Data Mining Lab (CSEN 3155)
Lab Assignments - Day 4

1. **Read the file "GrocBinary.csv" and answer the following questions:**

   (a) **Identify all frequent item set with support at least 30%.**
   (b) **Find out all association rules with support 40% and confidence 60%.**
   (c) **Find out all the with support 30%, confidence 70% and lift > 1**

2. **Read file "auto-mpg.csv" into a data frame named "auto". Display the file. And also display the number of rows. columns and summary of the dataset. What are the column names?**
3. **Display the content of cylinders column.**

R Factors

Factors in R Programming Language are data structures that are implemented categorize the data or represent categorical data and store it on multiple levels. They can be stored as integers with a corresponding label to every unique integer. Try the following use of Factor

```
auto$cylinders <- factor(auto$cylinders,
                         levels = c(3,4,5,6,8),
                         labels = c("3cyl", "4cyl", "5cyl", "6cyl", "8cyl"))
```

4. **Display how the content of cylinders column has changed.**
5. **Draw a histogram of acceleration with axis labels and titles.**
6. **Draw a histogram of mpg with axis labels and titles.**
7. **Draw a barplot of mpg with axis labels and titles.**
8. **Display the frequency count of cylinders and plot the frequency count.**
9. **Try these two box plots and interpret**

```
boxplot(auto$mpg, xlab = "Miles per gallon",
      col = "lightgray")

boxplot(mpg ~ cylinders, data = auto)
```

10. **What happens if you try the following**

```
pairs(~mpg +displacement, data=auto)

pairs(~mpg +displacement+horsepower, data=auto)
```