HA 2.3

Jean Jimenez

### HA 2.3

Download some monthly Australian retail data from [the book website](https://otexts.com/fpp2/extrafiles/retail.xlsx). These represent retail sales in various categories for different Australian states, and are stored in a MS-Excel file.

#### A

Read the data into R with the given script

##### Work

After downloading the excel file from the book website, I imported the retail data using readxl package’s function read\_excel(), which functions similarly to read\_csv() but for excel files. As stated in the question problem, the skip=1 is necessary because the excel file has 2 headers.

retaildata = readxl::read\_excel("retail.xlsx", skip=1)

After importing the retail data, I went ahead and observed the data set. It has 380 observations/ rows and 190 columns. It is time series data.

#### B

Select one of the time series as follows (but replace the column name with your own chosen column)

##### Work

From the retail data set, I selected the A3349719C column to put through the timeseries ts() function. frequency argument of 12 was used because the data is monthly. The start argument of (1982,4) was set because April 1982 was the first data point. myts variable holds the time series object.

myts = ts(retaildata[,"A3349719C"],  
 frequency=12, start=c(1982,4))

#### C

Explore your chosen retail time series using the following functions:

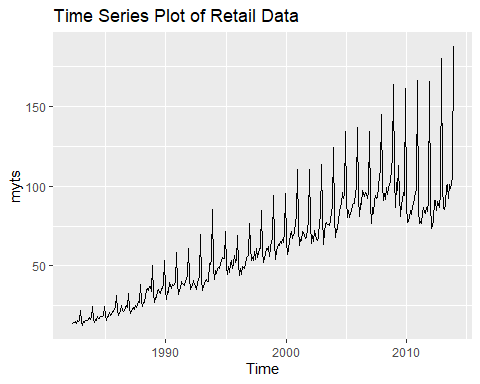
autoplot(), ggseasonplot(), ggsubseriesplot(), gglagplot(), ggAcf()

Can you spot any seasonality, cyclicity and trend? What do you learn about the series?

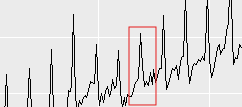
##### Work

###### Autoplot

First, I plot the retail data with autoplot() . This function is nice because it automatically plots a line graph for time series object.



The retail data has seasonality. You can almost see the yearly pattern clearly, which I can best describe as one wavelength:

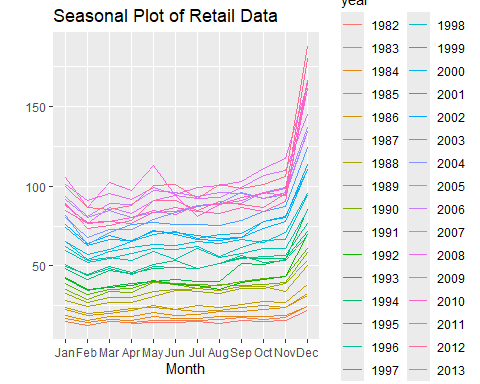


I can’t really tell anything about cyclicity using this plot, there might be some.

The trend of the graph is increasing. There are more sales over time.

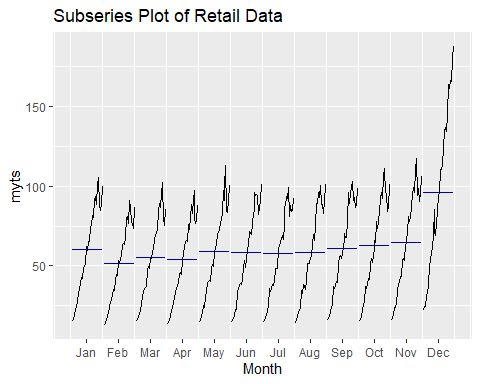
###### ggseasonplot

Next, I use ggseasonplot().



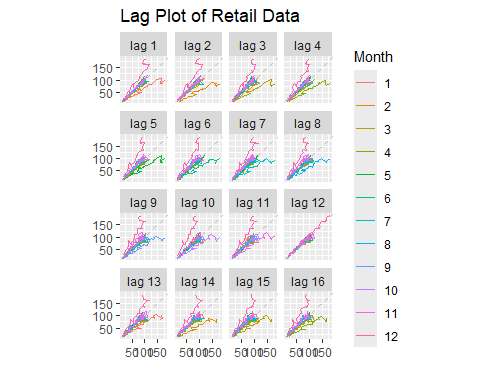
ggseasonplot() shows the distribution of the data year by year. This better shows seasonality, with the increase being obvious around November of each year.

###### ggsubseriesplot



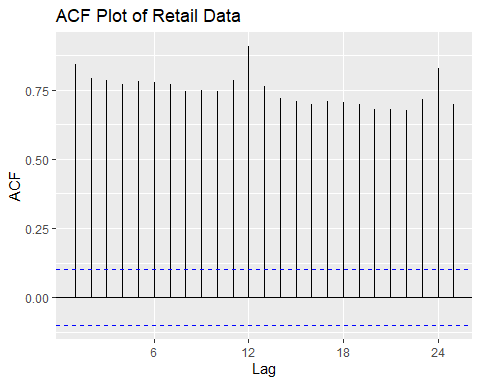
ggsubseriesplot() shows the distribution of the data for each of the months every year. The month of December’s mean is much higher than the mean of the other months, meaning December always had significantly increased sales than the earlier months.

###### gglagplot



gglagplot() produces a lag plot. For each lag plot, the pattern is repeating which supports seasonality.

###### ggACF



ggacf shows the autocorrelation plot. It supports seasonality with peaks every 12 months. It also shows significant autocorrelation since the bars are above the blue dotted lines.