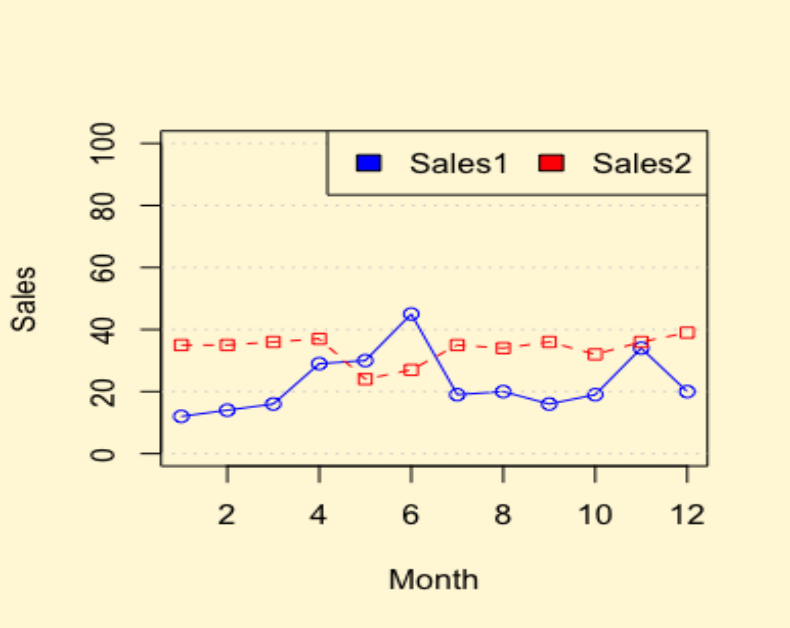
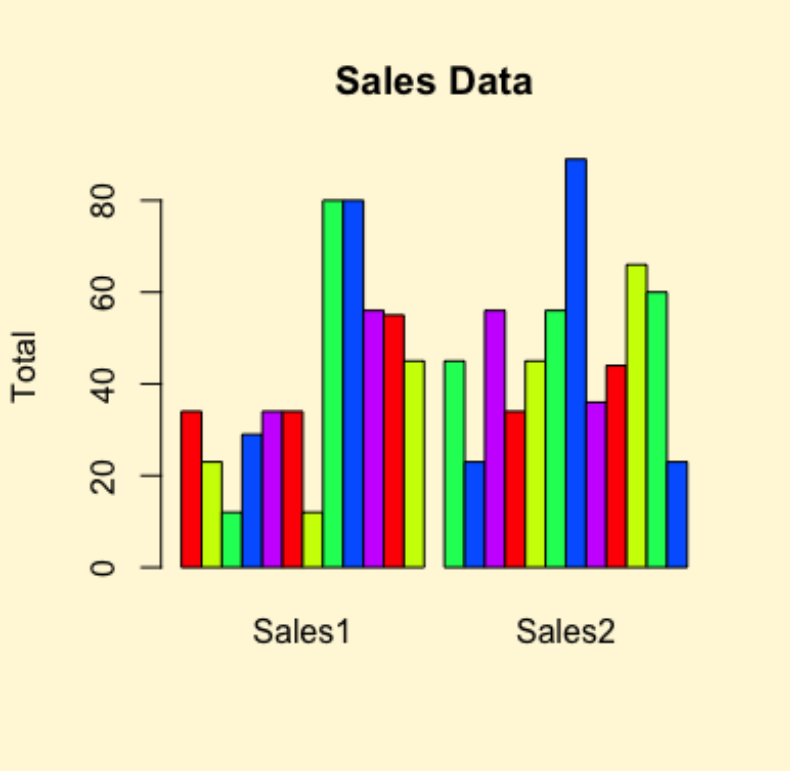
Lab1 Part 1: Exploratory Data Analysis & R Basics:

Problem 1: Output



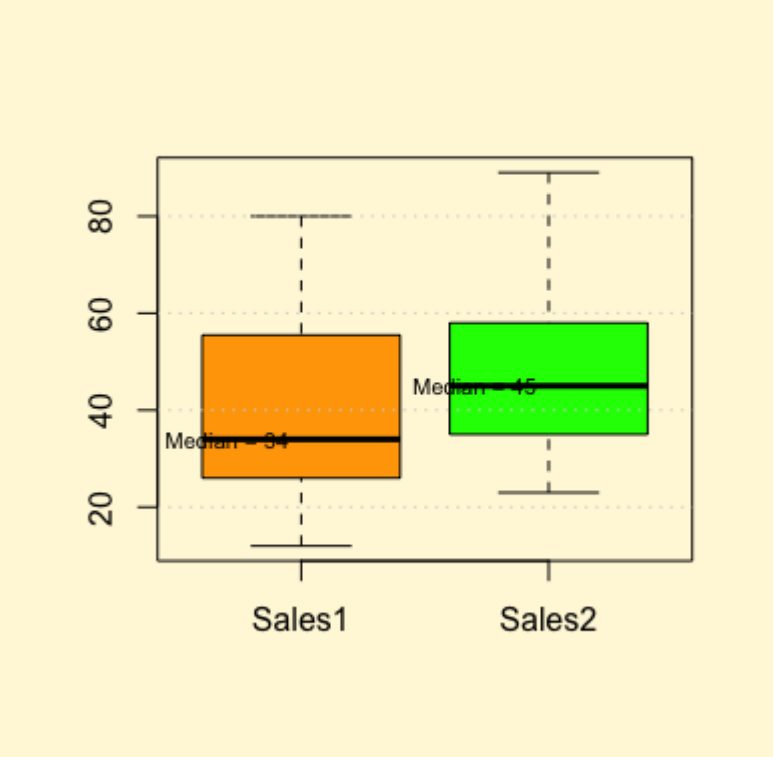
Running the code multiple times, generate different line charts for Sales2 data. This happens because the sales 2 data frame is set to generate random numbers between the ranges of 12 to 34.

Problem 2: Output



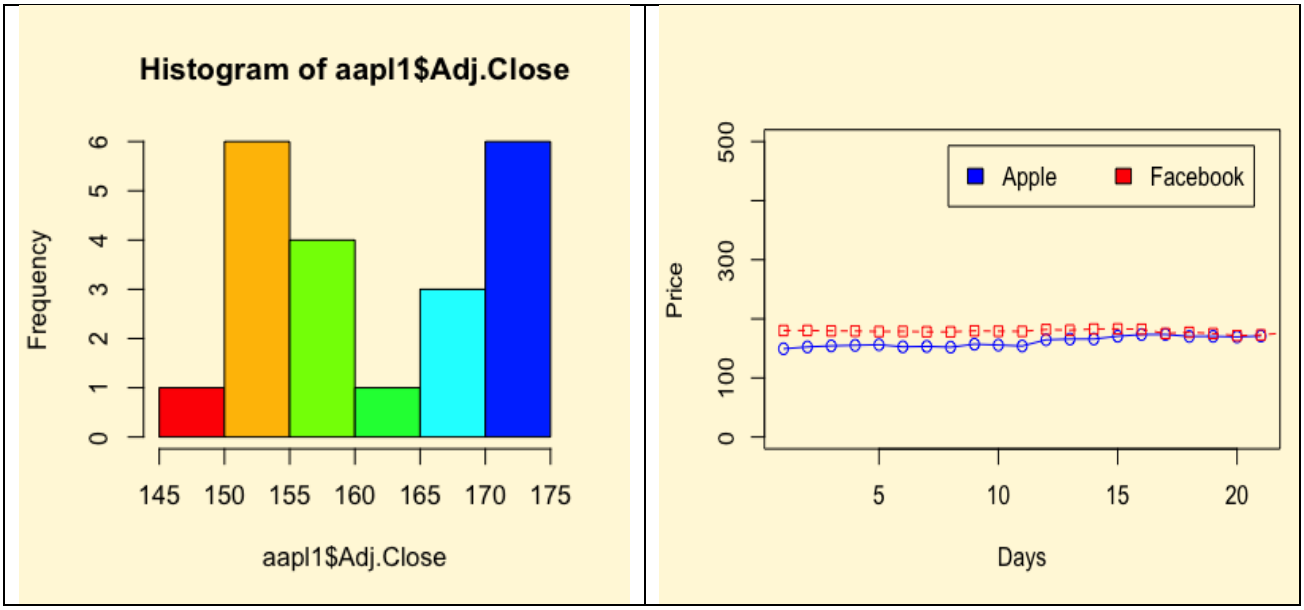
By Plotting the Sales1 and Sales2 data as a histogram, we can easily compare the performance of sales1 and sales2 for every individual month of a year. The color has been set to rainbow which produces such vibrant colors in the histogram.

Problem 3: Output



The graph visualization we can say that the range of sales1 is greater than range of sales2. The difference between upper quartile and lower quartile for sales 1 is more, indicating that 50% of the sales in Sales1 data have more variation than the sales data for sales2. The amount of data which falls under the lower quartile is more for Sales1. Hence it can be interpreted that sales 2 has a better performance than Sales1.

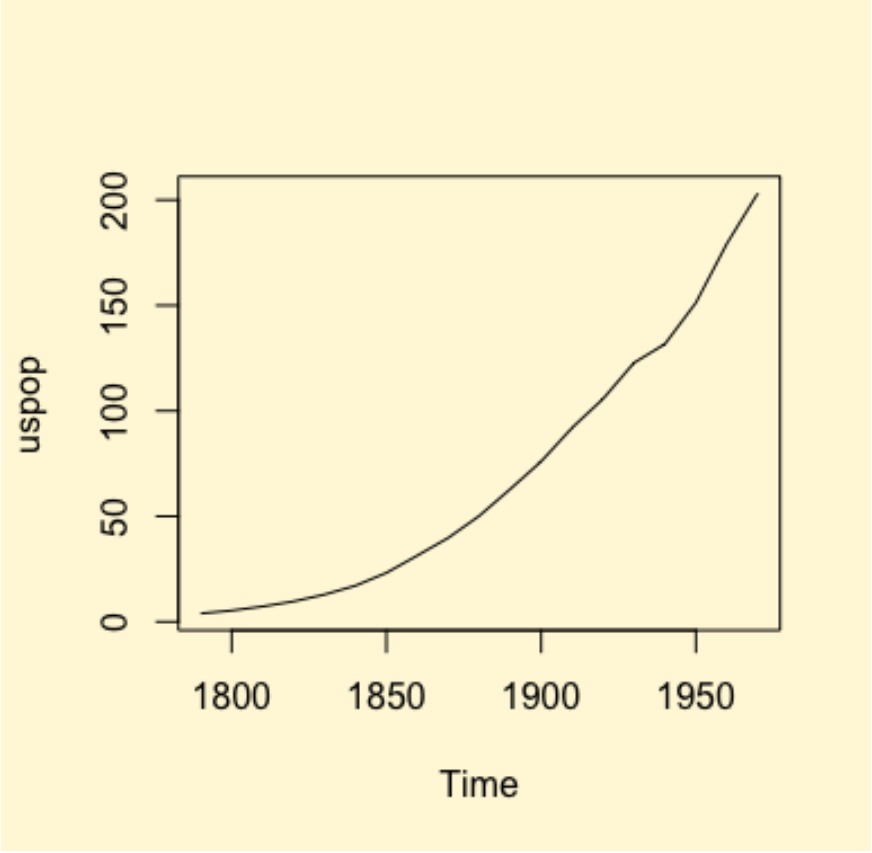
Problem 4: Output



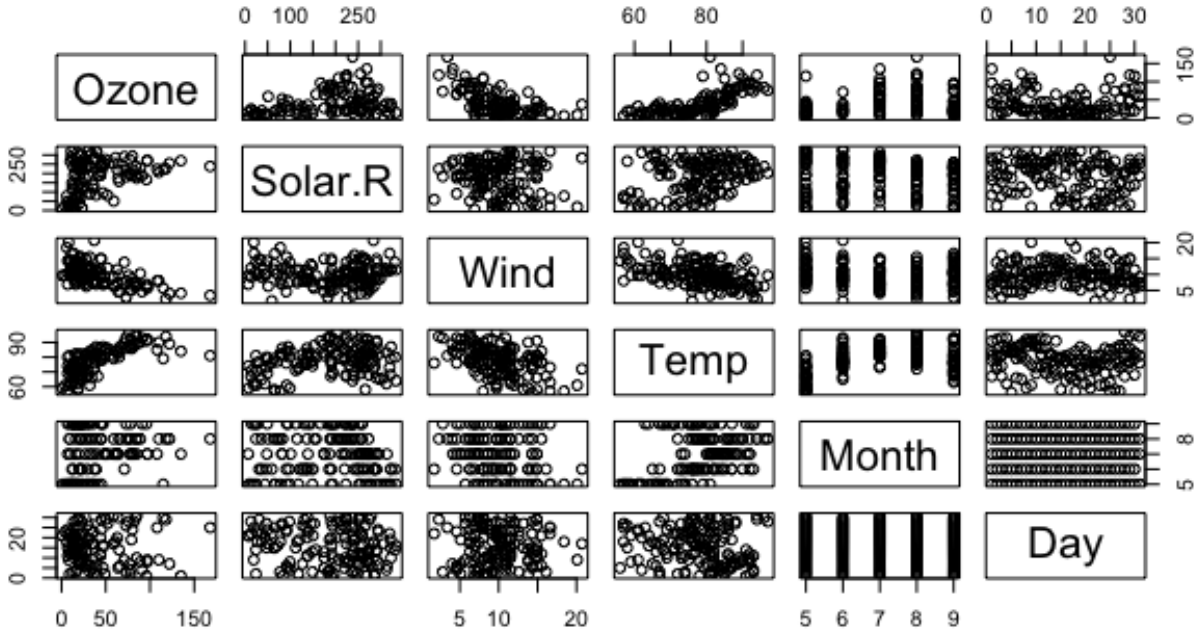
It is clearly evident from second graph that stock prices of Facebook where greater than Apple's prices in the initial period. As timeline progresses, the Apple price increases and catches up to Facebook stock prices.

Problem 5:

Output1: - Below graph shows the US population dataset present in R library in relation to timeline of years.

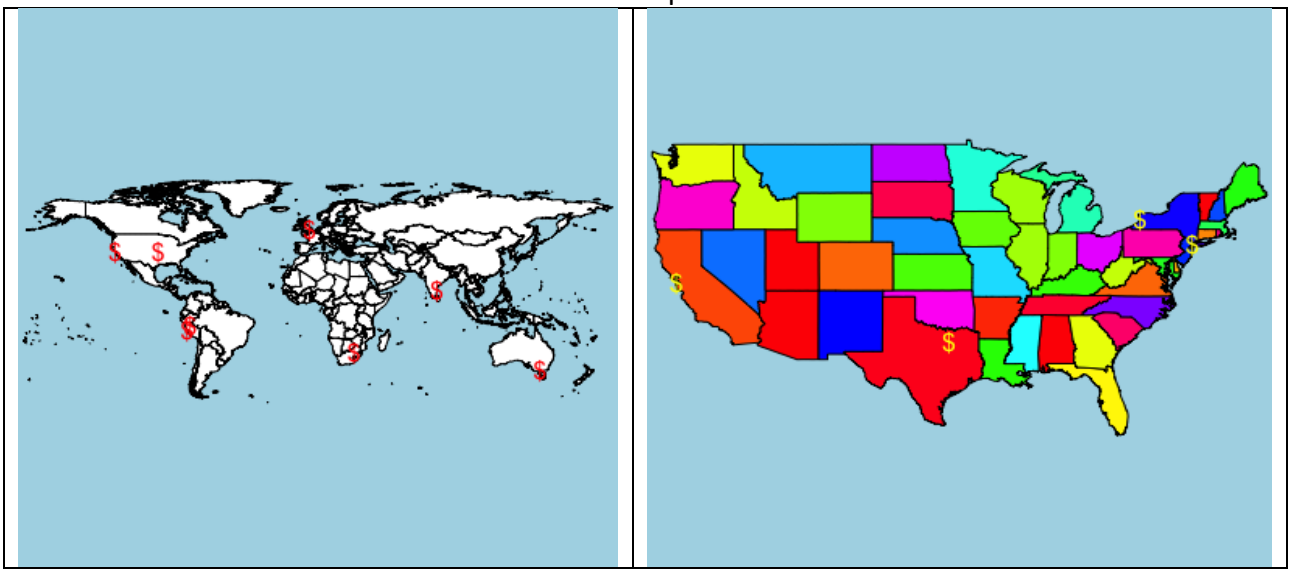


Output 2: - Following figure shows the plot of airquality library which is present in R. The library contains data of six parameters.



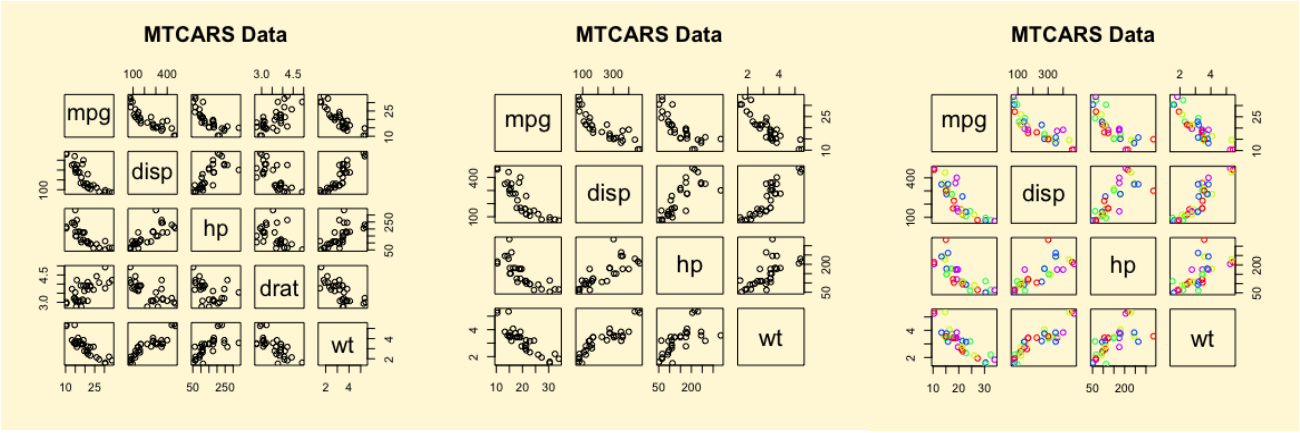
Problem 6:

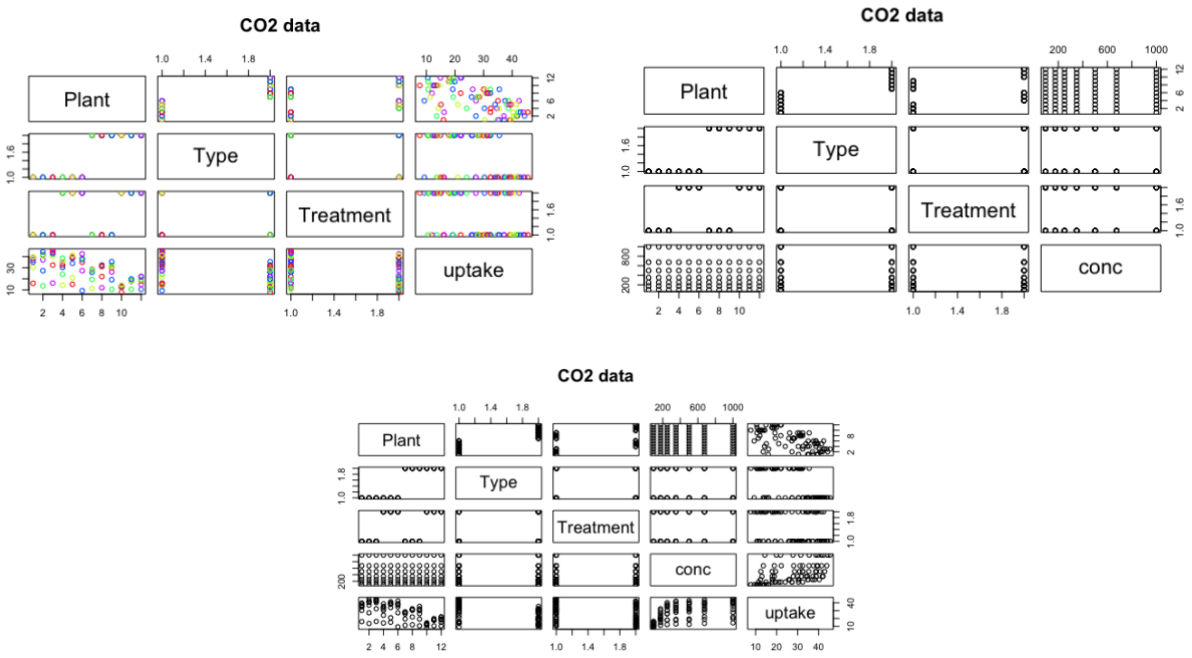
Output



For the above plots, Google API was used to find the geolocation of specific places from a user defined data frame and plot using the maps library present in R.

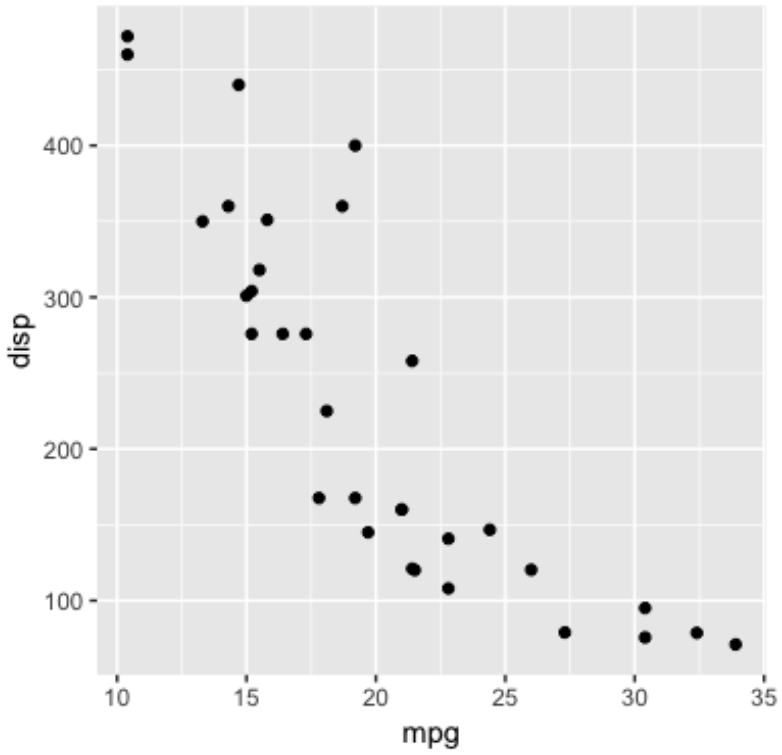
Problem 7: Output





CO2 is the rich dataset present in R which has used in problem 7. The second and third plot related to CO2 in the document contain 5 dependent variables and 4 dependent variables. The first plot is a color mapping of the scatter plot with 4 dependent variables.

Problem 8: Output



The above plot has been rendered from data present in the mtcars R library.