SPSS Assignment #1 - Graphs

The main goal of this assignment is to give you some experience on how to use SPSS to create some statistical graphs. This particular SPSS assignment is worth 20 points. I want you to work in pairs for this assignment. You only need to submit one report for your group. Your report should be type-written and well organized.

- Open the SPSS data file health_exam_results.sav. Check your email or you can download it from my web site at http://websites.uwlax.edu/storibio/ (then click STAT 245, then Health Exam Data). Clicking on it should start the SPSS program automatically and your data will show up in an SPSS window.
- II. This data are from the U.S. Department of Health and Human Services, National Center for Health Statistics, Third National Health and Nutrition Examination Survey. It has a total of 80 cases (40 males and 40 females) with each case having values for 14 variables. These variables are listed below. For this assignment, you will only work with 3 variables (*Gender, Height, and Weight*).
 - 1. Gender
 - 2. Age (in years)
 - 3. *Height* (in inches)
 - 4. Weight (in pounds)
 - 5. *Waist* (circumference in cm.)
 - 6. *Pulse* (pulse rate in beats per minute)
 - 7. *SysBP* (systolic blood pressure in mmHg)
 - 8. *DiasBP* (diastolic blood pressure in mmHg)
 - 9. Cholesterol (in mg)
 - 10. BodyMass (body mass index)
 - 11. *Leg* (upper leg length in cm)
 - 12. *Elbow* (elbow breadth in cm)
 - 13. Wrist (wrist breadth in cm)
 - 14. Arm (arm circumference in cm)

Note that in the lower left corner of your SPSS spreadsheet you will find two options: *Data View* and *Variable View*. Clicking on the "Variable View" will show you the list of variables in the data with their characteristics. This is where you can modify the characteristics of variables. Clicking on the "Data View" will take you back to the actual values of the variables. This is where you can change the values of the variables. In either view, you should be able to see the "Analyze" and "Graphs" buttons in the topmost toolbar. Take note of these two buttons as we will be using them very often.

- III. Assume that these 80 individuals comprise a simple random sample taken from the United States. <u>The goal of this study is to obtain meaningful information about the whole U.S.</u> <u>population based on this data set</u>. As in any data analysis, the first thing that you should do is to explore the data graphically. So, construct 2 boxplots for each of the two quantitative variables (*Height* and *Weight*) – one for males and another for females. You can do this by
 - 1. Selecting "Graphs" \rightarrow "Legacy Dialogs", then choose "Boxplot".
 - 2. Choose "Simple" and "summaries for group of cases", then click "define".
 - 3. Choose a quantitative variable that you wish to work with and then press it to move your selected variable into the "variable" space. Next move *Gender* to the "Category Axis". Then hit "ok".
 - 4. You can then copy and paste these box plots into Microsoft Word

OR, you can also

- 1. Select "Graphs", choose "ChartBuilder" and "boxplot", then drag one of the available boxplot icons to the Chart preview space above.
- 2. Use the pointer to grab the variable and drag them to the appropriate space.
- IV. Next, study the distribution of these variables by looking at their histograms. The "Histogram" can be found under "Graphs" → "Legacy Dialogs". Describe what you see for each variable.
- V. Here are the things that I want you to address and include in your report.
 - 1. Copy and paste all relevant boxplots into a word document. There should be a total of 4 boxplots (male height, female height, male weight, and female weight).
 - 2. Compare and contrast the 40 males and 40 females, with respect to both height and weight, based on their corresponding boxplots. Are these results consistent to what you know about males and females?
 - 3. Are there outliers? If there are, what do you think caused this value to be an outlier? Is it likely to be due to human error in recording or is it a rare event? Explain your answer. [Hint: *Check the other attributes of these individuals.*]
 - 4. Copy and paste all relevant histograms (*again, there should be 4 of them*). Describe the distributions of these two variables for males and females. Are they more or less symmetric? Or are they skewed? Based on these histograms, estimate where the means are and compare between the two genders.
 - 5. Finally, comment on how useful these plots were in your study of the data.

Your score for this work will depend on the following criteria:

- 1. Completeness of the report.
- 2. Organization and neatness of the report.
- 3. Accuracy and quality of discussion.