.8 52.9 to 60.8 25.8 60.8 plus 27.3
3.436%	1612 to 1868	Up to 1408	Up to 39.4 8.7 39.4 to 45.9 16.7 45.9 to 52.9 24.2 52.9 to 60.8 23.3 60.8 plus 27.1
9.461%	1868 to 2095	Up to 1408	Up to 39.4 9.0 39.4 to 45.9 21.2 45.9 to 52.9 25.4 52.9 to 60.8 27.2 60.8 plus 17.1
12.07%	2095 to 2340	Up to 1408	Up to 39.4 8.6 39.4 to 45.9 15.4 45.9 to 52.9 25.4 52.9 to 60.8 23.9 60.8 plus 26.6
6.503%	2340 plus	Up to 1408	Up to 39.4 8.6 39.4 to 45.9 17.8 45.9 to 52.9 18.9 52.9 to 60.8 23.5 60.8 plus 31.1

DecisionQ sells predictive modeling software and complementary professional services. Alternatively, components from FasterAnalytics can be integrated into third party applications as part of broad data management and analysis platform.

If you have any further questions or would like to schedule a more detailed demonstration in person or over the web, please contact us.

DecisionQ Corporation
 3726 Connecticut Ave. NW, Suite 519
 Washington, D.C. 20008
www.decisionq.com
 Phone: 415-254-7996
 Fax : 415-276-6356
 Email: info@decisionq.com