

Visualization with JMP® – The Overlay, Spinning, Surface, and Contour Plots

By Mark Anawis, Abbott, Abbott Park, IL, U.S.A.

Abstract

The starting point for analysis should always be a graphical representation of the data. Graphics are a valuable tool for exploring and summarizing data. They often uncover patterns (e.g. clusters and outliers) and provide strategies for further analysis. These strategies may involve segmentation and/or analysis with and without outliers. Several of the plot platforms found in JMP® v.8¹: the Overlay, Spinning, Surface, and Contour Plots are useful in creating an understanding and conveying an understanding of data patterns.

Introduction

The starting point for analysis should always be a graphical representation of the data. Graphics are a valuable tool for exploring and summarizing data. They often uncover patterns (e.g. clusters and outliers) and provide strategies for further analysis. These strategies may involve segmentation and/or analysis with and without outliers. This brief presentation will serve as an introduction to visualization using several of the plot platforms found in JMP®: the Overlay, Spinning, Surface, and Contour Plots. These plots are found under the ‘Graph’ menu. We will be using version 8 of JMP® for this presentation, and all manipulations described here may be found in the JMP® online manual: the User’s Guide (predominantly in the ‘Graph Menu’).

The **Overlay Plot** creates plots of one X variable and multiple Y variables shown on a single X axis or multiple X axes. If two different scales are used for the Y axis, they can still be overlaid by using a right axis. Although the X variables can be Numeric or Character Model Types, the Y variables must be Numeric. It is important to understand that regression can not be performed using this platform. The ‘Fit Y by X’ should be used instead.

The **Spinning Plot** displays a 3-D view of points for three variables at a time from the columns selected. It is located under the Scatterplot 3D platform. It also performs an approximation of higher dimensions through principal components. It gives factor analysis style rotations of principal components which correspond to variable direction clusters.

The **Surface Plot** displays a 3-D view of surfaces for two variables at a time from the columns selected. It can display sheets, points, isosurfaces, or density grids. It also functions as an option in model fitting.

The **Contour Plot** shows contours of a response in a rectangular coordinate system. It allows several options including specifying the number of contour levels, line or filled contours, show points, label contours with response values, coloring schemes, and saving contour construction information in a data table.

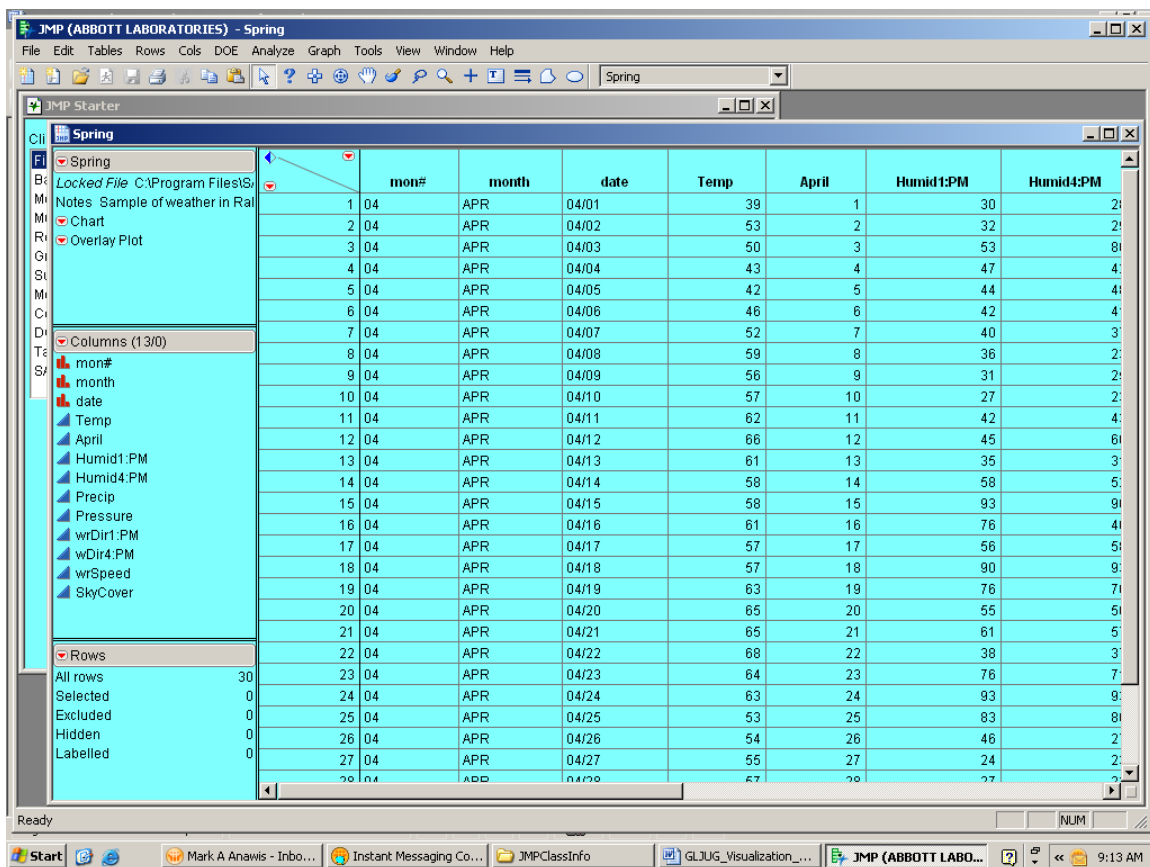
Overlay Plot

Two Variables on a Single Y-axis

We will use the Spring.JMP®² file. The data table is a sample of weather in Raleigh NC for the month of April, obtained from the weather service. It has a row for each day in April which is denoted in the column 'April'. We are interested in comparing humidity at different times of the day during the course of the month. Launch JMP® and select 'File', 'Open', and 'Spring.jmp'. It is located in the folder:

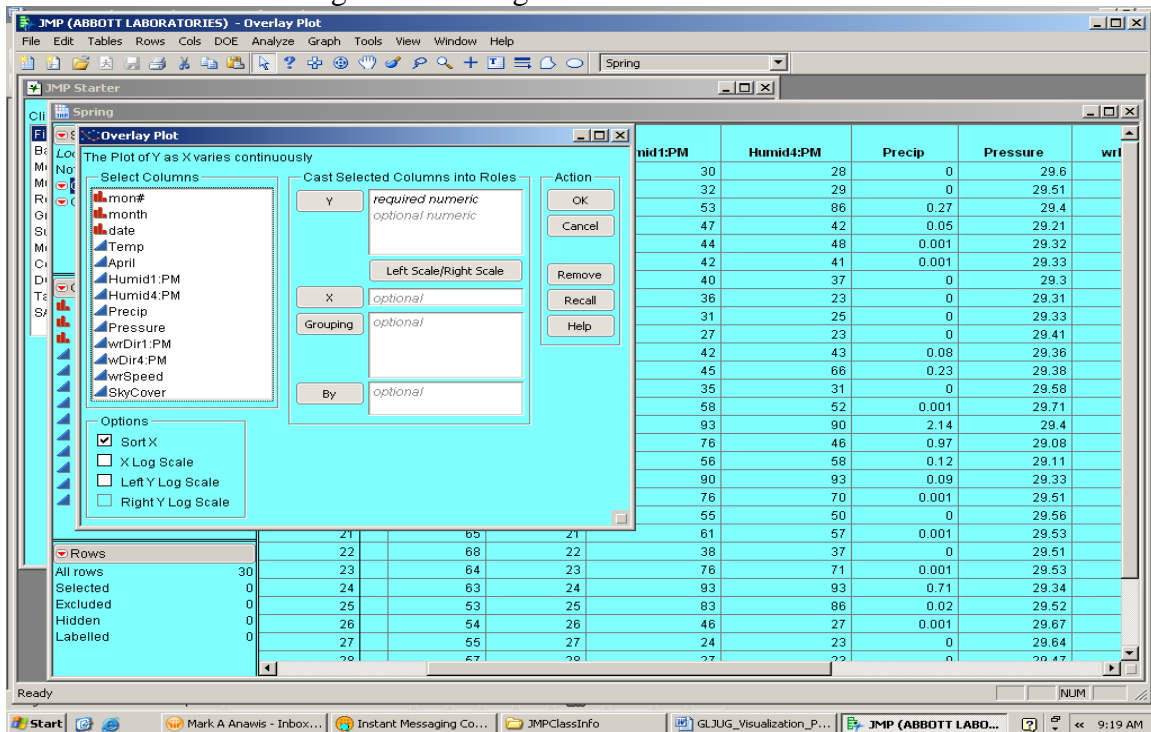
C:\Program Files\SAS\JMP\8\Support Files English\Sample Data.

The data file appears as follows:

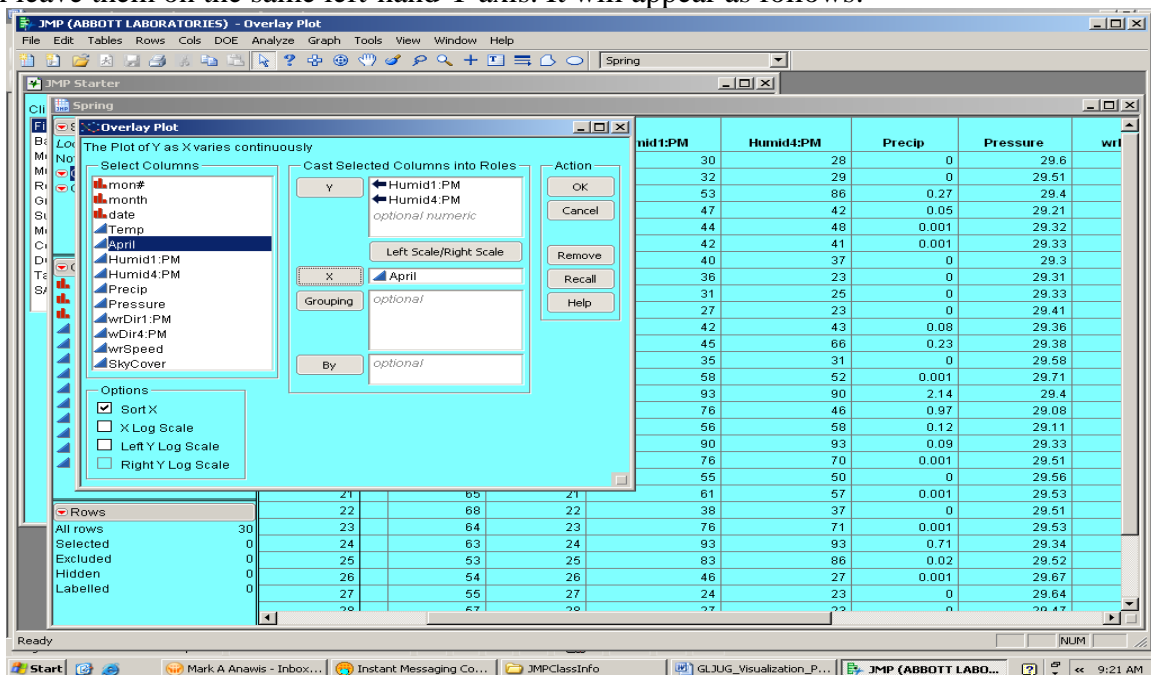


Select 'Graph' and 'Overlay Plot'.

You will see the following launch dialog:

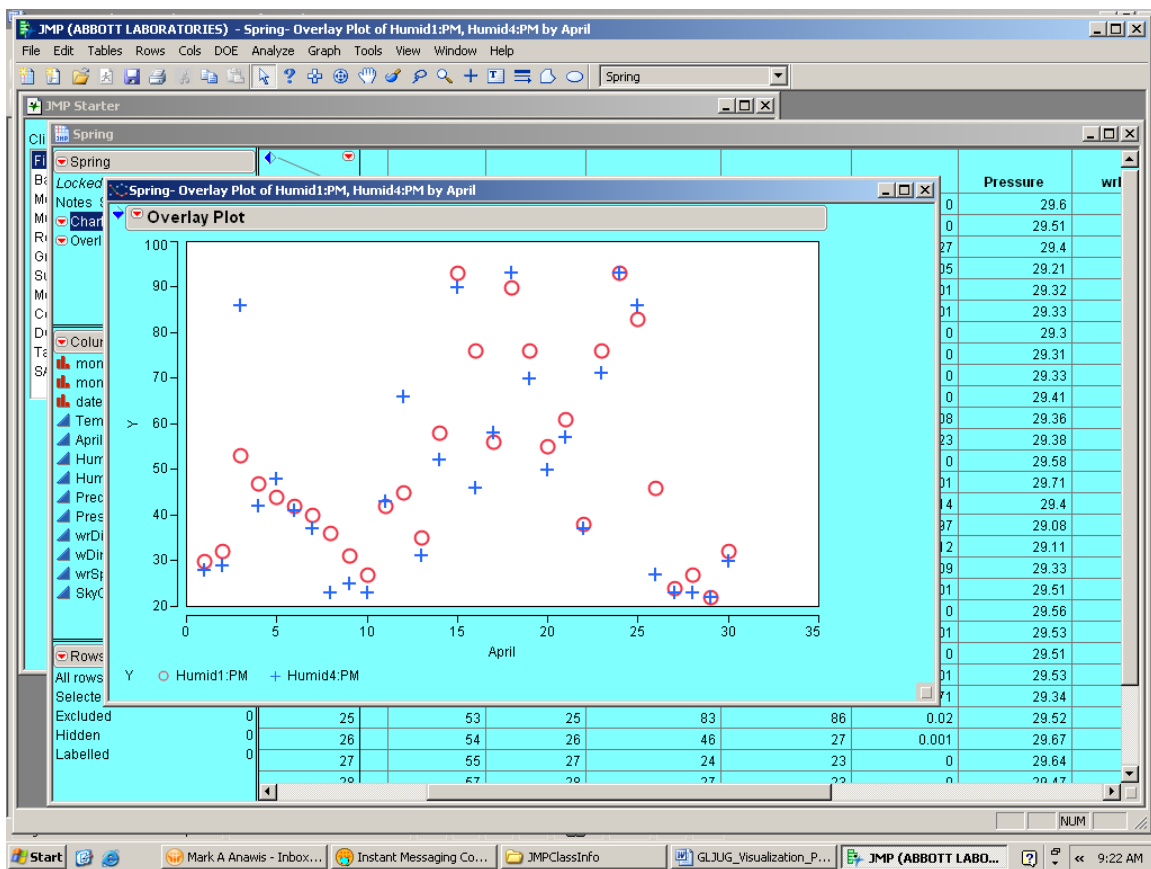


For Y, select 'Humid1:PM' and 'Humid4:PM'. For X, select April. The arrows signify on which X axis the variable appears. To put one of the Y variables on the right-hand axis, click on it in the Y role and click the 'Left Scale/Right Scale' button. Since these use a similar scale, we will leave them on the same left-hand Y axis. It will appear as follows:



Select 'OK'.

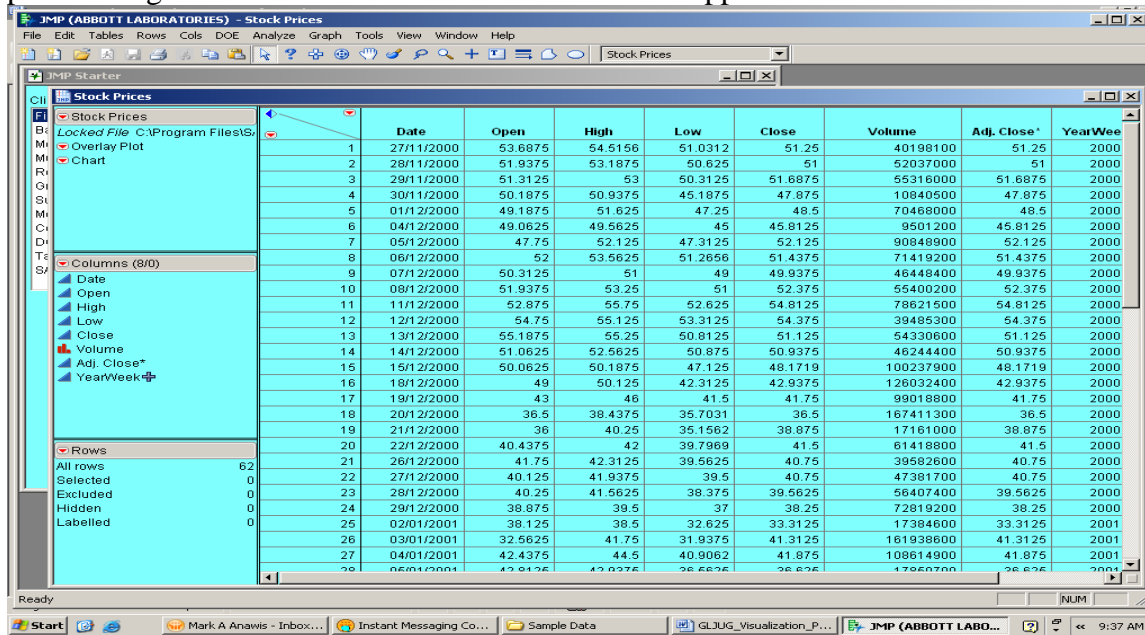
It will appear as follows:



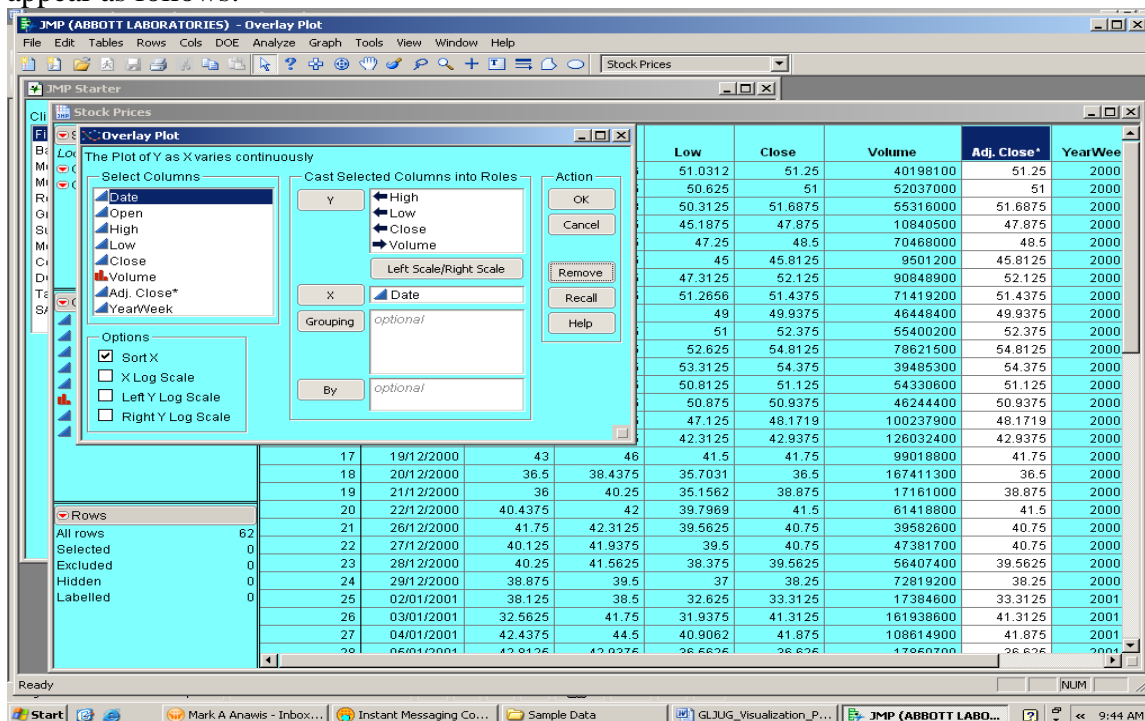
There is a legend at the bottom of the plot which identifies the markers and colors which identify each Y variable. From a review of the plot, we can see that the two humidity readings 3 hours apart generally track together and that one reading is not consistently higher than another. Most of the changes in humidity are due to daily differences.

Two Variables with a Second Y-axis

Open the 'Stock Prices.jmp²' file. This data contains a stock's opening, closing, high and low prices along with volumes for different dates. It will appear as follows:

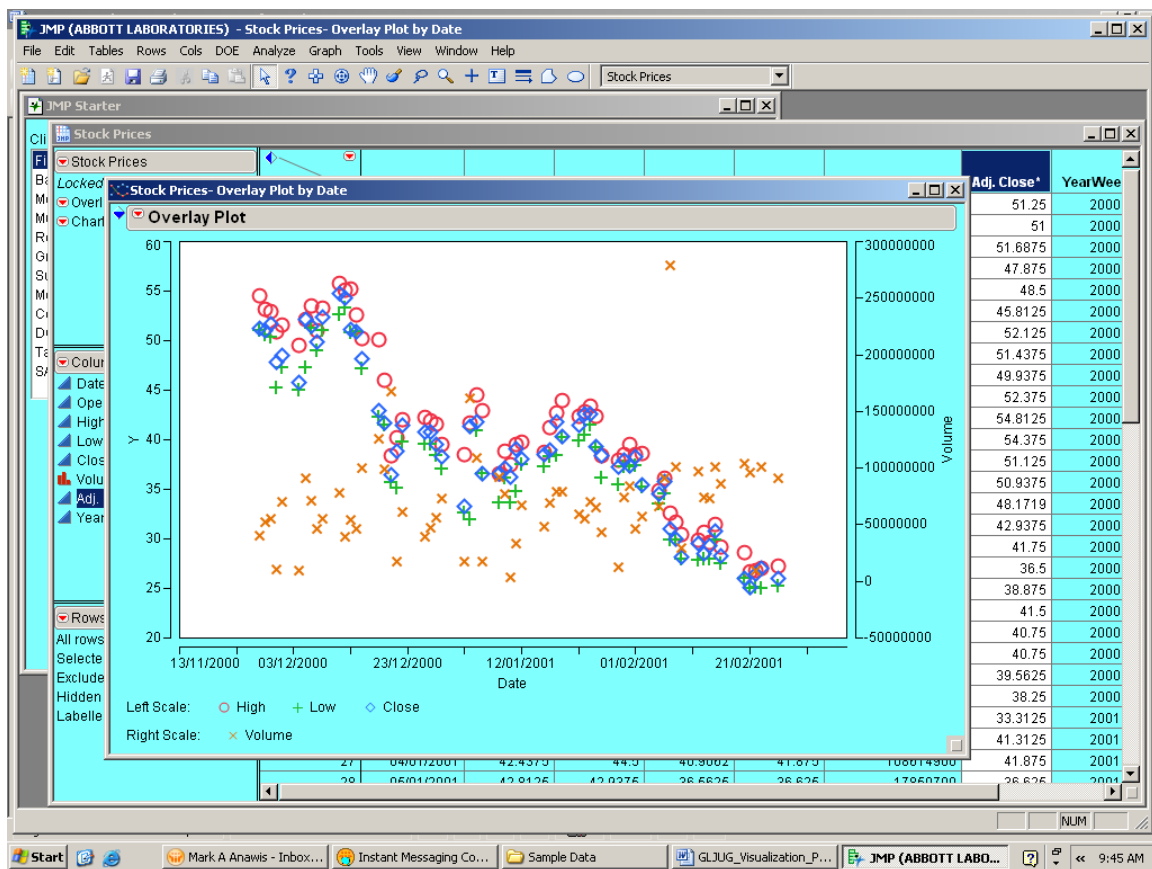


Although the prices and volumes are of different scale, we can still use the overlay plot, but with two different axis. Select 'Graph', 'Overlay Plot' and assign 'High', 'Low', 'Close', and 'Volume' as Y. Select 'Volume' and click 'Left Scale/Right Scale'. Select 'Date' as X. It will appear as follows:



Select 'OK'.

It will appear as follows:

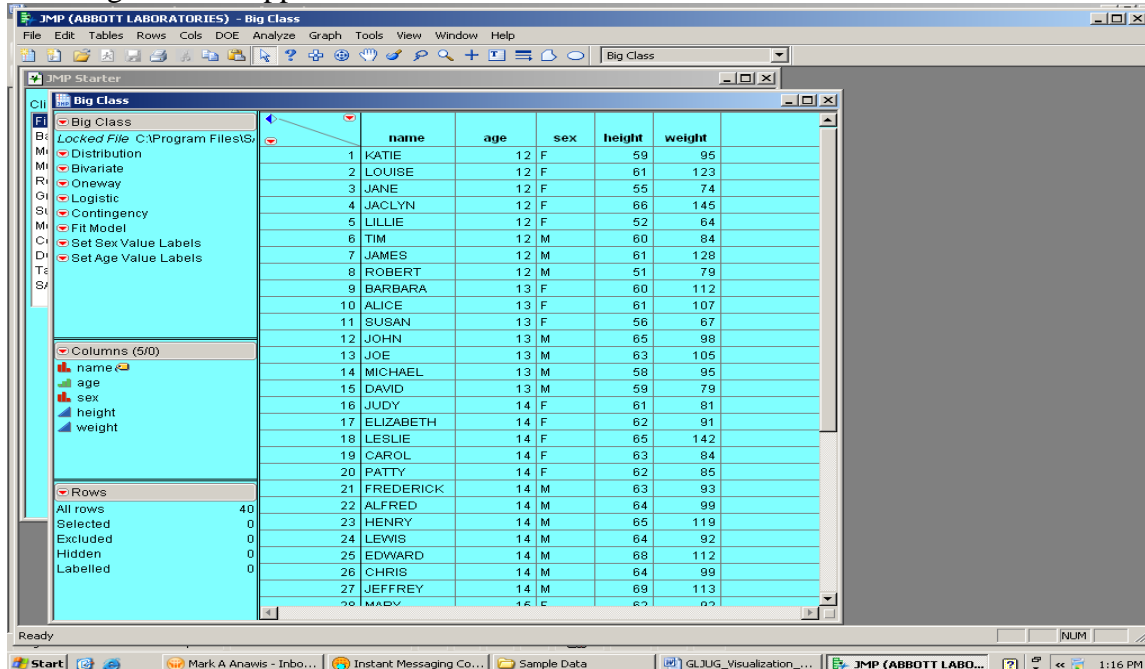


This plot show two stock price drops whereas the volume remains consistent over the duration of the plot.

Grouping Variables

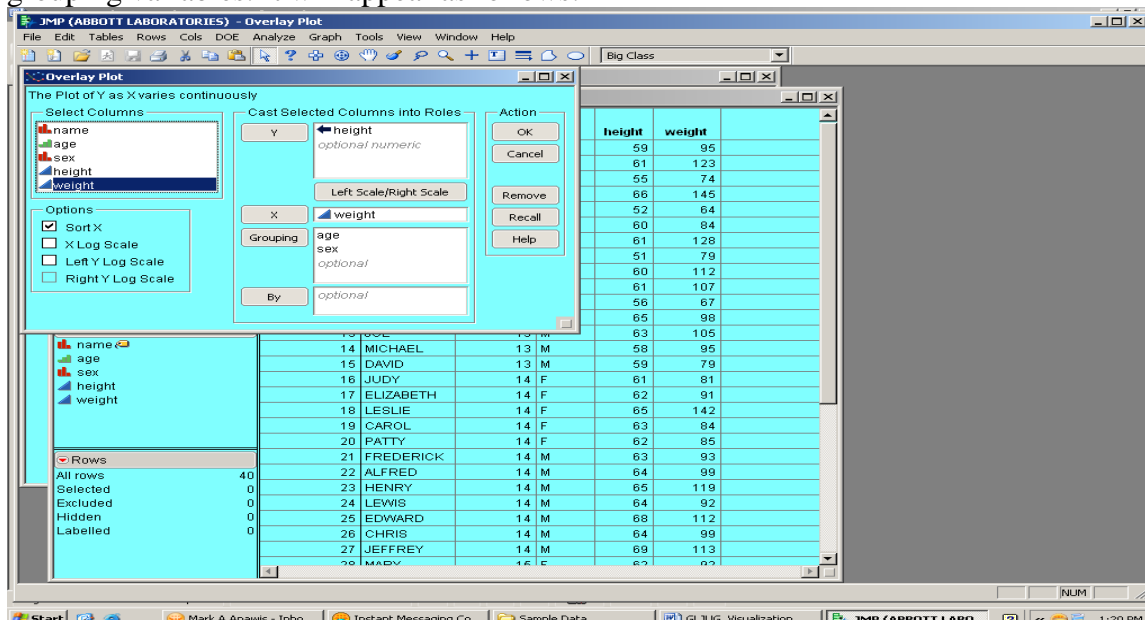
Several plots can appear in one window through the use of grouping variables. Using one grouping variable, the plots will appear stacked. Using two grouping variables, the plots will appear as a matrix. The 'Separate Axes' option shows plots with their own axes.

Open file 'Big Class'. This data set contains class information for students including sex, height, and weight. It will appear as follows:



	name	age	sex	height	weight
1	KATIE	12	F	59	95
2	LOUISE	12	F	61	123
3	JANE	12	F	55	74
4	JACLYN	12	F	66	145
5	LILLIE	12	F	52	64
6	TIM	12	M	60	84
7	JAMES	12	M	61	128
8	ROBERT	12	M	51	79
9	BARBARA	13	F	60	112
10	ALICE	13	F	61	107
11	SUSAN	13	F	56	67
12	JOHN	13	M	65	98
13	JOE	13	M	63	105
14	MICHAEL	13	M	58	95
15	DAVID	13	M	59	79
16	JUDY	14	F	61	81
17	ELIZABETH	14	F	62	91
18	LESLIE	14	F	65	142
19	CAROL	14	F	63	84
20	PATTY	14	F	62	85
21	FREDERICK	14	M	63	93
22	ALFRED	14	M	64	99
23	HENRY	14	M	65	119
24	LEWIS	14	M	64	92
25	EDWARD	14	M	68	112
26	CHRIS	14	M	64	99
27	JEFFREY	14	M	69	113
28	MARY	15	F	62	92

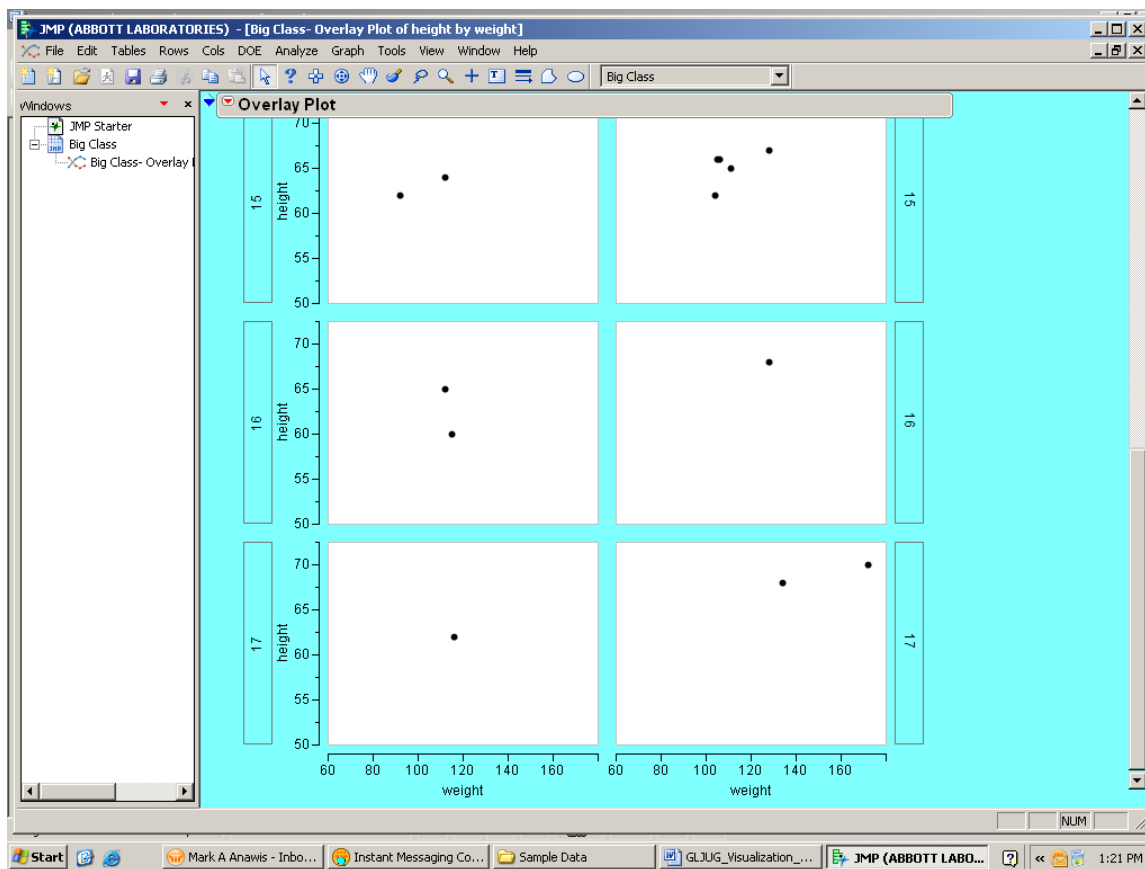
Select 'Graph', 'Overlay Plot' and assign 'height' as Y, 'weight' as X, and 'age' and 'sex' as grouping variables. It will appear as follows:



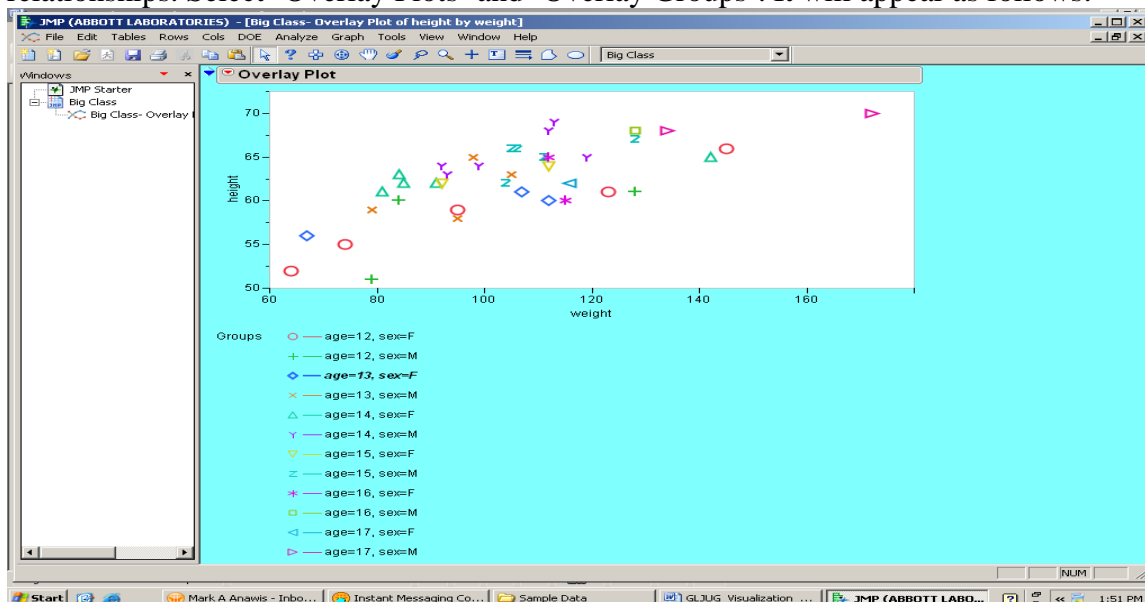
	height	weight
59	95	
61	123	
55	74	
66	145	
52	64	
60	84	
61	128	
51	79	
60	112	
61	107	
56	67	
65	98	
63	105	
58	95	
59	79	
61	81	
62	91	
65	142	
63	84	
62	85	
63	93	
64	99	
65	119	
64	92	
68	112	
64	99	
69	113	
62	92	

Select 'OK'.

It will appear as follows (Note: only a portion of the plot is displayed):



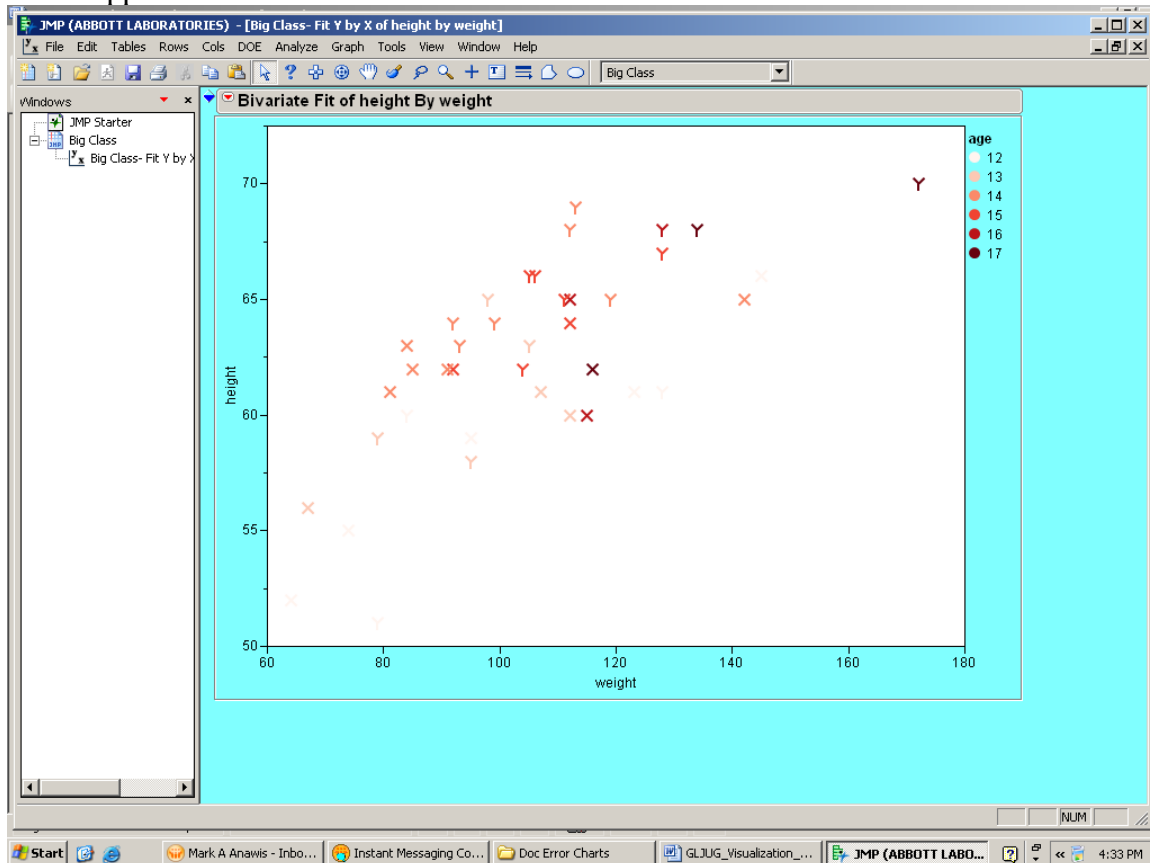
Although this may be useful in some situations, where there are too few points and many graphs such as in this situation, it would be better to have all the points in one graph to identify relationships. Select 'Overlay Plots' and 'Overlay Groups'. It will appear as follows:



This is more informative, but it would be better if sex and age were more easily identifiable. Let's go back to the data table. We will give males and females different markers and create a gradient for age. Select a cell with 'M'. Next to Row red arrow, select 'Row Selection' and 'Select Matching Cells'. Next to Row red arrow, select 'Markers' and 'Y'. All males will be marked by a 'Y'. Next to Row red arrow, select 'Row Selection' and 'Invert Row Selection'.

Next to Row red arrow, select 'Markers' and 'X'. All females will be marked by an 'X'. To create a graph, select 'Fit Y by X'. Select 'height' as 'Y, Response' and 'weight' as 'X, Factor'. Select 'OK'. To add a gradient of colors by age, right-click inside the graph and select 'Row Legend'. Select 'age' and for 'Colors:', select 'White to Red'. Select 'OK'.

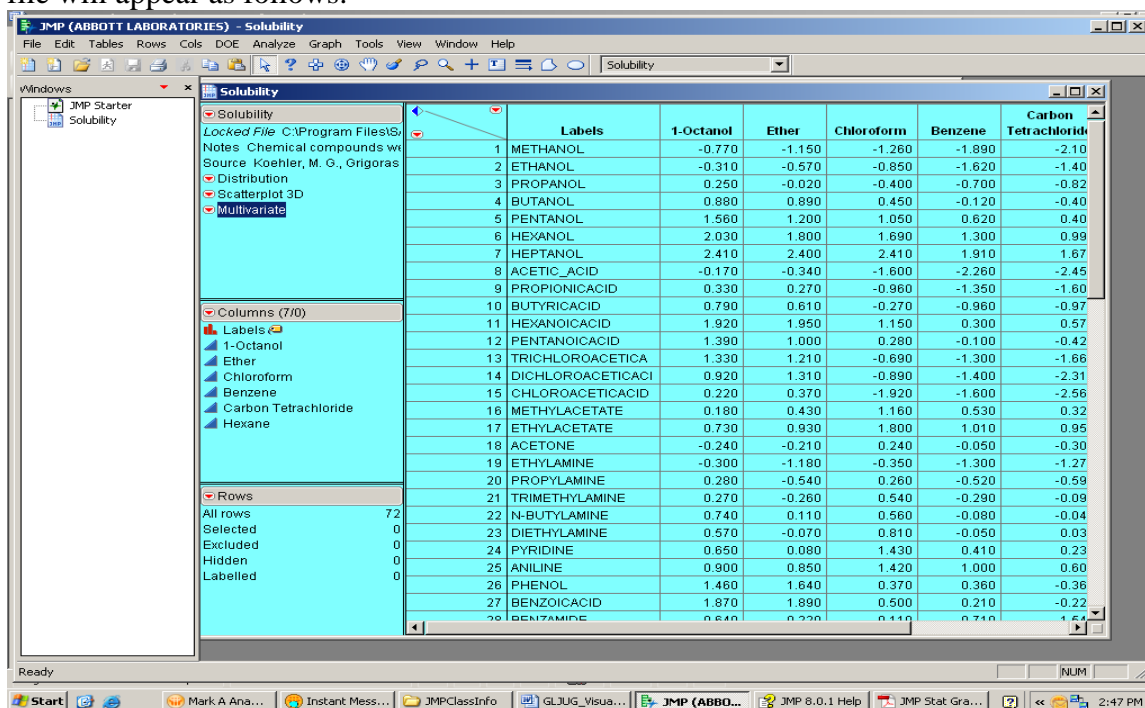
It will appear as follows:



We can see that among students over 120 lbs and over 65 inches, most are males and that the older students are also taller and heavier as we would expect.

Spinning Plots

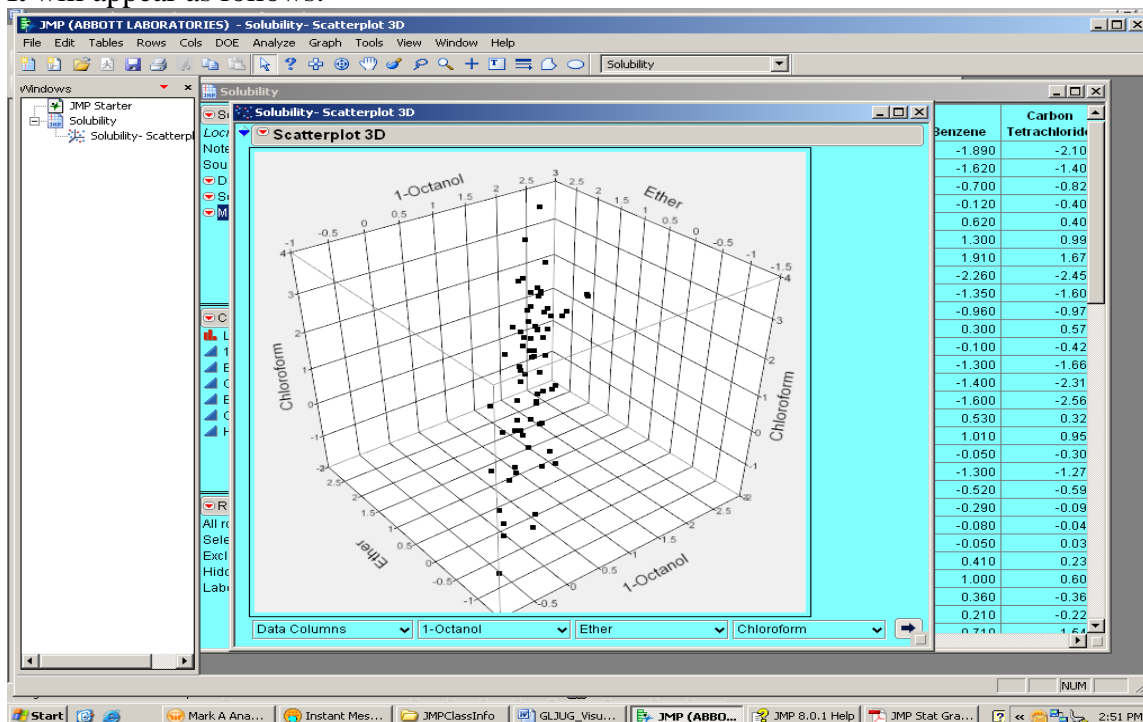
We will use the Solubility.jmp² file. Select 'File', 'Open', and 'Solubility.jmp'. For this data set, chemical compounds were measured for solubility in different solvents. This table shows the log of the partition coefficient (logP) of 72 organic solutes in 6 aqueous/nonpolar systems. The data file will appear as follows:



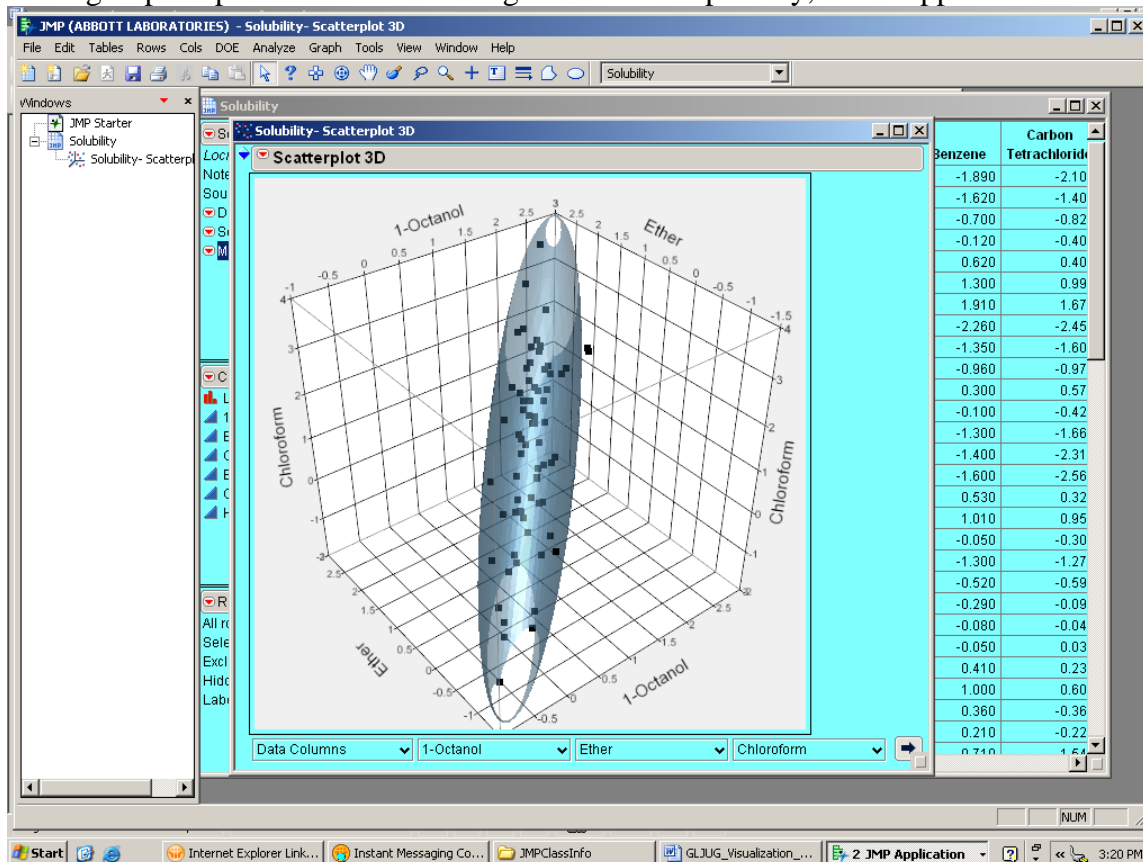
	Labels	1-Octanol	Ether	Chloroform	Benzene	Carbon Tetrachloride
1	METHANOL	-0.770	-1.150	-1.260	-1.890	-2.10
2	ETHANOL	-0.310	-0.570	-0.850	-1.620	-1.40
3	PROPANOL	0.250	-0.020	-0.400	-0.700	-0.82
4	BUTANOL	0.880	0.890	0.450	-0.120	-0.40
5	PENTANOL	1.560	1.200	1.050	0.620	0.40
6	HEXANOL	2.030	1.800	1.690	1.300	0.99
7	HEPTANOL	2.410	2.400	2.410	1.910	1.67
8	ACETIC ACID	-0.170	-0.340	-1.600	-2.260	-2.45
9	PROPIONICACID	0.330	0.270	-0.960	-1.350	-1.60
10	BUTYRICACID	0.790	0.610	-0.270	-0.960	-0.97
11	HEXANOICACID	1.920	1.950	1.150	0.300	0.57
12	PENTANOICACID	1.390	1.000	0.280	-0.100	-0.42
13	TRICHLOROACETICA	1.330	1.210	-0.690	-1.300	-1.66
14	DICHLOROACETICACI	0.920	1.310	-0.890	-1.400	-2.31
15	CHLOROACETICACID	0.220	0.370	-1.920	-1.600	-2.56
16	METHYLACETATE	0.180	0.430	1.160	0.530	0.32
17	ETHYLACETATE	0.730	0.930	1.800	1.010	0.95
18	ACETONE	-0.240	-0.210	0.240	-0.050	-0.30
19	ETHYLAMINE	-0.300	-1.180	-0.350	-1.300	-1.27
20	PROPYLAMINE	0.280	-0.540	0.260	-0.520	-0.59
21	TRIMETHYLAMINE	0.270	-0.260	0.540	-0.280	-0.09
22	N-BUTYLAMINE	0.740	0.110	0.560	-0.080	-0.04
23	DIETHYLAMINE	0.570	-0.070	0.810	-0.050	0.03
24	PYRIDINE	0.650	0.080	1.430	0.410	0.23
25	ANILINE	0.900	0.850	1.420	1.000	0.60
26	PHENOL	1.460	1.640	0.370	0.360	-0.36
27	BENZOICACID	1.870	1.890	0.500	0.210	-0.22
28	BENZAMIDE	0.640	0.220	0.110	0.710	1.54

Select 'Graph', 'Scatterplot 3D' and all the variables as 'Y, Columns'. This allows you to select different groups of 3 later. If you specify a 'Weight' variable, it will draw the points as balls with their volumes scaled by the 'Weight' variable. Select 'OK'.

It will appear as follows:



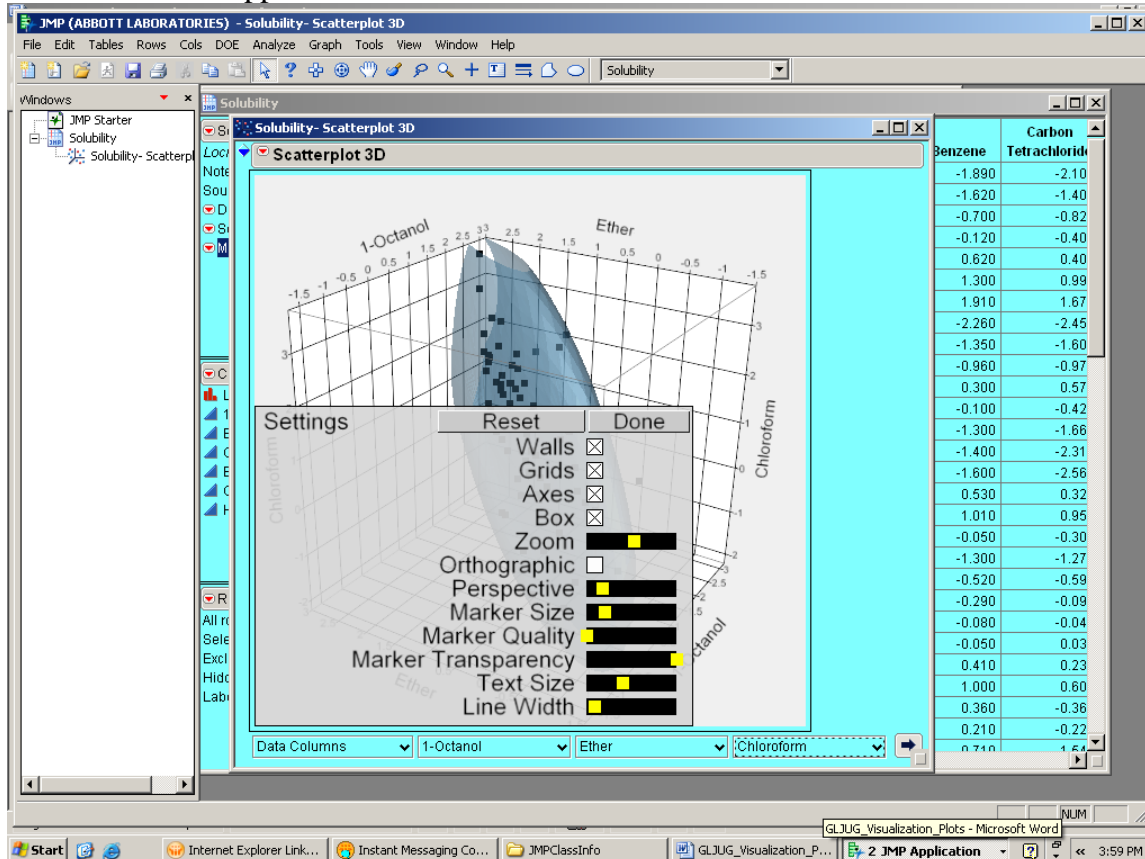
Initially, the first three variables are plotted, but others may be selected from the bottom axis controls. Normal Contour Plots can be drawn at a specified percentage of coverage and with a specified transparency (0 = transparent, 1 = opaque). Either a single shell can be selected (ungrouped option) or multiple contours based on the levels of a column (grouped option). Using the ungrouped option with 0.95 coverage and 0.5 transparency, it will appear as follows:



This plot is useful for identification of potential outliers for further investigation.

A 'Nonpar Density Contour' is also available which draws a 95 % kernel contour around the points for non-normally distributed data. It also has the grouped and ungrouped option, but also a control panel with three levels of contour, transparency, bandwidth, and resolution.

To change display options, right-click inside the plot and select 'Settings' to produce a popup menu which will appear as follows:



The settings can be reset to their original settings by selecting 'Reset'. 'Done' closes the dialog box. The walls, grids, axes, and box can be removed. Marker, graph, and text size can be enlarged or diminished using the appropriate sliders.

Spinning the plot can be accomplished in three ways:

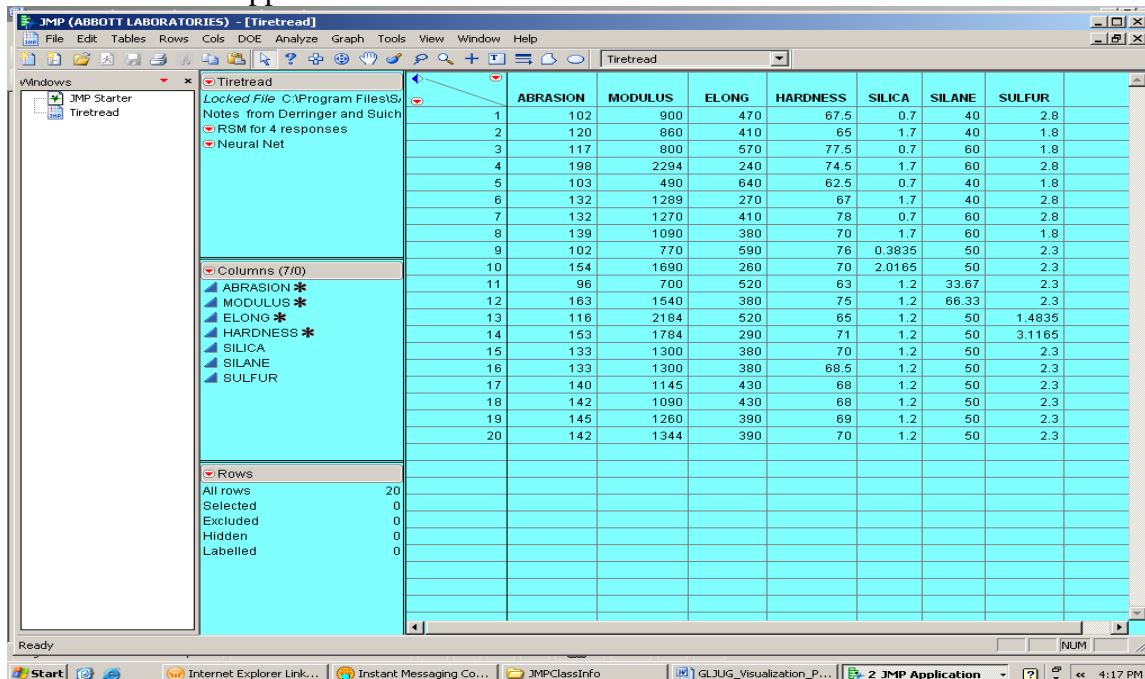
- Click and drag the mouse
- Keypad arrows
- Arrow keys

To spin, hold the mouse button or keys down and the spinning continues as long as the mouse button or keys are held down. To maintain a continuous spin, Shift-click. To stop the spinning, click inside the plot.

Surface Plots

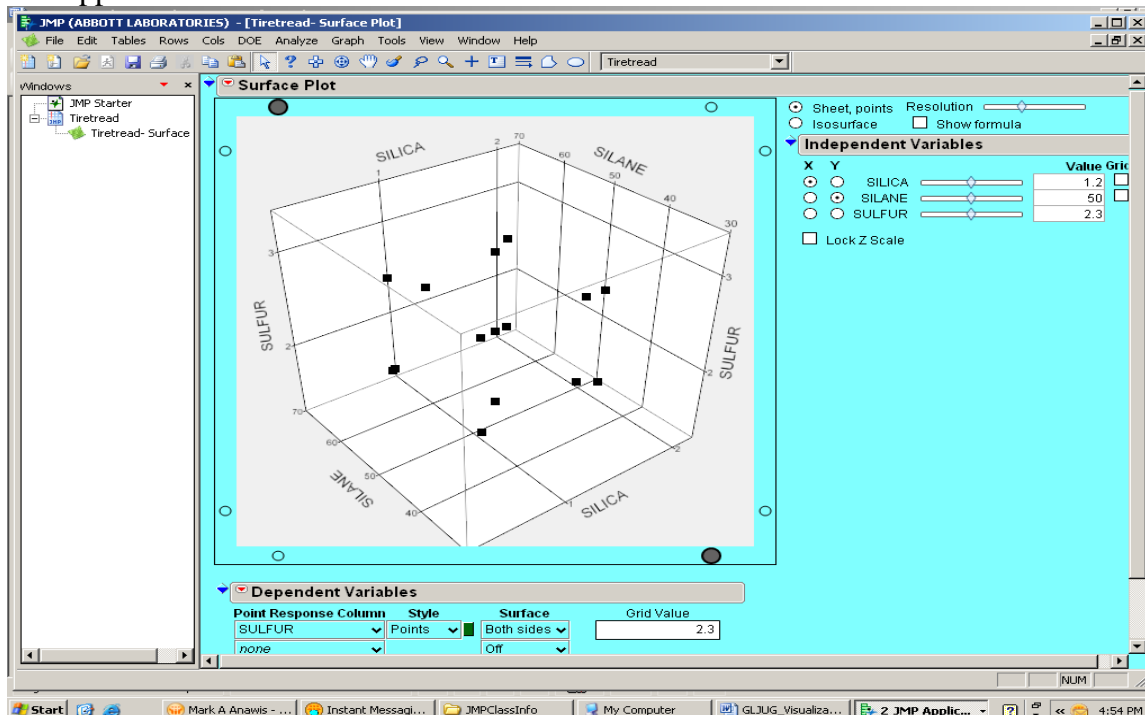
We will use the Tiretread.jmp^{2,3} file. Select 'File', 'Open', and 'Tiretread.jmp'.

The data file will appear as follows:

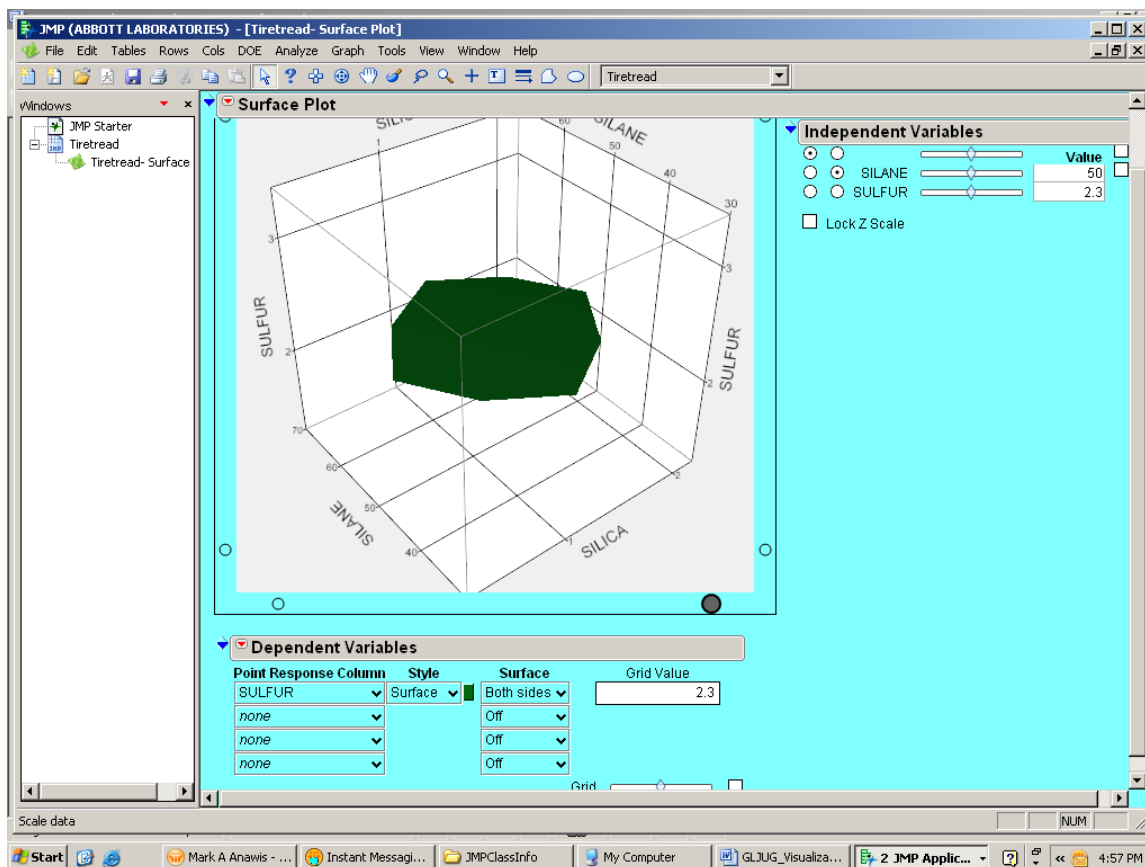


	ABRASION	MODULUS	ELONG	HARDNESS	SILICA	SILANE	SULFUR
1	102	900	470	67.5	0.7	40	2.8
2	120	860	410	65	1.7	40	1.8
3	117	800	570	77.5	0.7	60	1.8
4	198	2294	240	74.5	1.7	60	2.8
5	103	490	640	62.5	0.7	40	1.8
6	132	1289	270	67	1.7	40	2.8
7	132	1270	410	78	0.7	60	2.8
8	139	1090	380	70	1.7	60	1.8
9	102	770	590	76	0.3835	50	2.3
10	154	1690	260	70	2.0165	50	2.3
11	96	700	520	63	1.2	33.67	2.3
12	163	1540	380	75	1.2	66.33	2.3
13	116	2184	520	65	1.2	50	1.4835
14	153	1784	290	71	1.2	50	3.1165
15	133	1300	380	70	1.2	50	2.3
16	133	1300	380	68.5	1.2	50	2.3
17	140	1145	430	68	1.2	50	2.3
18	142	1090	430	68	1.2	50	2.3
19	145	1260	390	69	1.2	50	2.3
20	142	1344	390	70	1.2	50	2.3

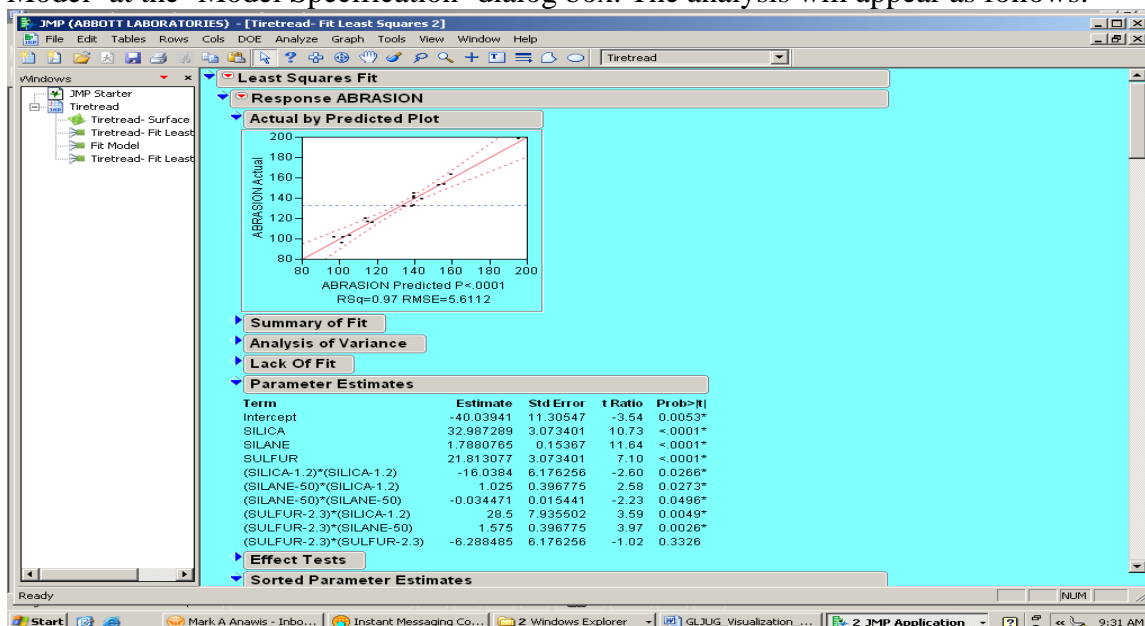
Select 'Graph', 'Surface Plot' and 'Silica', 'Silane', and 'Sulfur' as 'Columns'. Select 'OK'. It will appear as follows:



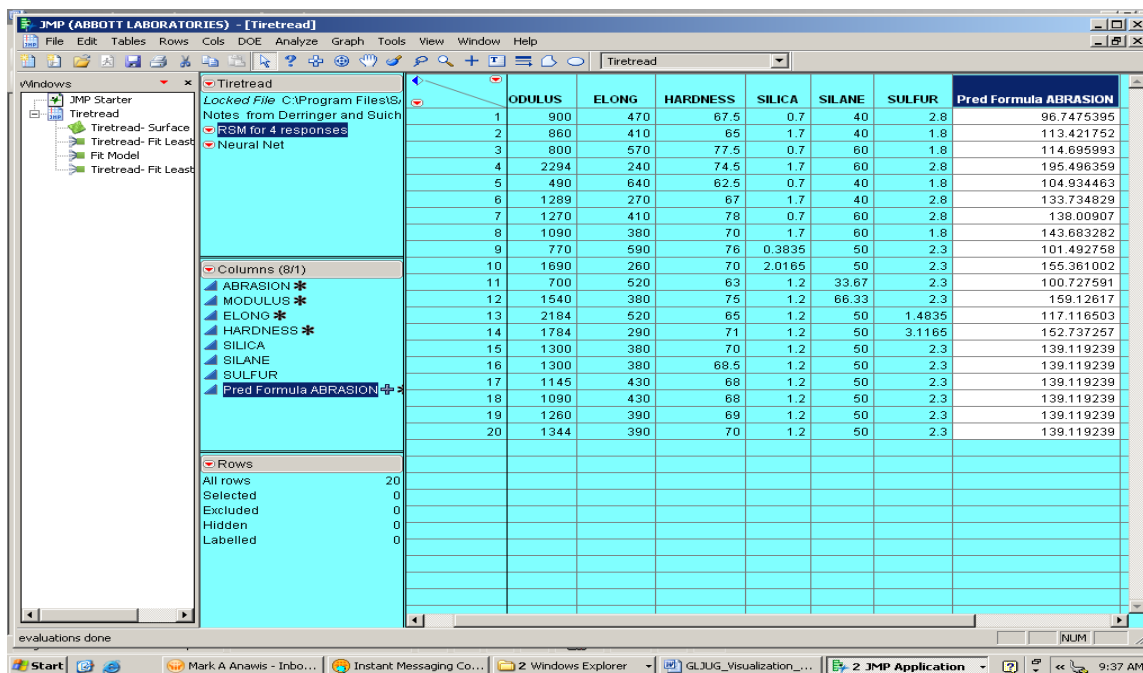
Notice that this looks a lot like the Scatterplot 3D at this point. Under 'Dependent Variables' and 'SULFUR', select 'Surface'. It will appear as follows:



However, a surface plot of a prediction formula with the data as points is more informative. To generate this, go to the data table and run the script for 'RSM for 4 responses'. Select 'Run Model' at the 'Model Specification' dialog box. The analysis will appear as follows:

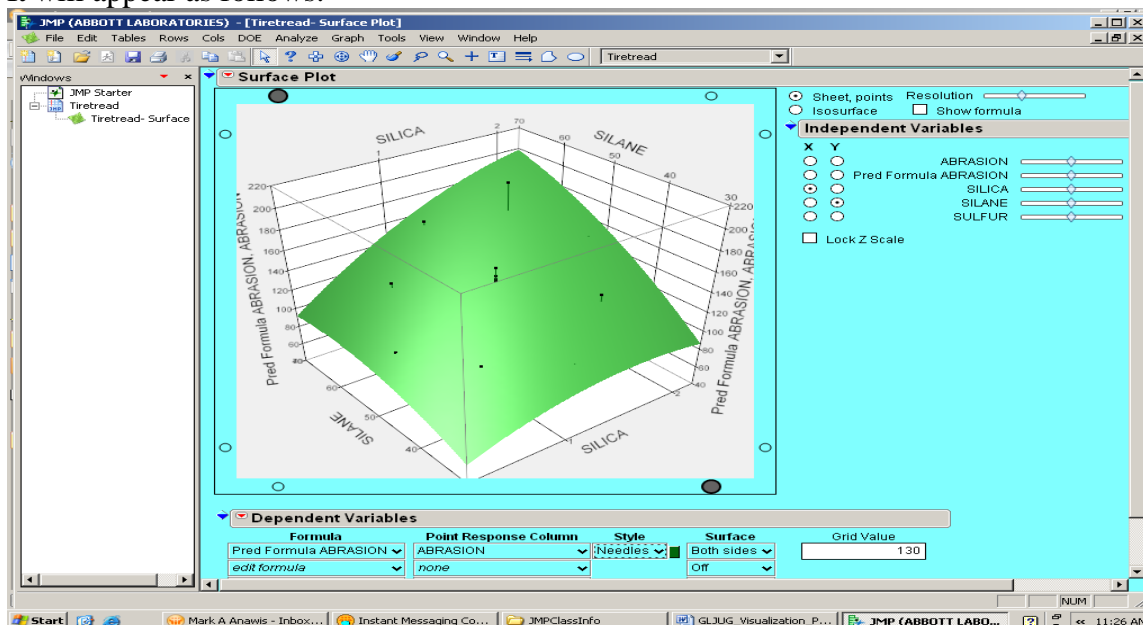


Although we have one model for each of four responses, we will create a surface plot for the first model – Abrasion. Select the red arrow next to 'Response ABRASION', 'Save Columns', and 'Prediction Formula'. There is now a new column called 'Pred Formula ABRASION' in the data table. It will appear as follows:



Select 'Graph', 'Surface Plot' and enter 'ABRASION' and 'Pred Formula ABRASION' and 'OK'. Under 'Dependent Variables', select 'Needles' for the 'Style'.

It will appear as follows:

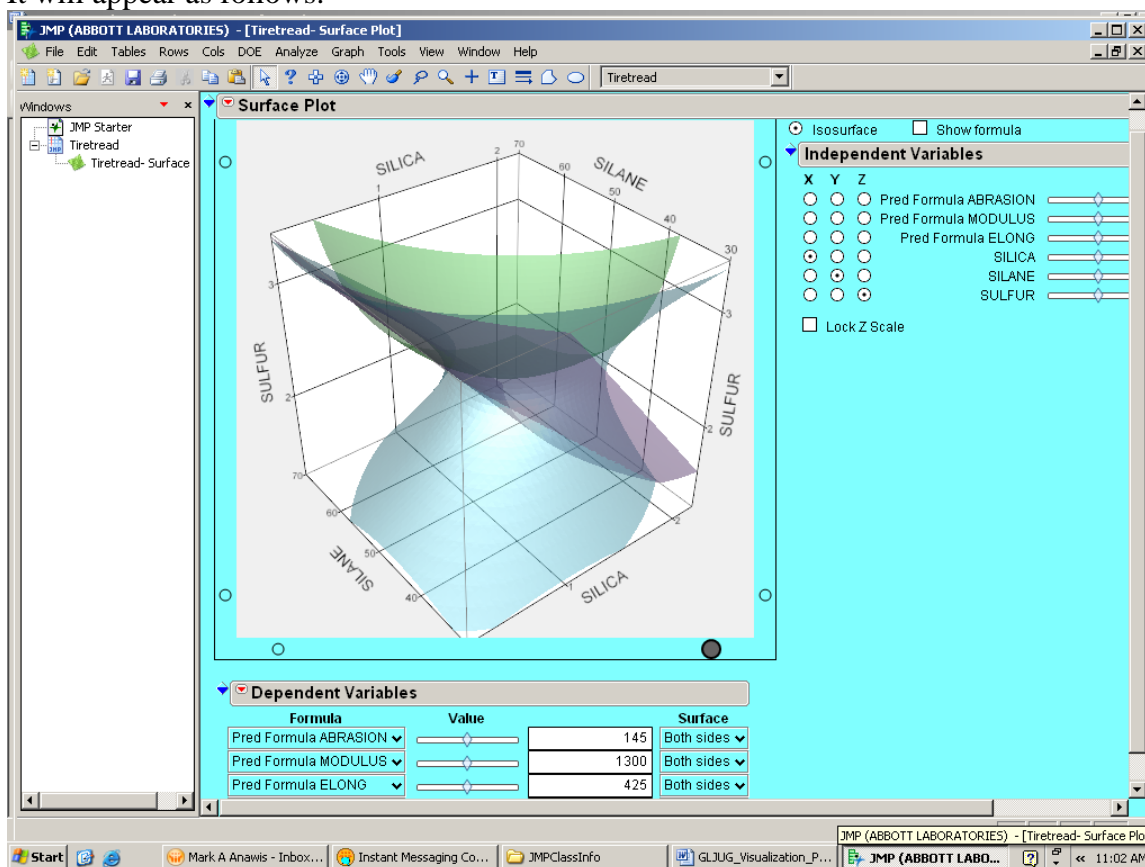


This plot provides a larger surface and a good comparison of the predicted values with the actual values. With the 'Needles' style, lines are drawn from the x-y plane to the points. If extrapolation outside the data points is not desired, right-click on the surface, select 'Sheet Properties' and check the 'Limit X and Y to Point Response Column'. Notice that although we have chosen only 'ABRASION' and 'Pred Formula ABRASION' columns, 'Silica' and 'Silane' are also plotted automatically since they appear in the prediction equation along with 'Sulfur'. To display the prediction equation, click the 'Show Formula' checkbox in the upper right hand area of the surface plot. Different variables may be selected by clicking on the desired radio button. Grids can also be selected for certain variables to aid in visualization. Sliders and edit boxes can be used to set the current value of a variable.

Isosurfaces

We can extend the concept of plotting one dependent variable against two independent variables to plotting one or more dependent variables against three independent variables. These are referred to as isosurfaces and require a formula with three independent variables for each isosurface desired. Let's create three isosurfaces. Since we already have a prediction formula for ABRASION, let's create two more. Re-run the script for 'RSM for 4 responses'. Select the red arrow next to 'Response MODULUS', 'Save Columns', and 'Prediction Formula'. There is now a new column called 'Pred Formula MODULUS' in the data table. Select the red arrow next to 'Response ELONG', 'Save Columns', and 'Prediction Formula'. There is now a new column called 'Pred Formula ELONG' in the data table. Select 'Graph', 'Surface Plot', and 'Pred Formula ABRASION', 'Pred Formula MODULUS', and 'Pred Formula ELONG' as 'Columns'. Select 'OK'. When the surface plot appears, select the 'Isosurface Plot' radio button. Select all three formulas under 'Dependent Variables' and 'Both sides' for 'Surface'.

It will appear as follows:



To optimize, the 'ABRASION' can be set at minimum and the 'ELONG' to maximum, and the 'MODULUS' slider moved back and forth to determine which values are inside the other limits. Note: this takes some visual-spatial abilities on the part of the user!

The plot can also be rotated by placing the cursor in the plot where a circular arrow appears, then click and drag to rotate. Double clicking on an axis allows you to change minimum, maximum, increment, and tick mark label format. This can also be changed by placing the cursor in the plot until the grabber hand appears. Moving along the axis, changes the minimum and maximum, but not the overall range. Moving out, zooms out, while moving in, zooms in. There are also eight control knobs for position and color of lights. Right clicking on a knob to turn on a light allows the user to highlight different parts of the plot.

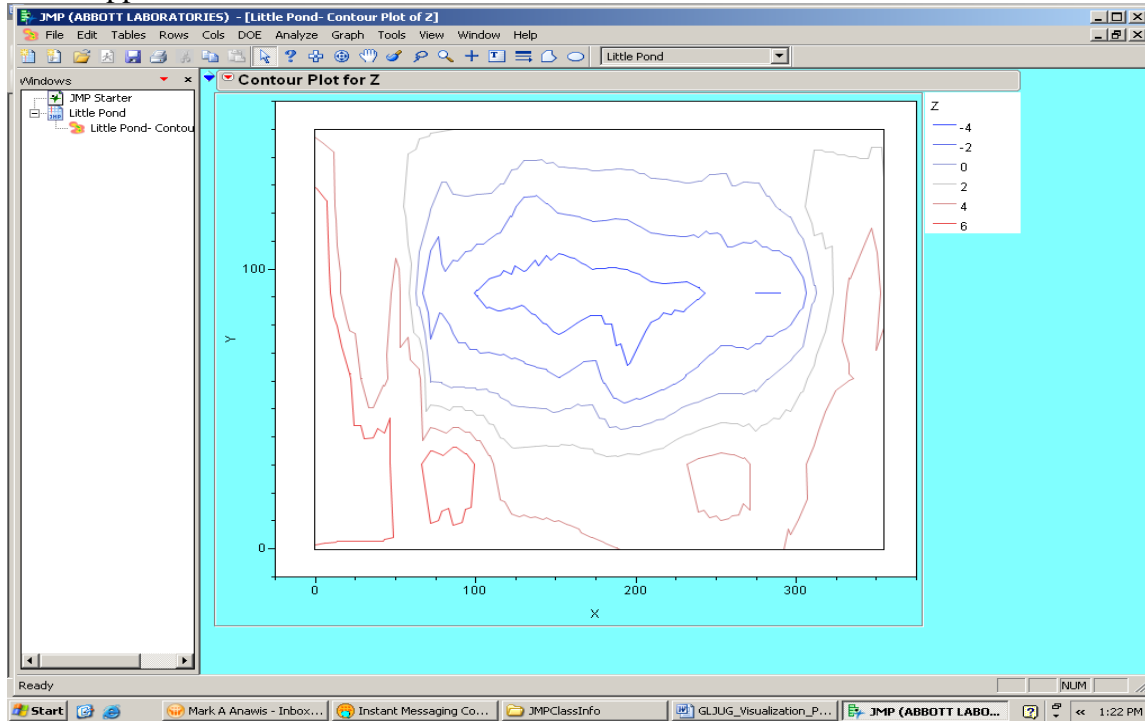
Contour Plots

We will use the 'Little Pond.jmp²' file. select 'File', 'Open', and 'Little Pond.jmp'. In this example X and Y are coordinates of a pond. Z is a response, depth. The data file will appear as follows:

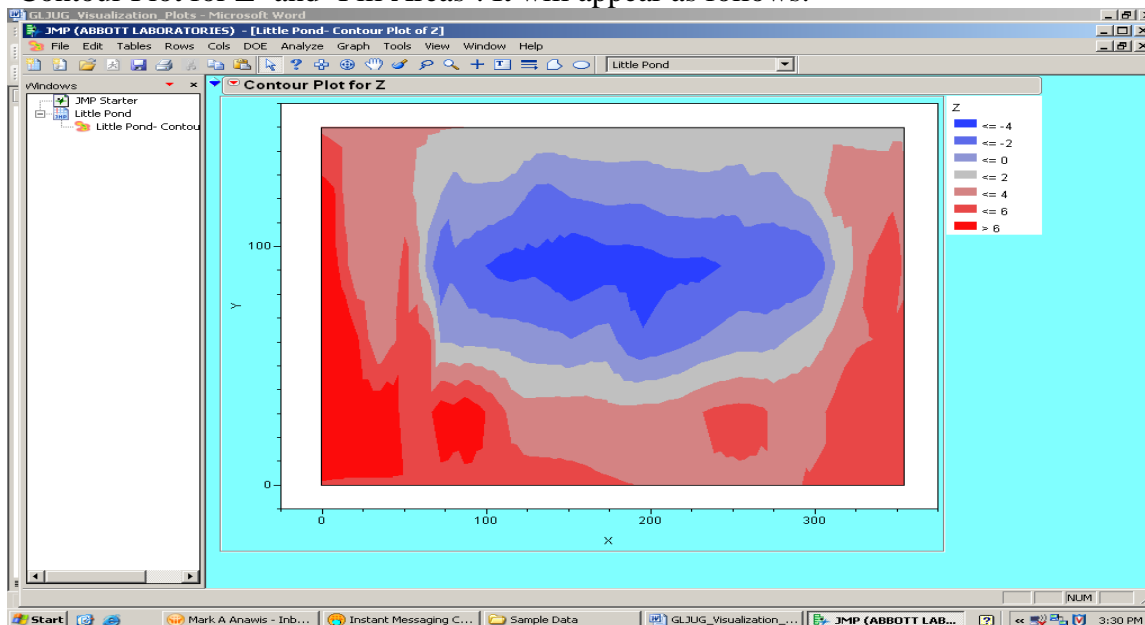
	X	Y	Z
1	0	0	5.90722656
2	0	30.609375	7.78417969
3	0	61.21875	9.61132813
4	0	91.828125	10
5	0	122.4375	6.80957031
6	0	150	3.72167969
7	7.24414063	0	5.85742188
8	7.24414063	30.609375	7.734375
9	7.24414063	61.21875	9.58007813
10	7.24414063	91.828125	6.84179688
11	7.24414063	122.4375	6.26464844
12	7.24414063	150	3.40478516
13	14.4882813	0	5.828125
14	14.4882813	30.609375	7.68457031
15	14.4882813	61.21875	8.203125
16	14.4882813	91.828125	4.58007813
17	14.4882813	122.4375	3.515625
18	14.4882813	150	3.08740234
19	21.734375	0	5.828125
20	21.734375	30.609375	7.63476563
21	21.734375	61.21875	6.21484375
22	21.734375	91.828125	2.22412109
23	21.734375	122.4375	2
24	21.734375	150	2.77050781
25	28.9765625	0	5.828125
26	28.9765625	30.609375	7.5859375
27	28.9765625	61.21875	3.99853516
28	28.9765625	91.828125	2
29	28.9765625	122.4375	2
30	28.9765625	150	2.453125
31	36.21875	0	5.828125

Select 'Graph', 'Contour Plot', and enter 'Z' as 'Y' column, and 'X' and 'Y' as X' column. Select 'OK'.

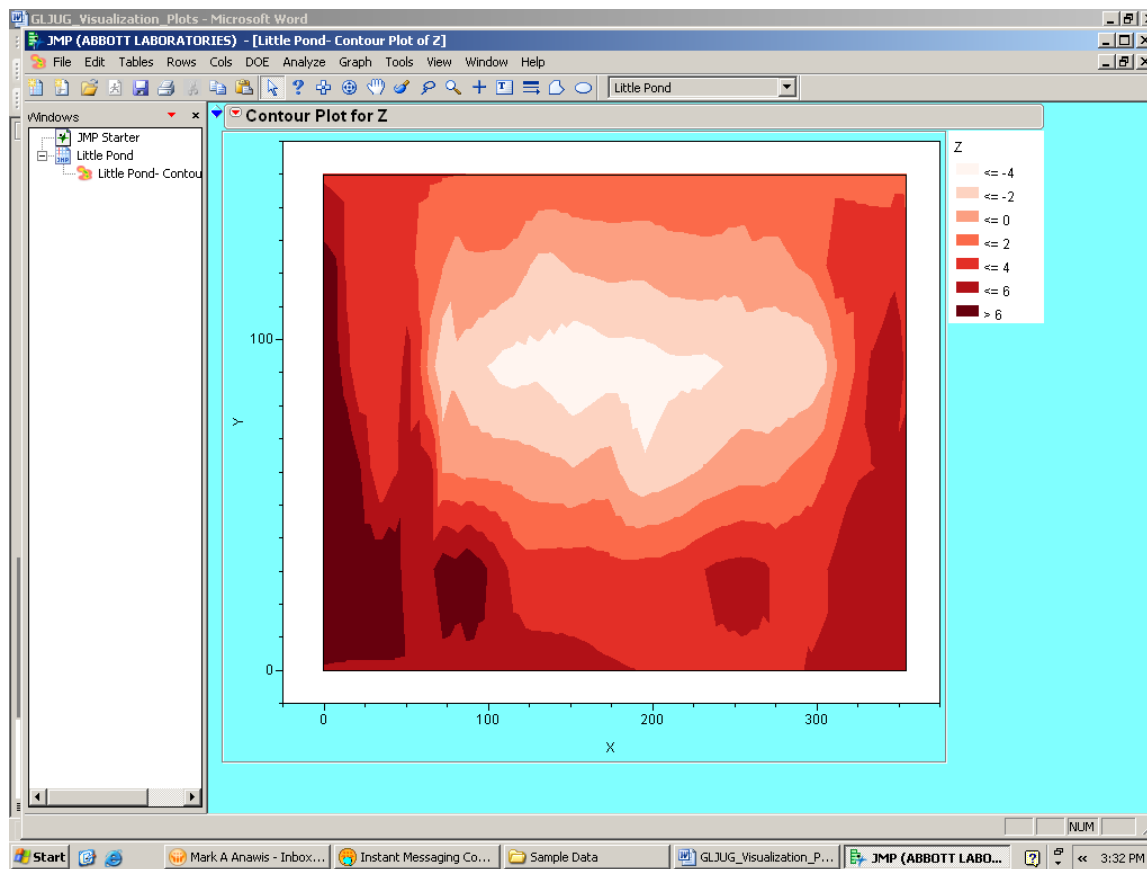
It will appear as follows:



Although the default contour levels are calculated to be equally spaced from the data, they can be changed by selecting red arrow next to 'Contour Plot for Z', 'Change Contours' and 'Specify Contours'. The # of contours as well as the minimum and maximum can be changed. The fill areas can also be filled in for a better visualization (in my opinion) by selecting red arrow next to 'Contour Plot for Z' and 'Fill Areas'. It will appear as follows:



However, a three color scheme for a continuous gradient is one too many. To change to a two color scheme, select red arrow next to 'Contour Plot for Z', 'Color Theme' and 'White to Red'. It will appear as follows:



We can see those combinations of X and Y values which contribute to large Z values (dark red) more easily. The color scheme can be easily reversed such that large Z values have light red by selecting red arrow next to 'Contour Plot for Z' and 'Reverse Colors'.

Conclusion

- The starting point for analysis should always be a graphical representation of the data for exploring, summarizing, and uncovering patterns.
- The **Overlay Plot** creates plots of one X variable and multiple Y variables shown on a single X axis or multiple X axes.
- The **Spinning Plot** displays a 3-D view of points for three variables at a time from the columns selected.
- The **Surface Plot** displays a 3-D view of surfaces for two variables at a time from the columns selected.
- The **Contour Plot** shows contours of a response in a rectangular coordinate system.

References

1. SAS® Institute, Inc., 2009, JMP® v.8 Statistics and Graphics Guide
2. SAS® Institute, Inc., 2009, JMP® v.8 sample data sets.
3. Derringer and Suich ,1980, "Simultaneous Optimization of Several Response Variables" Journal of Quality Technology, 214-219.

"Copyright 2009, SAS Institute Inc., Cary, NC, USA. All Rights Reserved. Reproduced with permission of SAS Institute Inc., Cary, NC"

Contact Information

Your comments and questions are valued and encouraged. Contact the author at:

Mark Anawis, M.A., ASQ CSSBB
Abbott
100 Abbott Park Road, Building AP8B
Abbott Park, IL, 60064
Phone: 847-937-4347
Fax: 847-938-2219
E-mail: mark.anawis@abbott.com

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration. Other brand and product names are trademarks of their respective companies.