**CSC583/783 Data Visualization**

**Homework 5**

**Due: Saturday, May 9.**

**(30 points)**

**Please answer the following questions on the usedcars.csv dataset which contains actual data about used cars advertised for sale. Use R and RStudio. For full credits, you should contain screenshots (outputs only) from RStudio.**

1. (3 points) Create a variable to store the usedcars.csv file.
2. (3 points) Show the structure of used car data.
3. (3 points) Summarize two features, price and mileage simultaneously.
4. (3 points) Calculate the median price.
5. (3 points) Calculate the 10th and 90th percentiles of a price.
6. (3 points) Draw a boxplot of used car price and mileage, respectively. Add a title and a label to each box plot.
7. (3 points) Draw a histogram of used car prices and mileage, respectively. Add a title and a label to each histogram.
8. (3 points) What is variance and standard deviation on price and mileage variable?
9. (3 points) Run table() on year, model, and color variable, respectively. Interpret each result.
10. (3 points) Create a new variable (i.e., usedcars$conservative) which indicate conservative colors as TRUE or FALSE. Divide nine colors into two groups (label TRUE for conservative colors: Black, Gray, Silver, White, and label FALSE for nonconservative colors: Blue, Gold, Green, Red, and Yellow) and use %in% operator.