**Exam II**

*MAT 135*

*Introduction to Statistics*

*Total Points :*

Instructions: The first part of the exam is hand written. You should be able to fit all your work on this exam. Show all of your work for partial credit. After you finish the first part, begin work on the Excel portion of the exam.

For the last three parts, you will be provided with a raw data set in the form of an Excel spreadsheet and asked to perform some statistical analysis. Please label all cells you use for any calculations or short answers. Label and format all graphs. When you are done, upload the spreadsheet to the Exam II drop box on *D2L*.

**Part 1: Hand Written Calculations**

Instructions: For the first two problems, you are given the following dataset that shows the number of minutes a student named Snuz Malsado was late for class each class day of *MAT 135* the first month of the semester,

{ *1.3 min, 2.1 min, 0.8 min, 1.5 min, 2.1 min, 0.75 min, 1.7 min, 6.2 min* }

1. (5 pts) By hand, using only a calculator as necessary, calculate:

i. The sample mean.

ii. The median.

iii. The deviation of each observation from the sample mean.

iv. The sample variance.

v. The standard deviation.

2. (3 pts) Using the results from #1, calculate the z-score of each observation in this dataset.

i. Are there any potential outliers in this dataset?

**Part 2: Descriptive Statistics**

Instructions: The American Racing Pigeon Union holds annual pigeon races and records the race times of all participating pigeons. You will now perform some analysis on this dataset. Go to *D2L > Course Materials > Content > Exams > Exam II* to find the Excel spreadsheet entitled "*Pigeon\_Racing\_YOURNAME.xls*". Replace the '*YOURNAME'* with your own name and save it to your computer. All speeds are measured in miles per hour (*seriously*). **Use the pigeon speed as the variable of interest; you may ignore all the other columns!**

You will need this data set for the rest of the problems on the exam. *Make sure to save frequently so you do not lose any of your work by accident!* Any time you finish a problem, click the save button, just to be sure!

3. (3 pts) Create a relative frequency distribution. Use 7 evenly spaced (*equal class width*) classes. Start the first class at 76 (i.e. *round the minimum observation down to a whole number*).

i. Create a histogram for the relative frequency distribution.

4. (10 pts) Using the appropriate Excel function, calculate the following descriptive statistics for the data set. Save your answer in a cell and clearly label it:

a. The Mean

b. The Median

c. The Minimum

d. The Maximum

e. The Range

f. The 25th Percentile

g. The 75th Percentile

h. The Interquartile Range

i. The Sample Variance

j. The Standard Deviation

5. (2 pts) Refer back to the answers for *3a, 4a* and *4b.* Is this distribution skewed left, skewed right or approximately normal?

i. Based on the skew, where will we most likely find outlying observations, in the left or right tail of the distribution?

**Part 3: Box Plot**

6. (5 pts) Using the appropriate results from #4, construct a box plot for the pigeon racing speeds. Make sure to calculate the *differences* before stacking your bar chart.

i. Based on your box plot, around where are most of the pigeon speeds clustered?

**Part 4: Outliers & Statistical Theorems**

7. (2 pts) Find the upper and lower bound for outliers with 95% certainty using the Empirical Rule. Save your results in a cell.

8. (2 pts) Create a column next to the *Speed* column and entitle it *Z-Scores*. Use this column to calculate the z-score of each pigeon speed.

i. Based on the results of your calculations, identify any outliers in this distribution. Highlight or label these observations in some way.

ii. Verify the actual value of observations with outlying z-scores also satisfy the Empirical Rule conditions for outliers calculated in part 7.

9. (2 pts) Use the Empirical Rule to estimate the 97.5th percentile of this distribution. Save your answer in a cell and label it.

i. Calculate the actual 97.5th percentile of this distribution.

ii. Refer back to your answer to #5. Is the Empirical Rule a good approximation for this distribution? Why or why not?