Lab 2 – Pets

In Lab 2, you're doing to look at the distribution of pets, in this case, cats and dogs, in different regions of the continental United States. You're going to generate some graphs of pets by region, and generate a table of statistics that you will analyze.

First, let's look at the data.

There are 3 files for you to use. Pets.mat is an array that is 48 by 6, with each state, except Alaska and Hawaii each having their own row. The first column is the total number of households surveyed in that state. The 2nd column is the number of households reporting any cat or dog pets. Column 3 is the number of households that reported having at least one dog as a pet, and Column 4 is the total number of dogs reported in that state. Column 5 and 6 are the same, but for cats. Households with no cat or dog but did have other pets were not included in these numbers. In thinking about this data, let's imagine what a household with different collections of pets would report. If you have a single dog, you would be included in Column 1, 2 and 3 and you would contribute 1 to Column 4. If you have 2 dogs, you would be included in Column 1, 2, 3, and contribute 2 to Column 4. Same for a household with only cats, except you would contribute to Columns 1, 2, 5 and 6. If you have 1 dog and 2 cats, you would contribute to Columns 1, 2, 3, 4, 5 and 2 to Column 6.

US_states.csv is a data file that lists each state (except Alaska and Hawaii) and what row it is, it's index. You do not need to load this, but you will need to open it for Part II. Lab2.m is a script file we started for you, and should make this lab easier to understand.

For Part 1, do your normal setup and the only difference is that you need to extract the zip file into your Lab2 directory – you can't work with it zipped.

In part 2, you create the variables you are going to work with. Just like in Lesson 2, you are going to extract certain rows, based on the instructions. Open us_states.csv, and determine what rows you are going to put into these variables, based on what region they are grouped into. For example, Alabama is row 1, and it is in the South, so you will need to include row 1, plus the 15 other states associated with the South when you create the variable South.. You will want to include all 6 columns with each variable.

Region_households is the number of households surveyed in each region, so you'll have to use some of your math skills. Hint: <tt>sum</tt> or a <tt>+</tt> symbol.

For Part III, you're going to generate a pie chart of your survey results, by region. A pie chart is appropriate for comparing overall numbers, so generate a pie chart of how many households were surveyed, by region. Since we have broken it out into 4 regions, you should have 4 pie pieces. Make sure the legend is right.

For Part IV, you're going to generate two bar charts, side by side, plotting the number of cats by region on one and number of dogs by region on the other. Use the set(gca) code to label graphs.

For Part V, you're going to do something very similar to the end of Lesson 5, where you generate a table. We have already provided the format, all you need to do is the calculations into the right variables. To do this, figure out what should go into each row, and do the calculations. Get a row right, and then move on. Don't be afraid to

comment out some variables you haven't finished yet to see if you get the results you're expecting.

As always, make sure you have organized your script into cells.

For the analysis, again, you're writing bullet points. Remember the data triangle – data is the numbers, information is the data in some context, knowledge is what the data means, and what it might be caused by, and wisdom is understanding everything. Your bullet points need to emphasize the information to knowledge transition.

Your first set of bullet points are on why one region has the highest number of pets. Your second set of bullet points are on why there is a higher number of cats than dogs. Your third set of bullet points are to analyze the survey itself – what problems are there in this survey.

Then you get to write a paragraph about how you would write and conduct this survey to make it more robust.

Remember, the course homepage, under Resources, has sample bullet points and paragraphs to make sure you understand the expectations.

Good luck!