CADS Project 1: General R Skills

This project will use the dataset based on the survey that was conducted in our class. 15 questions-survey was administered.

Using the R-Studio, import the Class Survey data file:"ClassSurveyData.csv"

```
Install.packages()
library(dplyr)
library(ggplot2)
# Make sure the data file is in the proper folder. Use getwd(). Then you can easily pull in the file.
```

Answer the following questions:

A.

1. Obtain a summary of the data.
2. How many Females are in the class?
3. How many Males are in the class?
4. How many Freshmen are in the class?
5. How many Females are older than 25?
6. What is the highest GPA reported? Hint: Check for missing data first.
7. Find the percentile for Exam1. Hint: The percentiles are 10,20,30,.....90%
8. How many different majors are reported?
9. Find the average age of the Males
10. Find the average age of the Females
11. Create an histogram for the Age

B.

1. Find the mean of the Homework grades.
2. Create a new column for mean_Homework and append to the data frame.
3. Give a new name to the updated data frame: New_ClassSurveyData.csv
2. Create a Scatter plot of mean_Homework and Exam1 grades.
3. Is there a correlation (positive/negative) between mean Homework and Exam1 grade?
4. Calculate the strength by computing r**2.
5. Create a regression line for y (Exam1 grade) and x (mean_Homework grade).
6. For the equation, what is the slope and intercept?
7. Using the equation, can you predict the Exam score for someone who averages 75% on homework?