

## CS130 Assignment #3

**Date assigned:** Monday, October 1, 2012

**Date due:** Wednesday, October 10, 2012

**Points:** 75

Goals: SPSS data definition, type, measure, descriptive statistics, charts, hypothesis testing

Create a Word document **WordAnswers3PUNuetID.docx** using your PUNetID that will contain answers to each of the following questions in order.

Answer each question in your Word document under the header **Problem# - Question #**.

### **Problem #1 Use SPSS to solve this problem.**

A pretty comprehensive list of caffeinated drinks can be found at <http://www.energyfiend.com/the-caffeine-database>.

1) Import the table of *Drink*, *Fluid Ounces*, *Caffeine (mg)*, and *mg/oz*, and *category* (Energy Drink, Energy Shots, ...) from the Web into an Excel. Name your new worksheet **Caffeine Drinks**. Get rid of ALL extraneous text such that your worksheet has a proper title, column headings for the data, and the full list of data under the column headings.

Note: As this worksheet will be opened in SPSS, you want the column headings to have good names with no spaces in the names as the Excel names will become SPSS variables.

2) Using SPSS, create a dataset **Caffeine.sav** from the Excel data in 1). Make sure that you set the type, label, values, and measure correctly for your data.

3) Find the Mean, Median, Std Deviation, Minimum, and Maximum values for the column of data representing mg/floz. Paste the table of values created by SPSS in your Word document.

4) Using SPSS, create a bar graph that shows the number of Coffee, Energy Drinks, Energy Shots, Other, Soda, and Tea totals in the entire dataset. Display the bars in blue and display the total number of each category on the bar as we've discussed in class. Make sure your bar graph has a proper title and appropriate axis names. Paste your graph in the Word document

5) Using SPSS, create a pie graph that shows the number of Coffee, Energy Drinks, Energy Shots, Other, Soda, and Tea totals in the entire dataset. Display the total number in each category on each piece of the pie as a percent as we've discussed in class. Make sure your pie graph has a proper title. Paste your graph in the Word document.

6) Using Excel, create the same pie graph as in 4). There is no need to display the total number in each category on each piece of the pie. Make sure your pie graph has a proper title. Paste your graph in the Word document. In Excel, the process is a little more tedious as you need to list each category and then in a cell next to each category, use the COUNTIF function to count the number of entries in each category.

7) Is the milligrams per fluid ounce higher for energy drinks as opposed to soda. In your Word document, answer each of the following questions:

- a) What hypothesis test will you use to answer this question? Explain in detail why you selected the test you did.
- b) State the Null Hypothesis.
- c) Perform the correct hypothesis test.
- d) Paste in the results produced by SPSS.
- e) Do you accept or reject the Null Hypothesis? Why?
- f) State your conclusion.

7) After finding the mean of the milligrams per fluid ounce for ONLY the sodas, you are to determine whether the mean milligrams per fluid ounce is greater than 2.75 in a statistically significant way. Answer each of the following questions in order in your Word document:

- a) What hypothesis test will you use to answer this question? Explain in detail why you selected the test you did.
- b) State the Null Hypothesis.
- c) There are a couple of ways for you to use only the soda observations. You can use Transform->Compute Variable or you can grab only the soda observations from the Excel file and place these observations in SPSS. Save the SPSS dataset as **Soda.sav**, then perform the correct hypothesis test.
- d) Paste in the results produced by SPSS.
- e) Do you accept or reject the Null Hypothesis? Why?
- f) State your conclusion.

## **Problem #2**

We sometimes use a correlation analysis to indicate the degree of a linear relationship that exists between two variables. Typically the researcher will gather two or more observed values from each subject (or sampling unit) and then try and see if there is a linear relationship between any of the variables. SPSS computes the Pearson correlation coefficient which has an absolute value between 0 and 1 where a 1 indicates a perfect linear relationship and a 0 means that no linear relationship exists whatsoever. If two variables increase or decrease proportionately, then we say a positive correlation exists between these two variables. If one variable increases while another decreases, then there is a negative correlation.

In the CS130 Public folder on Turing is a data file called WesternStates.txt that is a "tab" delimited data file. The file consists of scores for several different variables where the higher the score, the better, except for the variables Housing and Crime. In these two cases, the lower the score, the better.

Create an SPSS dataset from the WesternStates.txt by importing the data into SPSS as we have done in class several times. Save the dataset as **WesternStates.sav**.

In your Word document, report on each of the following again placing the bold heading before each answer:

- a)** What two different variables have the highest correlation and what is the Pearson correlation coefficient?
- b)** Create and paste into your Word document a Scatterplot of these two variables.
- c)** Perform a linear regression on the two different variables with the highest correlation and paste in the Coefficients table produced by SPSS.
- d)** Perform the same linear regression using Excel and paste in a graph of the linear regression where the graph includes the linear equation and  $R^2$  value.

Parts a) through c) are to be done using SPSS while part d) is to be done using Excel.

### **Submitting your work:**

Submit your correctly named Word document to CS130 Drop by 4:45pm on the day that it is due.

Grading will be based on:

- Correctness of your results.
- Completeness of your results.
- Professional look of your Word document.