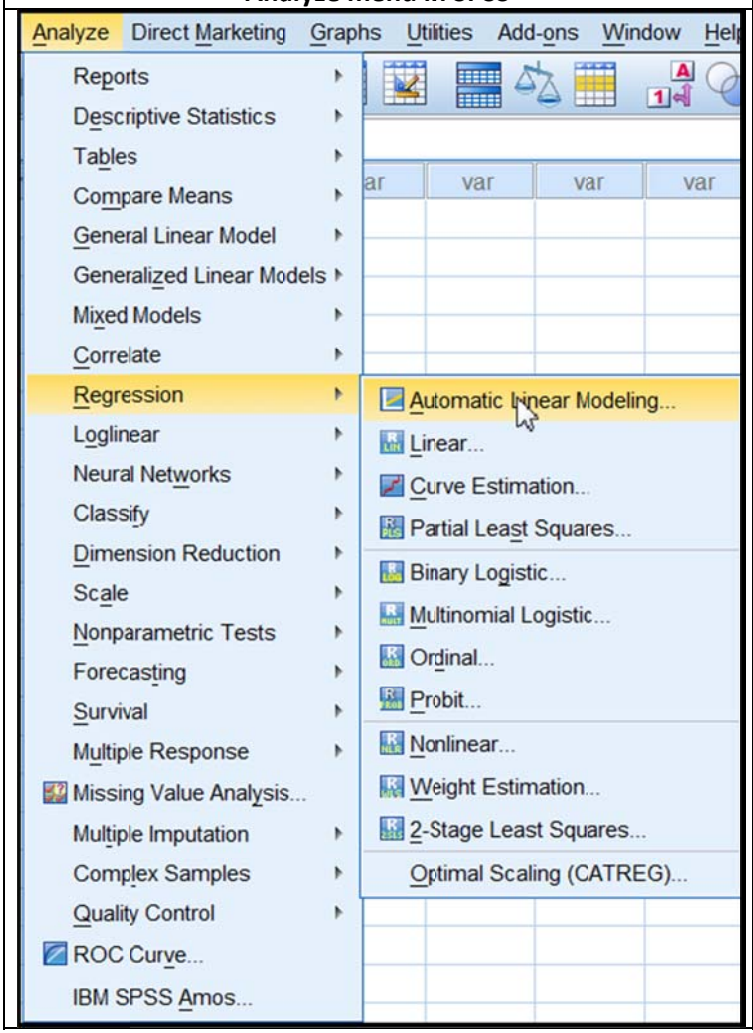
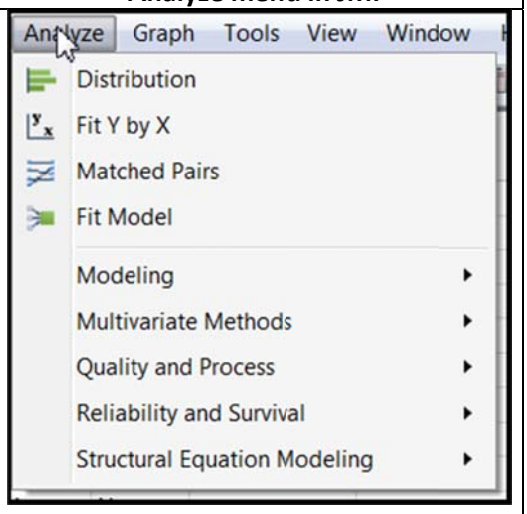
JMP Interface: iPod of statistical software

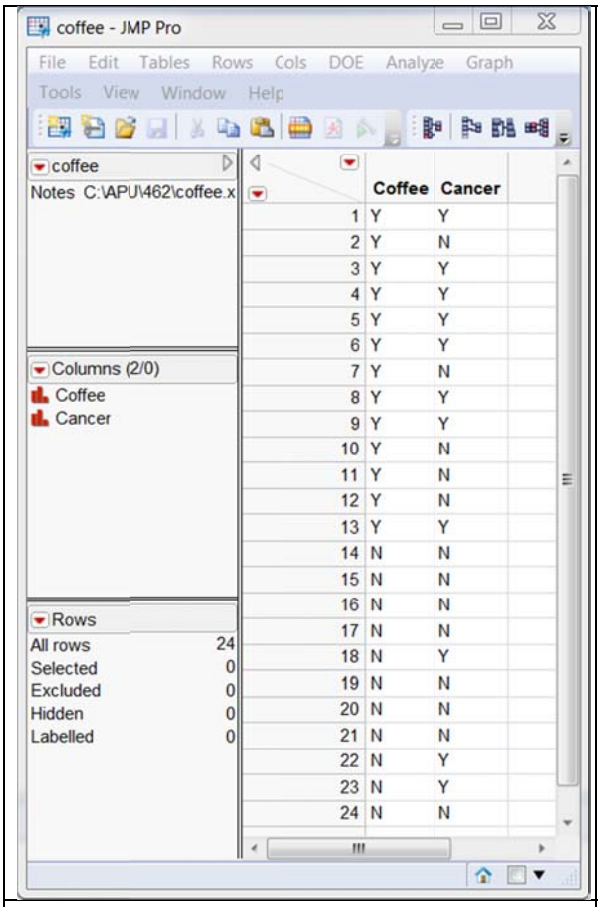
Chong Ho Yu, Ph.D. (2012)

cyu@apu.edu

www.creative-wisdom.com

JMP is software package created by SAS Institute for data visualization and exploratory data analysis. Its design principle is vastly different from that of SPSS. In SPSS the user must know exactly what to do while JMP does not require a solid foundation of prior knowledge. For example, when you access the pull down menu "Analyze" in SPSS, you can see many available options, but for novices this may be intimidating. On the other hand, JMP is like an Apple's iPod, which has a very clean and simple interface. In JMP you don't even need to know the name of the procedure. As long as you know what your dependent (X) and independent (Y) variables are, you can simply use "Fit Y by X" to solve most of your problems.

Analyze menu in SPSS	Analyze menu in JMP
	

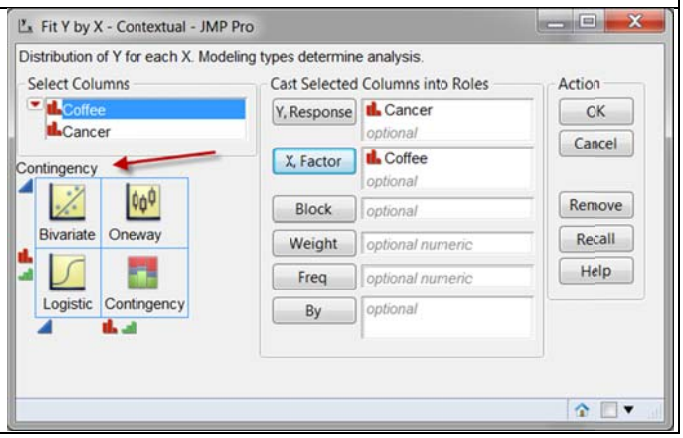


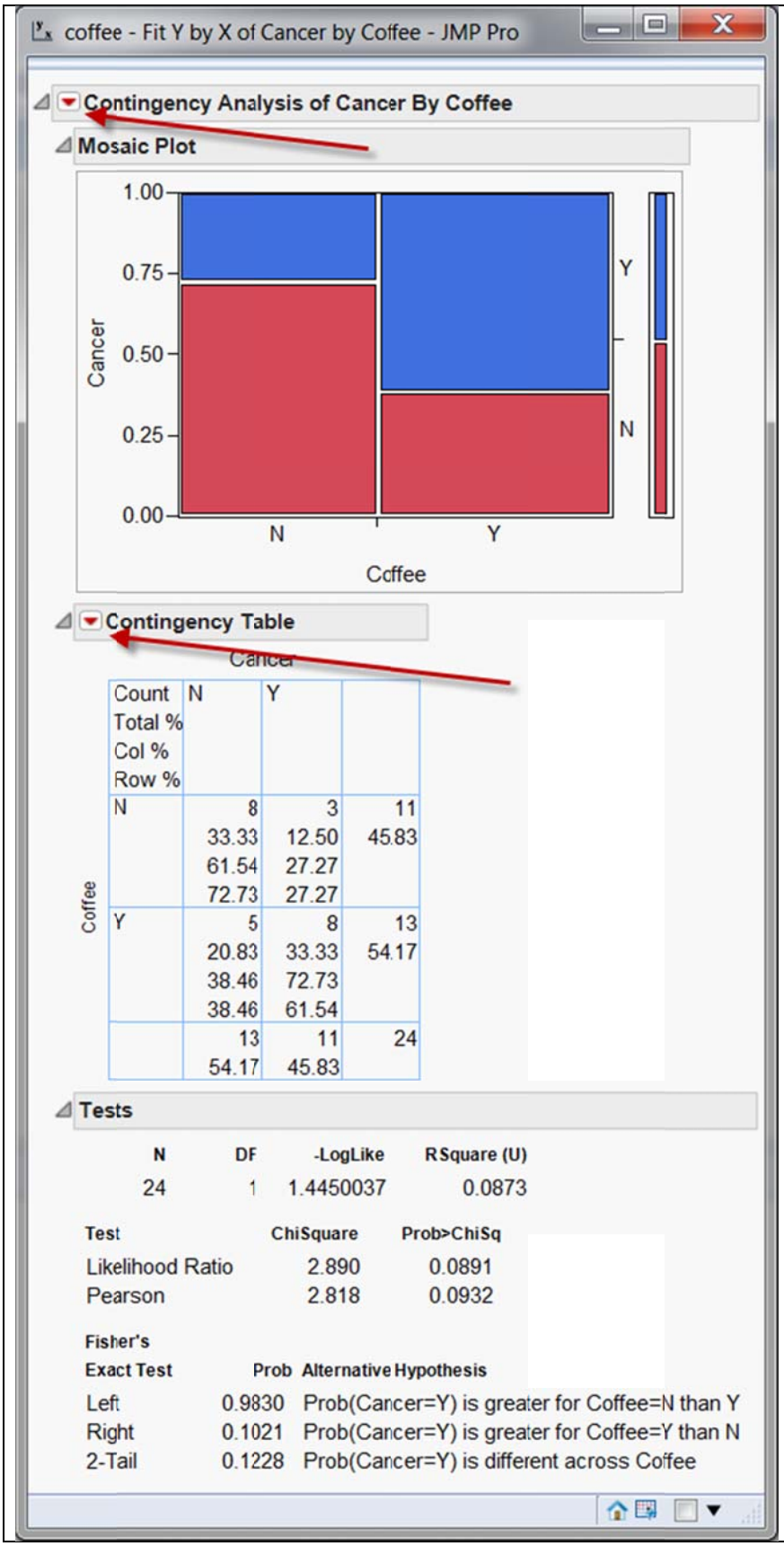
For example, in the data set shown on the left there are two nominal variables: Coffee consumption and cancer (Yes or No). I would like to know whether drinking coffee is related to having cancer or not. This is a typical Chi-square problem. Nonetheless, I don't have to know the term "Chi-square in order to proceed with data analysis.

I can select "Fit Y by X" from "Analyze," and then drag Coffee into X and Cancer into Y. In the pop-up window below, you can see the word "contextual" next to the phrase "Fit Y by X." In a contextual menu system, what the user sees in the next step depends on the previous step.

There are three types of variables according to the measurement level. A variable symbolized by a set of red bars is nominal while a variable represented by a set of green bars is ordinal. If the icon is a blue triangle, then the variable is continuous-scaled. There is no distinction between interval and ratio scales in computation (see the picture below).

After the variables are selected, JMP could recognize the data type and select the appropriate analysis. In this example because both variables are categorical, JMP realizes that it should be a contingency table (see the red arrow).

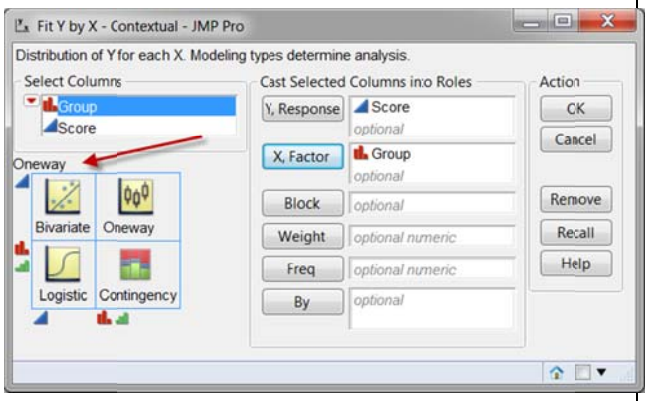


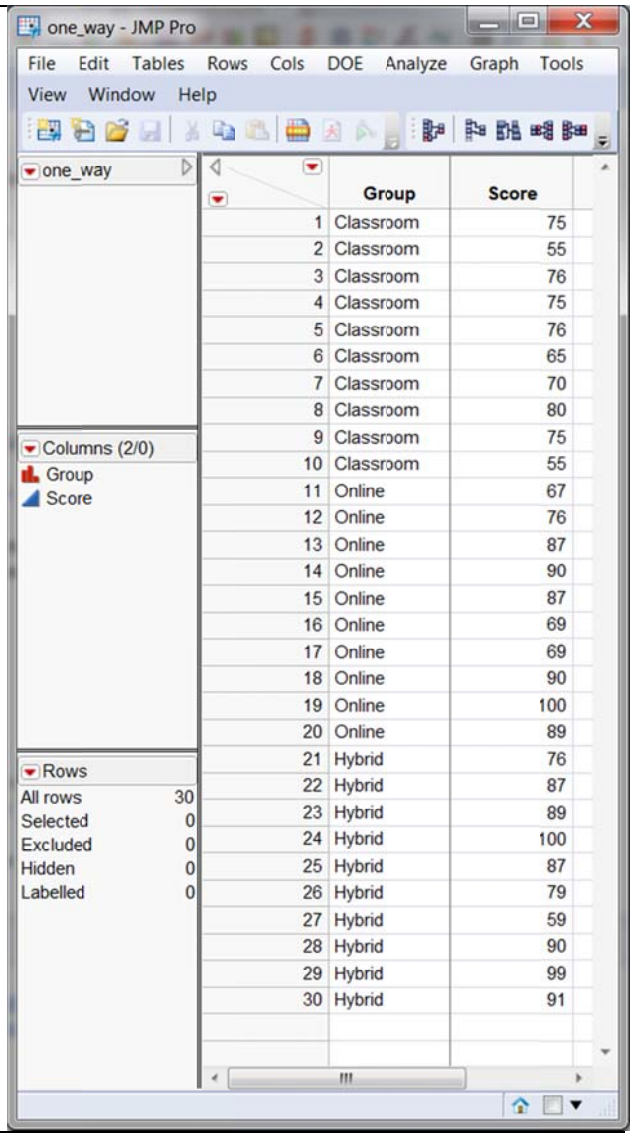


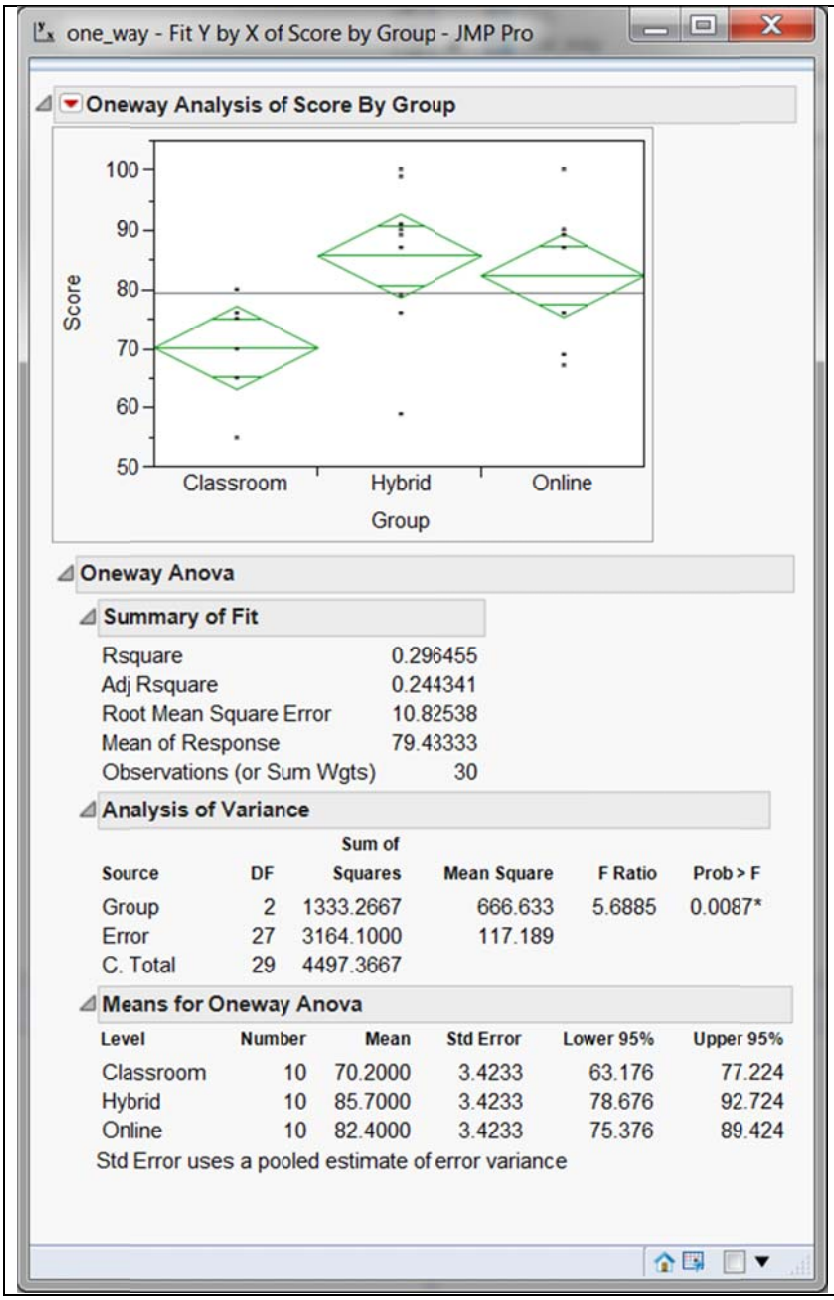
After clicking on the OK button, you can see a Mosaic plot, a contingency table, a Chi-square test result, and a Fisher's exact test result. In the contextual menu system, the subsequent menu options are tailored to this specific context. The user can request additional analysis by moving the mouse over the inversed red triangle.

The purpose of this write-up is not to explain the tests or to illustrate further analysis. Rather, the focus is on the user-friendly interface. The take home message is: using statistical software package does not necessarily require a steep learning curve.

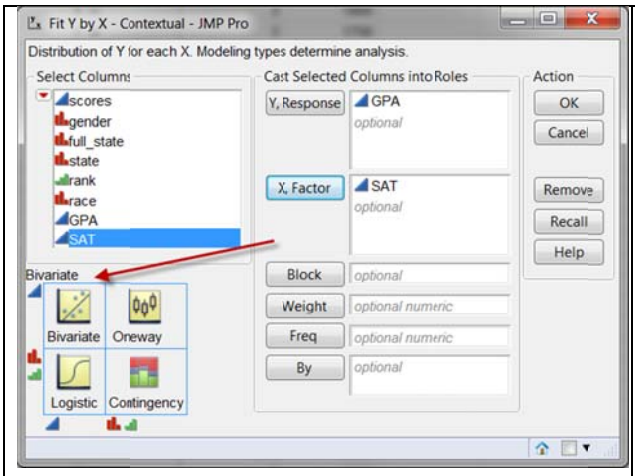
Now, let's look at another dataset. In this hypothetical data set there are two variables. One of them is a grouping factor: Group. The other is a numeric variable: test scores. I taught three classes using three different modes: Conventional classroom, online, and hybrid. I am interested in investigating whether different teaching methods would lead to different learning outcomes, as measured by test performance. This is a typical ANOVA problem. Again, I don't even need to know the name "ANOVA." What should I do next? "Fit Y by X," of course. I drag Group into X and Score into Y, and then push the OK button. By default JMP selects "Oneway" for this dataset.



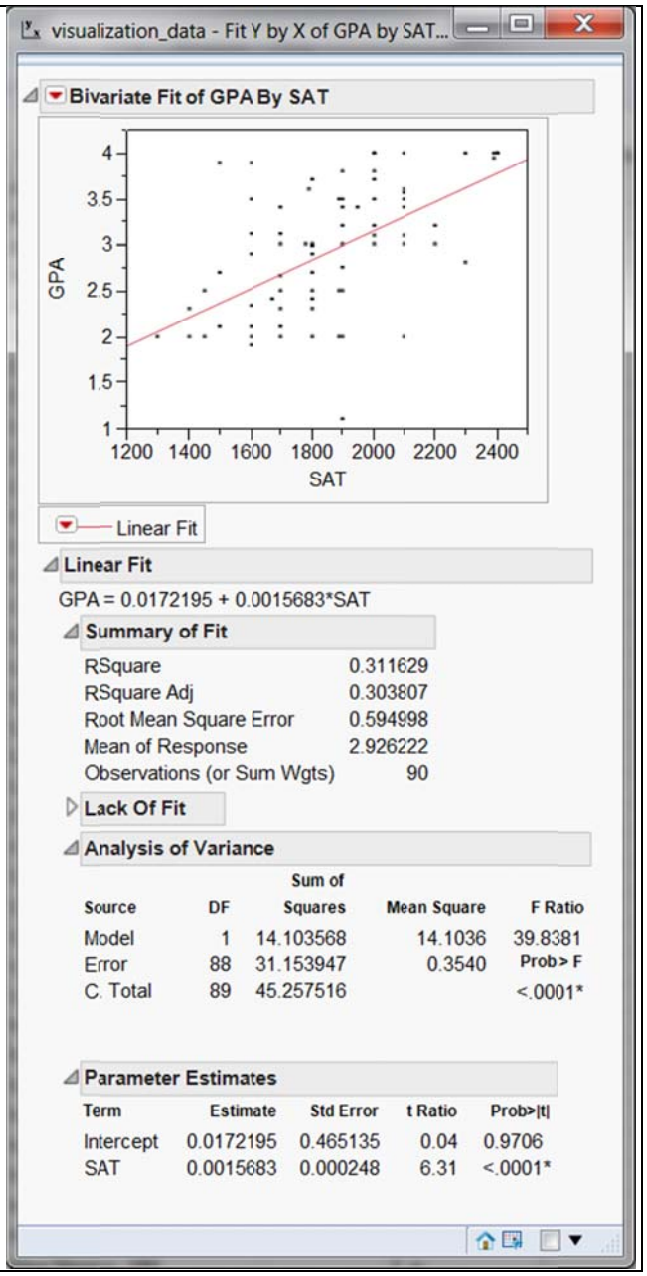


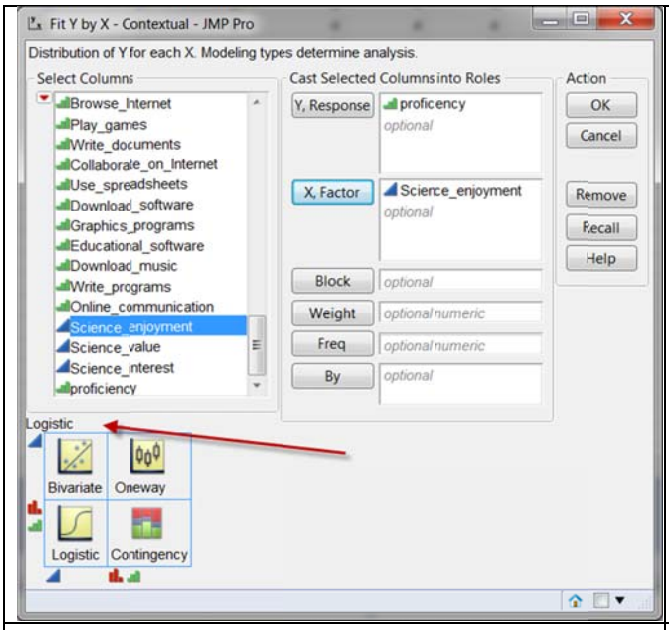


The panel on the left is the output. As mentioned before, the focus is about the interface and thus I will not discuss the meaning of the output. In this dataset there are three groups and thus ANOVA is used. But if there are only two groups, where should the user go to look for a t-test solution? Not surprisingly, the answer is still the same: Fit Y by X.

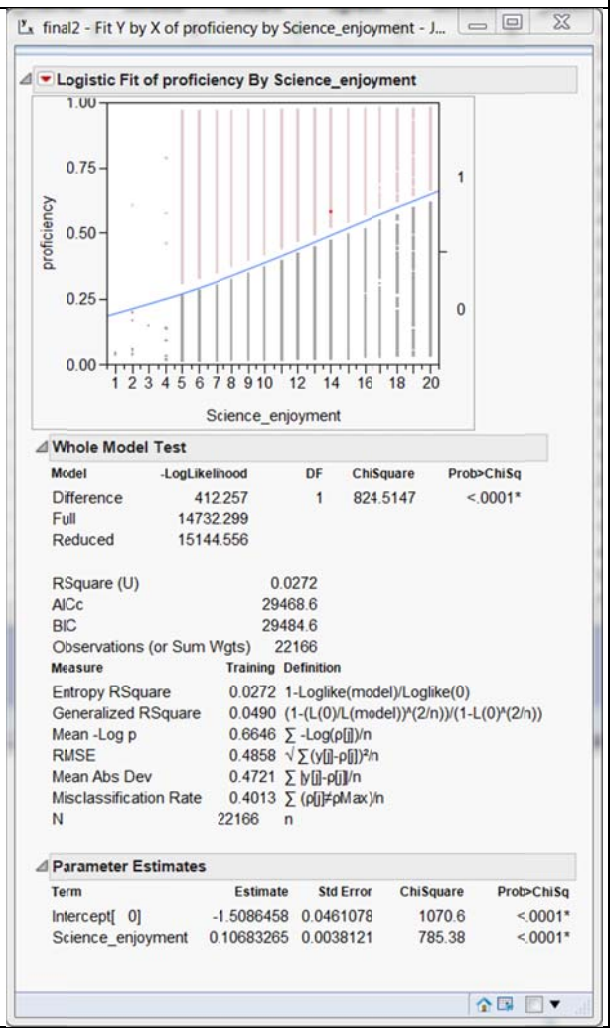


In another data set I have two continuous variables: GPA and SAT. I would like to use SAT to predict GPA. The problem can be solved by regression analysis. How can I run a simple regression in JMP? You know the answer: Fit Y by X. In this case JMP recommends a bivariate correlational approach. The panel on the right is the simple regression output. By default it shows the scatterplot only. You can request a linear fit or other procedures by using the inversed red triangle.





Last but not least, in another dataset I classified the observations into the proficient group (1) and non-proficient group (0) based on their science test scores. I would like to know whether the degree of enjoying science could predict their status. In this case, logistic regression is the best option, of course. Needless to say, JMP recommends it when "Fit Y by X' is used.

The right panel shows the output of the logistic regression model. JMP is like an Apple's iPod, but it is not really an iPod. IPod has one button only but in JMP you still need other items on the menu for more complicated research problems. Nevertheless, many basic computations can be performed via Fit Y by X. In addition, you may notice that in every output panel there are graphics and tables. JMP subscribes to the philosophy of data visualization: a picture is worth more than a thousand words. Novices may be confused by jargon such as sum of squares, root mean squared error, and likelihood ratio, but graphics are almost self-explanatory.

Next time when your friend asks you a statistical question, you can simply reply, "Fit Y by X!"