EXPLORATORY DATA ANALYSIS

1. Create a csv file with the following columns. Name, Age, Gender, Height, Weight. Populate the csv with 50 values.
    a. Calculate the mean age of the population
    b. Calculate the median age of the population
    c. Calculate the minimum and maximum value in Age
    d. Calculate the dispersion of Age of your population set
    e. What is the 75% percentile of the Age
    f. What is the standard deviation in weight and height ?
    g. Develop a linear regression model for height and weight(dependent) and get the coefficients. Perform the significance test for linear regression at a significance level of 0.05.
    h. Given a height predict the weight value.
    i. Plot the regression model.

2. You are entitled with the work of predicting the likelihood of occurrence of heart attack in people. You have collected a dataset which has the following features. Age, Gender, Bp, Blood Sugar level, Cholesterol and whether they have suffered a heart attack. Read the dataset and use R to create a logistic regression model for this dataset with the independent variables being age, BP, Blood Sugar level, Cholesterol level and the dependent variable being whether they have had a heart attack.

    a. Also calculate the median age of persons who have suffered heart attack.
    b. Range of Blood Sugar and cholesterol level