

JMP Density Ellipse Script Code Made Easy

By Charles Edwin Shipp and Kirk Paul Lafler

Abstract

We show the use of scripts to meet two useful objectives: (1) display a graphic of correlation between all selected continuous variables: using an $(n \times n)$ matrix of density ellipses, and (2) display a selected independent variable vs. a selected dependent variable for different classes to see which class is the best, and visually, by how much. We also show and discuss the elements of the 50% density ellipse and its usefulness. To conclude, we show how to store the script code for dissemination and reuse.

Introduction: Creating Scripts

1. Write from scratch
2. Create with interactive data discovery
3. Modify an existing Template

There are three ways to create a JMP script to store for later use. If you have had the JMP Scripting Class, you can create scripts. But even as an expert you usually want to begin by creating the script interactively using data discovery and then modifying it, saving the resulting JMP script.

JMP Processing

Input >> Program >> Output

JMP Table >> JMP Script >> JMP Journal

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 JMP table and JMP script.

What a JMP Script Looks Like

```
Multivariate(Y(:Loss, :Fitness, :Severity, :KPS,
:Longevity, :R%, :Benefit), Scatterplot Matrix
(Density Ellipses(1), Ellipse Color(4)));
```

This program performs correlation display using density ellipses for the continuous variables mentioned in the argument. Files have the extension '.jsl' such as STUDY02.JSL. (Graph not shown.)

The JMP Script Language creates JMP programs and is concise yet powerful and is easy to use.

Another Script Listing

```
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)));
```

This script generates 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 ellipse methods are some of the best ways to show correlation with a graph, but not the only ones you can use within JMP.

VA Lung Cancer (Case Study)

In the graphics at the end, you can see that in this Veterans Administration data, there is correlation between "loss" and "KPS", but there is little or no correlation in "loss" due to the variable, "age".

The results of interactive JMP analysis are called JMP Journals and can be put into MS Word or PowerPoint final reports for dissemination.

Cancer vs. KPS: positive correlation

As the ellipse approaches a straight line the correlation number goes to 1.0 and as it goes to a circle the statistical graphic is showing there is little or no correlation.

Interactive Steps

- Basic Statistics
- Bivariate
- Identify X as KPS (or AGE, or another independent variable)
- Identify Y as LOSS (or another dependent variable)
- Click OK
- Group By: Cell-Type (or another class variable)
- 50% Density Ellipses
- Click on OK

The results of doing these steps in JMP are shown in the graphs at the end of the paper.

Circles that are squashed vertically or horizontally are the same as perfectly round circles. This is because the axis scaling is dependent on the choice of physical units to measure the variable on that axis.

VA Cancer: "Drill Up"

Interactive Steps

- JMP Starter
- Multivariate (twice)
- Identify Y as all numeric variables
- Click OK (shows the Matrix of Correlation, with visual density ellipses.)

Notice that in this slide of all continuous variables, if we know the independent variable we want to minimize is one of the "loss" variables, then we can look in that column for the independent causes to analyze and report. We choose the log-loss for the dependent variable and look for the contributors that will interest us.

In this study, it seems significant to show that the loss due to cancer is independent of age. This concludes the example case study, and we now draw conclusions on the presentation on making JMP script code easy.

Summary of Script Usage

1. Copy/Modify/Reuse Code
2. Same Script with Different Data
3. Example Correlation Discovery/Display
4. JMP Density Ellipse Code is: EASY!

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.

Summary

The JMP density ellipse applications allow both (1) looking at all continuous variables; and (2) contrasting cause and effect in classes using the same two variables. We show finding both positive correlation, which turns out to be very significant, and also little or no correlation. The ease of using JMP is shown.

Conclusions

There are two main ways to approach finding cause and effect in important studies, and JMP helps you in both approaches. You can first look at correlation among all important dependent and independent variables, and then focus on analyzing further the most promising. The second approach is when you have specific variables that you need to examine, as in the case study in this paper. We look specifically at the variables requested by management or lead personnel, and contrast cases for a given independent cause and a dependent effect. After this is completed, then we may choose to "drill up".

Trademark Citations

SAS, JMP, and SAS Certified Professional are registered trademarks of SAS Institute Inc. in the USA and other countries. The © symbol indicates USA registration.

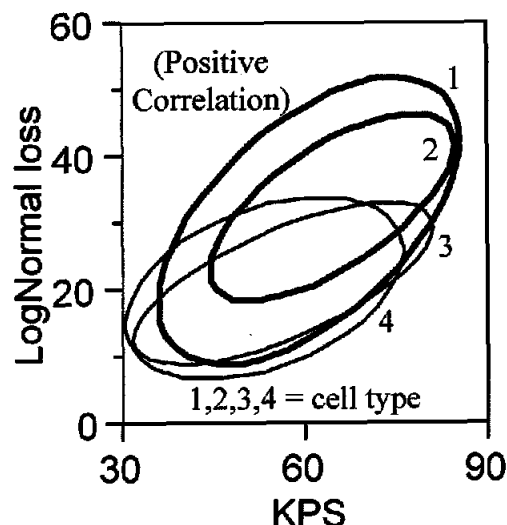
About the Authors

Charles Edwin Shipp is a programmer and consultant with over 25 years experience working with the SAS System. His current interest is JMP applications and instruction.

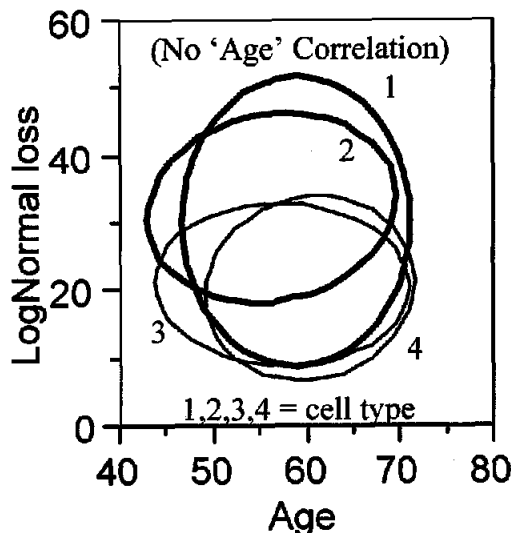
E-mail: CharlieShipp@AOL.com

Kirk Paul Laffer is a SAS® Consultant and SAS Certified Professional® with over 25 years of SAS software experience. He has written four books and over one hundred articles for professional journals and SAS User Group proceedings. His popular SAS Tips column appears regularly in the SANDS and SESUG Newsletters.

Results of JMP Script Code Runs



VA Lung Cancer Data from Kalbfleisch and Prentice, 1980, (JMP Sample Folder.) The density ellipses show the correlation for the cases more clearly than the scatter plot of data. (Cases are the four cell types.)



This important graph shows there is no correlation in age. Patients of all ages suffer equal loss in cells due to cancer.