AMS 7L **LAB #1** Spring, 2009

Data Types and Introduction to JMP

Name: _____

Lab Section: _____

Instructions:

- The TAs/lab assistants are available to help you if you have any questions about this lab exercise. If you have any questions please raise your hand and they will get to you as quickly as possible.
- At the end of class, you will need to turn in this cover sheet to your lab instructor. *If you do not turn it in, you will not get credit for this lab.* Be sure to write your name and section above.
- The following symbol ↬ at the beginning of a question means that after you answer that question you should raise your hand and have a TA or lab assistant review your answers up to that point. Once they have reviewed your work they will initial in the appropriate space in the table below. The purpose of this check is to be sure you have answered the questions correctly.
- Be sure to take the rest of the lab handout with you when you leave. It contains your answers and *JMP* instructions which you may find useful for doing homework assignments.

Check-Problem ↬	Lab Instructor's Initials
2	
8	

AMS 7L **LAB #1** Spring, 2009

Data Types and Introduction to JMP

Objectives:

1. To introduce the use of the statistical software package JMP
2. To review data type classifications
3. To practice basic data display methods

Getting Started: Log onto your machine using your ITS login (the same as for your @ucsc email). Note that unless otherwise specified, mouse clicks are left-clicks.

Part I. Starting JMP and Entering Data

This quarter, we will be using the statistical software package JMP, which is distributed by the company SAS (who make one of the most widely-used statistical software packages).

To start JMP, click on the "Start" button in the lower left corner of the screen, go to the "Math, Statistics and Graphing" menu item, go to the right to "JMP" and then to the right of that, click on "JMP 7". If it asks you for a directory for your preferences file, you can choose a directory by navigating to it and clicking on `Select`, or else you can just `Cancel` and not save any preferences. (Note that files 'saved' on the Desktop will likely be deleted when you log out.) This should now bring up two windows. The top one is the tip of the day, which you may or may not find useful. For today, it is probably too advanced a tip, and you can just click on the `close` button. This should leave you with the JMP Starter window.

Creating a new data set: In the JMP Starter window, click on `New Data Table`. A new window called `Untitled` will appear, with rows and columns for you to manually enter a data set. Click in the first row under Column 1. That cell should turn white and a cursor will blink, indicating that it is waiting for you to type in some data. We'll enter the following data on milk production in Tulare County (in the central valley of California). For the years 1989 through 1998, the annual gross value of milk produced (in millions of dollars) in the county was:

287 363 413 411 455 477 547 569 712 718

Enter these 10 datapoints in column 1. You can hit `Enter` to move down the rows. (If you want to use the numeric keypad to enter numbers, you may first need to hit the `Num Lock` key on the keypad.) If you make a mistake, you can double-click on the entry to highlight it and then correct it (hit `Enter` after you've fixed it). If you enter something that is not a number, JMP will give you a dialog box that says `Non numeric data entered. Change column to character column?` Click on the `Try Again` button and fix your mistake. When you're done entering the data, double-check your entries. If the data list were longer, you could use the scroll bar in the right to move up and down the list. Also note that if needed, you could re-size the window by moving the mouse to the lower right corner of the window and then clicking and dragging to change the window size.

It can be helpful to **label the variables** with names, because "Column 1" isn't terribly informative. The easy way to do this is to double-click where it says `Column 1`, which brings up a new window that will allow you to change the column name to something like "Milk Production" by

clicking in the box to the right of `Column Name` (where it says "Column 1") and then typing in the new name, and then clicking on the `OK` button. Another way to change attributes of the column is by left-clicking (the usual click) on the header where it says `Column 1`, and then right-clicking in the same box (which should now be blue). That will bring up a new menu, and left-click on `Column Info` to get the same window as above.

Similarly, you can **label the dataset** by double-clicking in the box in the upper left where it says "Untitled" (with the red hot spot button) and changing the name to something more helpful such as "Milk".

Question #1 What type of variable is milk production: quantitative or qualitative?

↬ **Question #2** Is milk production continuous or discrete?

On the left side of the window, the second box down should be labeled "Columns (1/1)" along with a red hot spot (triangle) button. Under the label you should see the name you gave your variable (e.g., "Milk Production") with a blue triangle on the left of the name. Click the blue triangle and you will see that you can specify whether the variable is continuous, ordinal, or nominal. Check that it is correctly labeled (with the check mark to the left of the appropriate type). Note that if you are entering a column of numbers, JMP will automatically guess that the variable is continuous. So if you mean the variable to be some other type, you will need to specify it in this way.

Part II. Data Summaries and Graphical Displays

At this point, we are ready to do some simple analysis. Move the "Milk" data window over to the side so that you can see the "JMP Starter" window again. You can move the "Milk" window by clicking and holding down the left mouse button on the blue bar at the top of the window, dragging the window over by moving the mouse, and then letting go of the mouse button when the window is in the right place. Now click on the "JMP Starter" window and notice the list of modes on the left. We'll first get some simple summary information. Click on the `Basic` menu item and see that the set of buttons in the window changes. The first button is `Distribution`. Click on this button, which will open up a new window labeled "Report: Distribution". Since we only have one variable (Milk Production), it should already be highlighted on the left under "Select Columns". Since this is the only variable, you can just click on the `OK` button. You should get a histogram, a boxplot and a bunch of numerical summary information below those.

Question #3 Looking at the histogram, near what value are most of the observations in this dataset?

Question #4 Look down in the summary information under "Moments". What is the mean (average) value of annual Milk Production over this time period?

Go back to the "Milk" window and click on one of the rows or data entries. Notice how JMP highlights that observation in the histogram. Try clicking on several different rows in the data table and seeing JMP show the corresponding points on the histogram.

Let's make one more plot. Since the data are in time order, it might be interesting to see if there is a trend over time. Close the "Distribution" window by clicking on the `X` button in the upper right of that window. Click on the "JMP Starter" window to re-activate it. On the menu list on the left, click on `Graph`. Then choose the `Overlay Plot` button. Again, since we only have one variable, JMP can easily figure out what we mean and you can just click on `OK`. If we had more than one variable, we'd have to give JMP more instructions. You should get a new window with a plot of milk production over time. Again, clicking on a point in the data ("Milk") window will highlight that point on the plot. Selecting several rows (click and drag) in the data window will highlight all of the corresponding points in the plot.

Question #5 What can you say about how the value of milk produced has changed over time?

Part III. More Practice with Data Types

This section is just paper and pencil; you don't need JMP for this section. For each of the following variables, state whether it is quantitative or qualitative. If it is quantitative, specify continuous or discrete. If it is qualitative, specify nominal or ordinal.

Question #6 Whether a gallon of milk is organic or regular.

Question #7 The amount of milk produced by a particular cow in one day.

↬ **Question #8** The number of cows on a particular farm.

Quit JMP by clicking on the `File` menu in the upper left corner and choosing `Exit`. We won't need this milk data again, so when JMP asks "`Save changes to the JMP Data Table Milk?`" you can click on the `No` button. If JMP asks you about saving preferences, you can skip that by clicking on the `Cancel` button. Please remember to **Log Off** (from the "Start" menu in the lower left of the screen).