

Introduction to JMP® under Windows: Hands-On Workshop Outline

Jenny Aquino Kendall, Data Partners, Earlham IA
Al M. Best, Data Partners, Richmond VA
Charles E. Shipp, Consultant, Los Angeles CA

Abstract

JMP under Windows is highly graphical, inviting us to interact with our data. Along with graphical tools for data visualization, JMP provides us with sophisticated statistical analyses. This combination of visualization and statistical analysis helps us to make good decisions grounded by a sound knowledge of our data. Because JMP is highly interactive, the best way to learn JMP is to use JMP, so in this hands-on workshop, we will learn how to get around in JMP by

- working with the JMP data table
- investigating variables using the Distribution of Y platform and the Correlation of Y's platform
- analyzing relationships such as ANOVA and simple linear regression with Fit Y by X
- learning about additional JMP features.

Welcome to JMP

Since the best way to learn JMP is to use JMP, we'll start by showing you an example based on actual analyses that are typical in the semiconductor manufacturing industry. Then, we'll give you a chance to use JMP and work with that same data.

Working with the JMP data table

The data consist of a variety of measurements obtained on a wafer which will eventually be broken down into semiconductor chips. Four sample measurements have been taken during all three shifts over three days. The temperature may vary within these samples. A resistivity test is made, an intervention occurs, and then at seven different sites on each wafer a measurement is made. Figure 1 shows the data as it appears in the JMP data table under Windows.

	Date	Shift	Temp	test1	site1	site2	site3	site4	site5	site6	site7	test2
1	3/1/96	1	500	0.000906	136.3	133.7	110.7	98.7	127.7	159.3	117.0	0.001986
2	3/1/96	1	510	0.000990	113.6	113.4	96.3	89.1	100.7	142.5	120.9	0.001298
3	3/1/96	1	520	0.000291	111.8	110.9	102.8	98.0	100.0	142.9	131.9	0.001410
4	3/1/96	1	580	0.000949	90.1	91.9	90.2	87.2	86.6	128.4	111.2	0.000695
5	3/1/96	2	500	0.000692	133.1	133.9	109.9	98.0	124.3	156.9	113.2	0.001914
6	3/1/96	2	510	0.000351	118.0	117.7	101.7	92.3	105.0	141.8	121.7	0.001618
7	3/1/96	2	520	0.000685	119.0	119.1	108.6	103.0	109.7	150.1	137.7	0.001781
8	3/1/96	2	580	0.000675	98.6	100.3	94.5	91.9	92.1	134.2	120.3	0.001157
9	3/1/96	3	500	0.000645	131.5	131.9	108.5	96.5	123.5	159.6	112.6	0.002084
10	3/1/96	3	510	0.000654	110.8	110.7	95.3	86.0	99.1	135.8	121.1	0.001361
11	3/1/96	3	520	0.000947	111.5	110.7	100.7	95.3	102.2	142.0	129.3	0.001506
12	3/1/96	3	580	0.000977	89.1	90.5	87.0	83.8	83.2	125.3	111.9	0.000854
13	3/2/96	1	500	0.000972	135.8	135.5	110.9	100.0	127.7	158.9	115.1	0.001811
14	3/2/96	1	510	0.000287	118.4	116.9	101.5	90.1	105.1	142.6	125.7	0.001558
15	3/2/96	1	520	0.000168	118.8	120.0	110.4	102.5	110.1	149.0	136.2	0.001907
16	3/2/96	1	580	0.000731	98.0	100.2	95.5	94.5	93.9	136.9	120.7	0.001180

Figure 1. The `tesdata.jmp` data table.

A JMP data table is organized like a spreadsheet. Columns correspond to variables and rows correspond to observations. There are 36 rows in the data table and 12 columns. We can manipulate and rearrange the data in the JMP data table with Tables, Rows, and Columns menus.

As we investigate and analyze these data, we will learn about key JMP concepts, including

- Data type (are the data character or numeric?)
- Modeling type (Nominal, Ordinal, or Continuous?)
- Column role (X or Y?)

Analyzing the Data

Rather than discuss in words how data type, modeling type and column role work, we will demonstrate by putting our columns into action by using the analysis platforms available to us in the **Analyze** menu.

Using the Analyze menu commands

One visual analysis that can help us is a scatterplot matrix, which we can easily obtain as an option using the **Correlation of Y's** platform in the **Analyze** Menu. The platform name gives us a hint as to what column role will be needed.

Select **Analyze >> Correlation of Y's** and then
 Scatterplot Matrix See Figure 2.

The engineer in charge of this manufacturing process would like to know if there are differences in wafer measurements for different days and/or different shifts. We can use Colors and/or Markers to distinguish different

groups. We'll use the **Colors** and **Markers** options in the **Rows** menu to assign a different color to each **Date** and a different marker to each **shift**. For these data, while we don't see any particular patterns associated with **Date** or **shift**, we do notice a very strong correlation between site1 and site2 that may bear further investigation.

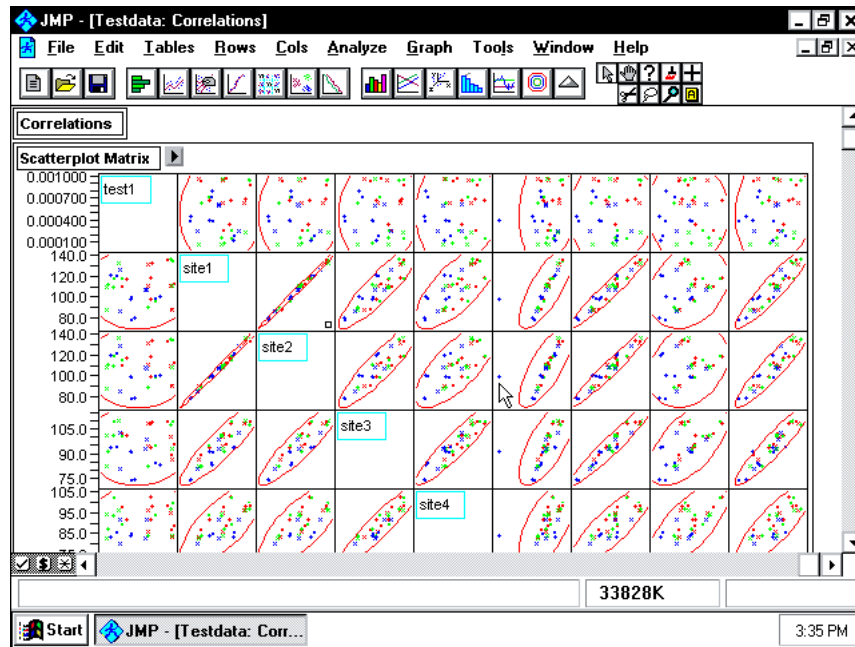


Figure 2. A Scatterplot Matrix from the Correlation of Y's Platform.

The columns labeled **site1**, ... **site7** represent different measurements taken on the same wafer. If the engineer wants to examine differences among the sites, then we will need to rearrange this table so the site columns become rows. We can easily do this using the **Stack**

Columns option in the **Tables** menu. We'll call this new JMP data table **teststk.jmp**. Now we use the same data, rearranged (Figure 3). We can use either table, depending on the analysis we want to do although in this workshop all remaining analyses use this data table.

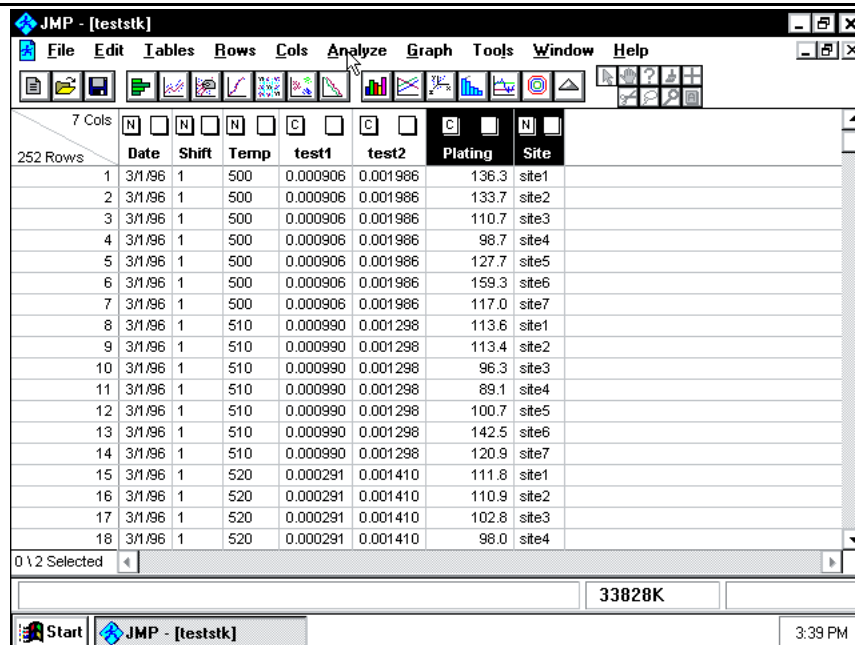


Figure 3. After using the **Stack Columns** Command.

Analyze Menu >> Distribution of Y

Since our engineer is interested in examining these data in detail, we'll continue our investigation by examining the distribution of values for the plating and test measurements. This time, we will assign the **Y** column role to all of the columns in our JMP data table, **teststk.jmp**. Once we have assigned roles (assigning roles prior to using an analysis platform is optional), all we

need to do is click on the **Distribution of Y** tool button, or click on the **Analyze** menu and select **Distribution of Y**.

JMP produces a report for each column (Figure 4). The type of report will be different, depending on whether we have identified our column as Nominal or Continuous. (An Ordinal column will have the same type of report here as the Nominal column).

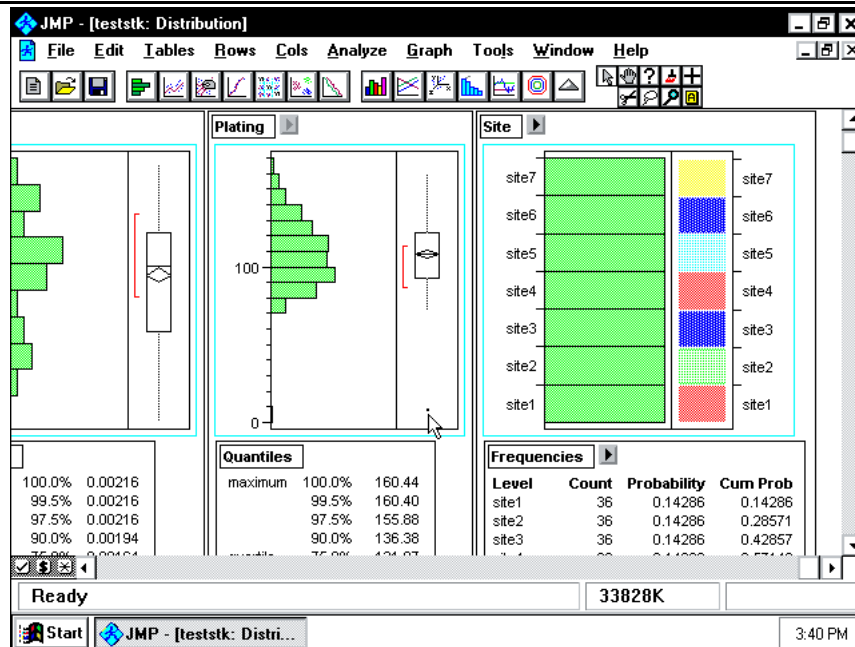


Figure 4. The Distribution of Y Window.

What first appears in each report depends upon the preferences that are set for each platform. What we see here are the default preferences, which we can change by using the **Preferences** option in the **File** Menu.

For Continuous variable columns, we see a histogram and an outlier box plot. For the Nominal variable columns, we see a bar chart, a mosaic chart, and a frequency count table. If an outlier box plot, mosaic chart, or any other plot or report is unfamiliar, then we can ask JMP to tell us what the plot means by using JMP's context-sensitive help tool.

The outlier box plot will point out to us if there are any extreme values in that column. Highlighting an extreme value in the outlier box plot will highlight the row that corresponds to that point in the JMP data table. This is because the JMP data table and the Analysis reports are linked.

We use the histogram to help us describe the shape of the distribution of values. By using the hand tool, we can

change the size and endpoints of the histogram bars to gain a better idea of how our values are distributed.

Not only are the rows in the JMP data table linked to the reports, but the data in the reports are also linked. Highlighting data in one report will highlight the same data in the other reports. This gives us a preliminary look at relationships among column variables.

Statistical Reports

The Moments report provides us with common statistics for continuous data, such as the mean and standard deviation. Additional statistics may be had by clicking on the arrow to the right of the report name. Any time we see a small arrow to the right of a report, JMP has another layer of information to offer. The Frequencies report provides us with statistics common for nominal or ordinal data.

We can modify our report displays using the check mark () and asterisk options () (located in the bottom left corner of the JMP window). We can also save statistics to our JMP data table using the options.

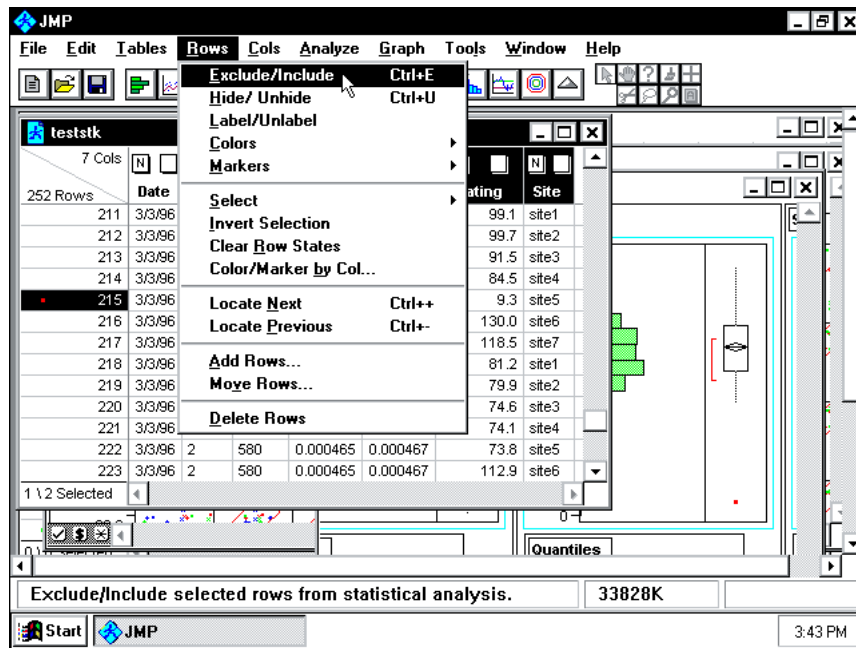


Figure 5. Excluding an Outlier from Future Analysis.

Analyze >> Fit Y by X

While we can gain a preliminary idea of relationships in data using the tools and options in the Distribution of Y platform, we can obtain more in-depth analyses on the nature of the relationships between variables using Fit Y by X. We can obtain four types of analyses with this platform, depending on whether the modeling type is Nominal or Continuous, with column role X or Y. We'll investigate and discuss ANOVA and regression JMP analyses and options using our **teststk.jmp** data table.

Each column's modeling type is shown in the Fit Y by X dialog. Analyses types are also displayed (Figure 6).

Modeling Type	Column Role	JMP analysis
Continuous	Y	ANOVA
Nominal	X	
Continuous	Y	simple linear regression
Continuous	X	
Nominal	Y	contingency table analysis
Nominal	X	
Nominal	Y	logistic regression
Continuous	X	

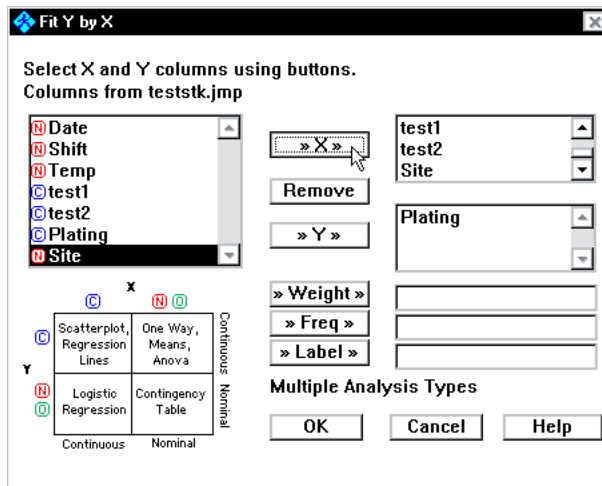


Figure 6. Selecting One Y Variable and All the Rest as X Variables.

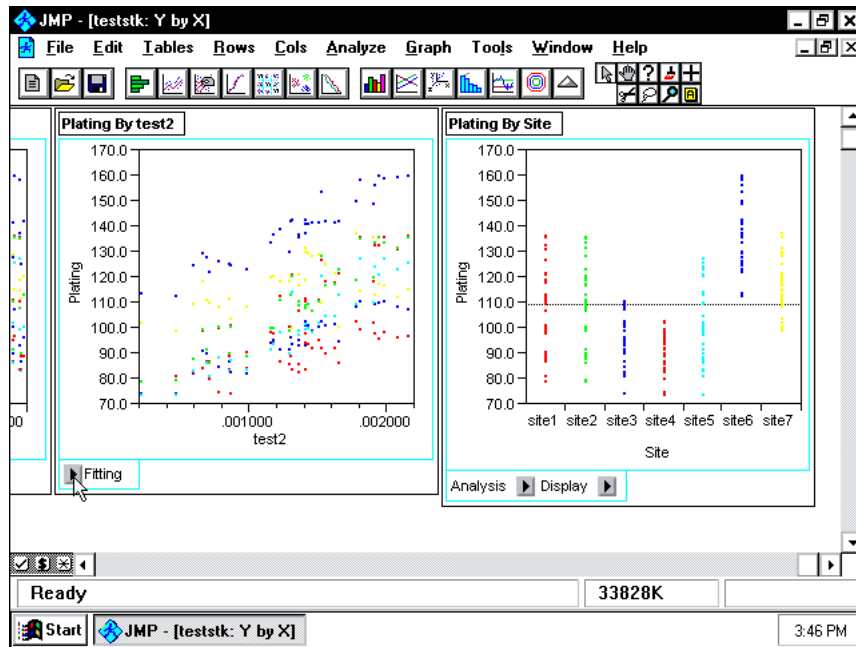


Figure 7. A Portion of the Fit Y by X Window.

Additional JMP Features

There are numerous report and graphical options available specific to each type of analysis.

More complex analyses are also available through the **Analyze** platforms, including complex linear modeling through Fit Model, Nonlinear, and Survival Data analyses. Additionally, there is a **Graph** menu with many graphic platforms including contour plots, Shewhart Control Charts

among others. Extensive data summarization can be handled via the Group/Summary option in the **Tables** menu. Screening, response surface, and other types of experiments are easily designed with the Design Experiment option in the **Tables** menu.

JMP has a complete, context sensitive help system (Figure 8).

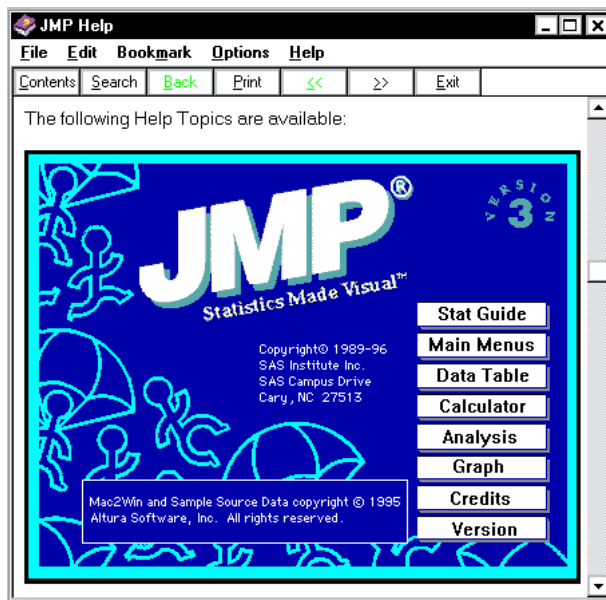


Figure 8. The Top Level of the JMP Help System.

Additionally, JMP under Windows also makes it easy to interact with SAS® software, as JMP can read SAS transport files into a JMP data table, as well as save data in a JMP data table into a SAS transport file. JMP can also import text files in addition to SAS Transport files. In addition, JMP provides a journaling facility for saving JMP reports in rich text format which can then be imported into a word processing package.

Acknowledgments

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

A copy of this paper and related resources are at:
<http://views.vcu.edu/~albest/sugi97.html>

The Institute maintains a related site at:
<http://www.sas.com/jmp/>

Jenny Aquino Kendall may be reached through the internet at: zenomind@earthlink.net or by mail at: Data Partners, 1275 Hogback Bridge Road, Earlham IA 50072. Phone: 515-834-2462. Data Partners is a private company primarily involved with JMP related consulting, training, and publishing.

Al Best may be reached through eMail at: Best@richmond.infi.net or on the internet at: <http://views.vcu.edu/~albest> or by mail at: Data Partners, 12467 Ashland Vineyards LN, Ashland VA 23005. Phone: 804-730-2700.

Charlie Shipp may be reached through eMail at: Macfamily@AOL.com or by mail at: 1650 Barrywood, San Pedro CA, 90731-1254. Phone: 213-847-7266.

© Copyright 1997, Data Partners

10 Jan 97