Week 2 -2 Assignment – 35 Points

**Part A (15 Points)**

**You need to do 5.1 to be able do the rest. These are part of the same program. Do not code them separately**

1. Do the exercise in page 135, figure 5.2. Explain the outcome
2. Do the exercise in page 145, figure 5.5. Explain the outcome
3. Do the exercise in page 152, figure 5.7. Explain the outcome

**Part B (20 Points)**

Load usedcar.csv in RStudio. Perform the following initial data exploration and preparation procedures.

1. Get summary statistics of the dataset
2. Get summary statistics of the price and mileage columns
3. Get the .25, .50, and .75 quantiles of the Price column
4. Get the 0.0, 0.25, 0.50, 0.075, and 0.100 quantiles of the Price Column
5. Check the dataset for any 0, missing value, and n/a value. Remove if there is any
6. Get the frequency of each model, color, and transmission
7. Check for outliners. Can you explain each outlier (if any)
8. Create the histogram for Price and Mileage columns. Interpret the result.
9. Add another column containing the categorical data of the Price column.

For all the above work, you must provide a **short** explanation, description, or interpretation. Just R code is not enough