HW3Part2

Shanglun Li

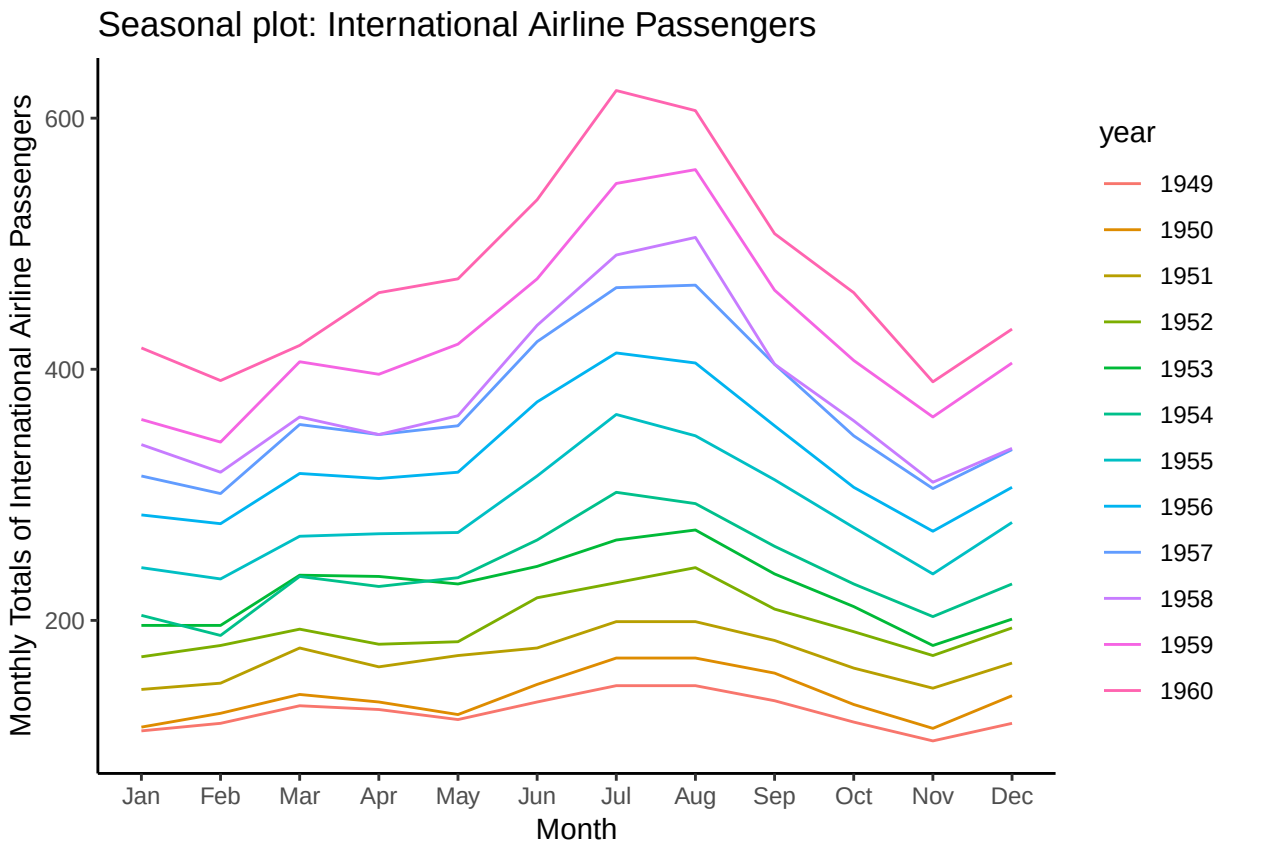
5/29/2019

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

```
## Warning: package 'forecast' was built under R version 3.5.2
```

```
theme_set(theme_classic())

# Plot
ggseasonplot(AirPassengers) + labs(title = "Seasonal plot: International Airline Passengers", y = "Mont
```



Here I used the AirPassengers dateset in R. This dataset is the monthly totals of international airline passengers from 1949 to 1960. The number of passengers is counted in thousands. Here I used ggplot2 and forecast package. Then I used the ggseasonplot command to draw a yearly seasonal plot of the international airline passengers. There are 12 years in total. Each year has its own line plot to demonstrate the pattern of the number of passengers for each month. In terms of truthfulness, I think the graph is truthful. There is no missing value in the dataset. In addition, the dataset is officially saved in R. Thus, the data is truthful. I separate the dataset by year without changing any part of the dataset. Therefore, the data visualization should be truthful in this case. In terms of functional, my data visualization is functional. It compares 12 years' number of passengers by demonstrating them all in one graph. We can clearly see that the number of international airline passengers is increasing by year. In addition, my visualization also shows that there are more passengers in July and August, and less passengers in November and February. In addition, the graph is

also clearly labeled. The approximate number of each point can be easily figured out. My data visualization uses simply white background, which is clean. In addition, its line has different colors to extinguish each years' records. Furthermore, the color representation of each year is showed in the side bar. Audience can refer to it easily. Thus, my data visualization is beautiful and clean. My data visualization not only gives the comparison of the number of international airline passengers among 12 years, but also gives the audience an idea which month has the most international airline passengers and which month has the least, which cannot be detected simply from the dataset. Thus, it is insightful. Finally, my data visualization gives the audiences some enlightening information. For example, after detecting that the peak month of passengers is July and August, we can buy the plane tickets earlier or later to avoid the huge amount of passengers and get a cheaper ticker. We can also be informed that it is good to buy plane ticket in November and February, since there is less passengers in these months and the plane ticket should be relatively cheap at these months. Thus, my data visualization is enlightening.