AMS 7L: Lab 1

Today is about getting your first practice with the software we’ll use this quarter in AMS 7L: the statistical package JMP, distributed by SAS (one of the most widely-used data analysis environments in the world). We’ll look at a variety of descriptive methods (graphical and numerical) for examining one variable at a time, and you’ll spend some time with the JMP tutorial and help files, to which you can turn any time you need help and the instructors of AMS 7 and 7L aren’t around.

In what follows, all click instructions are left mouse clicks unless otherwise noted.

1. Getting started. Log on to your machine using your ITS login (the one you usually use for UCSC email). Double-click on the icon that says PC Server on ‘be-lab’ (F), double-click on Math, Statistics and Graphing, and double-click on JMP6 shortcut. You’re now in JMP; it should immediately offer you two windows: the Tip of the Day, which will in the early going be too advanced to be of use to you (when you’re done looking at it, Close this window by clicking on the X in the upper right corner of the window), and the JMP Starter window, which is where everything useful begins.

2. Creating a new data set. Click on New Data Table; a new window called Untitled will appear, with rows and columns for you to manually enter a data set (soon we’ll bring data into JMP by other means, but in this first lab it’s good practice [to get ready to work with small data sets of your own] to enter the data set manually inside JMP.

3. Manual data entry. Click in the first row under Column 1; that entry in the data table will turn white and a cursor will start to blink, indicating that you can now type in some data.

Zar (1999) reports on data from a study of the wing length (in cm) of immature monarch butterflies (Danaus plexippus). A representative sample of \( n = 24 \) young monarchs was taken; their wing lengths, sorted from smallest to largest, were

\[
\begin{align*}
3.3 & \quad 3.5 & \quad 3.6 & \quad 3.7 & \quad 3.8 & \quad 3.8 & \quad 3.9 & \quad 3.9 & \quad 3.9 & \quad 4.0 & \quad 4.0 & \quad 4.0 & \quad 4.1 & \quad 4.1 & \quad 4.1 & \quad 4.2 & \quad 4.2 & \quad 4.3 \\
3.3 & \quad 3.5 & \quad 3.6 & \quad 3.7 & \quad 3.8 & \quad 3.9 & \quad 3.9 & \quad 3.9 & \quad 4.0 & \quad 4.0 & \quad 4.0 & \quad 4.1 & \quad 4.1 & \quad 4.1 & \quad 4.2 & \quad 4.2 & \quad 4.3
\end{align*}
\]

Take a minute or two to type this data set into the column; you can hit Enter to go from one row to another or you can click into the next row after typing each number in.

If you type in the wrong number, double-click on the wrong entry and it will be highlighted; you can then type the right number, which will replace the wrong one. JMP will think you want to continue editing the data; to get out of that mode you can just click in an empty row at the bottom of the data set.
If you type something that’s not actually a number (for example, at one point by mistake I typed 4.w2), a window will appear that says Non numeric data entered. Change column to character column? and you’ll be given three buttons to choose among: Try Again, Change, and Revert. Click on Try Again and JMP will let you have another chance at typing the right data in.

When you’re done typing in the data, double-check it by using the slider or the up and down arrows on the right of the Untitled window to go back to the top of the data, and make sure that all of the data entries are correct; at the end of this you should have the data set above, with 24 observations, correctly entered into JMP.

Another thing to try is to Resize this window by putting the mouse arrow on the exact lower right corner of the window (the mouse arrow will turn into a little sideways arrow pointing from northwest to southeast, so to speak); then hold down the left click button and use the mouse to make the window wider or narrower, taller or shorter, whatever you want — in this case it’s natural to make it tall and narrow so that all 24 data values are displayed, and then to check the entire data set you won’t need to scroll up and down.

4. **Naming a variable.** I’d like to name this column Wing-Length and I don’t know how.

This is a good moment to start getting some help — click on the Help button at the far right of the top of the JMP window; click on Tutorials; and click on Beginners Tutorial.

A new window will appear called Beginner’s Tutorial — Move this window (by clicking on the bar on top of it to turn the bar blue; then hold down the left click button and use the mouse to move the window where you want it to go) so that it doesn’t obstruct your view of the Untitled window, and follow the instructions in this window until you get to the point where it says

In an analysis, a hot spot appears on an Outline item. The hot spot gives commands that extend the analysis. You can also access these commands with a right-click inside the outline item’s territory.

This suggests that if I put the mouse arrow in the Untitled window in the cell labeled Column 1 and right-click, I might get a menu that will allow me to name this column — try this; you should get a menu with Column Info ... as the first entry; (left) click on this option and a new window will appear called Column 1; click in the box called Column Name, type Wing-Length, and click on the OK button: the variable has been correctly named.

5. **Naming a data set.** Using the same idea, put the mouse arrow in the box called Untitled (with the red hot button in it) in the upper left corner of the Untitled window and right-click; type in Butterfly and hit enter; now the data set has been named Butterfly and the variable measured on the butterflies has been named wing length.

6. **Data types.** In the left part of the Butterfly window toward the middle there’s a box called Columns (1/1) with a red hot button in it; (left) click on the little blue triangle next to Wing-Length in this box and a menu of data types will appear with three possibilities: Continuous (which has a check mark next to it), Ordinal, and Nominal. Evidently JMP has inferred from the numerical flavor of the Wing-Length variable that it’s continuous,
which is true, so in this case I leave this aspect of things alone by clicking somewhere else in the JMP window and the data types menu disappears (if my variable had been ordinal, for example, this would have been a good time to tell JMP by selecting that data type instead of Continuous).

7. **Graphical and numerical descriptive methods.** To focus on the Butterfly window and the analysis we’re about to do on the Wing-Length variable, you can [Minimize](not kill) the Beginner’s Tutorial window by clicking on the underscore _ button in the upper right corner of it.

Look again at the JMP Starter window — on the left side the menu offers 11 options: File, Basic, Model, Multivariate, Survival, Graph, Surface, Measure, Control, DOE, and Tables.

In later labs we’ll explore some of options 3–11; for now, click on Basic, and JMP offers you eight things to consider doing: Distribution, Two-Sample t Test, Matched Pairs, Fit Y by X, Oneway, Bivariate, Contingency, and Logistic.

Again, in later labs we’ll examine some of options 2–8 on this list; for now, click on Distribution and a new window called Report: Distribution will appear (you can minimize the JMP Starter window at this point to concentrate on what follows).

Click on Wing-Length and click the OK button to get our first results: a new window called Butterfly - Distribution of Wing Length will appear — you may want to move this window as close to the top of the JMP window as you can because it’s going to get longer, and you may as well resize it so that all of its contents (including the summary called Moments at the bottom) are visible.

Under Wing-Length there are three types of summary: a histogram and boxplot, a table with the variable’s Quantiles, and a table with the variable’s Moments.

In week 2 of the lectures we’ll talk about histograms and quantiles, and moments is just a fancy word for things like the mean and the standard deviation of a variable (which will also be covered in week 2); boxplots will be covered in lab next week.

At this point your lab instructor will briefly mention the main descriptive features of this variable, as summarized by the first window produced by the Distribution command in JMP.

Notice that clicking on one of the blue half-diamonds makes the corresponding part of the results display disappear.

After talking through the main features of the descriptive summary that appear in the default JMP analysis, click on the red hot spot next to Wing-Length: a new menu with 15 options will appear (only one of which has a check-mark next to it: Outlier Box Plot, which was described above).

Clicking on Display Options shows that Quantiles and Moments are default options (which we already knew by the fact that their results are displayed); clicking on More Moments produces additional summaries like the sum of the data values and their variance (as well as some other things you can read about in T&T).
We’ll come back to this example in next week’s lab, when we’ve talked more about graphical and numerical descriptive methods in class.

8. **Beginner’s Tutorial.** You can finish out the rest of this lab by going back to the Beginner’s Tutorial and working your way through it to the end.

Before quitting JMP, at this point you would ordinarily save your data set if it’s newly created or if it has been modified; we’ll talk about how to do that next week.

There’s no lab assignment to work on this week — see you next week.