Assignment By: Salwa Niaz (sxp240086)

Business Analytics with R

Assignment 1

Answer 1

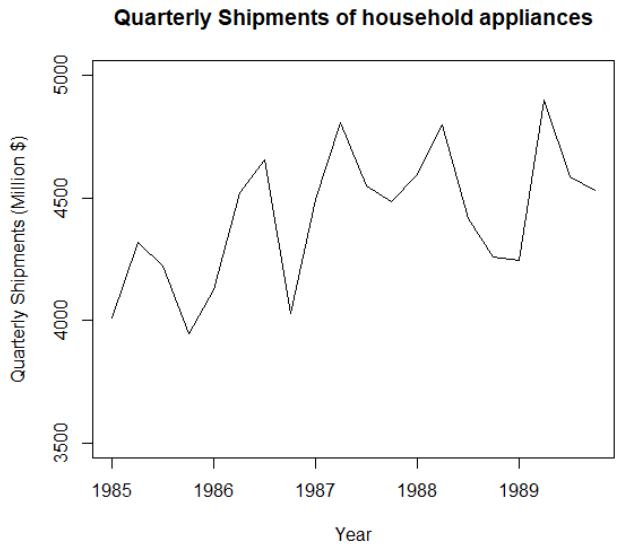
a. **Deciding whether to issue a loan to an applicant based on demographic and financial data (with reference to a database of similar data on prior customers):**
   Supervised Learning
b. **In an online bookstore, making recommendations to customers concerning additional items to buy based on the buying patterns in prior transactions:**
   Unsupervised Learning
c. **Identifying a network data packet as dangerous (virus, hacker attack) based on comparison to other packets whose threat status is known:**
   Supervised Learning
d. **Identifying segments of similar customers:**
   Unsupervised Learning
e. **Predicting whether a company will go bankrupt based on comparing its financial data to those of similar bankrupt and nonbankrupt firms:**
   Supervised Learning

Answer 2

a. Predicting the price of automobiles based on the features like make, engine-type, number of doors, fuel-type, etc.
Regression
b. Predicting the income of people based on the features like occupation, age, gender, education level, marital status, etc.
Regression
c. Predicting whether income is above or below 50K based on the features like occupation, age, gender, education level, marital status, etc.
Classification
d. Predicting the average life expectancy of different countries based on their GDP, population, schooling, and health-related metrics.
Regression
e. Predicting whether a customer would cancel their hotel booking or not based on the features like when the reservation was made, how many rooms were reserved, how the rooms were reserved, etc.
Classification

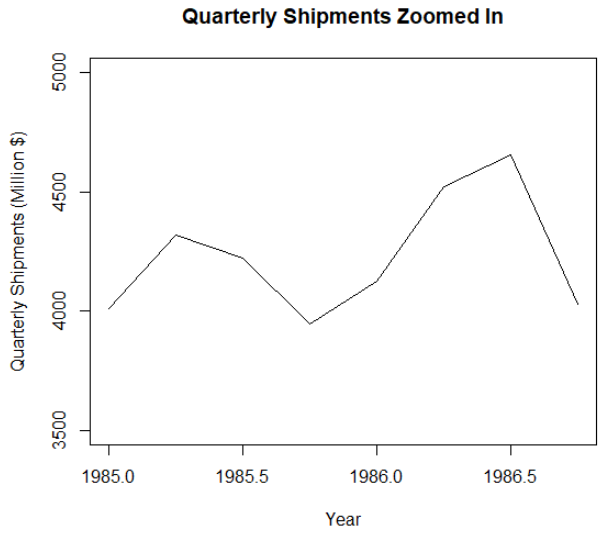
Answer 3

(a)



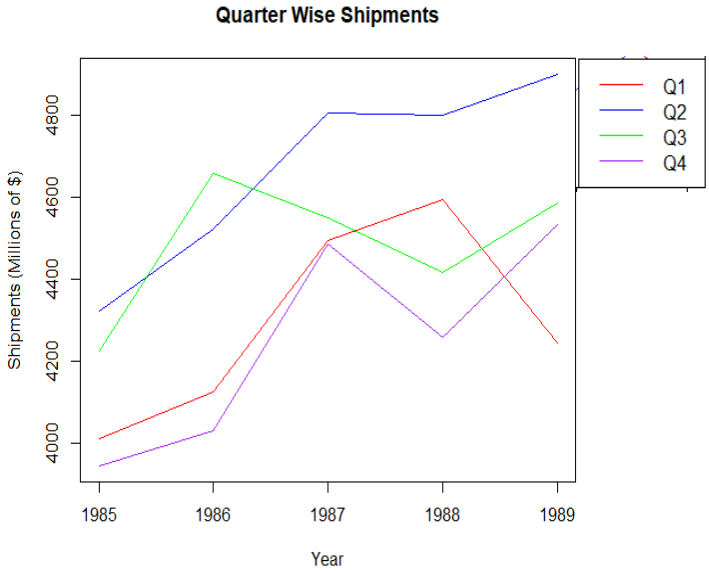
In this graph, pattern isn't quite visible. It does not seem like shipments have a quarterly pattern.

(b)



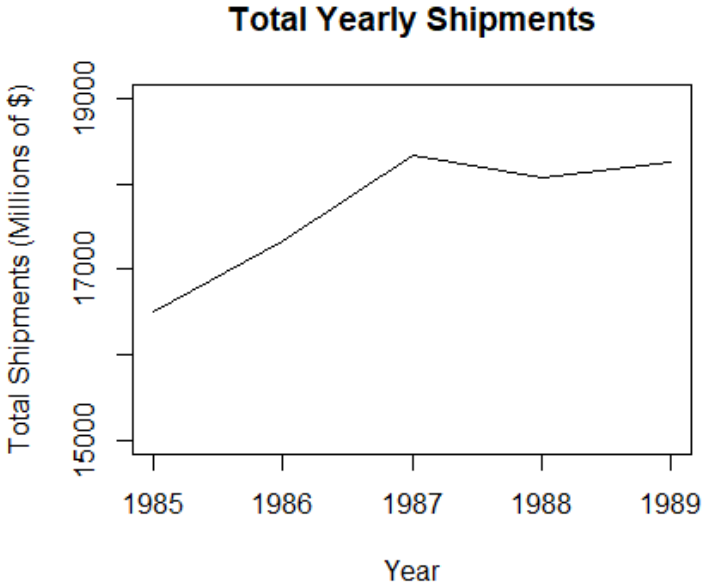
In this graph, it is a bit clearer to understand quarterly shipments. In 1985, shipments increase from Q1 to Q2 then slightly decrease in Q3 and sharply fall in Q4. In 1986, Q1 sees an increase in shipments from Q4 of the previous year, they increase sharply in Q2 and then slightly more in Q3. Shipments decrease in Q4 of 1986 also.

(c)

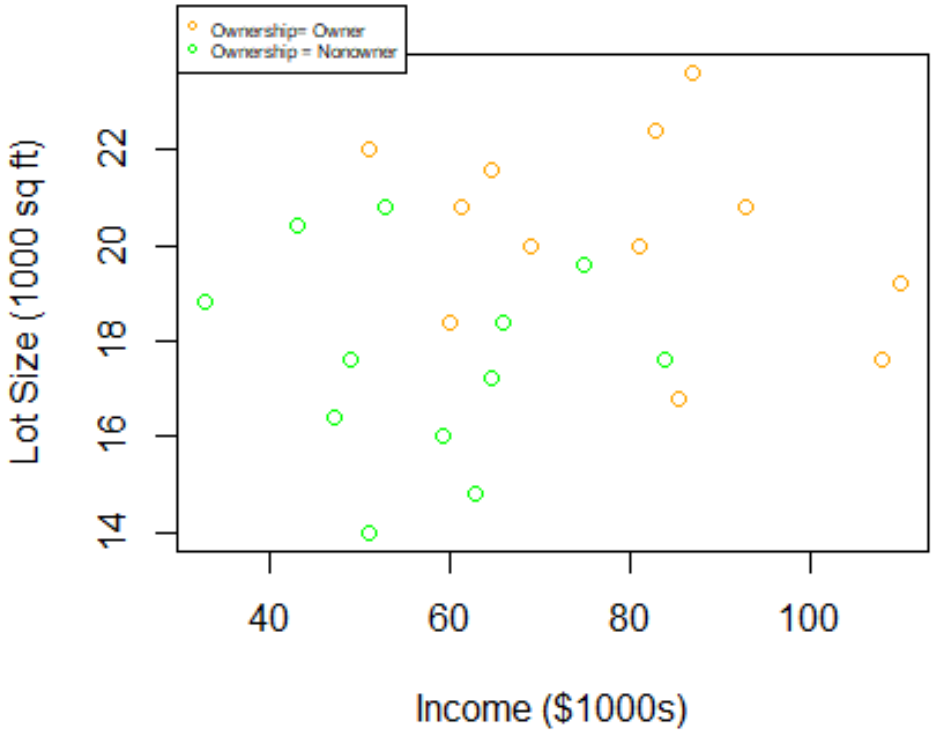


Yes, there is a difference between quarters. Generally, shipments in Q1, Q2 and Q4 have increased except drops in 1989 for Q1 and in 1988 for both Q2 and Q4. Q3 has shown an unpredictable pattern.

(d)



Answer 4



Answer 5

(a)

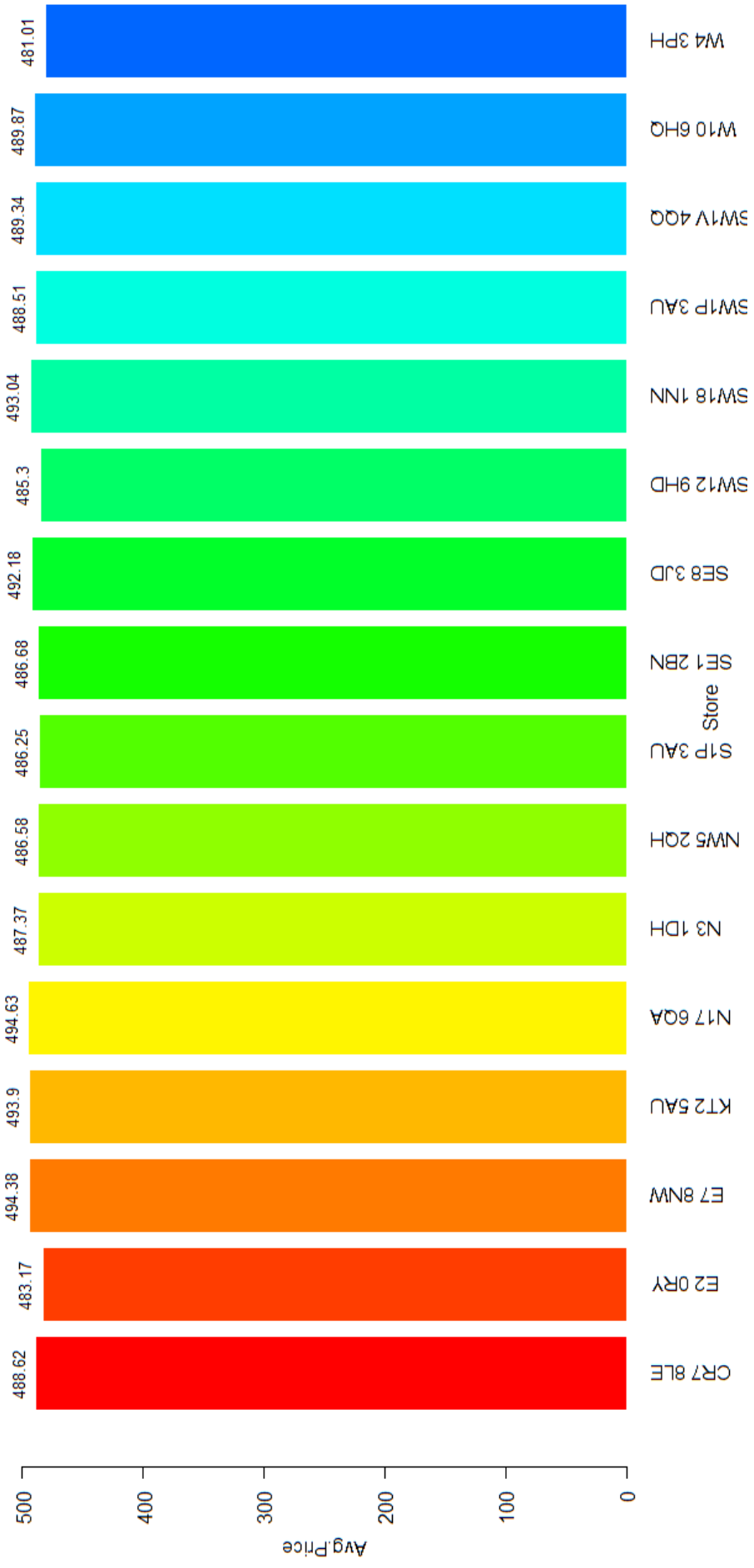
Highest average: Store N176QA with avg price of 494.63 (yellow bar)

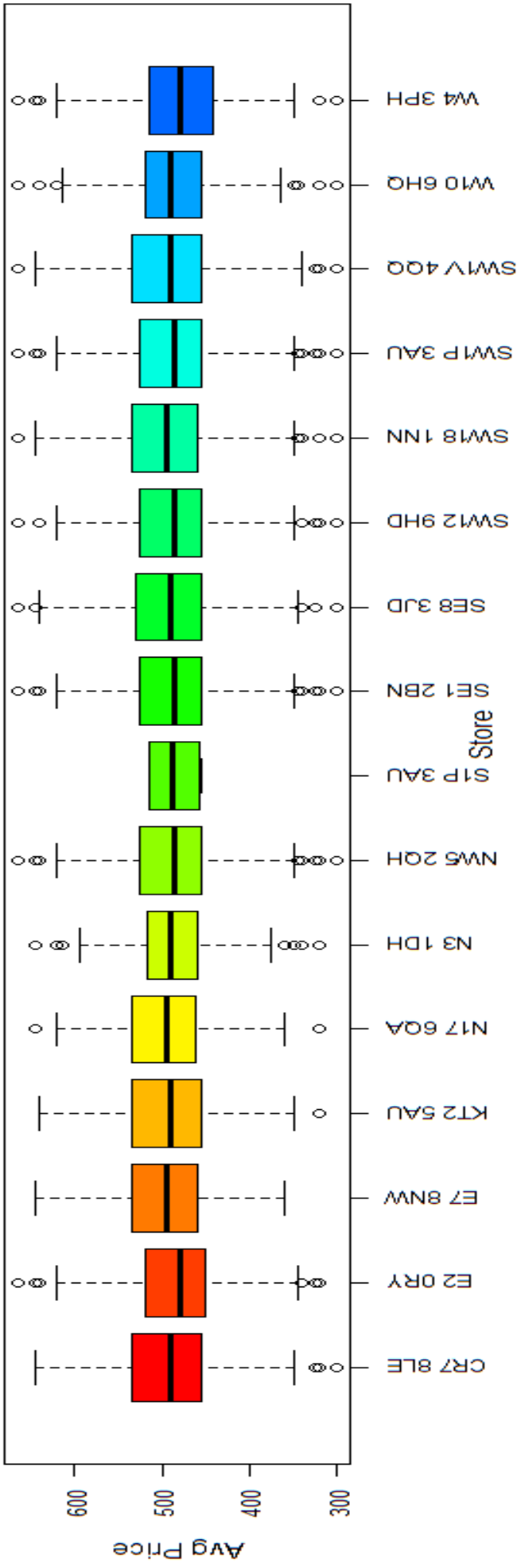
Lowest average: Store W43PH with avg price of 481.01 (dark blue – last bar)

(b)

The two shops have a similar price distribution. They are both almost symmetrical. The shop N176QA has a higher Q1, Q2, Q3 and minimum value than the other shop. The maximum values for both are almost equal. The shop W43PH has more outliers than N176QA.

(graphs on following pages)


Avg.Price
0
100
200
300
400
500
Store
CR7 8LE
E2 0RY
E7 8NW
KT2 5AU
N17 6QA
N3 1DH
NW5 2QH
S1P 3AU
SE1 2BN
SE8 3JD
SW12 9HD
SW18 1NN
SW1P 3AU
SW1V 4QQ
W10 6HQ
W4 3PH
488.62
483.17
494.38
493.9
494.63
487.37
486.58
486.25
486.68
492.18
485.3
493.04
488.51
489.34
489.87
481.01


Avg Price
600
500
400
300
Store
CR7 8LE
E2 0RY
E7 8NW
KT2 5AU
N17 6QA
N3 1DH
NW5 2QH
S1P 3AU
SE1 2BN
SE8 3UD
SW12 9HD
SW18 1NN
SW1P 3AU
SW1V 4QQ
W10 6HQ
W4 3PH