JMP® Coders 101 Insights

Charles Edwin Shipp, Consider Consulting, San Pedro, California
Kirk Paul Lafler, Software Intelligence Corporation, Spring Valley, California

Abstract
JMP Version 10 is a powerful data discovery tool that is used interactively on your personal computer. We illustrate finding cause and effect in Veterans Hospital cancer patient data and show the code JMP creates as we interact. We demonstrate and discuss briefly the derivation of the density ellipse and its value to show data position, extent, and correlation in cause and effect. We conclude with leads to JMP learning and user exchange interaction.

Introduction
SAS Institute recently released Version 10 of JMP, and with it, additional choices: in addition to JMP Version 10, there are three additional packages: JMP Genomics, JMP Clinical, and JMP Pro (for data mining and advanced predictive modeling). Most JMP users will settle on JMP Version 10, but specialists and institutions with needs in the other areas will adapt quickly to the more advanced packages. Visit www.JMP.com to learn the possibilities.

In our brief time together, we give an instructive overview and introduction to JMP script codes and how easy they are to use. Of course, as with SAS coding, more advanced customization is also possible.

Creating Scripts
There are three ways to create JMP scripts—the most common being interactive work with JMP, saving results.

- Write from scratch in the scripting language
- Create with interactive data discovery
- Modify an existing template (data/script)

Here is how to think about the windows used as JMP is running: the input in rows and columns are just like SAS tables and can be imported from flat files, excel files, SAS files and other sources. The controlling and processing programs are called JMP scripts. And the output from JMP processing can be held in a JMP journal.

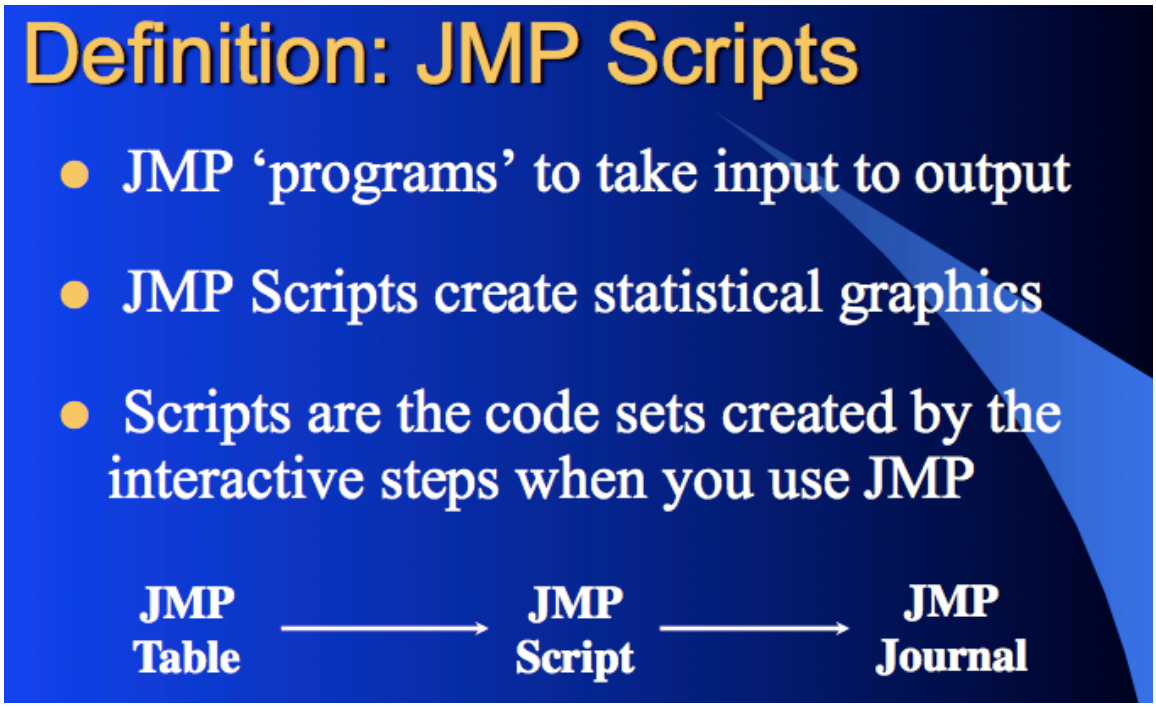

Figure 1. JMP Scripts allow you to program

New terms are used for JMP input, programs, and output, shown next in Figure 2.

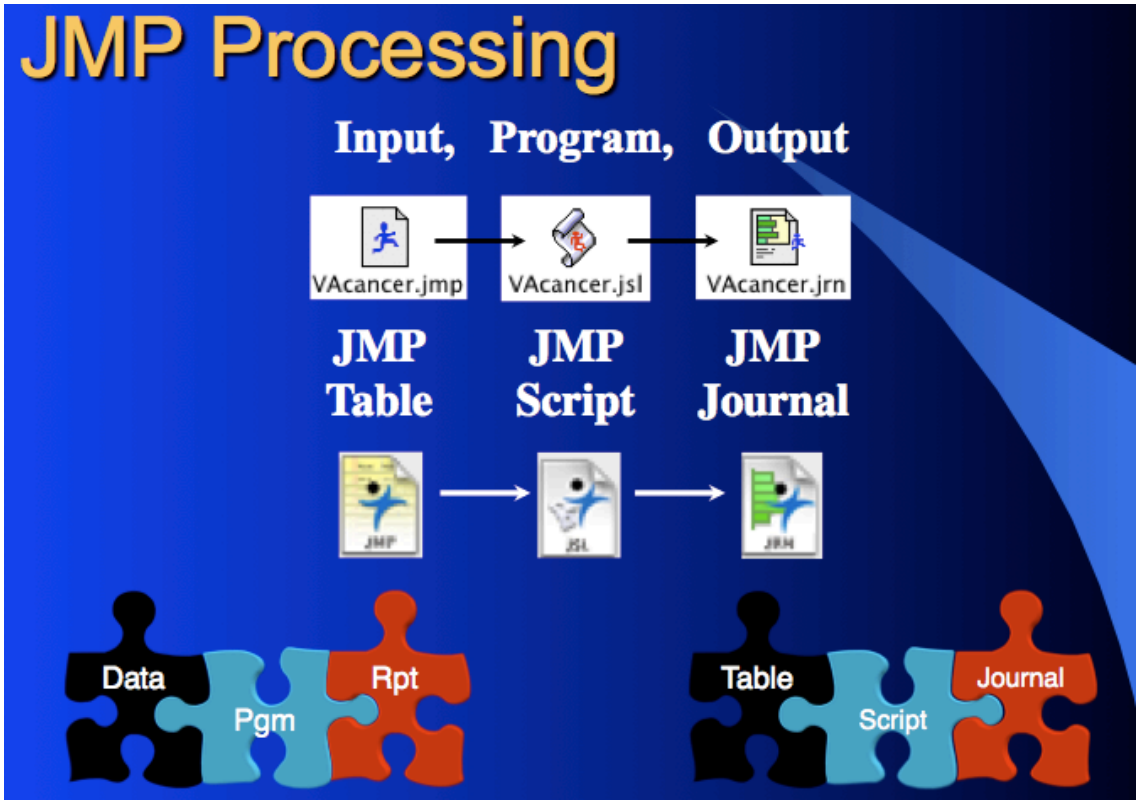

Figure 2. JMP input, program processing, and output

Options in JMP Processing

JMP processing uses three windows: Table window, Script window, and Result/Output window. These windows can be changed and then, when satisfied with results, saved to permanent files for distribution and/or reuse. In this regard, a Template, consisting of a table and script, can be held constant in script, and input data changed for many cases; or the data table can be held constant and the script modified over time. Of course, if variable names are changed then they need to be changed simultaneously to maintain correspondence between the JMP table and the JMP script..

A very wide selection of JMP categories

JMP scripting extends into general data reduction, data mining, genomics, clinical exploration and verification, data preparation, visualization, quality & reliability, modeling and design of experiment. What you learn in one area will help you in the other areas, such as linear regression, multiple regression, mixed models, nonlinear modeling, time series, control charts, bivariate analysis and analysis of variance. You will find well-constructed example scripts organized by discipline categories also. These programs are stored within the JMP software and run on demand as JMP runs on your personal computer at home or at work. You can also create a library of your own scripts and build enterprise data collections.

A first example code-set using density ellipse to illustrate

The JMP Script Language is concise yet powerful. Over 80 JMP platforms are like SAS procs that can be called and applied. Our first example program performs correlation display and table printout using density ellipses for the continuous variables mentioned in the argument.

JMP data files have the extension '.jsl' such as EXAMPLE.JSL and scripts plus references can be stored in the data files also.

Results of the first example

The results in figure 4 show the results of the script processing. It shows the correlation among all continuous variables. Look for a cell showing high correlation, and then look horizontally and vertically to find the two variables involved.

A sample JMP script and resulting output are shown in Figures 3 and 4.

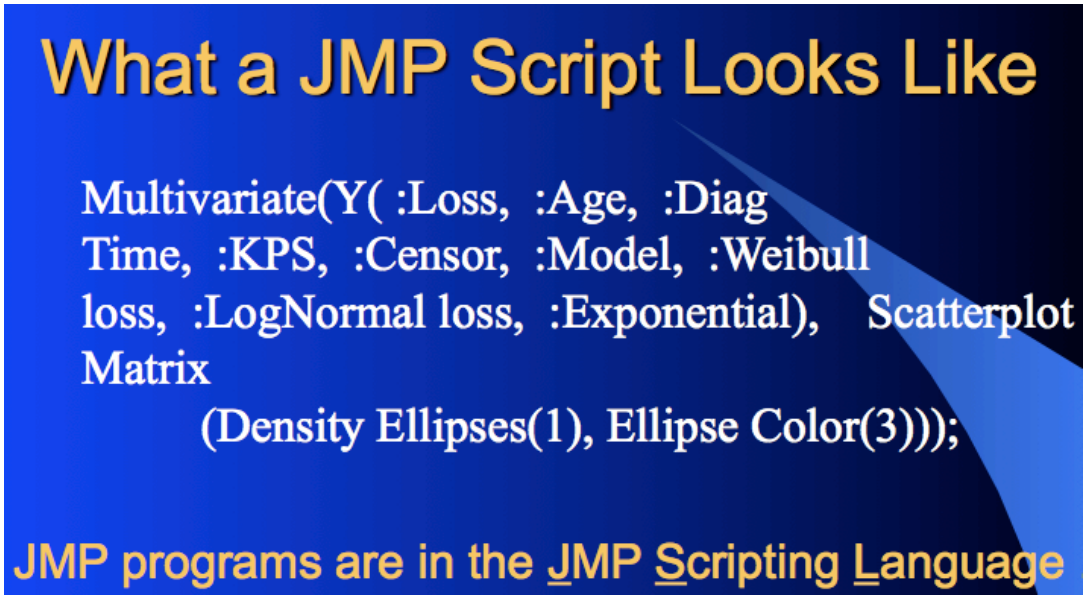

Figure 3. Sample JMP script using the density ellipse for the real variables in a JMP table

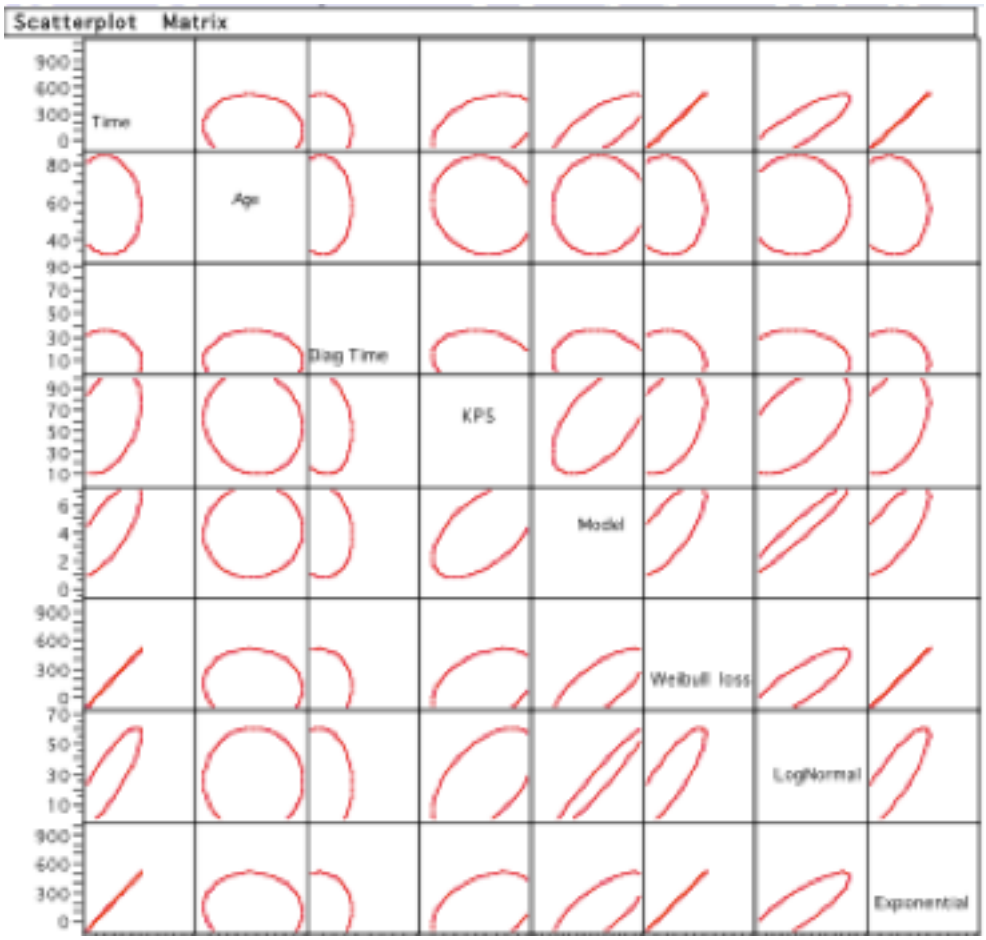

Figure 4. Sample output in a JMP journal, the results of running the JMP script

Two more similar scripts are now illustrated for density ellipse analysis, shown in Figure 5.

Two More Scripts

```
Bivariate(Y( :Loss), X( :KPS), Show Points(0), Fit Where( :Cell Type ==
"Adeno", Density Ellipse(0.5)), Fit Where( :Cell Type == "Large", Density
Ellipse(0.5)), Fit Where( :Cell Type == "Small", Density Ellipse(0.5)), Fit
Where( :Cell Type == "Squamous", Density Ellipse(0.5)));

Bivariate(Y( :Loss), X( :Age), Show Points(0), Fit Where( :Cell Type ==
"Adeno", Density Ellipse(0.5)), Fit Where( :Cell Type == "Large", Density
Ellipse(0.5)), Fit Where( :Cell Type == "Small", Density Ellipse(0.5)), Fit
Where( :Cell Type == "Squamous", Density Ellipse(0.5)));
```

NOTE: KPS measures the ability of cancer patients to perform ordinary tasks.
The Karnofsky Performance Scores range from 0 to 100.

Figure 5. Observe the conciseness these two JMP scripts

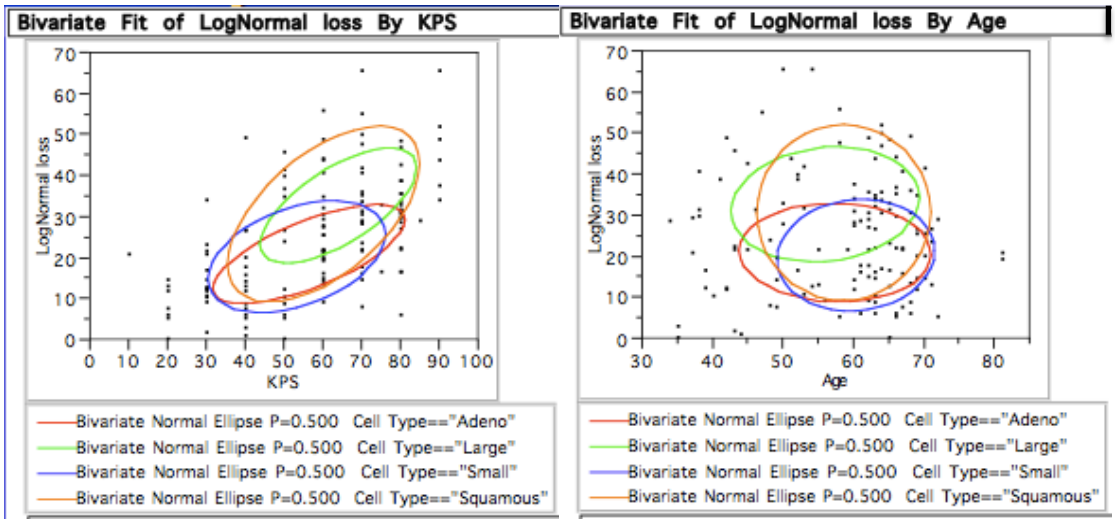

Figure 6. The results of the two scripts, showing KPS effects; but the independence from age.

These two scripts generate statistical graphics to show cause and effect as Y = f (X), where X is age or KPS, and Y is cell loss due to cancer. The two scripts are the same except for the definition of X. The method used is one of the best examples to show correlation in a graph, but not the only one you can use within JMP.

In this example, as KPS increases, say from 30 to 80, the measures for cell loss also increase, say from 10 to 50, hence there is positive correlation. For age (right graph) we see no correlation.

Figure 7 illustrates the concept of enclosing data points with a more visible density ellipse.

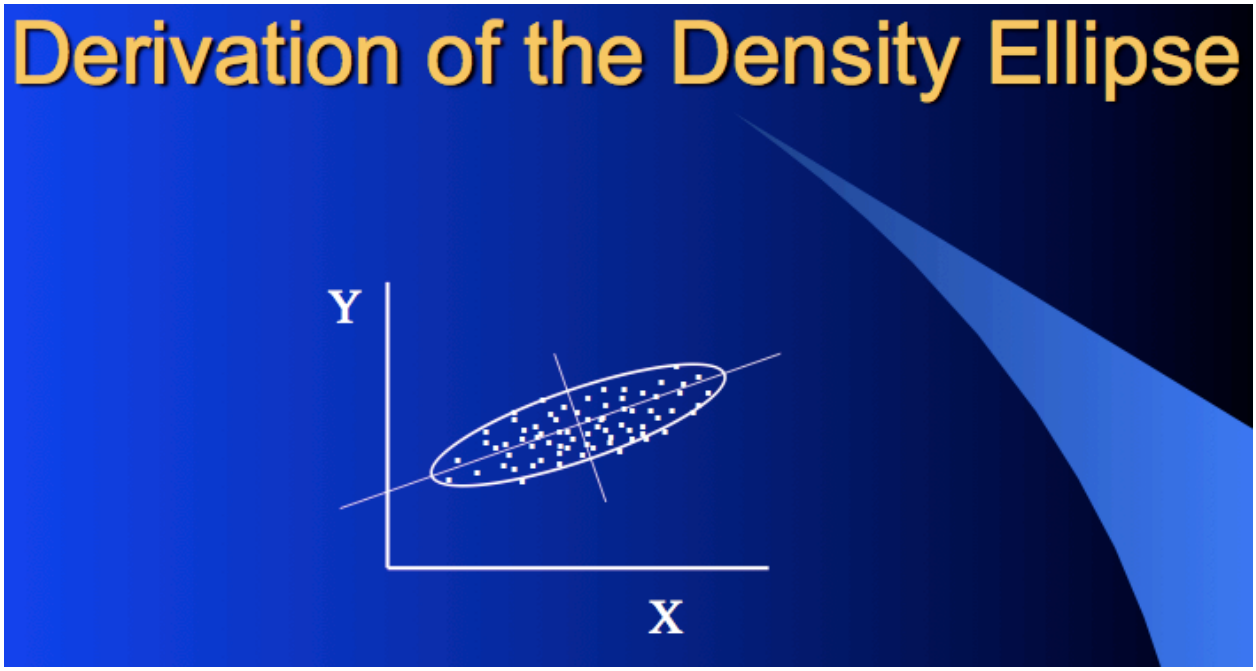

Figure 7. The density ellipse can replace the cloud of data points

Understanding the basis for density ellipse visualization

Just a brief word in a nutshell about the density ellipse calculation. First, the center of mass is calculated; then the principle axis and a line perpendicular through the center of mass. Then all the points can be regressed onto these two lines and standard deviation values measured from the center of mass so that an ellipse can be determined. Actually, to make it more robust, points on the ellipse are calculated around the center of mass, moving along the ellipse curve.

There is a multiplication factor involved, ALPHA, that determines the percentage of points to be statistically included. Common values are 0.50, 0.90, and 0.95 — they are concentric ellipses.

Summary of possible reusable script usages

- Copy/Modify/Reuse the JMP code
- Same script with current/different data
- Modify an existing template (data/script)

Once a snapshot is taken of the script and resulting journal results, the process window of the script and the output window can be saved into the data table file, and a single file can be sent with instructions to document results and findings.

Conclusion

JMP is the correct tool for discovering your data by exploring with statistical graphics and tables. It is also a great learning tool for school, government, and private industry. Help and tutorials are available, and the '?' tool can be put over any element of output or program controls for a description of statistics and usage.

There are many statistical platforms within JMP to use with robust and trusted methods. The intuitive interface is easy to learn and standardized throughout. We have given an introduction to understanding JMP scripting. Be involved in user exchange and JMP online and annual activities, both internationally and locally.

All you need to know (and learn) is in the software package as it runs; you can also learn online with http://jmp.com/academic/learning_library.shtml and other resources within www.JMP.com

References

Lafler, Kirk Paul and Charles Edwin Shipp (2009), *"Collaborate with SAS Users Around the World with sasCommunity.org," – Awarded "Best Contributed Paper",* Proceedings of the 2009 SAS Global Forum (SGF) Conference.

Hinrichs and Boiler, JMP Essentials – An Illustrated Step-by-Step Guide for New Users, SAS Press, 2012

Lehman, Ann, JMP for Basic Univariarte and Multivariate Statistics: A Step-by-Step Guide, SAS Press, 2005

Murphrey, Wendy, and Lukas, Rosemary, Jump into JMP Scripting, SAS Press, 2009.

Olsen, Chris, Teaching Elementary Statistics with JMP, SAS Press, 2011.

Ramirez, Jose, Analyzing and Interpreting Continuous Data Using JMP: A Step-by-Step Guide, SAS Press. 2009.

Sall, John, JMP Start Statistics: A Guide to Statistics and Data Analysis Using JMP, Fifth Edition, SAS Press, 2012.

Shipp, Charles Edwin and Kirk Paul Lafler (1995), *"Training in a World of Cost-Cutting and Downsizing," - Awarded "Best Contributed Paper",* Proceedings of the Twentieth SAS Users Group International (SUGI) Conference - 1995.

A wide selection of other JMP books in important areas can be viewed at http://jmp.com/academic/books.shtml

Acknowledgments

The authors thank Donalee Wanna, MWSUG 2012 Program Chair, Rex Pruitt, MWSUG 2012 Operations Chair, the MWSUG Executive Committee, SAS Institute, and Conference Leaders for organizing a great conference!

Trademark Citations

About the Authors

Charles Edwin Shipp is a programmer, consultant and author, and has been using the SAS and JMP software since 1980. He is credited in the original JMP manual for his roles in the early days. He has written more than one hundred papers and has been an invited speaker at more than one hundred International, regional, local, and special-interest SAS and JMP conferences and meetings, and is the recipient of 11 “Best” contributed paper awards. Charlie is the co-author of three books including the ever-popular Books by Users (BBU) book, Quick Results with SAS/GRAPH Software. Currently, Charlie is involved as an eBook author, sasCommunity.org Advisory Board member, app developer, consultant for Trivani Foundation International, and consultant in JMP and JMP Genomics.

Kirk Paul Lafler is consultant and founder of Software Intelligence Corporation and has been using SAS since 1979. He is a SAS Certified Professional, provider of IT consulting services, trainer to SAS users around the world, and sasCommunity.org Advisory Board member. As the author of four books including PROC SQL: Beyond the Basics Using SAS, Kirk has written more than five hundred papers and articles, been an Invited speaker and trainer at three hundred-plus SAS International, regional, local, and special-interest user group conferences and meetings throughout North America, and is the recipient of 19 “Best” contributed paper, hands-on workshop (HOW), and poster awards. His popular SAS Tips column, “Kirk’s Korner of Quick and Simple Tips”, appears regularly in several SAS User Group newsletters and Web sites, and his fun-filled SASword Puzzles is featured in SAScommunity.org.

Comments and suggestions can be sent to:

Charles Edwin Shipp
Consider Consulting
E-mail: CharlieShipp@aol.com
Twitter: @ShippAhoy
~~~
Kirk Paul Lafler
Software Intelligence Corporation
E-mail: KirkLafler@cs.com
Twitter: @sasNerd
~~~