Question1

library(fpp3)

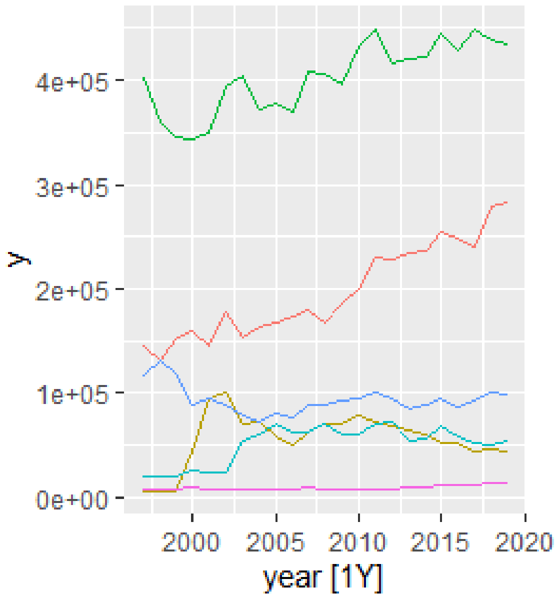
install.packages("USgas")

library(USgas)

usTotalChosen <- filter(us_total, state == "Maine" | state == "Vermont" | state == "New Hampshire" | state == "Massachusetts" | state == "Connecticut" | state == "Rhode Island")

usTotalChosenTsibble <- as_tsibble(usTotalChosen, index = year, key = state)

autoplot(usTotalChosenTsibble)



Massachusetts, Connecticut, Vermont all have seen a steady increase in gas consumption over the period chosen. In contrast, Rhode Island, Maine, and New Hampshire all saw either increases till approximately 2003 before dipping down to lower levels.

This contrast in gas consumption could be due to an increase in the population as in the case of the first example or a decrease in the population as with the second example.

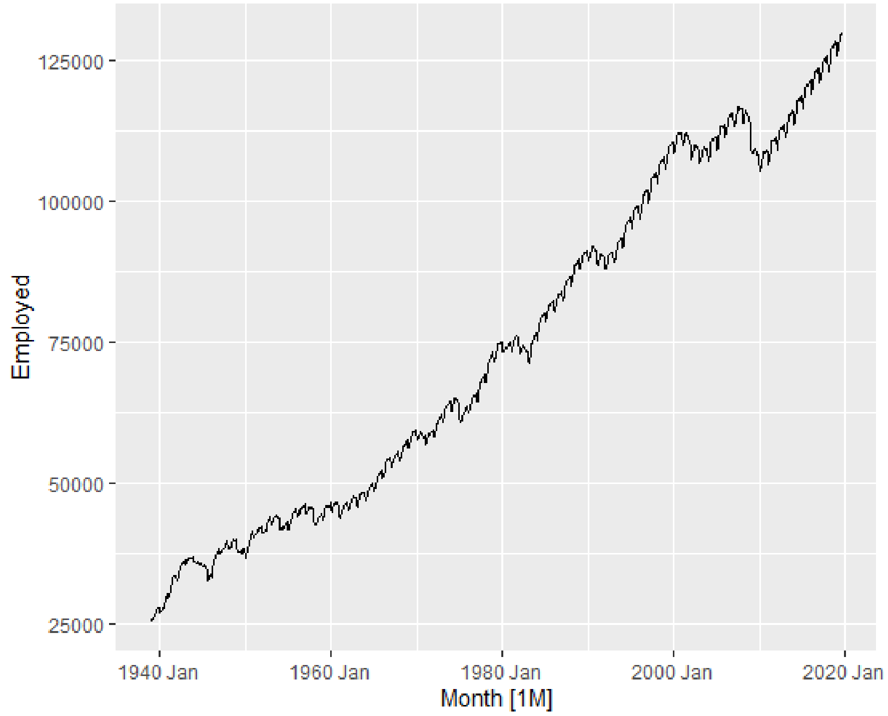
Vermont has the lowest gas consumption owing to its smaller population whilst Massachusetts has the highest gas consumption owing to its large population size in contrast to the others.

Question 2

usEmploymentChosen <- filter(us_employment, Title == "Total Private")

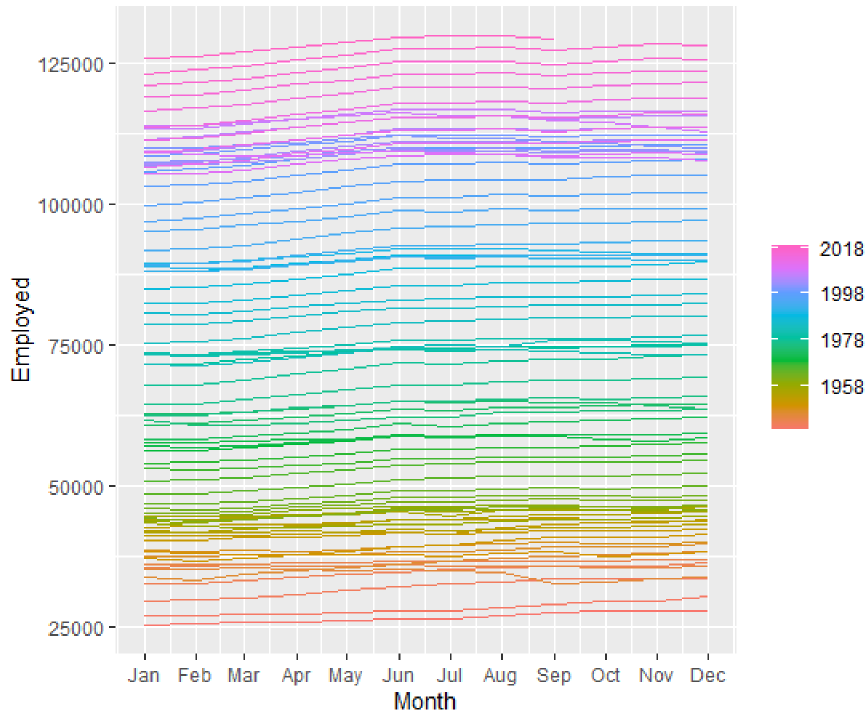
autoplot(usEmploymentChosen)

Figure 1.



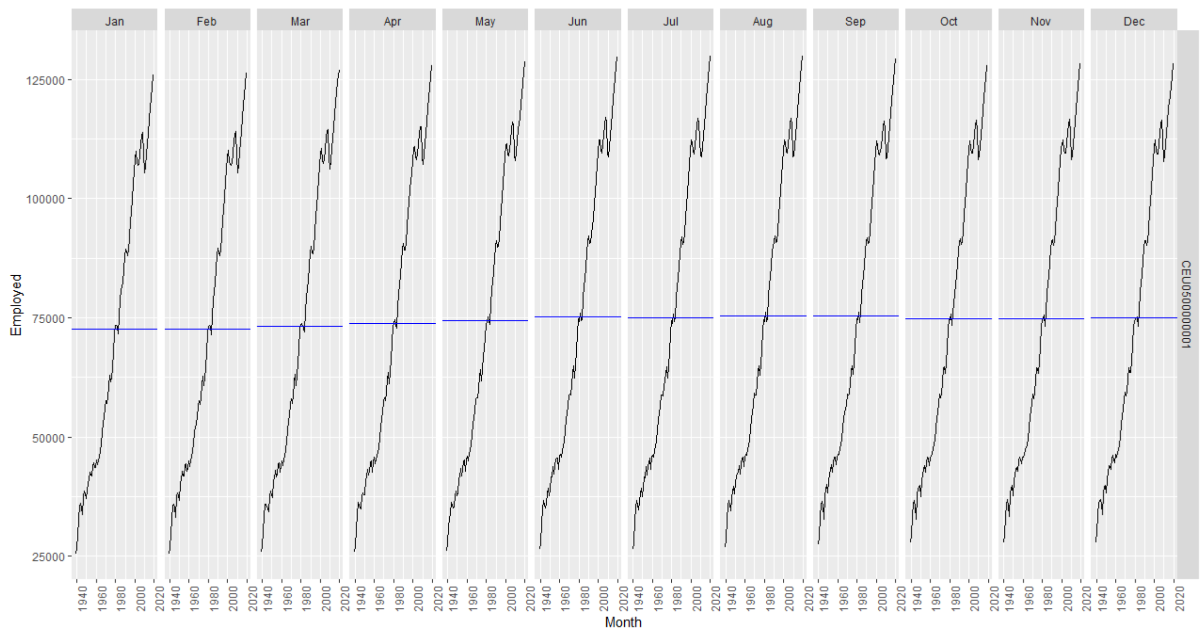
gg_season(usEmploymentChosen)

Figure 2



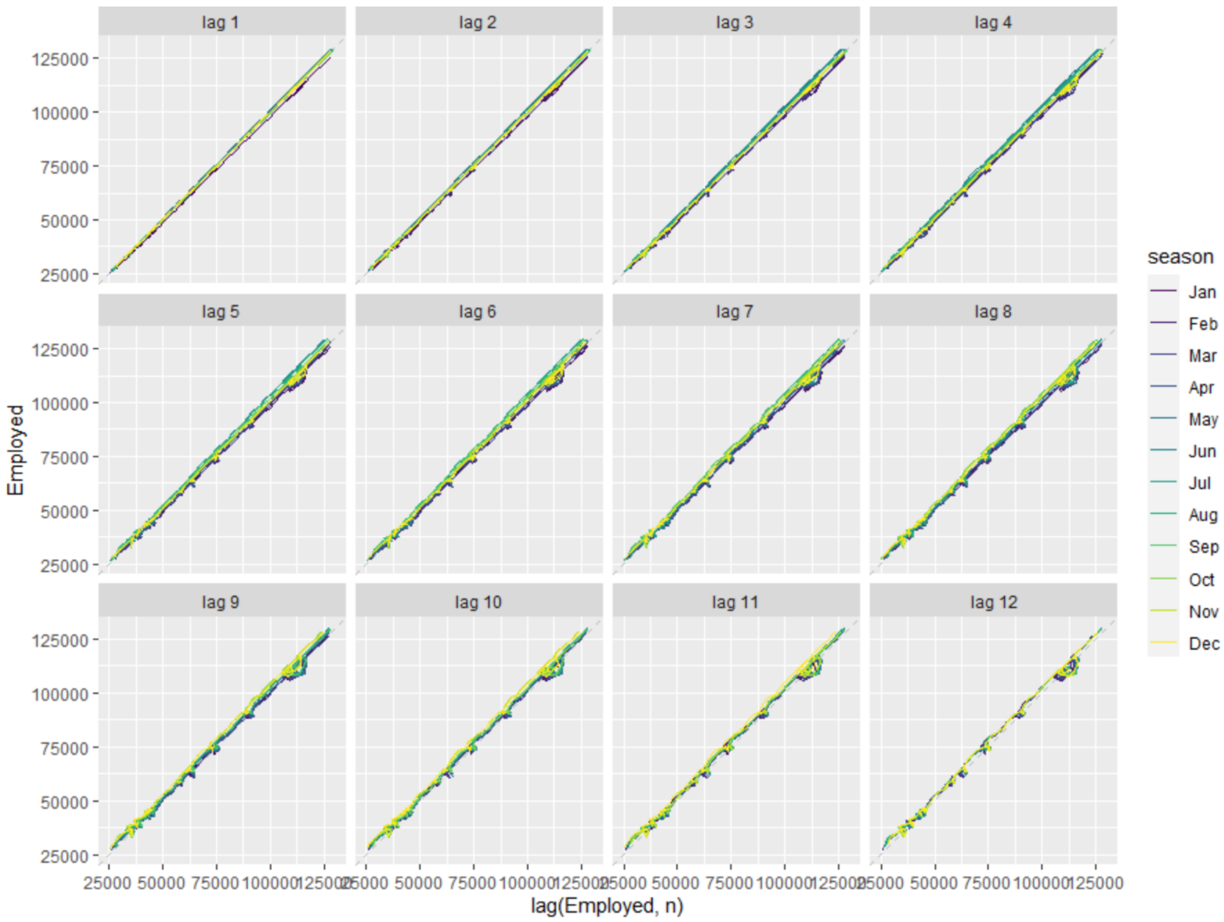
gg_subseries(usEmploymentChosen)

Figure 3



gg_lag(usEmploymentChosen, lag = 1:12) [due to 12 months in year]

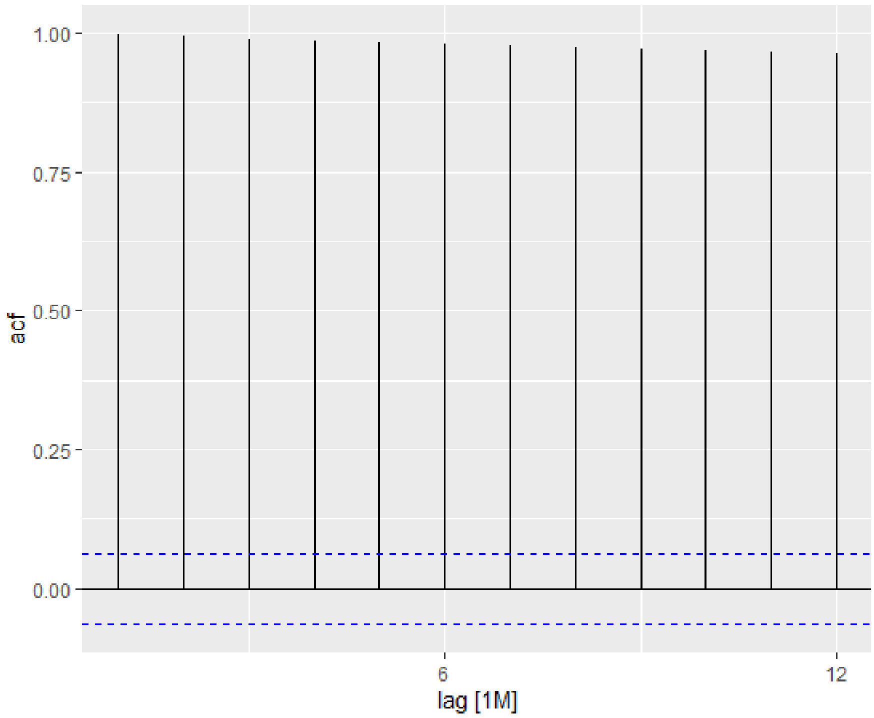
Figure 4



ACFmodel <- usEmploymentChosen %>% ACF(Employed, lag_max = 12) [due to 12 months in a year]

autoplot(ACFmodel)

Figure 5



Can you spot any seasonality, cyclicity and trend?

By analysing the autoplot results from figure 1, we can see the seasonality of employment over time due to the smaller microtrends within the larger time frame i.e. holiday employment, June peak in employment The ACF model also highlights the large seasonality over time due to the strong and consistent peaks which are above the blue 95% confidence intervals. However, this seasonality seems to be decreasing at a steady rate as seen by the lowering levels of ACF over time.

Similarly, the graph also showcases the cyclicity of the graph due to certain periods in which there are lowered rates of employments owing to economic issues i.e., GFC of 2008. Nevertheless, cyclicality is not repeated throughout the graph highlighting the existence of only low levels of cyclicality.

There has also been a general positive and upwards trend with the data as seen by the increased numbers of employment over time with the exception of certain periods of time i.e., GFC and economic downturns, especially in figure 1.

What do you learn about the series?

These series show insight into the trends of Total private employment in the US with clear indications of the general trends that the data seems to follow as well as the seasonality of the data. It also showcases the impact of cyclicality i.e., through the lowered rates of employment especially during GFC whilst also highlighting the strong and continued growth of the US employment figures.

What can you say about the seasonal patterns?

From figure 1 we are able to see the strong seasonality of the data due the micro cycles that exist within the graph i.e., within a year.

However, the subseries graph showcases that the variations and patterns are less obvious over months since there was only a slight increase in employment figures towards the middle of the year before lowering again.

The smooth nature of figure 2 also highlights the existence of seasonal patters although it is very insignificant and more a gradual change over time.

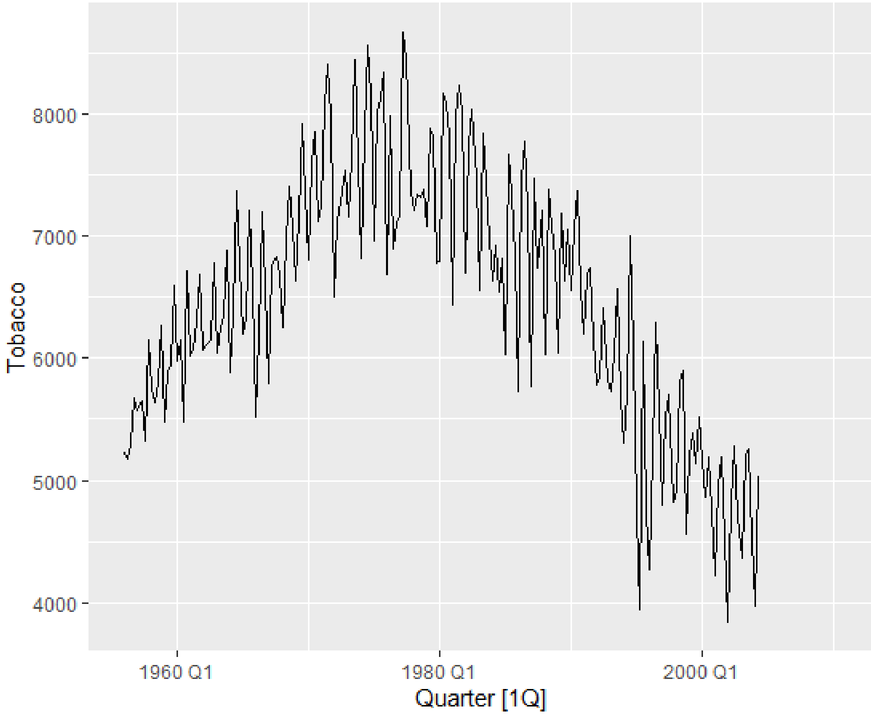
Can you identify any unusual years?

The main unusual year would be 2008 in figure 1 owing the 2008 Global Financial Crisis which especially impacted the US as seen by the significant drop in the employment figures. With the financial sector making up a large proportion of private employment along with small businesses which were also impacted, the ramifications were more severe for this component of the total employment, as seen by the data.

Question 3

a)

aus_production %>% autoplot(Tobacco)



aus_production %>% features(Tobacco, features = guerrero)

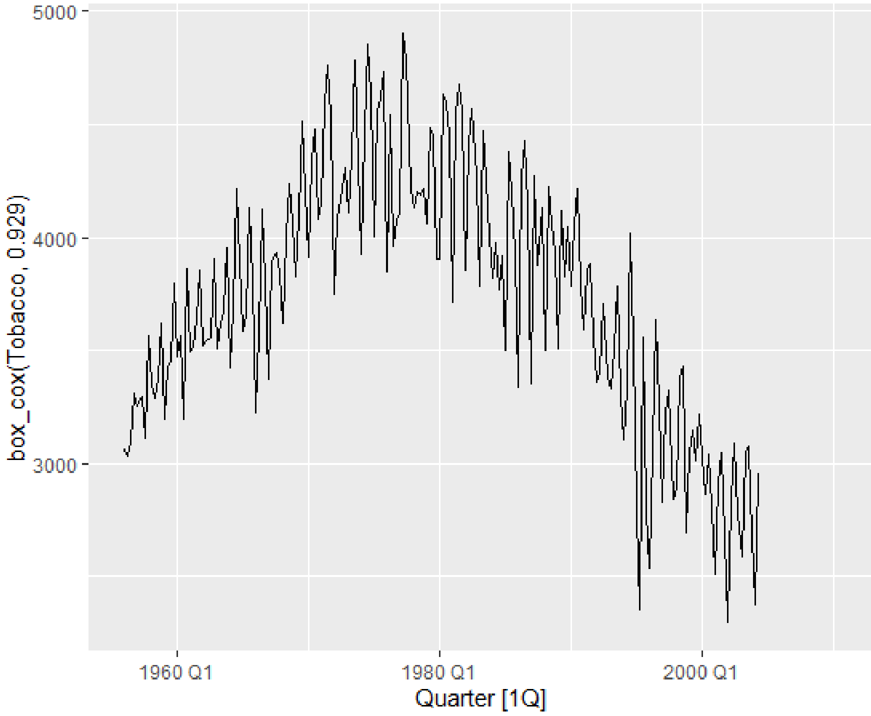
A tibble: 1 x 1

lambda_guerrero

<dbl>

1 0.929

> aus_production %>% autoplot(box_cox(Tobacco,0.929))

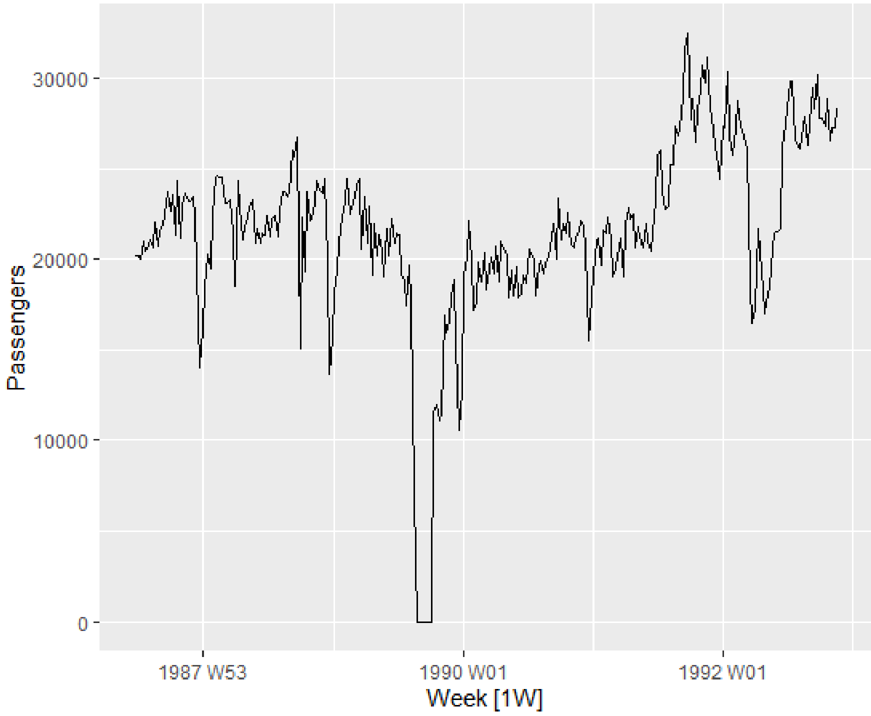


Therefore since the graphs variance and general trend stayed quite similar when transformed to pre-transformation, the transformation was not necessary since the data has not become easier to analyse.

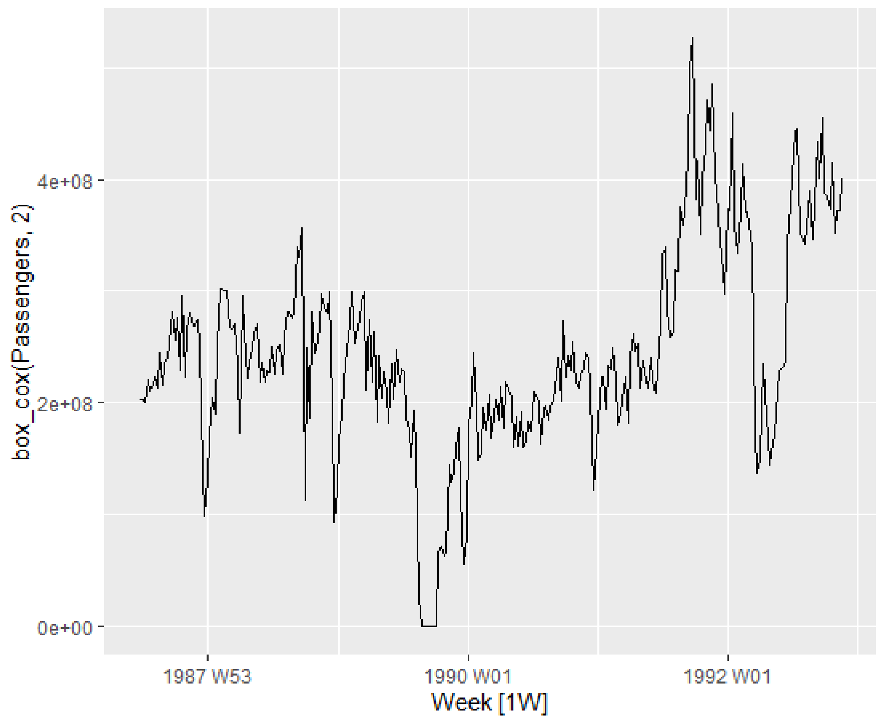
b)

ansettChosen <- filter(ansett, (Airports == "SYD-MEL" | Airports == "MEL-SYD") & Class == "Economy")

autoplot(ansettChosen)




```
> features(ansettChosen, features = guerrero)
Feature variable not specified, automatically selected `.var = Passengers`
# A tibble: 1 x 3
  Airports Class   lambda_guerrero
  <chr>    <chr>             <dbl>
1 MEL-SYD  Economy            2.00
> ansettChosen %>% autoplot(box_cox(Passengers, 2))
```



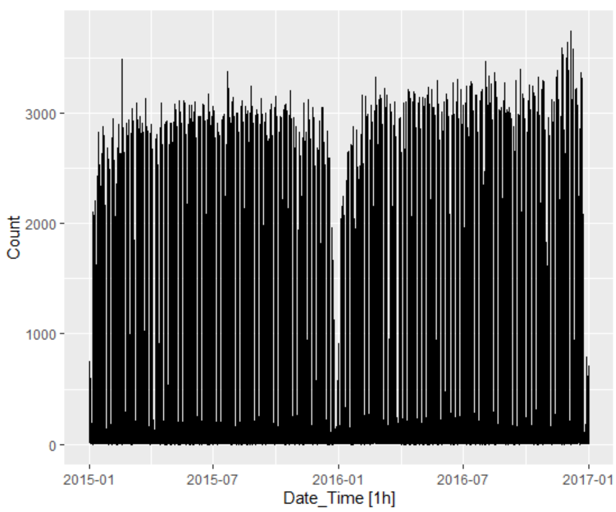
There has been significant change in variation when the original data was transformed with the dip in passengers just before 1990 being adjusted to the general trend and variation that the data has. Therefore in this case, the box cox transformation was helpful and useful.

The transformations lambda was figured out using guerrero as per the lecture example, allowing for an appropriate and valid lambda to be found and used which reduced the outlier and extreme values.

c)

chosenPedestrian <- filter(pedestrian, Sensor == "Southern Cross Station")

chosenPedestrian %>% autoplot(Count)



features(chosenPedestrian, .var = "Count", features = guerrero)

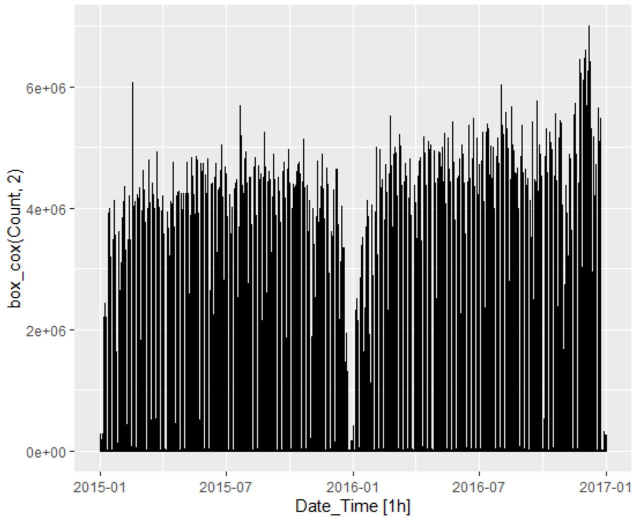
A tibble: 1 x 2

Sensor lambda_guerrero

<chr> <dbl>

1 Southern Cross Station 2.00

chosenPedestrian %>% autoplot(box_cox(Count, 2))



There was no significant change with the data and with the variance within the graph when transformed with box cox owing to large amount of tight data that was presented in the graph. Therefore, in the case, the transformation was not helpful.