BUAN 448, HW3

2022-09-09

Question 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. Answer the following questions using the dataset:

- Item a) Create a well-formatted time plot of the data using ts and plot functions. (Hint: follow the same way we did in class) By looking at the plot, does there appear to be a quarterly pattern?

```
library(ggplot2)
library(forecast)
```

```
## Registered S3 method overwritten by 'quantmod':
##   method            from
##   as.zoo.data.frame zoo
```

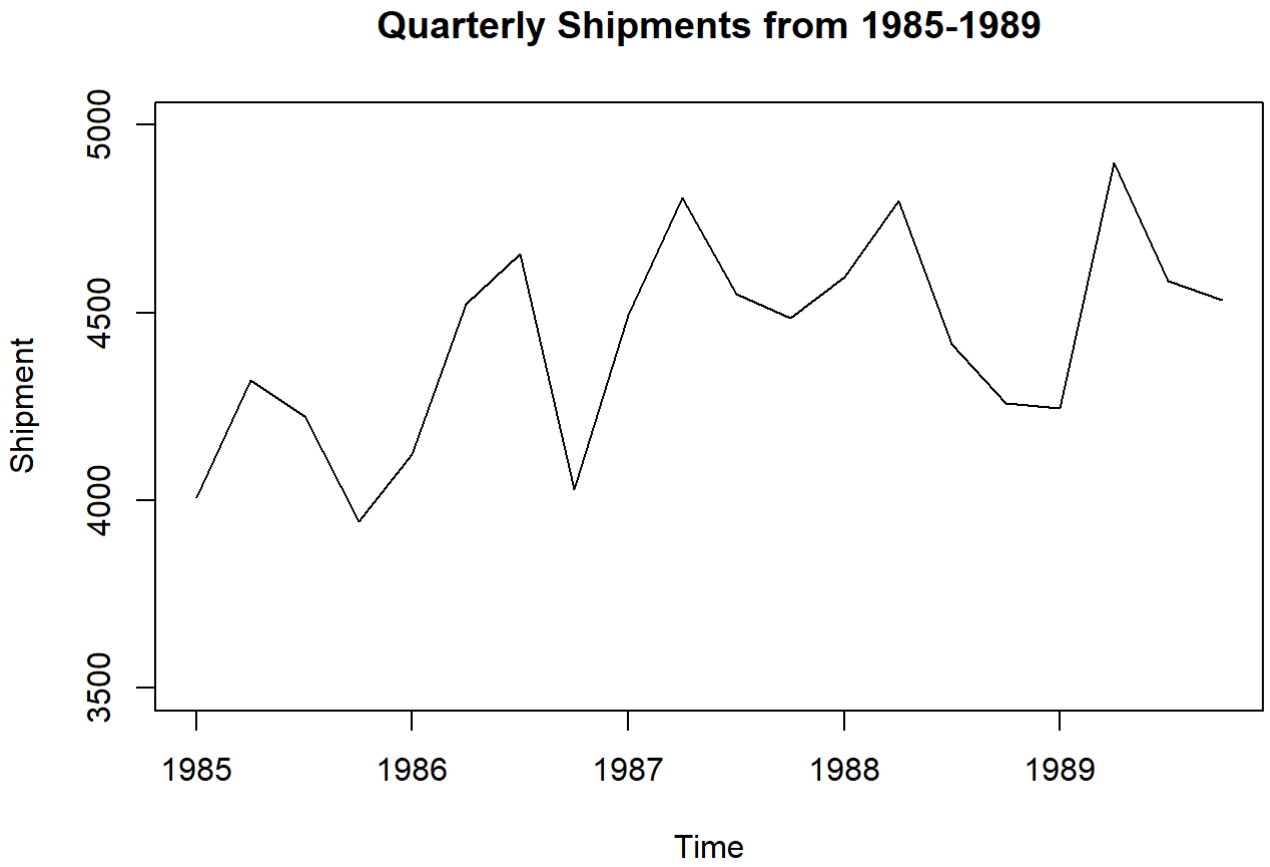
```
library(here)
```

```
## here() starts at D:/A_Lehigh/2022 Fall/BUAN 488 - Predictive Analytics/HW/HW3
```

```
shipment <- read.csv(here("ApplianceShipments.csv"))


shipment_ts <- ts(shipment$Shipments, start = c(1985, 1), end = c(1989, 4), frequency = 4)

plot(shipment_ts, xlab = "Time",
     ylab = "Shipment",
     ylim = c(3500, 5000),
     main = "Quarterly Shipments from 1985-1989")
```



Question 1 Answer

According to the previous plot, we can see that the shipment from 1985 to 1989 shows a clear quarterly pattern.

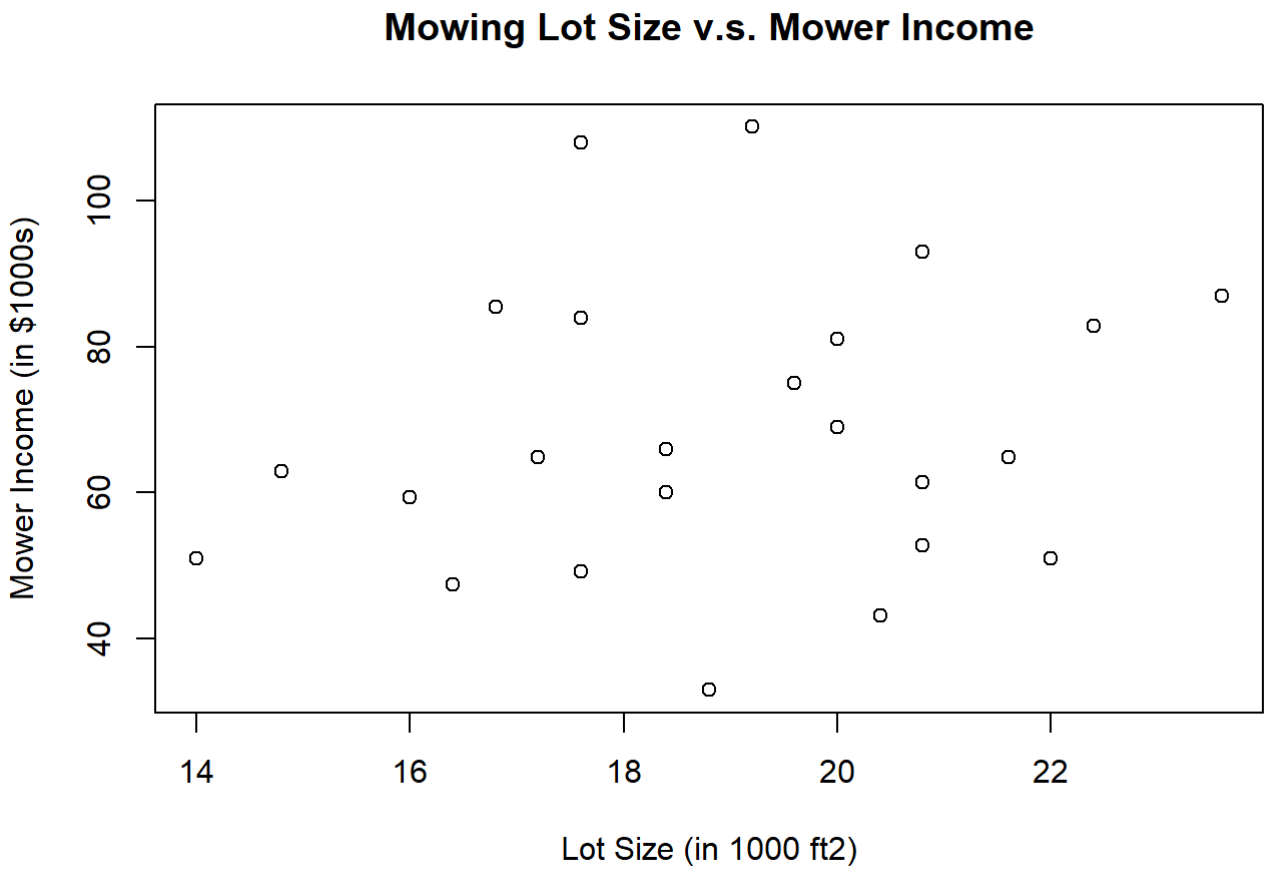
Question 2: Sales of Riding Mowers: Scatter Plots.

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 ft2). The marketing expert looked at a random sample of 24 households, given in the file RidingMowers.csv.

- Item a) Using R, create a scatter plot of Lot Size vs. Income by plot function. By looking at the plot, do you get an insight?

```
mower <- read.csv(here("RidingMowers.csv"))

plot(y = mower$Income,
     x = mower$Lot_Size,
     ylab = "Mower Income (in $1000s)",
     xlab = "Lot Size (in 1000 ft2)",
     main = "Mowing Lot Size v.s. Mower Income")
```

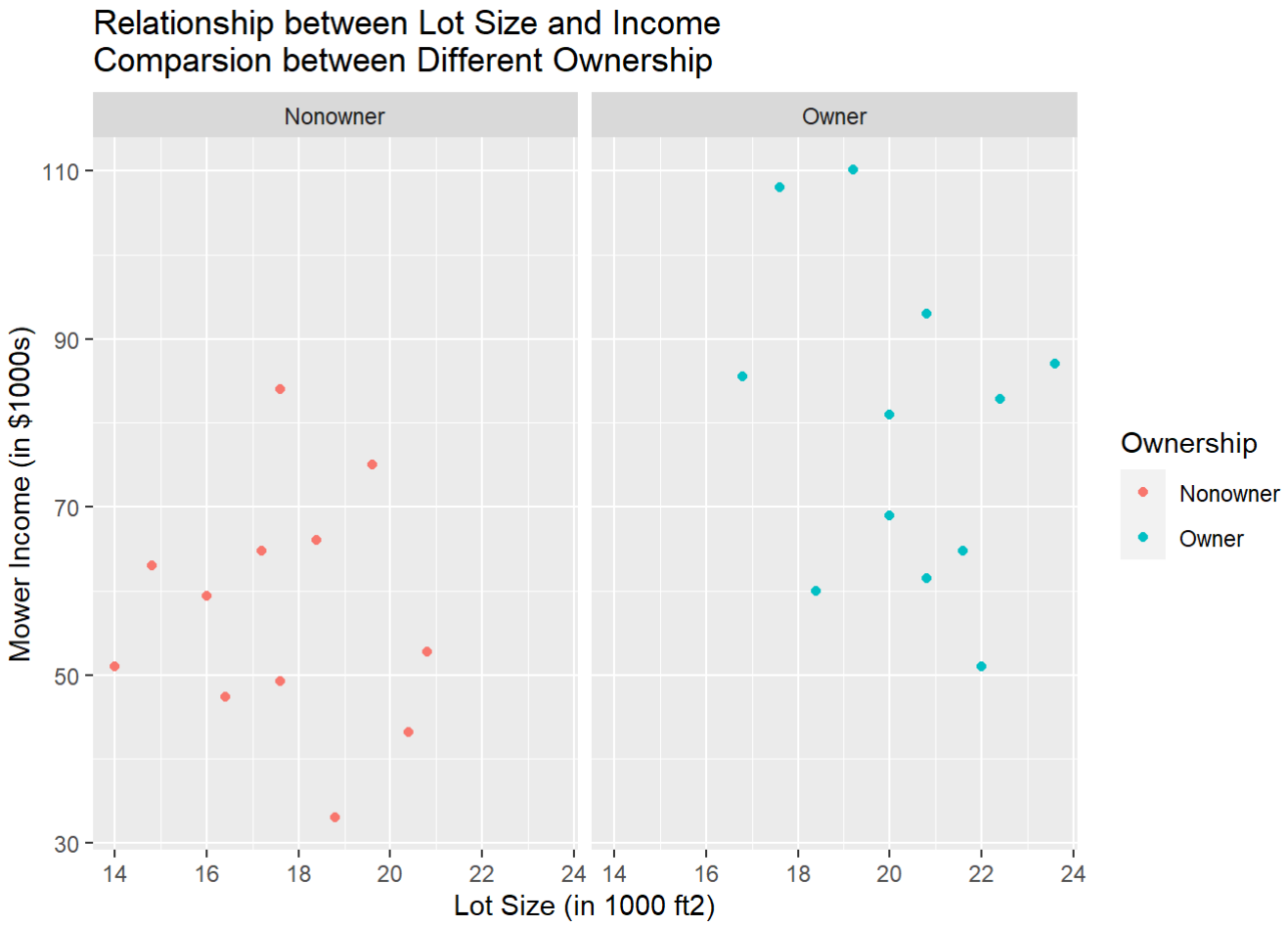


Question 2 (a) Answer

According to the previous scatter plot, we barely can see any clear relationship between lot size and income.

- Item b) Using R, create a scatter plot of Lot Size vs. Income, color-coded by the outcome variable owner/nonowner by ggplot function (There is an argument for “aes” object that you can color points based on the variable you choose). This time you color the points by Ownership. Now, by looking at the plot, would you get a conclusion that there is a relationship between Lot_Size and Ownership. Justify your answer.

```
ggplot(mower) +
  geom_point(aes(x = Lot_Size, y = Income, color = Ownership)) +
  facet_wrap(~Ownership, ncol = 2) +
  xlab("Lot Size (in 1000 ft2)") +
  ylab("Mower Income (in $1000s)") +
  ggtitle("Relationship between Lot Size and Income\nComparsion between Different Ownership")
```



Qestion 2 (b) Answer

According to the previous graphs, we can see that there's a relationship between lot size and ownership. When we are looking at Owner, we can see the data points drop at high lot_size and high income, which concentrate at the upper right corner of the graph. In contrast, the data points for non-owner concentrate at the lower left part of the graph, which illustrates that non-owners usually are not able to access to higher lot_size and have high income.