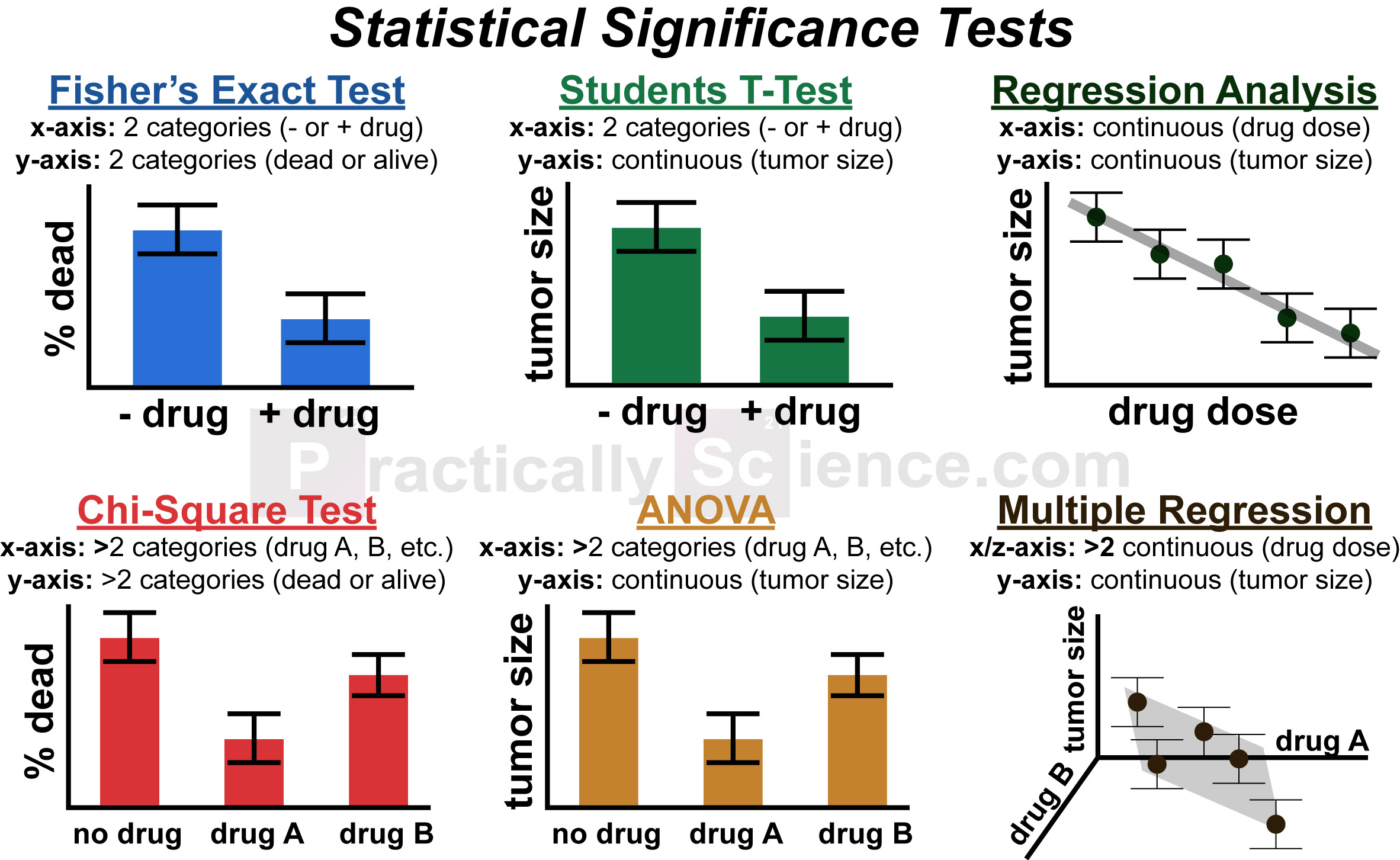
Some figures and statistics tips

**Figures should include** bar graphs or scatterplots (likely with regression lines fitted). Bar graphs should include well-labeled axes and error bars (+/- one standard error). Each figure should have a concise, descriptive caption and include any necessary legend, and a description of what the error bars are. The convention is to put a \* above the bars if they are significantly different at the p=0.05 level, but you should describe this in the caption.

**Notes on statistics**

Choosing the right test. For this assignment, we want you to use a test from the top row of the figure below or a chi square test unless you are already very comfortable with ANOVA. No multiple regressions, please. What’s the right test for a given analysis? As you can see, categorical vs. continuous data, and the number of categories if you have categories, are the key factors.

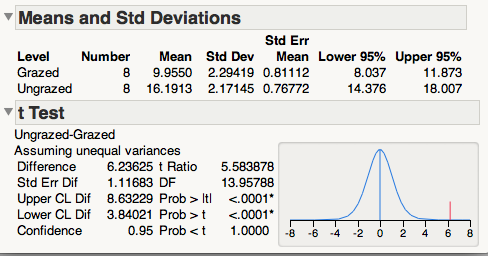
The t-test, ANOVA and regressions assume that you have normally distributed data. If you don’t, your first line of defense is to try a data transformation. If you can’t get a transformation to normalize your data, you have to use a nonparametric test. To check for normality of your data, see the instructions below.



**For t-tests, you need to** **check that your data are somewhat normally distributed**. Use the histogram function in JMP under “basic” in the JMP Starter Menu under “Window.” Ideally, the histogram is shaped like the blue curve in the screenshot below – peak in the middle, tailing off in either direction, and symmetrical. If it’s not symmetrical at all, has multiple peaks or no peaks, or otherwise looks nothing like a “Bell Curve,” you need to try transforming your data.

* **t-tests** in JMP: Window -> JMP Starter -> Basic -> 2-sample t-test

Output will look about like this below a crude graph of means and standard deviations:

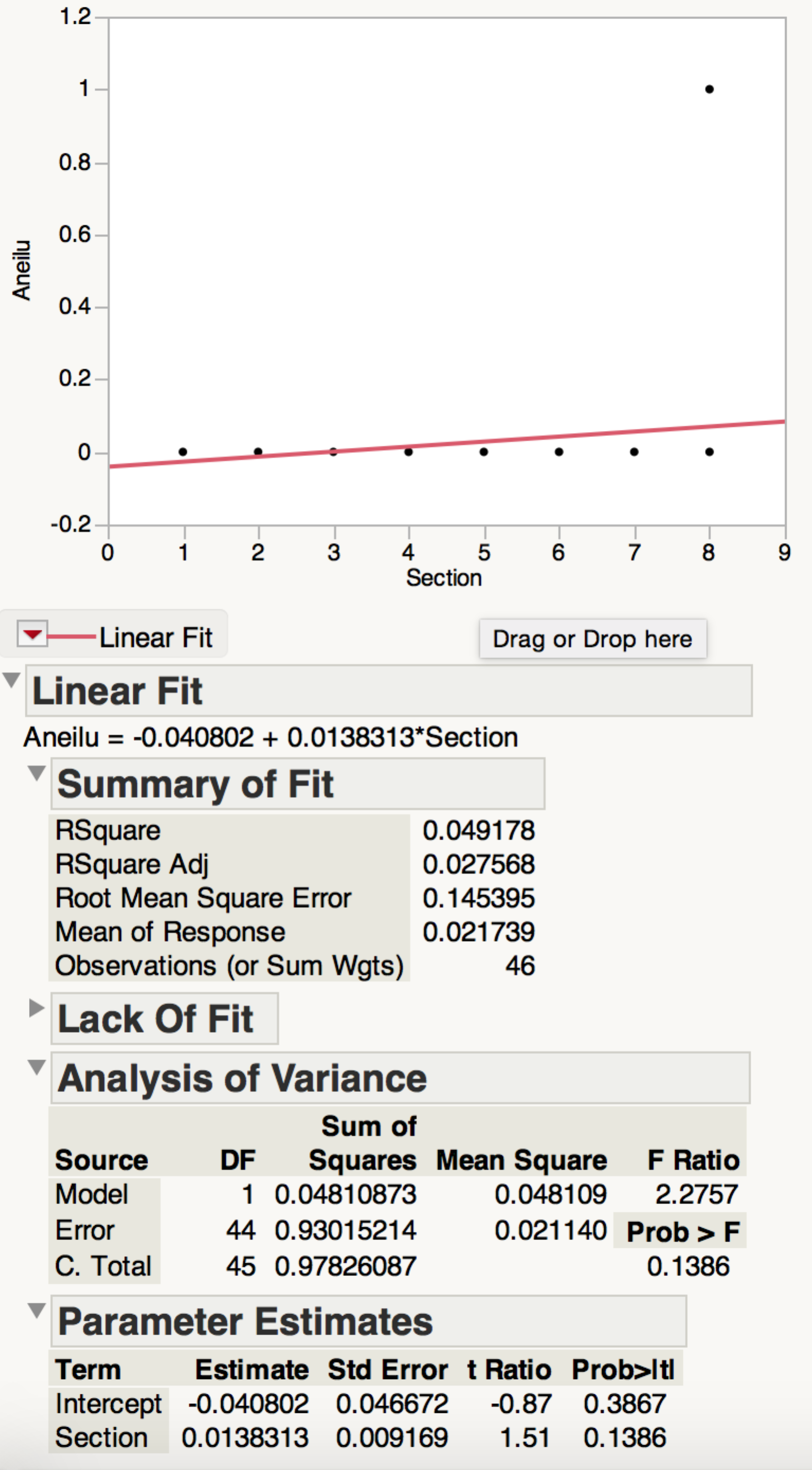
-Your two-tailed p-value is the Prob >|t|.

-Your t (test statistic) is the t Ratio

-“Difference” reports the difference between the means (listed for each Level in the top panel with standard deviations, standard errors of the mean, and 95% confidence intervals).

- The red line on the normal curve shows you where the difference between means falls (here, the red line is way out on a tail, indicating that it’s very unlikely to get sampled if the true difference between means is zero).

To create a figure showing a difference between two groups (as in a t-test), you must make a **bar chart** with standard error bars. Go to Graph-> Chart -> put your independent variable in the “Categories, X, Levels” box then select your y-variable and use the dropdown menu beside “Statistics” to select “Mean.” This way each bar height will be the mean across y values for the corresponding x category. Click in the lower left box to turn on “Add Error Bars to Mean” and check that it’s set at 1 standard error. Once the graph shows up, you can click on axis labels, etc. to make these informative and include units.

**For regressions**, Window -> JMP Starter -> Basic -> Fit Y by X. Your data need to be continuous (both y and x) for it to give you a bivariate scatterplot, so make sure you have assigned these correctly in the datasheet.

- Output will be a scatterplot with title “Bivariate Fit of…”

- This is a useful figure for just looking at your data. To fit a line, click the red arrow next to the title and select “Fit Line.”

Then you will see this output like the figure on the left. Below “Linear Fit” is the best-fit equation for the line. R2 Adj is the best measure of what % of the variance in y is explained by x (in this case, it is ~28%). At the bottom, below “Parameter Estimates,” you see the coefficients for the intercept and then for the independent (x-, predictor) variable. Each is accompanied by a p-value. In the example at left, neither is significant at the p<0.05 level. The predictor variable, Section, has a p=0.1386.

In this case, the **scatterplot with fitted line figure** can be used for a paper as long as the axes are labeled thoughtfully, and the figure includes a good caption. The y-variable axis should begin at zero, ideally. To fix this, select the black arrow tool above the graph title (on left), then double-click the y-axis on the graph and set “Minimum” to zero.

For chi-square tests, I recommend computing them by hand (with a calculator) – it’s fast, and then you know what on Earth is going on because you see the expected and observed values and can see where the proportions you measured depart from the null expectation. A decent guide is at: <http://davidmlane.com/hyperstat/B155367.html>

Another approach, with a video, is at <http://www.statisticshowto.com/what-is-a-chi-square-statistic/>. It is also possible in JMP to use Window -> JMP Starter -> Basic -> Contingency for chi-square tests.