Contents

[Problem 2.10.3 1](#_Toc536424582)

[Autoplot 1](#_Toc536424583)

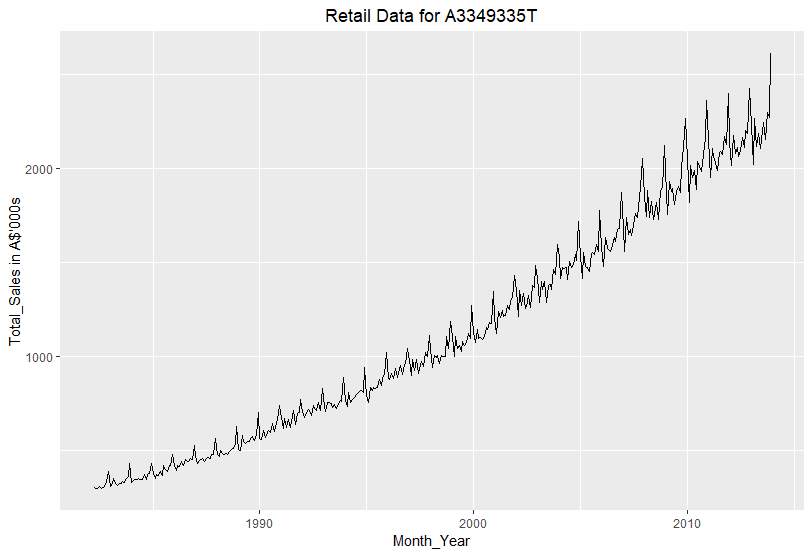
[Seasonplot() 1](#_Toc536424584)

[Lagplot() 3](#_Toc536424585)

[acf() 4](#_Toc536424586)

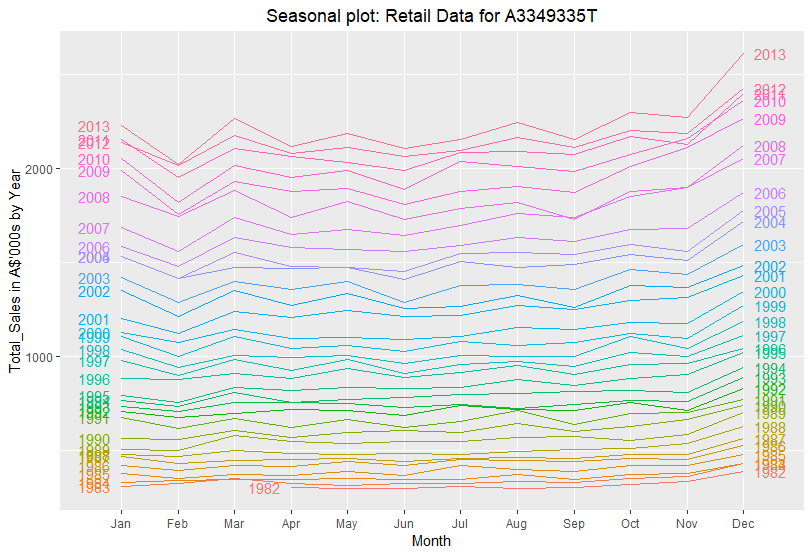
# Problem 2.10.3

## Autoplot

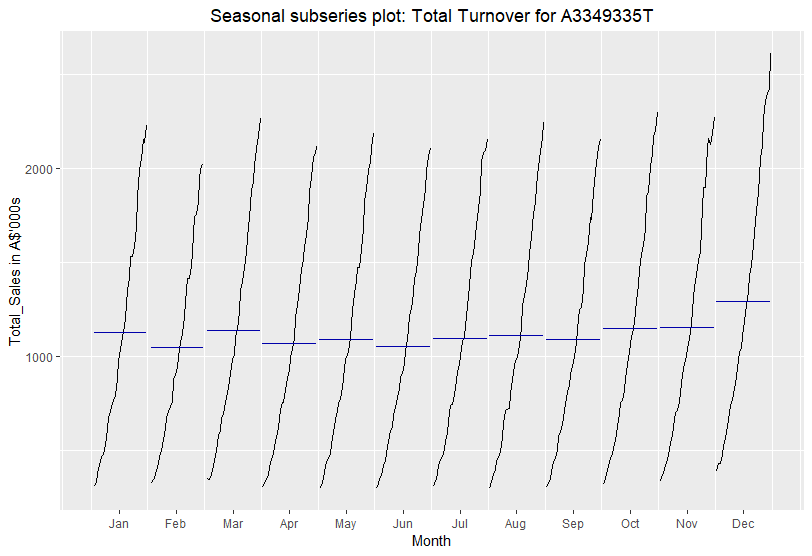


There is an obvious increasing trend with a clear seasonality whose magnitude increases as the year advances. However, there is no evidence of cyclic behavior in this chart. An interesting observation is that at the beginning of each year, there is a sudden drop in sales (which translates to “turnover” in Australian business terms means, as used in the data set). This appears to be related to the spike in shopping during holidays at year-end, which makes the comparative volume of sales at beginning of year look much less.

## Seasonplot

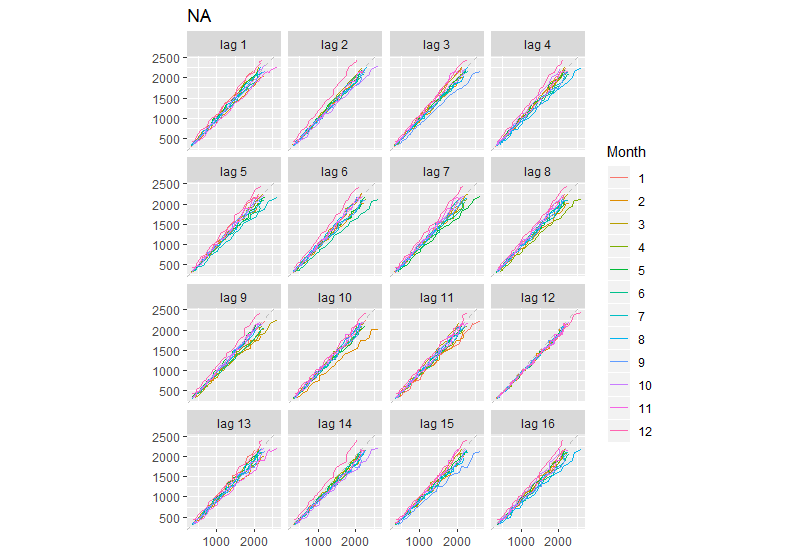


The seasonality appears to be quite similar across all years, with the seasonality being flat for most months except the beginning and the end of each year. For all the years in the data shown above, the sales spiked in December due to the shopping seasonality effect. Starting with year 1997, there appears to be a noticeable decline in sales in February of each year, which as noted above could be due to the spikes in sales during the holiday season comprised of the months of December and January. The data set appears to have started in April of 1982.



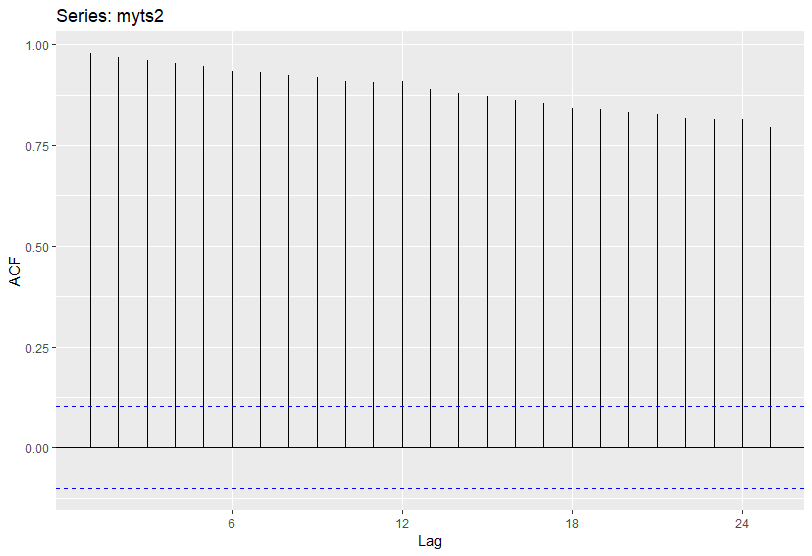
As previously observed, the mean for December is indeed higher than those of other months, with all months displaying increasing sales over time.

## Lagplot



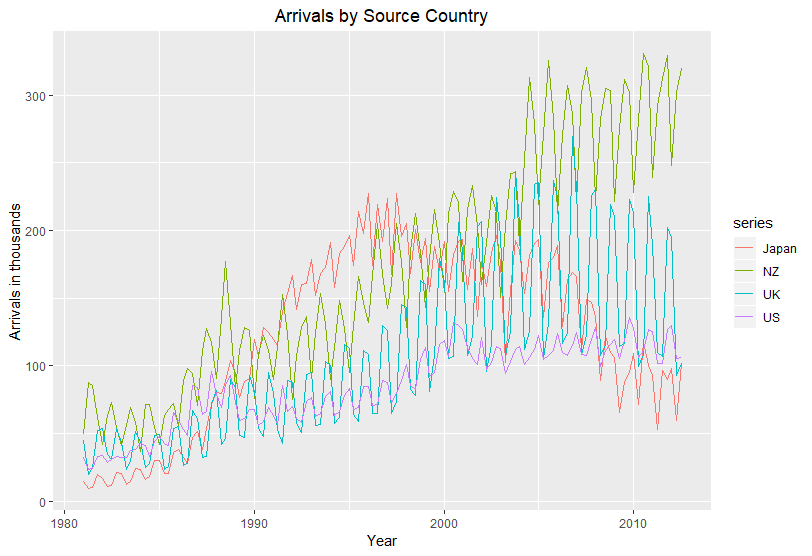
Using the window() function, I started with year 1983 since that is the first year with the fully available data, since the data collection appears to have started in April 1982. For all months, the relationship is positive, with lag 12 showing the strongest seasonality effect.

## acf



The autocorrelation measures for all lags are significantly higher than the dashed blue line, which indicates that the correlations are significantly different from zero. Small lags are larger than larger lags due to the increasing trend in the sales data. This is explained by the fact that observations nearby in time are also close to each other’s size, which leads to the autocorrelations of the data having positive values decrease slowly as lags increase. The data are seasonal, in particular, r1 is higher than that for other lags given the relative flat time series for most months for all years except spikes at the year-ends.

# Problem 2.10.7

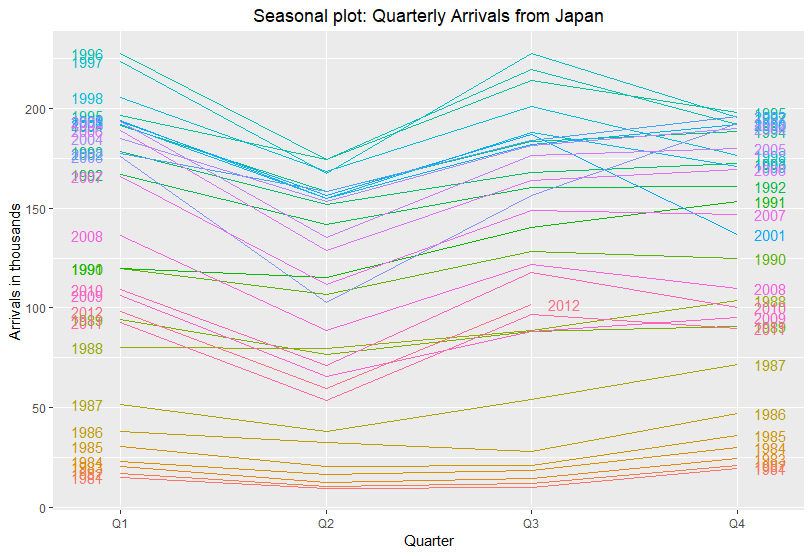


There appears to be a seasonality in the time series for all countries. With the exception of New Zealand, the peaks occur in the fourth quarter and subsequently falling in the first quarter of the following year. This makes sense given that Australia is in the southern hemisphere, and so December would be a summer month there, while it would be winter month in the three countries except New Zealand. Unsurprisingly, for New Zealand, the peak quarter is 3, which would be spring for both countries. There is a generally upward trend for all but Japan, for which the trend seems to have reversed in late 1990s. Historically, this makes sense since Japan was swept up in the Asian Financial Crisis that started in 1997 and lasted until late 1998, following which Japan began its longest-lasting stagflation in the first decade of 2000. There does not appear to be any sign of cyclicality, however.

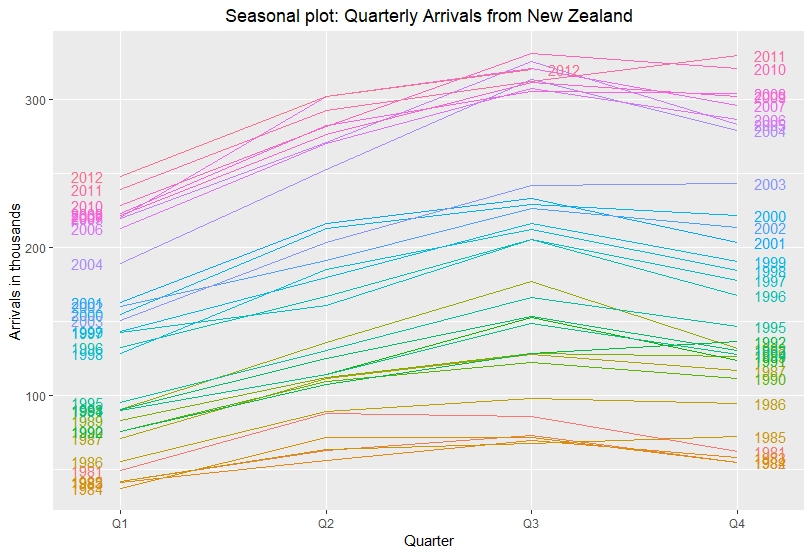
## Seasonal plots

Below is the seasonal plot for tourists arriving from Japan. Interestingly, the number of arrivals was more or less flat throughout the year for all years until 1987, after which the arrivals started displaying a zigzagging pattern, starting high in Q1, then falling in Q2, then rising again in Q3, then falling in Q4.

This could be explained by Japan’s aging population. In the 1980s, the working-age people comprised the majority of the population, and hence most likely came to Australia throughout all seasons, especially in May through August when it’s summer in Japan, but winter in Australia. Such months are typically not very popular in Australia, but for younger tourists, skiing could have been an attraction.

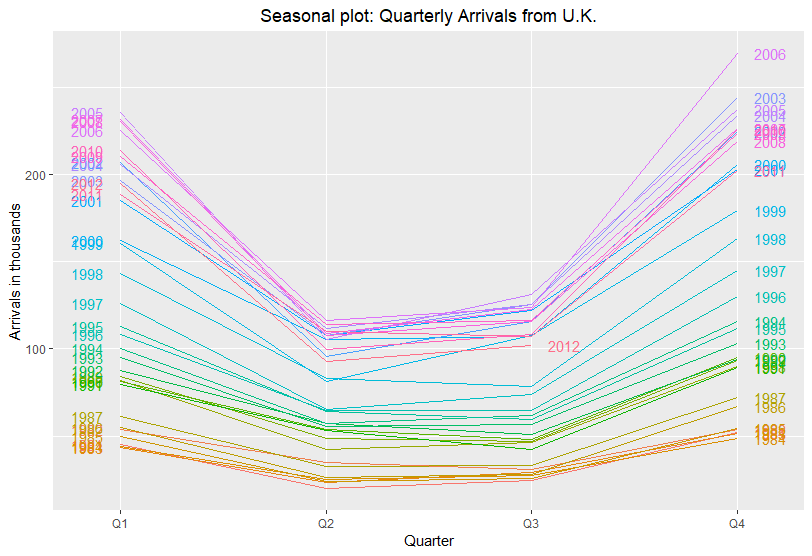


Below is the seasonal plot for tourists from New Zealand.

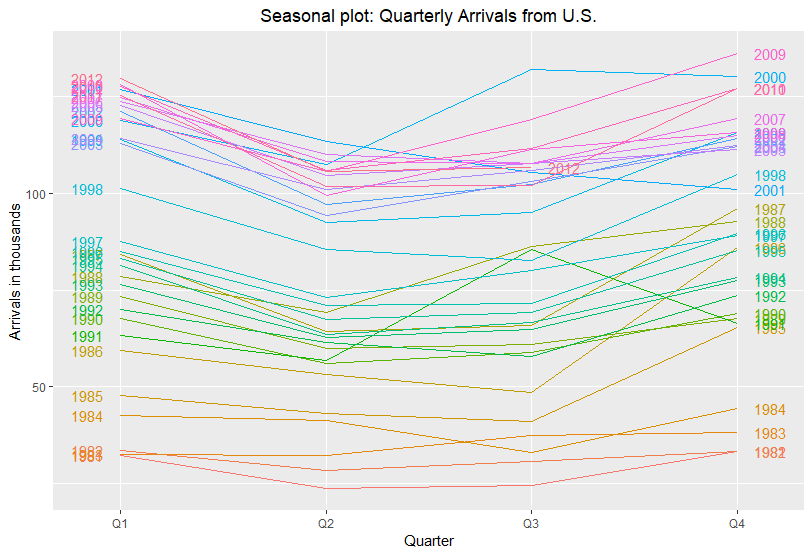


Unlike Japan, New Zealand tourists clearly exhibit a constant seasonality for all years, starting low in Q1 then gradually increasingly throughout the year for most years. For the first decade in 2000, this increasing trend in each year seems to have peaked in Q3 then falling slightly in Q4.

Below is the seasonal plot for the U.K. In stark contrast to the two earlier countries, the volumes of tourists from the U.K. seem to have followed a U-shaped pattern for all years, with the high number coming in Q1, then dipping sharply in Q2 and staying flat through Q3, then rising sharply in Q4. What is amazing is that this trend seems to have continued for all years in the sample, and hence makes U.K. a highly seasonal tourist country for Australia.

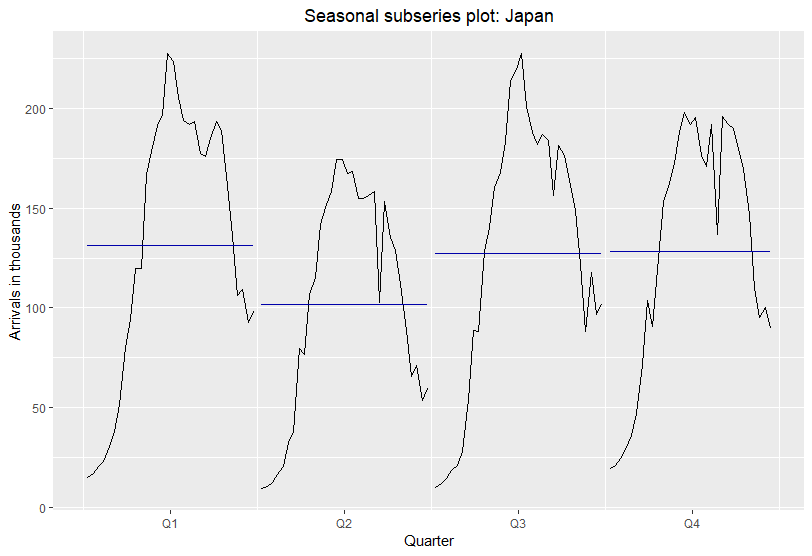


Below is the seasonal plot for tourists from the U.S.

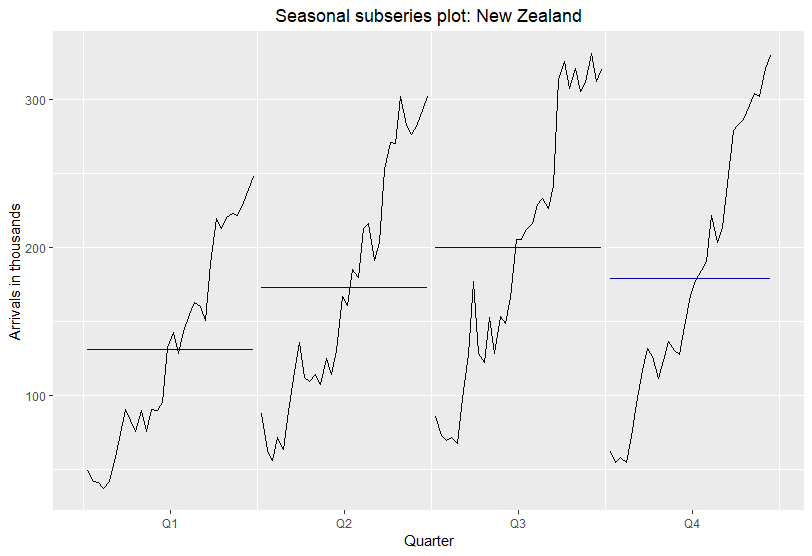


There is a clear upward trend in annual volume as seen by the clear shift between each line from earlier years to later years. The quarterly trend within each year is more varied than the one for the U.S., although the U-shaped density appears to have been followed for most of the years, with the exception of early 1990s. In particular, the sudden spike in Q3 of 1991 stands out as a clear anomaly. This is a bit strange given the fact that the oil price had surged due to the U.S. entering a war against Iraq during that time period. One would surmise that this would have made the cost of travel higher.

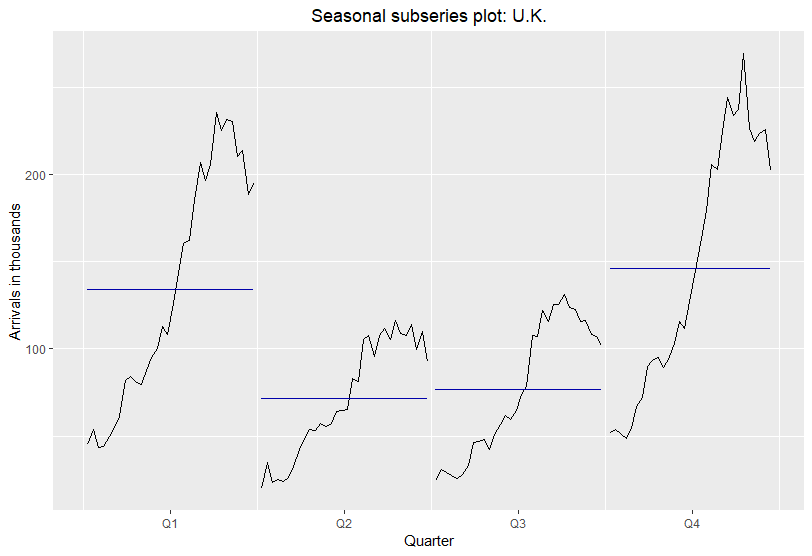
Below is the subseriesplot for Japan, which shows the inverted-U shaped arrival patterns for each quarter, peaking in the middle of each quarterly period.



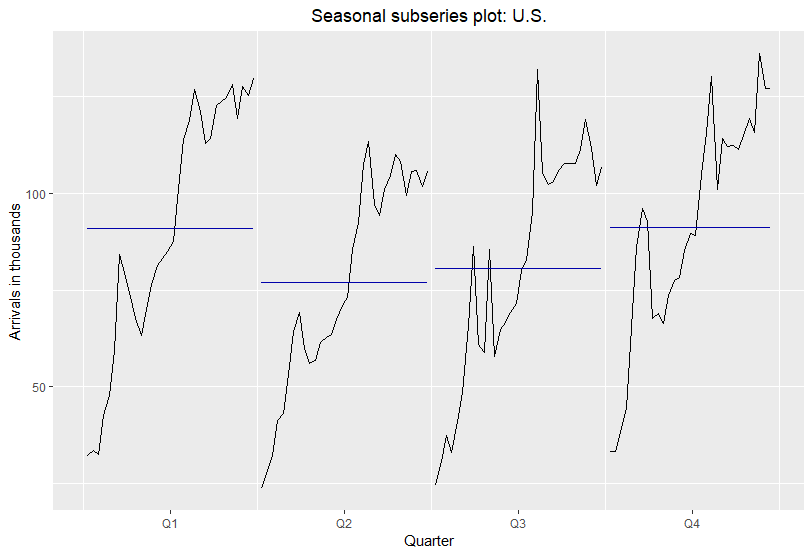
Below is the subseriesplot for New Zealand, which shows that in contrast to Japan, the trend within each quarter seems to be increasing, starting low in the beginning and peaking at the end of each quarter. This trend is a bit puzzling. While it is understandable that December would be the most popular month for Q4, it is less clear why there would be more tourists in June than in May or April in Q2, or in September than in August or July in Q3.



Below is the subseasonal plot for U.K., which is similar to the one above for New Zealand, except that the volume in each quarter peaks earlier than the end of the period. Again, it is not clear why this is the case for all quarters. For instance, it is certainly conceivable that tourists would prefer April over June in Q2 since it is warmer, but it is puzzling why there would be more tourists in November than in December, which is traditionally the most popular month in Australia for tourists.

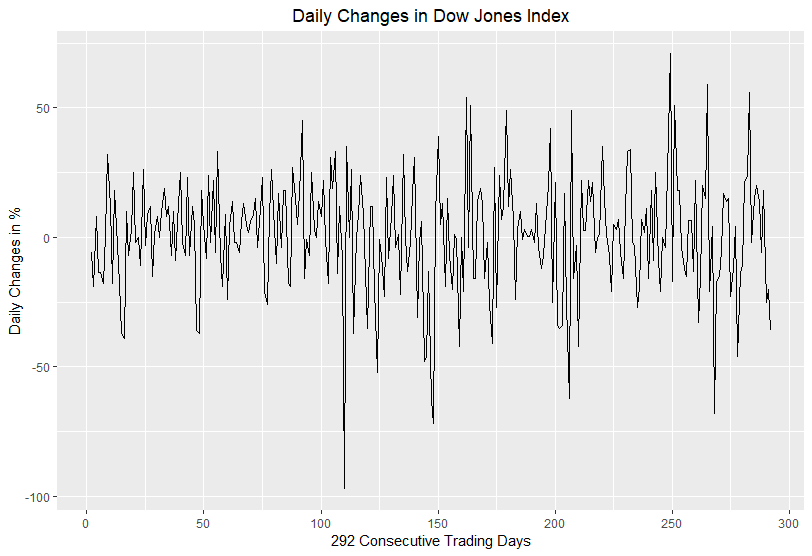


The U.S. subseries plot below shows more irregular pattern than the other countries. The peaks occur in the middle of each quarter for Q2 and Q3, but at the ends for Q1 and Q4. This makes sense for Q4, but not so for Q1 since one would expect January to be the most popular for U.S. tourists.

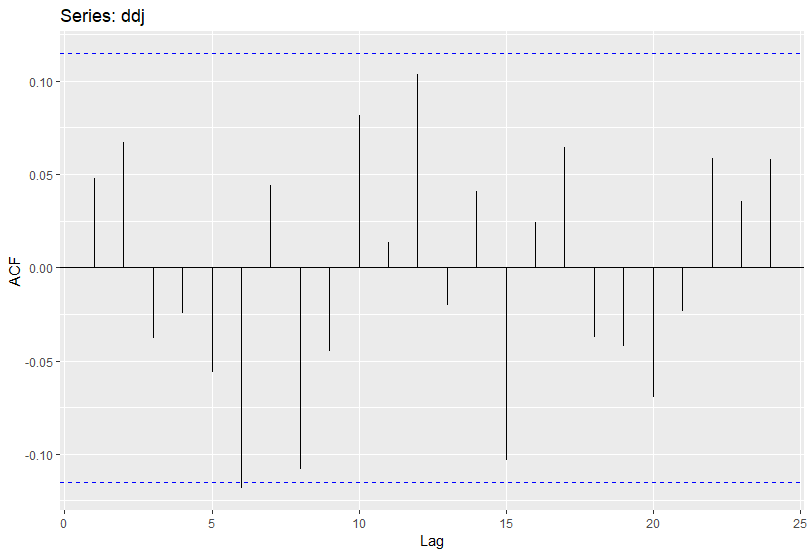


# Problem 2.10.10

Below is the autoplot of ddj, the daily changes in the Dow Jones index. The plot shows that the changes occurred in random fashion, with higher magnitude changes occurring through the period.



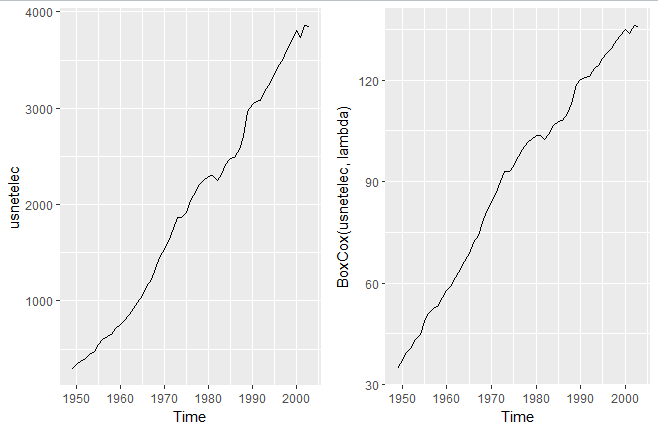
Below is the acf chart of ddj, which shows no apparent patterns. As expected, with the exception of r\_6, no other autocorrelation coefficients surpass the dotted blue line, which implies that the autocorrelations of daily changes are statistically zero at the 5% level. Regarding r\_6 lying outside the bound, it is expected given the 5% significance level that 1 or 2 of the 25 observations would lie outside the 5% bound. Hence, the daily changes appear to be white noise.



# 3.10.1

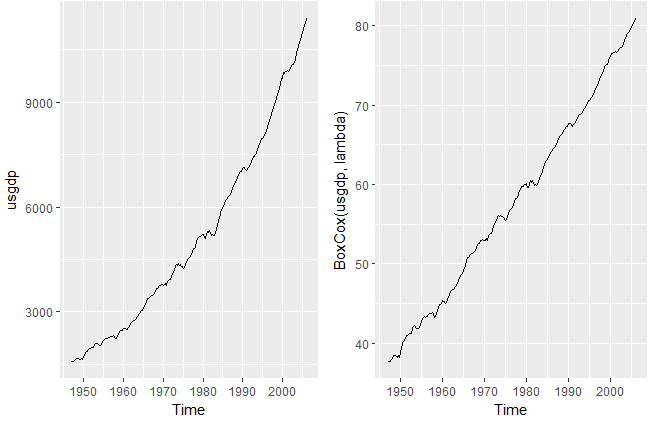
The Box-Cox parameter for usnetelec is 0.517 as shown below, which equates to the following square root transformation:

At the first glance, the differences between the two graphs are very slight and the vicissitude of variance in the first graph appears to be minor, suggesting the Box-Cox transformation may not be necessary. At any rate, we can see that the differences among the magnitudes of the variance at various points along the X-axis have become less pronounced and the response values now appear more normally distributed.



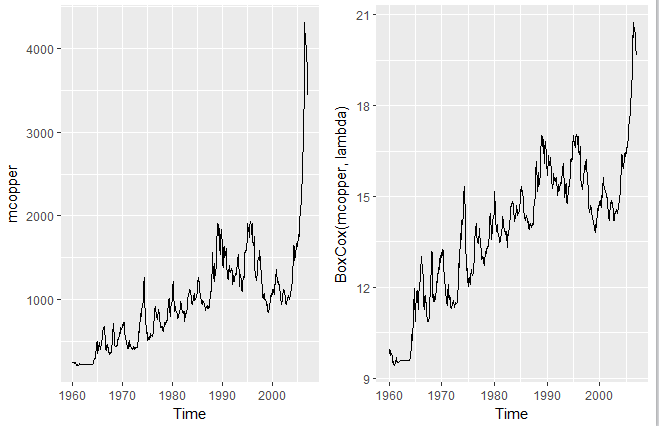
The lambda value for usgdp is 0.366, which equates to roughly a third root transformation:

Similar to the above, the vicissitude of variance in the original autoplot appears minor and the graph looks virtually linear, indicating that the Box-Cox transformation is perhaps unnecessary. At any rate, the variance among the magnitudes in differences in y-values are less pronounced, and the graph looks more linear, making it easier to approximate the y-distribution using the Gaussian assumptions.



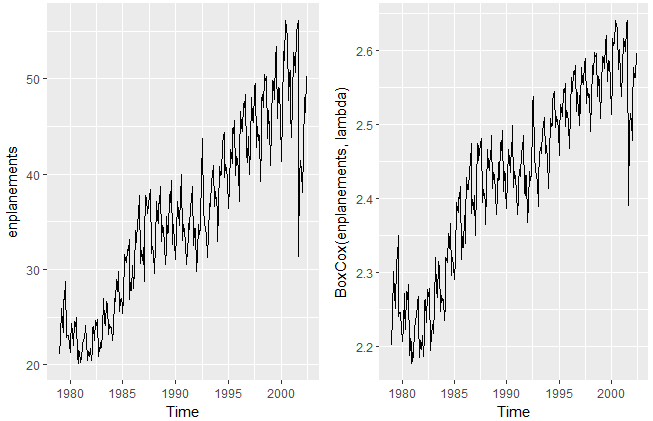
The lambda value for mcopper is 0.19, which equates to roughly a fifth root transformation:

To make response variable more amenable to normal distribution, the extreme values on the right tail have been reduced and the values on the left fat tail have been spread out more as shown below.



Finally, the lambda value for enplanements is -0.227, which roughly equates to an inverse fifth root as follows:

In the graph below, although not very obvious, the extreme values on the right tail have been brought closer together to the center.



# Problem 3.8

As can be seen below, the residuals have a unimodal distribution but appear slightly right skewed. From the ACF plot, we can see that the residuals have a significant autocorrelation across lags, so we cannot say that the residuals are uncorrelated.

