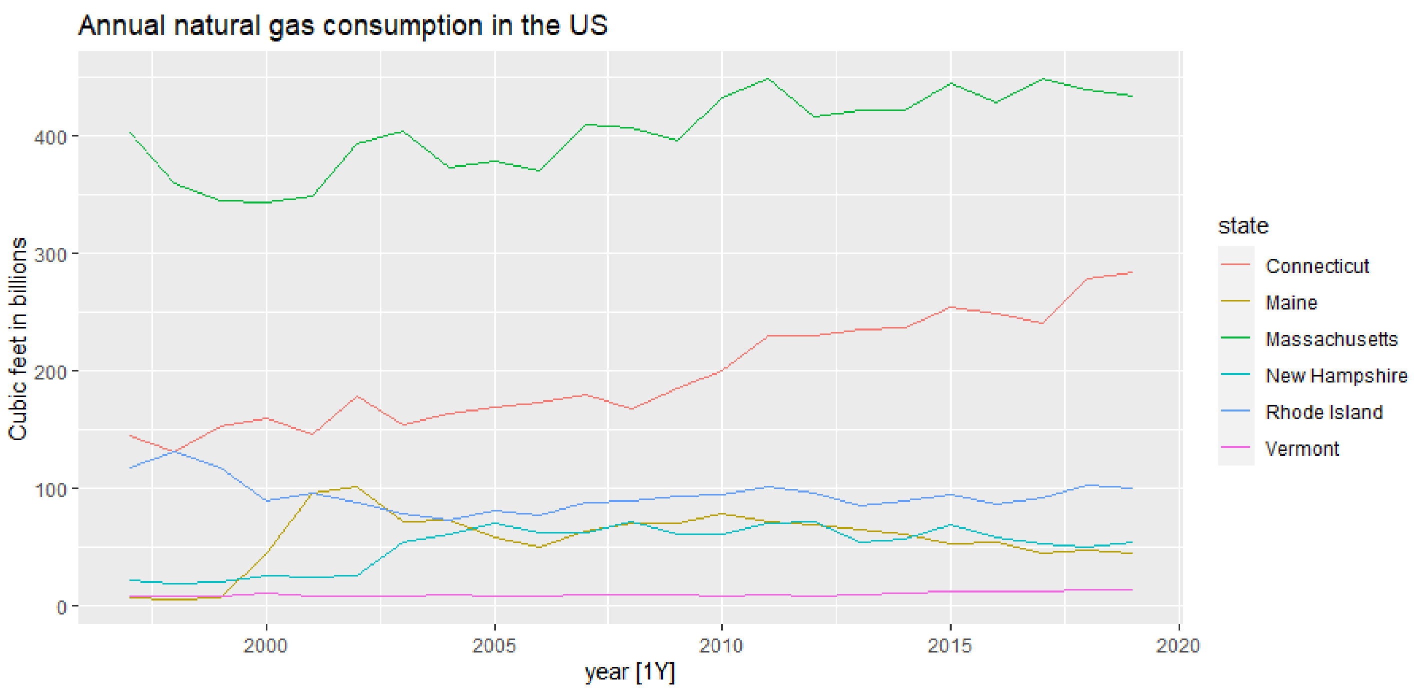
ECON2209 Problem Set 1

Problem 1:

Using Rstudio, the graph below was produced, showcasing the annual gas consumption for different states in the New England area in the US



Interpreting this plot, we can see that the annual gas consumption in Connecticut is following an increasing trend from 1997-2019, however the spikes across the years are following a part of the trend. Compared to other states, the gas consumption is higher throughout the years except Massachusetts. There is no appearance of seasonality in the data presented to us. Although, reproducing the same plot in a monthly or quarterly basis can reveal possible seasonality effects regarding gas consumption within each year. When analysing the annual gas consumption in Maine, we see that it follows a decreasing trend from 2002-2019. From 1997-1999, there is no evidence of gas consumption. The plot in Maine then has a drastic increase from 1999-2002. There seems to be a decreasing trend from 2002-2019. There is also some cyclicity in the data for Maine. The peaks in the cycle are in 2002, 2004, 2010 and 2016 and the troughs are in 2003, 2006, 2015 and 2017. In terms of seasonality, interpreting this plot gives us no evidence of seasonality in the data. Looking at Massachusetts, we can interpret that there is no increasing or decreasing trend. An obvious observation is that the annual consumption in Massachusetts is higher than other states, this may be due to the greater population compared to other states. The increase and decrease trend appear to be in random variation. New Hampshire has cyclicity from the trough in 2003 to the 2018, which is similar to Maine. Gas consumption becomes steady from 1997-2002. This is followed by an increase from 2002-2005. No apparent trend or seasonality appears in New Hampshire. Gas consumption in Rhode Island contains no trend, seasonality or cyclicity.

The data instead remains relatively consistent from 1997-2019 and at higher peaks the annual gas consumption of Rhode Island is larger than Vermont, New Hampshire, and Maine. Another obvious observation is the Vermont data, which remains consistent from 1997-2019, with no trend, seasonality or cyclicity, also it contains the lowest gas consumption compared to all the other states. Thus, we can confirm that there is a variety of trend, seasonality and cyclicity throughout the data for annual gas consumption for the different states in the US.

R code for Problem 1:

```
# Problem 1
# install USgas package
install.packages("USgas")
library(USgas)

# Create a tsibble from us_total with index year and key state
us_tsibble <- us_total %>%
  as_tsibble(index = year, key = state)
view(us_tsibble)

# Refine the tsibble to only states in the New England area
us_tsibble %>%
  filter(state %in% c("Maine", "Vermont", "New Hampshire", "Massachusetts", "Connecticut"
                      , "Rhode Island")) %>%
  autoplot(y/1e3) +
  labs(y = "Cubic feet in billions") + labs(title = "Annual natural gas consumption in the US")
```

Problem 2:

Figure 1: Autoplot of Total Private Employment in US

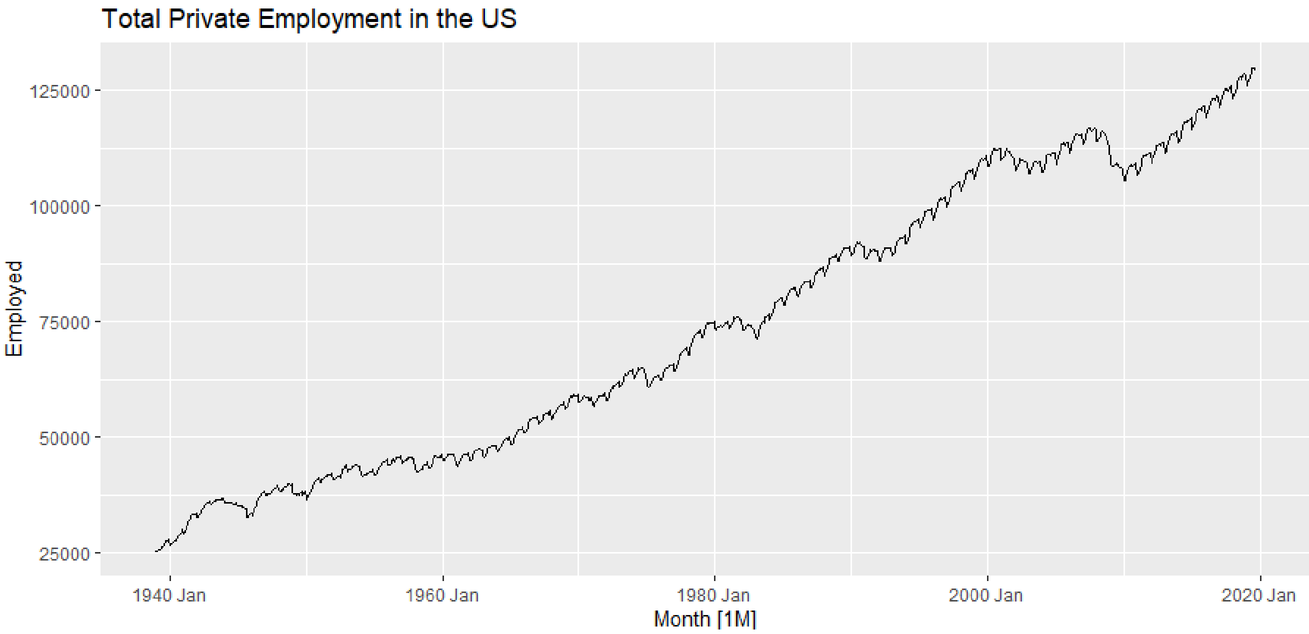


Figure 2: gg_season of Total Private Employment in US

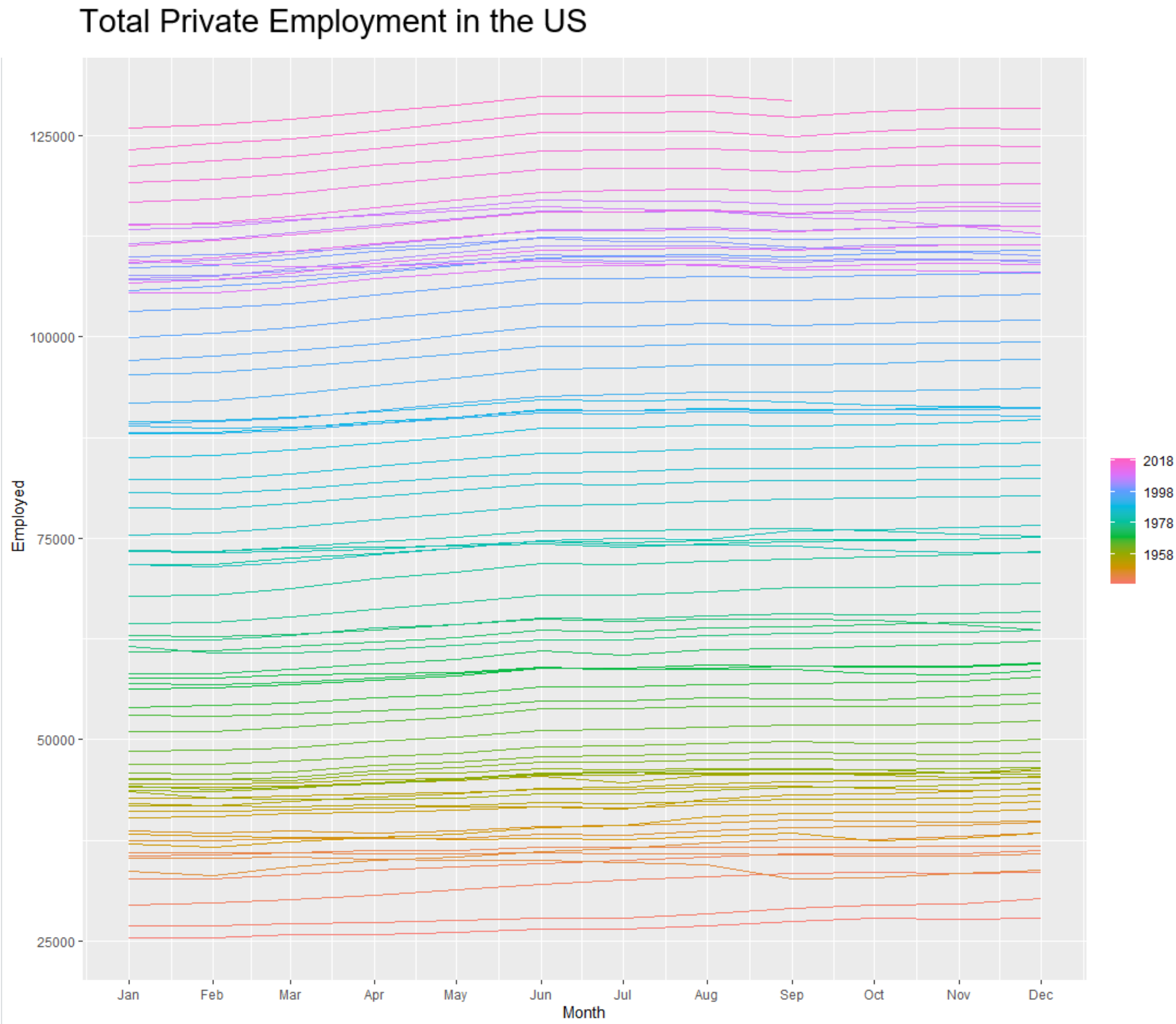


Figure 3: gg_subseries of Total Private Employment in US

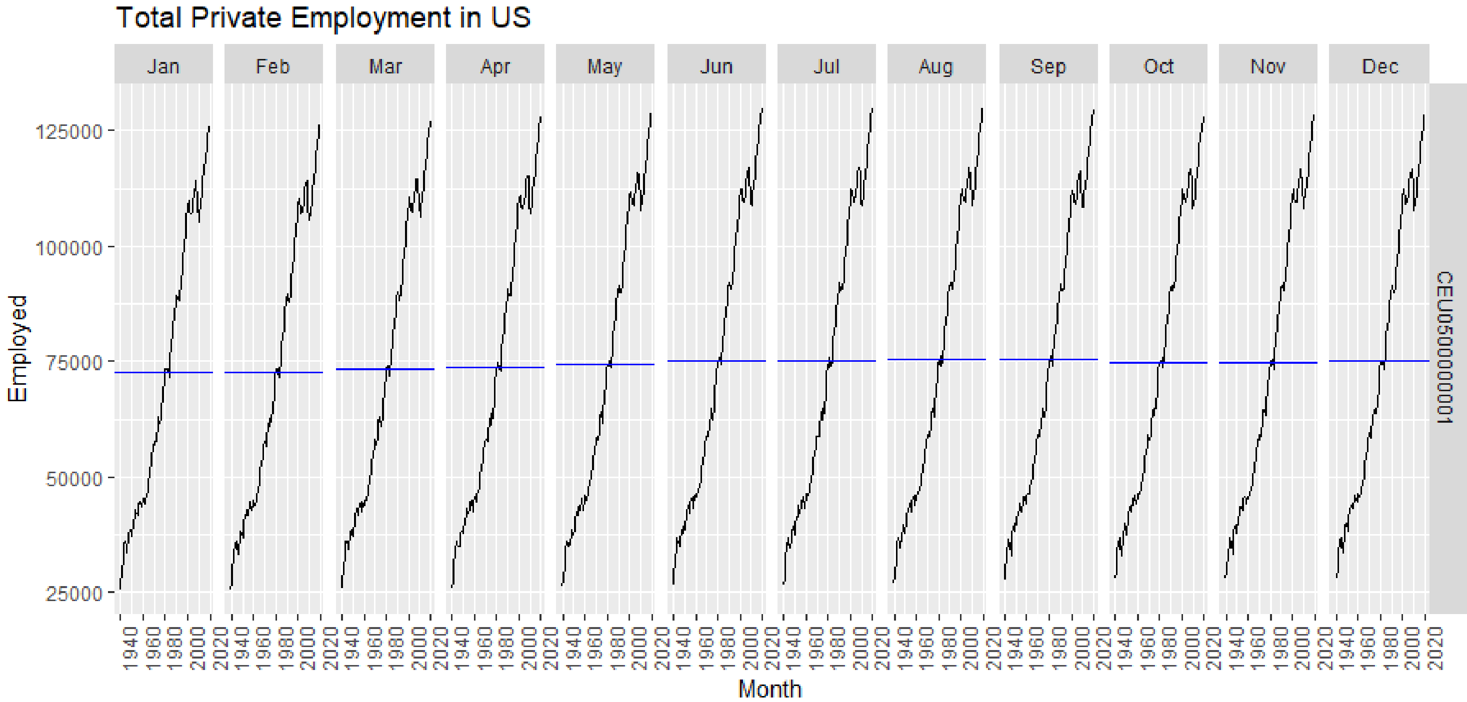


Figure 4: gg_lag of Total Private Employment in US

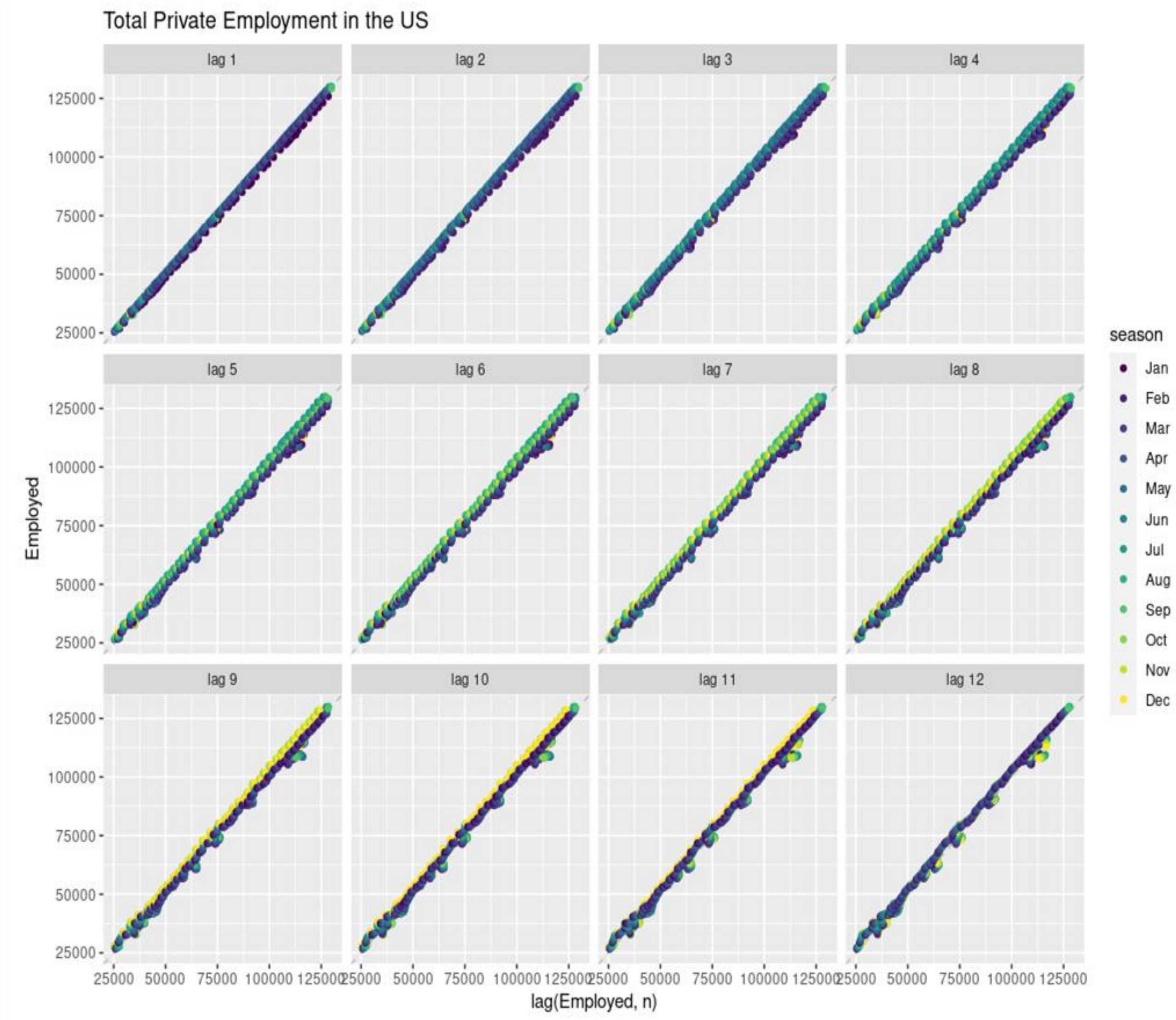
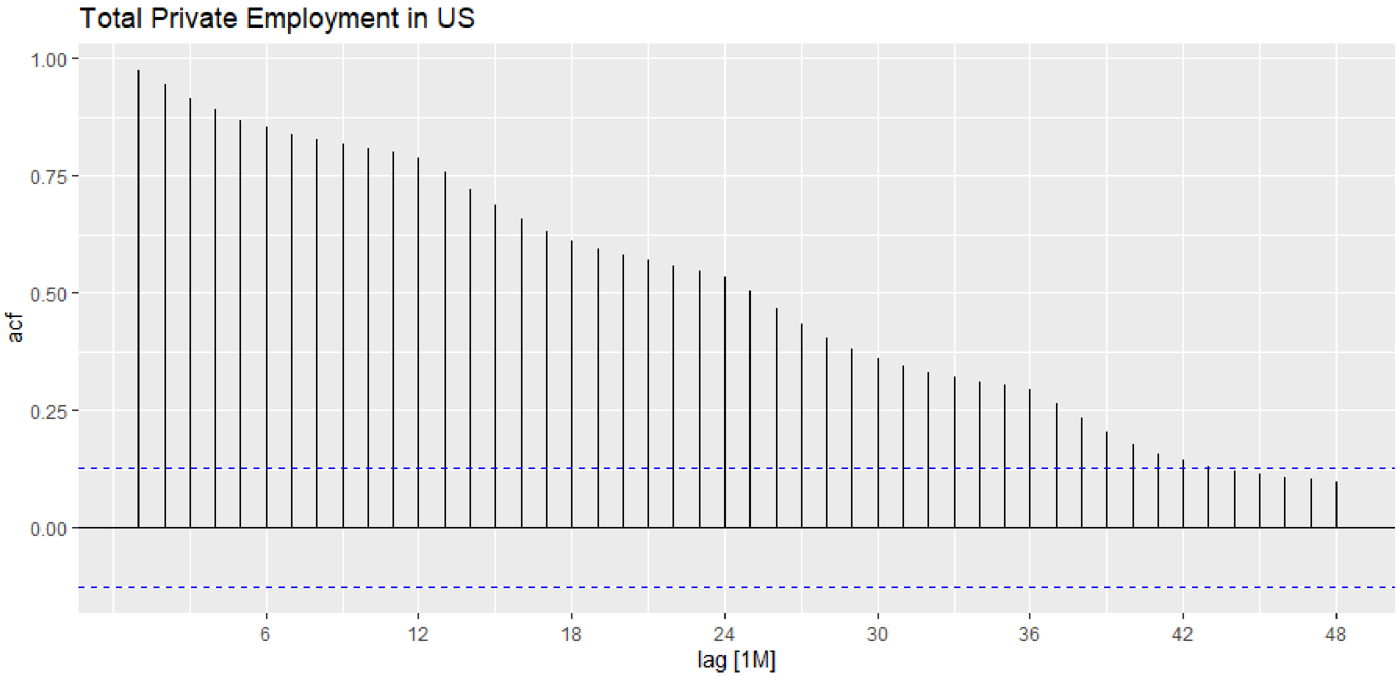


Figure 5: Autocorrelation function (ACF) of Total Private Employment in US



Note: this plot has been filtered to produce outputs from months of years from 2000 and greater.

Interpreting the line graph in figure 1 showcases that there is a clear increasing trend throughout the time of the data. The seasonal graph in figure 2 also indicates an upward trend across the months. As the US population increases throughout the period, we see a steady increase in the number of privately employed individuals. From figure 1, the regular peaks and troughs indicate that the total private employment is affect by seasonality in years. In addition, the seasonal plot in figure 2 indicates a slightly higher employment in the 3rd quarter of the year (from June to September). In figure 4, the lagged plot shows us a slight seasonal pattern, as lag 12 is the most linearly plotted lag compared to the others. In figure 5, the ACF plot shows a steady and decreasing relationship between y_t and the previous month however there are no seasonal effects. Going back to figure 1, the number of employed people in the US has increased steadily from 1939-2020. However, between 2008-2009, there is an enormous fall in employment. This decrease is likely due to the Global Financial Crisis which occurred in 2008. This rapid and enormous decline is also apparent in the seasonal plot in figure 2, where the pink and blue lines are overlapping around the 110000 level of employment. We can also interpret the employment drop due to the GFC in the subseries plot in figure 3.

R code for Problem 2:

```
# Problem 2
View(us_employment)

# Autoplot of Total Private Employment in US
employment <- us_employment %>%
  filter(Title == "Total Private")
auto <- employment %>%
  autoplot(Employed) + labs(title = "Total Private Employment in the US")
auto

# gg_season: Total Private Employment in US
us_employment %>%
  filter(Title == "Total Private") -> us_employment_priv
us_employment_priv %>% gg_season(Employed) + labs(title = "Total Private Employment in US")

# gg_subseries: Total Private Employment in US
us_employment_priv %>%
  gg_subseries(Employed) + labs(title = "Total Private Employment in US")

# gg_lag: Total Private Employment in US
us_employment_priv %>%
  gg_lag(Employed, geom= 'point', lags = 1:12) + labs(title = "Total Private Employment in US")

# Autocorrelation Function (ACF): Total Private Employment in US
us_employment_priv %>%
  filter(year(Month) >= 2000) %>%
  ACF(Employed, lag_max = 48) %>%
  autoplot() + labs(title = "Total Private Employment in US")
```

Problem 3:

Figure 6: Tobacco from aus_production (before transformation)

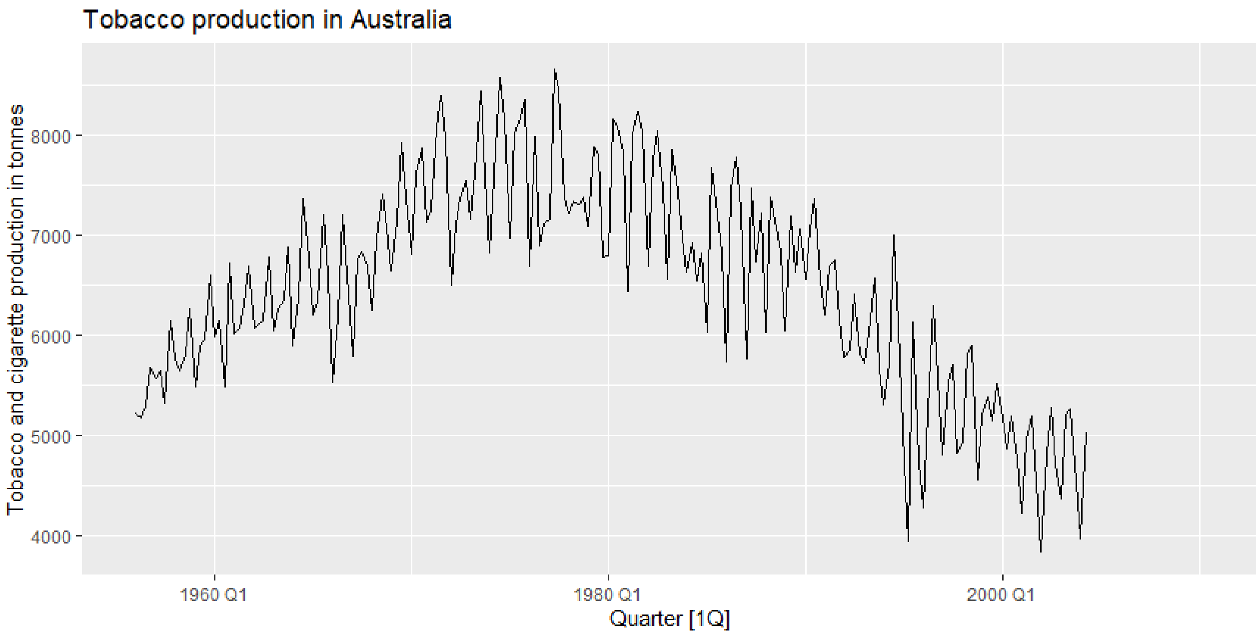
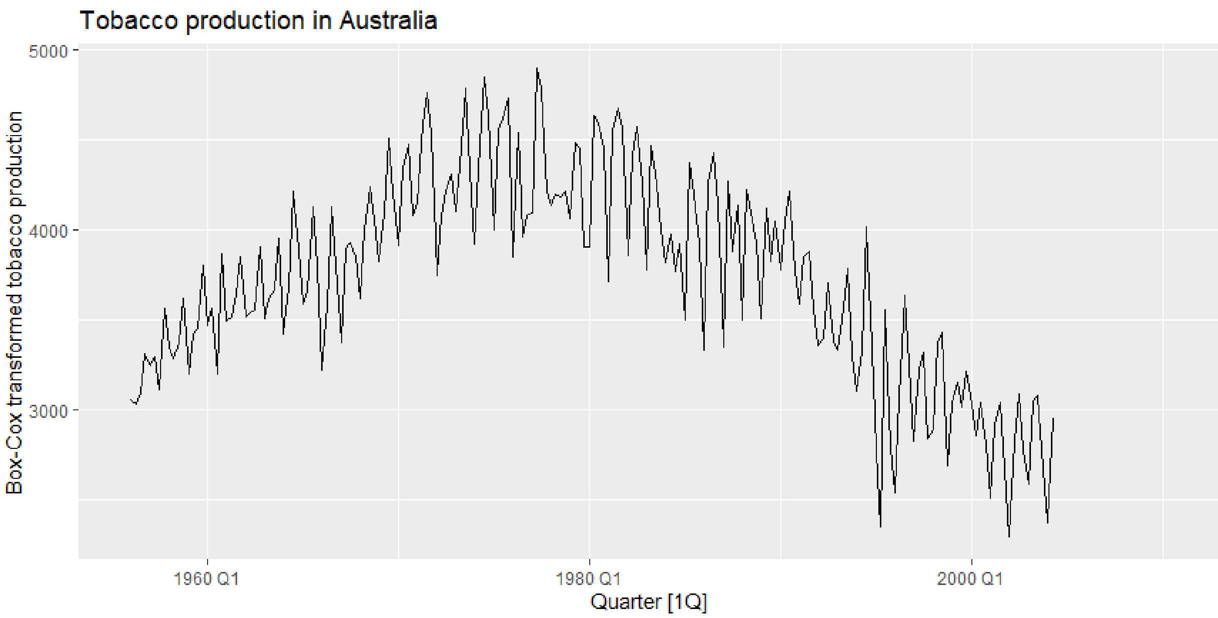


Figure 7: Tobacco from aus_production (after transformation)



We have implemented a box-cox transformation using Guerrero to get $\lambda = 0.929$. A λ value of 1 instigates no substantive transformation and since 0.929 is close to 1, there is no transformation needed. However, from figure 6 to the transformed figure 7, w can see that the transformed data has shifted downwards. However, there is no apparent change in the shape of the time series plot. The plot starts off with an increasing trend, then starts to decrease at around 1978. This is largely due to the numerous restrictions which the government and legislations have enforced on smoking which has affected Tobacco production. Smoking has been a major concern in society throughout the years which is why the decreasing trend shown in both figure 6 and 7 is consistent. There has been no substantive change while utilising other lambda values. The seasonality of the data is accurately displayed which is why there is no requirement to balance out random variation across the series.

Figure 8: Economy class passengers from ansett (before transformation)

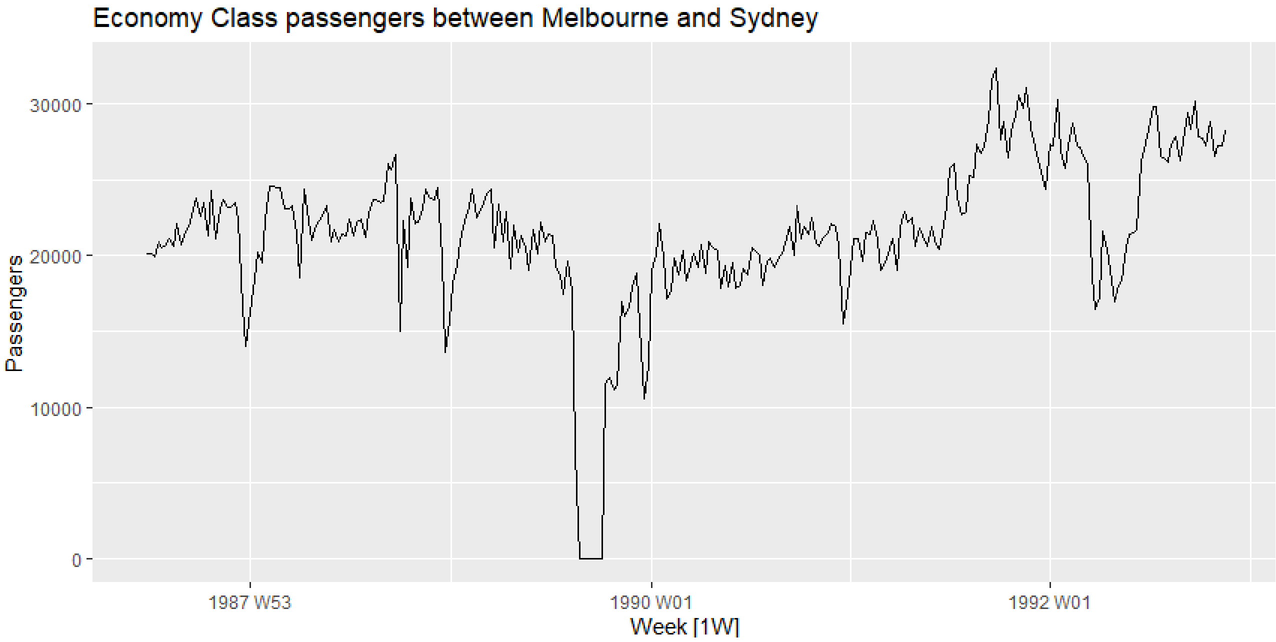
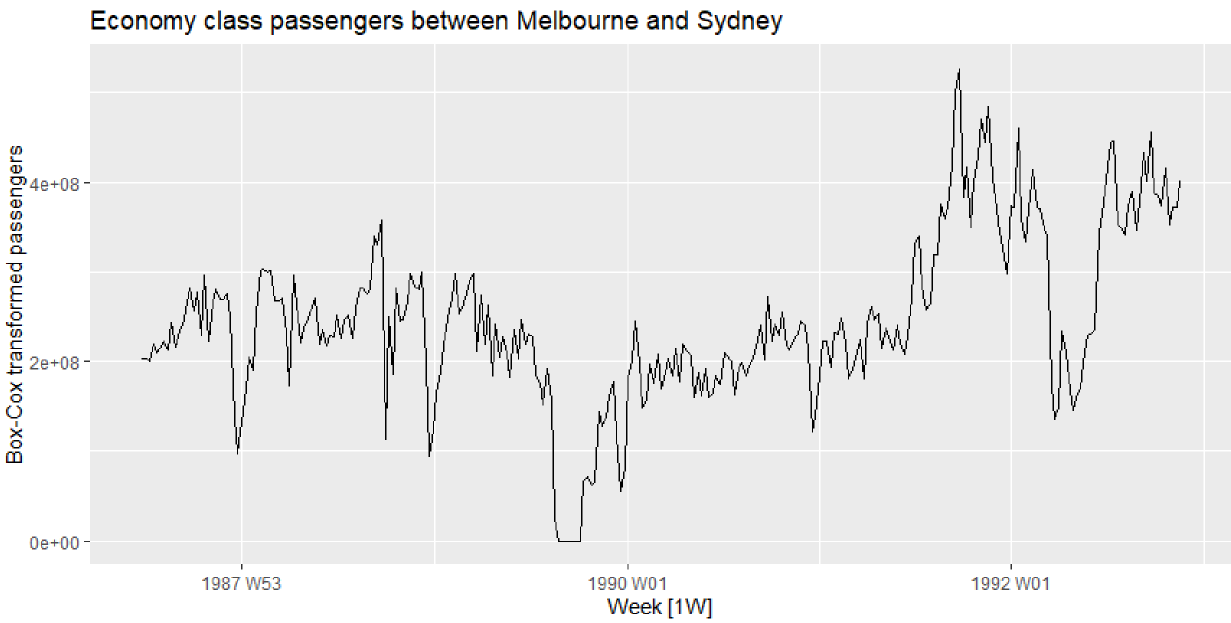


Figure 9: Economy class passengers from ansett (after transformation)



We have implemented a box-cox transformation, generated by Guerrero with $\lambda = 2.00$, however as seen from figure 9 there is not a substantial difference between this plot compared to figure 8 other than the chart being shifted up. One noticeable aspect is the recession in 1987 which occurred due to the New York Stock Exchange crashing (Black Monday), causing unemployment to rise significantly in Australia's economy, thus very few citizens were either unable to afford an economy plane ticket from Melbourne to

Sydney or they chose not to. The major decline in total passengers in 1987 is a lot less exaggerated in figure 9 compared to figure 8. The variance of this plot has been reduced through box-cox transformation, as seen from the smaller scale on the y-axis. The overall series contains less random variations as shown in figure 9, concluding that box-cox transformation was useful for $\lambda = 2.00$

Figure 10: Pedestrian counts from pedestrian (before transformation)

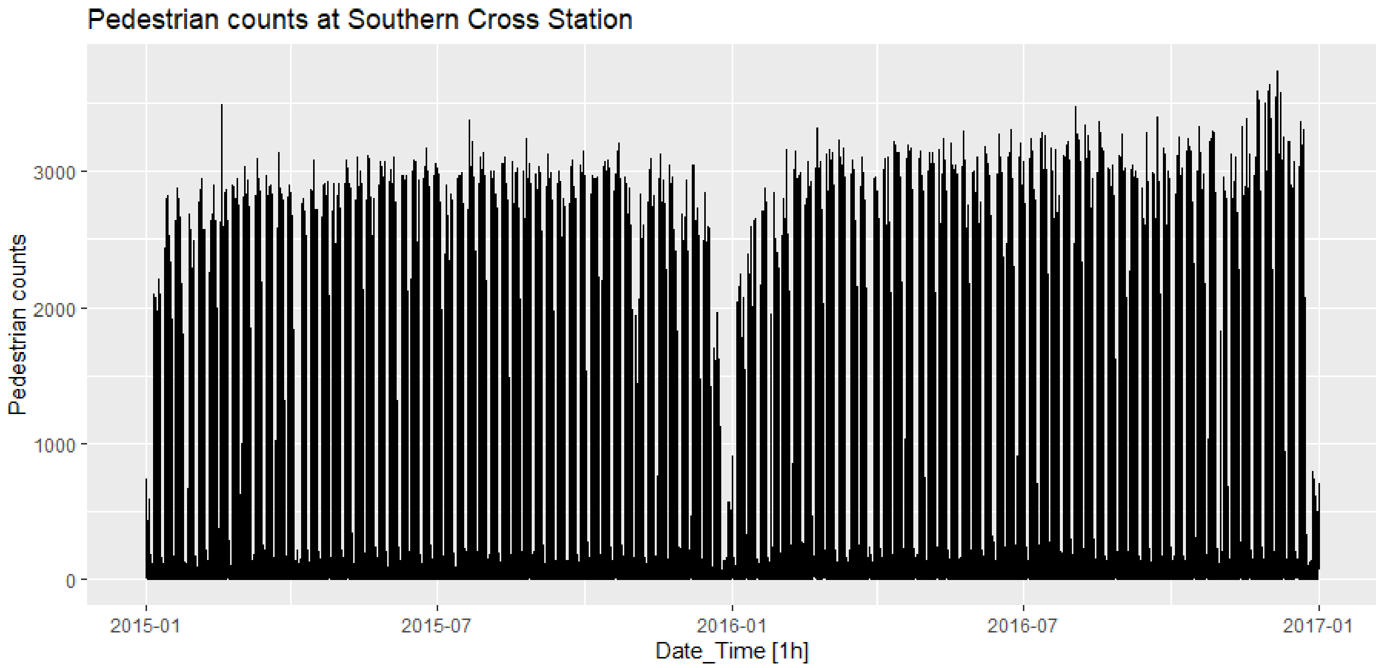
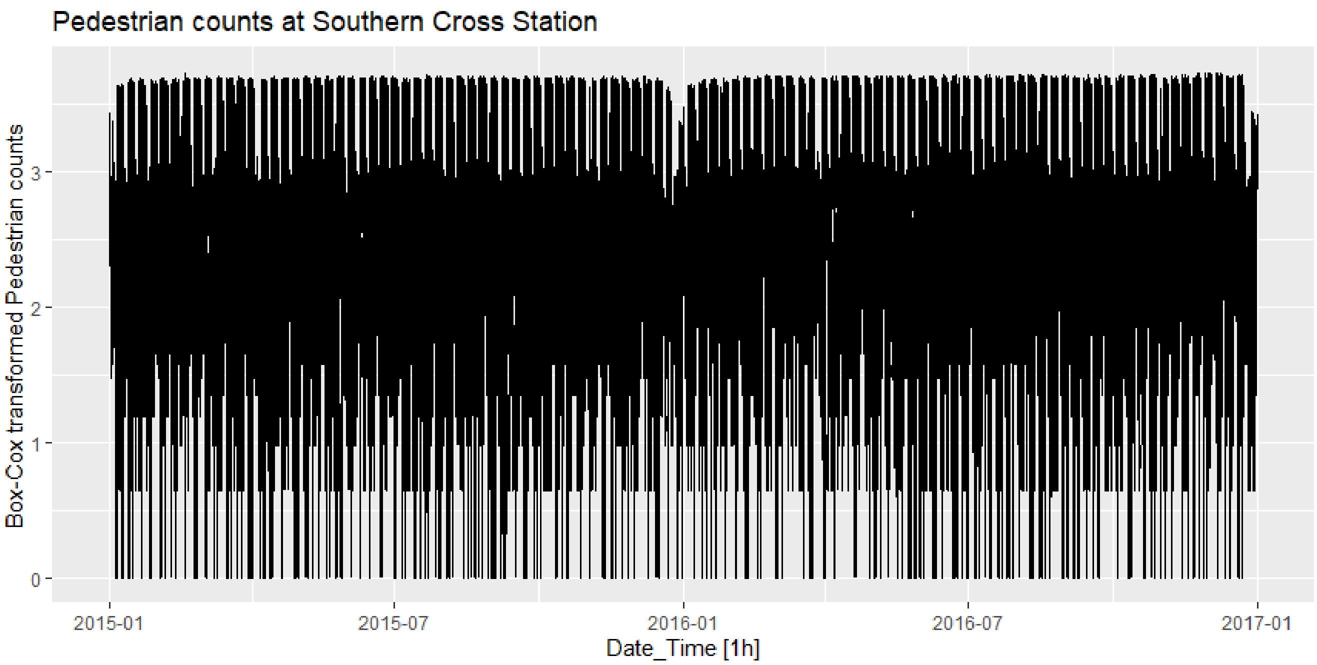


Figure 11: Pedestrian counts from pedestrian (after transformation)



The variance in figure 11 is relatively constant throughout, however the plot also shows strong daily and weekly seasonal fluctuations. An obvious observation in figure 10 is that in January 2016, we can see that there is a sudden decrease in pedestrians at southern cross station. This may have been due to a failure in the commuters from Melbourne to Sydney, which was immediately back to the normal pedestrian count, suggesting the situation had been resolved. The drop in pedestrian count is also apparent in figure 11 however it is less exaggerated, meaning the variance has been stabilised. We have implemented a box cox transformation with a $\lambda = -0.226$, which suggests an inverse transformation, as apparent from figure 10 to figure 11. The variance in figure 11 increases in an increasing rate. We can see that this transformation causes the plot to shift upwards. Other values of λ have been tested but do not have much significant changes, the Guerrero function gives us the best possible transformation for box-cox.

R code for Problem 3:

```
# Problem 3
# Tobacco from aus_production (before transformation)
aus_production %>%
autoplot(Tobacco) + labs(y = "Tobacco and cigarette production in tonnes", title
                        = "Tobacco production in Australia")

# Tobacco from aus_production (Box-Cox transformation)
tobacco_lambda <- aus_production %>%
  features(Tobacco, features = guerrero)
tobacco_lambda
tobacco_box <- aus_production %>%
  autoplot(box_cox(Tobacco, 0.929)) +
  labs(y = "Box-Cox transformed tobacco production", title = "Tobacco production in Australia")
tobacco_box

#Economy class passengers MEL-SYD (before transformation)
mel_syd_eco <- ansett %>%
  filter(Airports == "MEL-SYD", Class == "Economy")
mel_syd_eco %>%
  autoplot(Passengers) + ggtitle("Economy Class passengers between Melbourne and Sydney")

# Economy class passengers MEL-SYD (Box-Cox transformation)
mel_syd_eco <- ansett %>%
  filter(Airports == "MEL-SYD", Class == "Economy")
view(mel_syd_eco)
ansett_lambda <- mel_syd_eco %>%
  features(Passengers, features = guerrero)
ansett_lambda
ansett_box <- mel_syd_eco %>%
  autoplot(box_cox(Passengers, 2.00)) +
  labs(y = "Box-Cox transformed passengers",
       title = "Economy class passengers between Melbourne and Sydney")
ansett_box

# Pedestrian counts (before transformation)
ped_count <- pedestrian %>%
  filter(Sensor == "Southern Cross Station")
ped_count %>%
  autoplot(Count) + labs(y = "Pedestrian counts",
                        title = "Pedestrian counts at Southern Cross Station")

# Pedestrian counts (Box-Cox transformation)
pedestrian_lambda <- ped_count %>%
  features(Count, features = guerrero)
pedestrian_lambda
pedestrian_box <- ped_count %>%
  autoplot(box_cox(Count, -0.226)) +
  labs(y = "Box-Cox transformed Pedestrian counts",
       title = "Pedestrian counts at Southern Cross Station")
pedestrian_box
```

Reference list

- Press, A.A. (2016). *Victorian regional trains banned from Melbourne lines after safety scare*. [online] the Guardian. Available at: https://www.theguardian.com/australia-news/2016/jan/16/victorian-regional-trains-banned-from-melbourne-lines-after-safety-scare [Accessed 5 Mar. 2021].
- Tobacco in Australia (2020). *17.5 Impact of tobacco control strategies on the Australian economy - Tobacco in Australia*. [online] Tobaccoinaustralia.org.au. Available at: https://www.tobaccoinaustralia.org.au/chapter-17-economics/17-5-impact-of-tobacco-control-strategies-on-the-australian-economy.
- Federal Reserve History (n.d.). *Stock Market Crash of 1987 | Federal Reserve History*. [online] Available at: https://www.federalreservehistory.org/essays/stock-market-crash-of-1987.