

Stat 243 – Lab 2 -Descriptive Statistics

Background:

The goal of this lab is to successfully use Minitab to graph and analyze the data, copy this analysis back to a word processor, and finally to clean up the copied data into a neat report suitable for printing and handing in.

Follow these instructions:

- 1) Begin by downloading the MiniTab project labeled Lab 2 Data on the Lab 2 Assignment on Canvas. Open this project in MiniTab.
- 2) We will only use the “Weight” and “Height” variables, you may either keep the other columns or delete them.
- 3) Begin by creating histograms of both Weight and Height with the y-scale using percent.
 - Without clicking on the histogram, hover the cursor over the tallest bar in the histogram. After a few seconds, you should see the following information displayed on the graph “Value = 21.8254 Bin = 157.5, 172.5”. This means that the bin width covers the weights 157.5 Lbs to 172.5 Lbs and the height of the bar represents 21.8254% of data lies within these weights.
- 4) **QUESTION 1: What percent of the data lies within 187.5 and 202.5 lbs?**
- 5) **Copy this histogram for weight into a Word document.** To do this, click on the dropdown in the upper left of the histogram and select “Copy Graph”. Then right click on your word document and select paste.
- 6) Repeat the above steps to create a histogram for height. **Copy this histogram into a Word document** as well.
- 7) **QUESTION 2: What percent of the data lies within 66.25 and 68.75 inches?**
- 8) From the top menu select “Stat >Basic Statistics >Display Descriptive Statistics”
 - Now click on Weight and click the “Select” button then the “OK” button.

What Minitab has produced is a collection of statistics about the weights in the data set. Among the things it has computed are N (the number of data points), the Mean (Average), the Median, and the StDev (Standard Deviation). **Copy these descriptive statistics into your Word document** (being sure to include the label explaining what the numbers represent – the descriptive statistics for weight).
- 9) Use a similar command to obtain the descriptive statistics for height. **Paste these statistics into your Word document as well.**
- 10) Each of the two histograms appears to have an “outlier” point. One person on the “Weight” graph is particularly heavy. One on the “Height” graph is particularly short. We would like to see what effect removing these two people has on the histograms and on the descriptive statistics.
 - In the worksheet window, look at row 42 of the data. Notice that this person is particularly short.
 - Click on the row number 42. From the upper toolbar, click “Edit > Delete Cells”. Notice the result.

- Next select row 39. The person corresponding to this row is unusually heavy. Delete this row as well, just as you did the other.

11) Produce two new histograms together with descriptive statistics for the modified data sets. We want some way to distinguish the modified data from the original data. When you have produced the histogram, click on the title (“Histogram of Weight” for example) and change it appropriately (for example, replace it with “Histogram of Weight with Outliers Removed”). Be sure to **do this for both new histograms and both new sets of descriptive statistics.**

12) QUESTION 3: What effect did removing the outlier have on the average Weight of the data set? Why? What effect did it have on the Standard Deviation of the Weights? Why? Which was more affected by the removal of an outlier, the Mean (Average) or the Median? Why?

13) QUESTION 4: What effect did removing the outlier have on the average Height of the data set? Why? What effect did it have on the Standard Deviation of the Heights? Why? Which was more affected by the removal of an outlier, the Mean (Average) or the Median? Why?

Report Guidelines:

Begin your lab report by describing what data sets were analyzed. Include the reference to the web page where you obtained the data. Include all four histograms in your report together with the descriptive statistics for the variable the histogram describes.

On your two histograms with the outliers removed, **shade the regions corresponding to the points that are within 1SD of average.** Note that if you have done this correctly, the region should be roughly 68%, or about 2/3 of the data.

Be sure to **answer all questions** asked in the lab in the report.

Make the report as organized, as clear as possible, and stapled if exceeding 1 page.