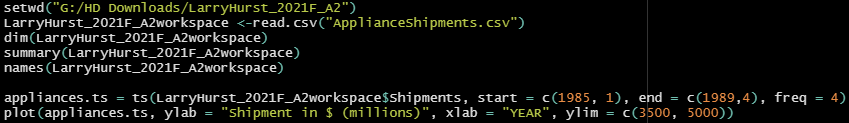
**BI Analyzation 1**

**1** **Shipments of Household Appliances: Line Graphs**

The file ApplianceShipments.csv contains the series of quarterly shipments (in millions of dollars) of US household appliances between 1985 and 1989.

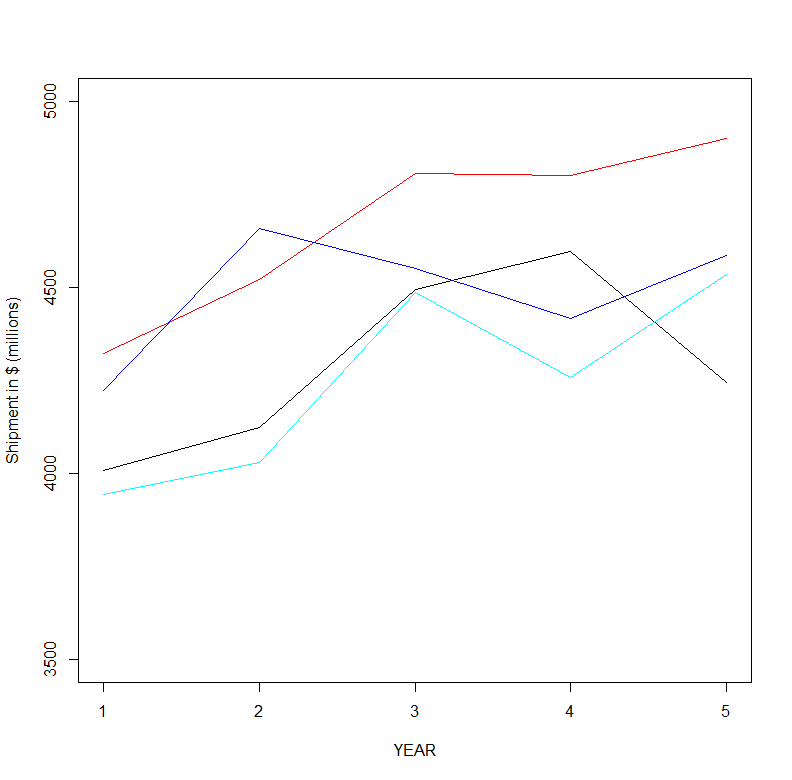
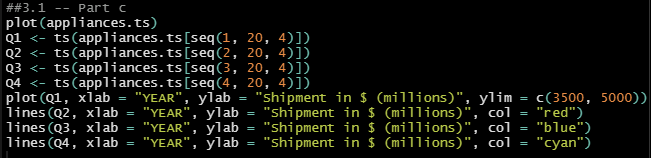
**(a) Create a well-formatted time plot of the data using R.**



**(b) Does there appear to be a quarterly pattern? For a closer view of the patterns, zoom in to the range of 3500–5000 on the y-axis.**

After specifically viewing the range between $3,500,000 to $5,000,000, shipments do appear to increase from the fourth quarter of one year to the first quarter of the next year. In the second and third quarters, almost every single year seems to have a decline between those periods.

**(c) Using R, create one chart with four separate lines, one line for each of Q1, Q2, Q3, and Q4. In R, this can be achieved by generating a data.frame for each quarter Q1, Q2, Q3, Q4, and then plotting them as separate series on the line graph. Zoom in to the range of 3500–5000 on the y-axis. Does there appear to be a difference between quarters?**

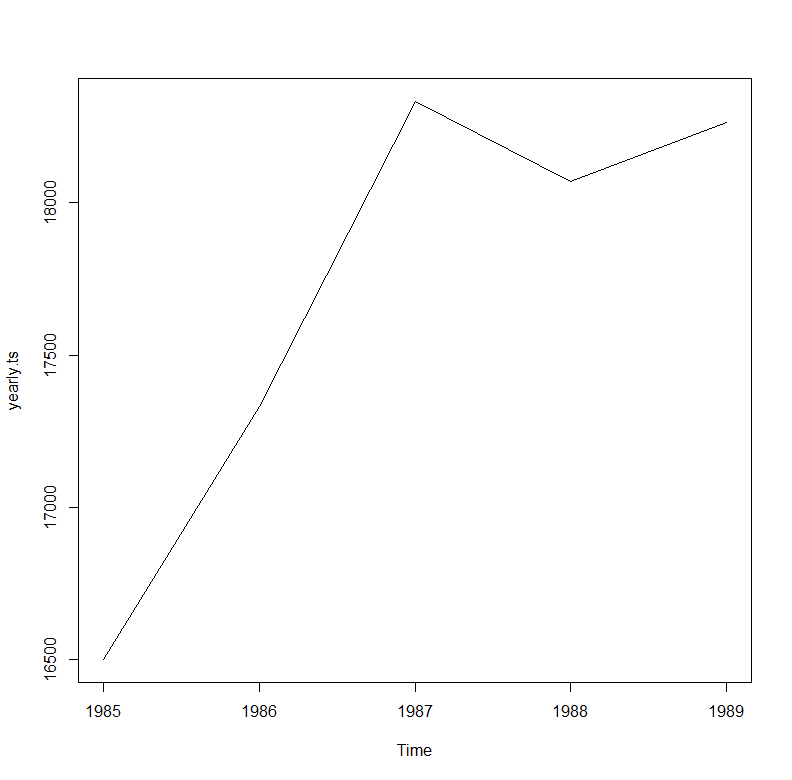


Judging by the chart, there seems to be a correlation between each quarter of each year. The only difference appears to be year 1 (black line). Q2, Q3, and Q4 all appear to have a slight increase from year 1 to year 2 and year 2 to year 3. There is a decline between year 3 and year 4 for Q2, Q3, and Q4; however, they all experience an increase from year 4 to year 5. Their line behaviors are all very similar in nature, despite the different amounts (in millions of dollars).

**(d) Using R, create a line graph of the series at a yearly aggregated level (i.e., the total**

**shipments in each year).**

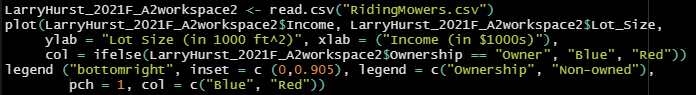
****

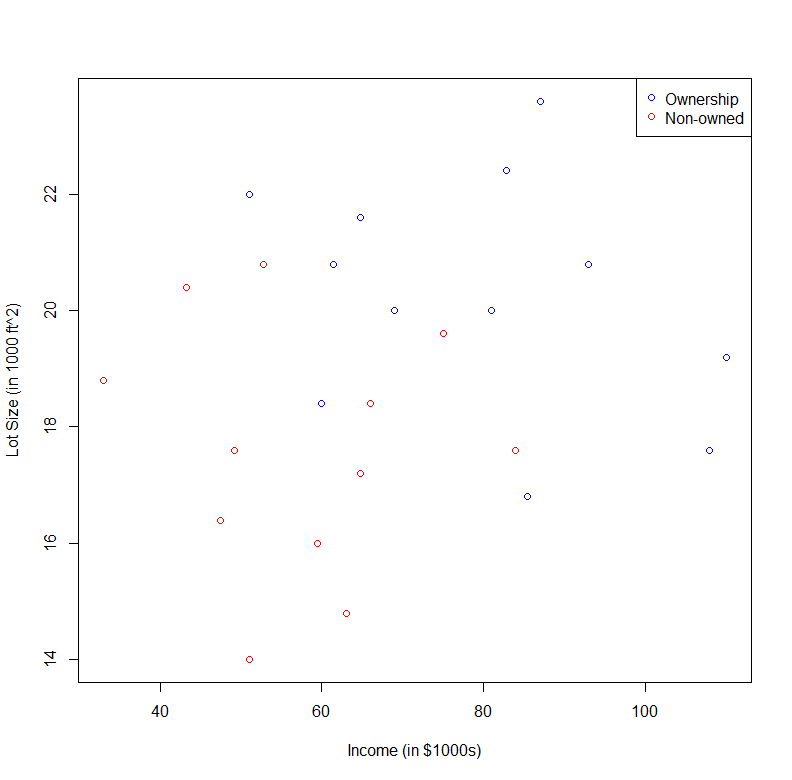
****

**2. Sales of Riding Mowers: Scatter Plots**

A company that manufactures riding mowers wants to identify the best sales prospects for an intensive sales campaign. In particular, the manufacturer is interested in classifying households as prospective owners or nonowners on the basis of Income (in $1000s) and Lot Size (in 1000 ft^2). The marketing expert looked at a random sample of 24 households, given in the file RidingMowers.csv.

**(a) Using R, create a scatter plot of Lot Size vs. Income, color-coded by the outcome variable owner/non-owner. Make sure to obtain a well-formatted plot (create legible labels and a legend, etc.).**

****

****

**3. Laptop Sales at a London Computer Chain: Bar Charts and Boxplots.** The file

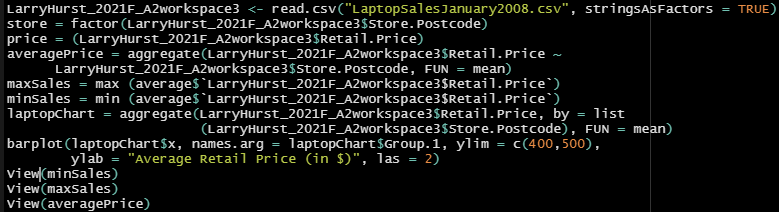
LaptopSalesJanuary2008.csv contains data for all sales of laptops at a computer chain in

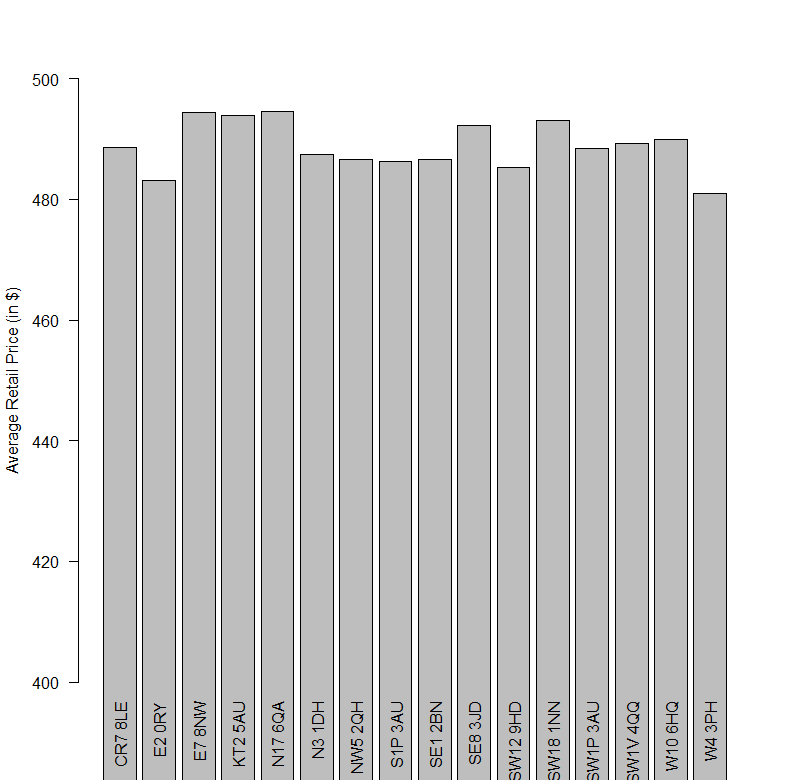
London in January 2008. This is a subset of the full dataset that includes data for the entire year.

**(a) Create a bar chart, showing the average retail price by store. Which store has the highest average? Which has the lowest?**

***Highest* Average Price**: Store **N176QA** (494.6341)

***Lowest* Average Price**: Store **W43PH** (481.0063)





**(b) To better compare retail prices across stores, create side-by-side boxplots of retail price by store. Now compare the prices in the two stores from (a). Does there seem to be a difference between their price distributions?**

There doesn’t seem to be a huge difference in the price distributions between store N176QA and store W43PH. The average price appears to be slightly higher in store N176QA. That being said, their minimum price and maximum price in each store appear to be similar judging by the boxplot below.



