Lab 2 (40 points) - Describing Distributions with Graphs and Numbers

Objectives: Numerical Summaries, Histograms and Boxplots

Remember to use the cleaned data set that you generated in Lab 1.

A (20 pts) Actual Elapsed time (Data Set: airline2008NovS cleaned) A flight analyst is interested in the graphical and numeric summaries for actual elapsed time (minutes) of the flights (ActualElapsedTime).

1. (5 points) Code
2. (2 points) Find the five-number summary for these data.
3. (3 points) Make a modified boxplot. Is the distribution symmetrical, left or right-skewed?
4. (3 points) Make a histogram of the data. Is the distribution symmetrical, left, or right skewed?
5. (2 points) Find the sample mean, $\bar{x}$, and sample standard deviation, s. Is the median close to the mean? Please explain your answer.
6. (2 points) If you had only one number to estimate the actual elapsed time of airline flights in November based on this dataset, what would it be? Please explain your answer.

B (20 pts). The time of Taxi in (Data Set: airline2008NovS cleaned) A flight analyst is interested in the graphical and numeric summaries of the time (minutes) that it takes for the flights to arrive at the gate after they land (TaxiIn).

1. (5 points) Code
2. (2 points) Find the five-number summary.
3. (5 points) Calculate the 1.5 IQR upper and lower limits for the outliers. Are there any outliers according to 1.5 IQR rule? Show the details of the calculation.
4. (3 points) Make a modified boxplot. Is the distribution symmetrical, left, or right skewed?
5. (5 points) Make a histogram of the data. Do you agree with the rule's suggestions about the outliers? Please explain your answer.