Assignment 3   Produce Descriptive Statistics with SPSS
Grade: 40/40

Preparation and Submission of This Assignment: There are 3 parts in this assignment. You will submit two files for this assignment, an SPSS data file and a word document file. Part 1 is to create the **SPSS data file** for the data used in this assignment using SPSS software. For naming the file for part 1 of this assignment, please use "A3" and your last name and first name initial plus with word "Data". So, if I were to do it, the file name would be **A3_ChangG_Data.sav**. Part 2 and 3 are to use SPSS to obtain descriptive statistics, and the answers should be put in the WORD assignment document that you download from the course page. For part 2 and 3 of the assignment, please **use this assignment Word document file** to do the work and name this file starting with **A3** plus your last name and first initial. If I were to submit this assignment, my filename would be **A3_ChangG.doc**. Please submit these two files through **Blackboard Assignment Dropbox** before deadline.

Part 1: Create SPSS Data (10 points)

Use the data provided in the **Assignment 2** to create an SPSS data file and save the data file with the word "**data**" as part of the file name, and also your last name and first name initials. So, if I were to do it, the file name would be **A3_ChangG_Data.sav**.

Data Recoding: Please enter numerical values for some of the categorical variables. For gender variable, please use value "1" for male and "0" for female. For Exercise per week variable, just enter the number. For instance, enter "2" for "2 days". For Daily hours of TV viewing variable, enter value "1" for "More than 2 hours" and "0" otherwise. And, then use the **Values** function in **Variable View** section to label these values with their actual meaning for each of the variables. For instance, label "1" with "Male" and "0" with "Female" for the gender variable. See Label Values for a Variable and Save Data File video in **Teach yourself SPSS** page (http://people.ysu.edu/~gchang/SPSS/SPSSmain.htm) to learn more about it.

Keep a copy of this data file in your disk, and submit a copy of the data file through Blackboard Assignment Dropbox along with your answers to part 2 and 3 of this assignment (Word Document file). See Blackboard instruction video for how to submit assignment through Assignment Dropbox.

In Teach yourself SPSS web site you can learn how to use SPSS software to create SPSS data file and make statistical charts and do statistical analysis. The web address for this SPSS instructions web site is: http://people.ysu.edu/~gchang/SPSS/SPSSmain.htm. (Please copy and paste the web address mentioned in this document to get on to those web sites.) There are text instructions and video instructions. You can view the video clips on **Data File Creation and Data Processing** section in the SPSS instruction web site mentioned above to learn how to create an SPSS data file. There are also other web sites that have SPSS videos for you to learn SPSS. Check SPSS References page.

Text instruction for doing SPSS exploratory data analysis in part 2 and 3 of this assignment can be found at the following web address: http://people.ysu.edu/~gchang/SPSSE/SPSS_EDA_16.pdf

To do this assignment, you need to have the SPSS software installed on your computer. If you do not have SPSS at home, SPSS is also available at Math computer lab in Lincoln Hall and Computer Lab in Cushwa. YSU bookstore also has Student Version of SPSS available. You may also lease a license from the following web site: http://www.onthehub.com/spss/ [look for **IBM® SPSS® Statistics Base GradPack 19 for Windows (06-Mo Rental))**.

Part 2: Exploratory Data Analysis (20 points)

In part 1 of this assignment, you have created an SPSS data file. This part of the assignment is to use that data file to perform the following tasks (make charts) using SPSS statistical software for exploring and understanding the data. For each chart below that you put in for answer, you must also label it with **figure number** and **title** and above each chart you need to write a sentence or two to describe what you see in the chart. (See Example of Assignment 3 in this document.)

1. Make a **histogram** for the **Weight variable** to display the distribution of this variable. (Use a class width of 10.)

   Figure 1 below is a histogram for displaying the frequency distribution of the quantitative variable weight (kg). The distribution for this graph is right skewed and multimodal.
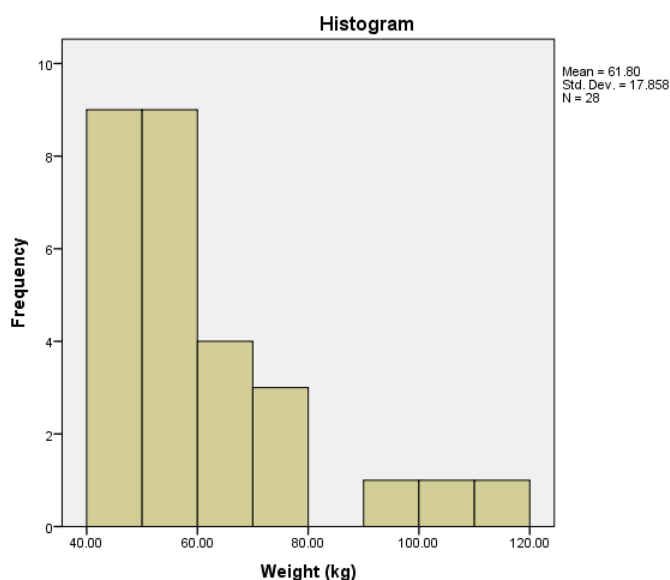

Figure 1: Histogram for Weight (kg)

2. Make a **frequency distribution table** for the **gender** variable to see the frequency distribution and then make a **bar chart**.

   Figure 2 below is frequency distribution table for displaying the frequency distribution of the qualitative variable gender. The table shows that the distribution of females and males are equal.

Gender

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Female	15	50.0	50.0	50.0
	Male	15	50.0	50.0	100.0
	Total	30	100.0	100.0	

Figure 2: Frequency Distribution Table for Gender

Figure 3 below is a bar chart for displaying the frequency distribution of the qualitative variable gender. The distribution of females and males is uniform.
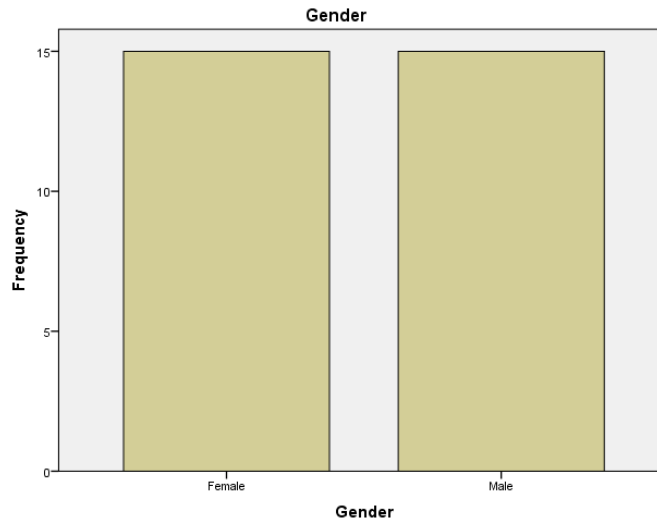

Figure 3: Bar Chart for Gender

3. Make a **cluster bar chart** to examine the correlation between **gender** and **Daily hours of TV viewing** variables. (Use the **Daily hours of TV viewing** variable as the category axis and **gender** variable as the cluster variable.)

    Figure 4 below is a cluster bar chart for displaying the correlation between the qualitative variable gender and the quantitative variable daily hours of TV viewing. The chart shows that females are more likely to watch 2 or fewer hours of TV and males are more likely to watch more than 2 hours of TV.
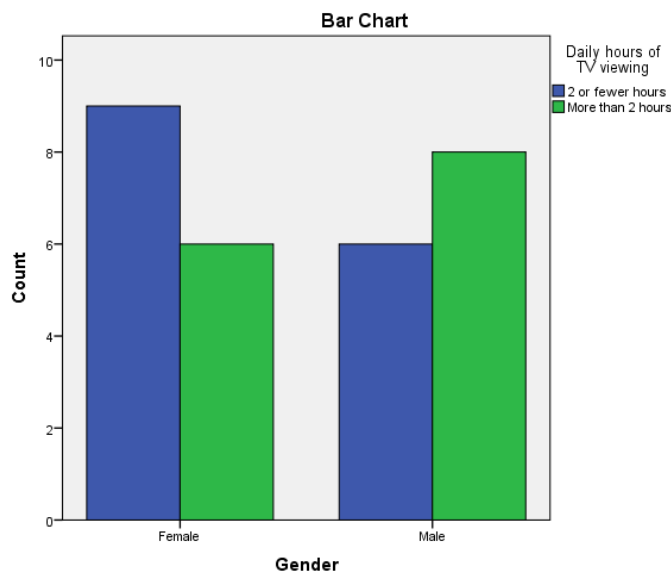

Figure 4: Cluster Bar Chart for Gender and TV Viewing

4. Make a **scatter plot** to examine the correlation between **weight** and **height** variables, and write a sentence to describe the trend you observed from the scatter plot.

Figure 5 below is a scatter plot that shows the correlation between the quantitative variables weight (kg) and height (m). The scatter plot shows that there is no correlation between these two variables.
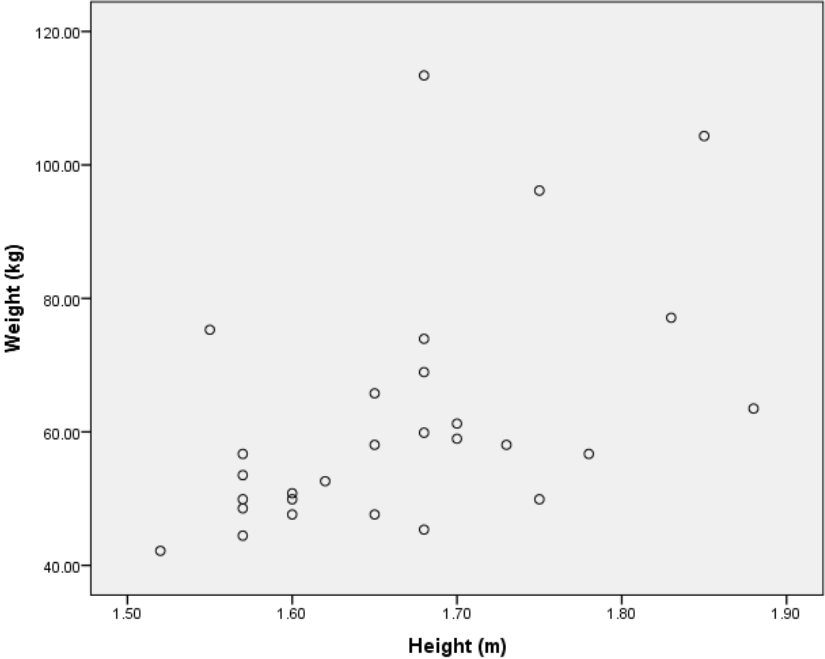

Figure 5: Scatter Plot for Height (m) and Weight (kg)

5. A quality control officer recorded the average length for a random sample of 10 of steel frames made from a production line in (inches). The sample was taken one every hour. Produce a time plot to display the trend.

Year	Average Length
8 am	5.1
9 am	4.9
10 am	5.1
11 am	5.2
12 am	5.0
1 pm	5.3
2 pm	5.5
3 pm	5.9
4 pm	6.5
5 pm	7.7
6 pm	9.6

Figure 6 below is a time plot showing the quantitative variable of average length (in) of steel frames every hour. The time plot line is ascending, which indicates an increase over time.
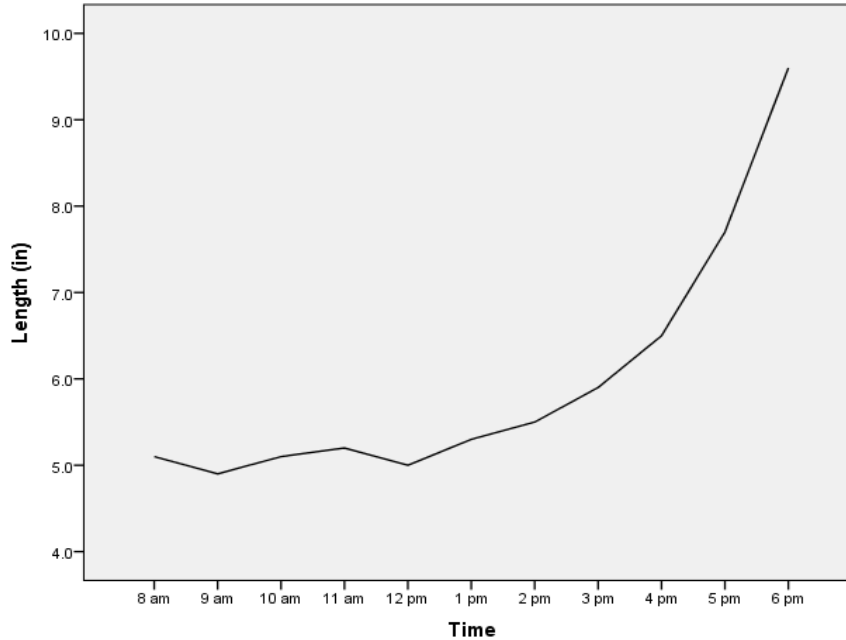

Figure 6: Time Plot for Average Length (in) of Steel Frames

Remember that all graphs in your paper should be properly labeled with figure number and title, see Example of Assignment 3. You should adjust the graph so that the graph is not too large in the document. To do so, after you have pasted the graph in MS-Word document, you can click on the graph and move the mouse pointer to a corner of the graph, and click and drag the corner to adjust the graph size.

Grading:

For each of the 5 questions above, if you

provide no answer	0 points
provide a chart with statement but are all not quite correct	1 points
provide a correct chart but no statement	2 points
provide a correct chart and statement but did not label with figure number	3 points
provide a correct chart with proper statement and proper label	4 points

Notes:
The web address for the SPSS text instructions on Exploratory Data Analysis is:
http://people.ysu.edu/~gchang/SPSSE/SPSS_EDA_16.pdf
Video instructions can be also found in **SPSS References** page for using SPSS and using MS-Word for typing report.
http://people.ysu.edu/~gchang/SPSS/SPSSmain.htm
If you have not used MS-WORD, this is the time to learn it. It is an important tool that you should know for many good reasons. Feel free to see me or contact me for assistance in learning this word processor.

Part 3: Descriptive Measures (Please fill in your answers in this document.)
(20 points)

1) Find the overall mean, median and sample standard deviation of **weight** variable in this data set.
   Sample Mean = __61.804_____
   Sample Standard Deviation = __17.858_____
   Sample Variance = ___318.892_____
   Sample Median = ___57.38_____

2) Does the distribution of the **weight** data for these children symmetrical belled-shape by looking at the histogram?
   (Circle or underscore or red colored your answer)   Yes   **No**

3) Report the percentage distribution of the **Daily hours of TV viewing** variable using the valid percentage distribution that do not include missing data.

Hours of TV Viewing	Relative Frequency
2 or fewer hours	__51.7____ %
More than 2 hours	__48.3___ %

4) Report the percentage distribution of **Exercise Per Week** variable using the valid percentage distribution that do not include missing data.

Exercise Per Week	Relative Frequency
0 Days	_20__ %
1 Days	__13.3__ %
2 Days	_13.3___ %
3 Days	_6.7____ %
4 Days	_20____ %
5 Days	_6.7____ %
6 Days	_6.7____ %
7 Days	_13.3___ %

5) Does the **weight** data suggest that it was from a normally distributed population? Perform a **normality test** and report the p-value of the test using .05 or 5% as the cutoff for decision making of the normality test.

   Report the p-value from the Shapiro-Wilk's normality test and it is: __0.000_____

   Your conclusion on the normality is (type your answer using less than 30 words):
   According to the normality test, since the p-value from the normality test is 0.000, the data suggests that weight data is not likely from a normally distributed population.

6) Report the mean, median and sample standard deviation of **weight** variable for **female subjects** in this data set.
   Sample Mean = __59.389_____
   Sample Standard Deviation = ___18.584_____
   Sample Median = ___52.16_____

Notes:

6

A **test of normality** video is in the following link:
http://people.ysu.edu/~gchang/SPSS/One_Quant.html

Remark: Charts and tables should always be properly numbered and labeled. See example below which will be the **format** of your all other future SPSS projects.

Example of Assignment 3: Descriptive Statistics

1. Figure1 below is a bar chart for displaying the distribution of the qualitative variable Gender. The frequency of the females is more than that of males.
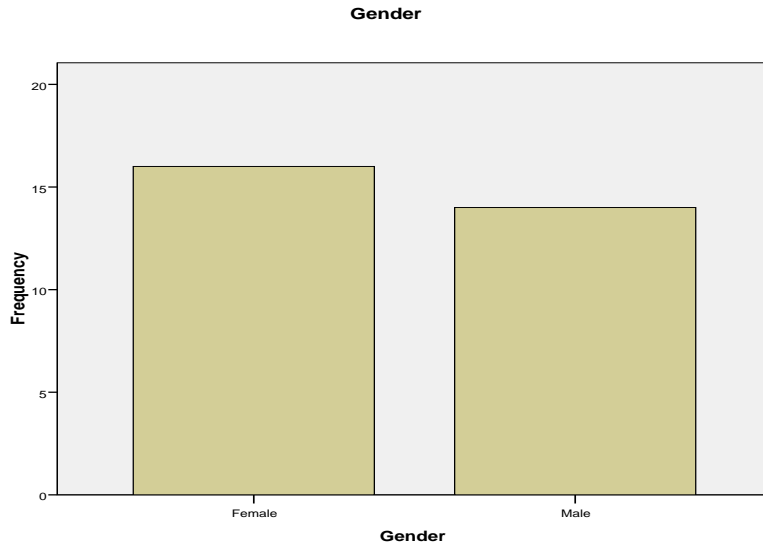
Gender

Figure 1: Bar Chart for Gender

2. The figure below is a histogram for the quantitative variable of height. The distribution seems slightly skewed to right.
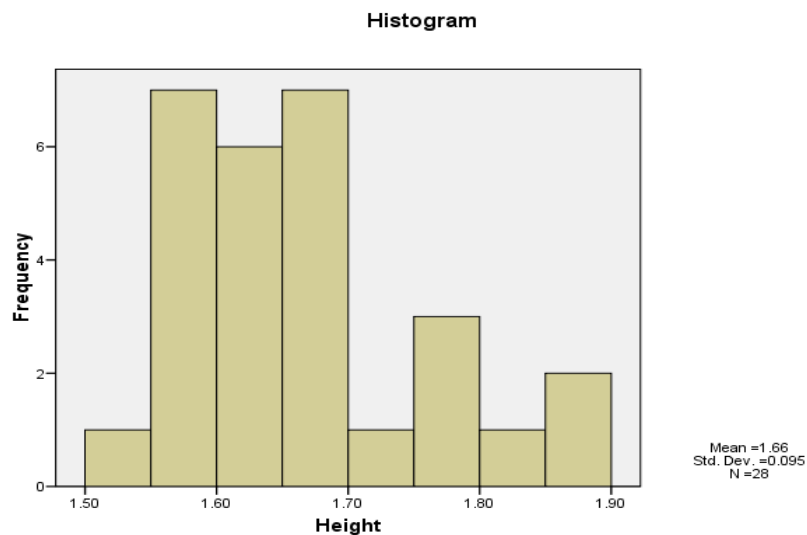
Histogram

Mean =1.66
Std. Dev. =0.095
N =28

Figure 2: Histogram for Height (m)

Few tips on MS-WORD

1) Use $\boxed{\text{Ctrl}}$ + $\boxed{\text{Alt}}$ + $\boxed{=}$ (press them at the same) to type superscript, and do the same to go back to normal text.   Example: X^2

2) Use $\boxed{\text{Ctrl}}$ + $\boxed{=}$ to type subscript, and do the same to go back to normal text.
Example: X_2

3) For Greek letters and math symbols, from the MS-WORD menu bar, click and select through the following sequence: **Insert / Symbol**. You can insert symbols like: μ σ Φ Ω $\neq$ $\approx$ $\otimes$ $\subseteq$ $\supset$ $\pm$ $\leq$ and more …

4) Click and select through the following sequence for inserting page number:
 Insert / Page Number …

5) Click and select through the following sequence to produce a mathematical equation with mathematical symbols: Insert / Object / Equation

   Example: $\sum_{i=1}^{n} \sqrt{x_i}$

There are more to explore in MS-WORD. You should start getting use to using a word processor to write your projects and papers.