Homework 1 – Exercise 2

2a) Download the file ‘tute1.csv’ and read the data into R

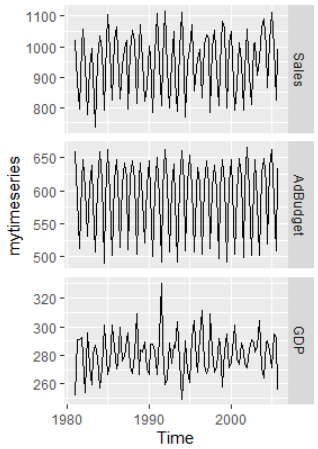
```
#Load data
tute1 <- read.csv("tute1.csv", header=TRUE)
View(tute1)
```

2b) Convert the data into timeseries

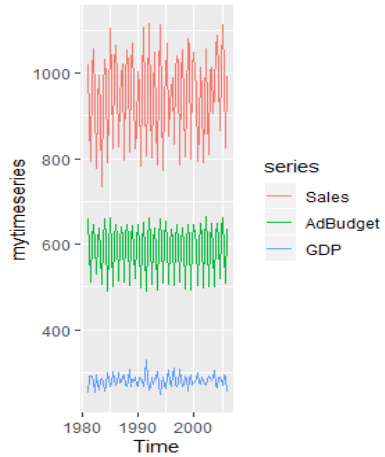
```
#Convert the data to time series
mytimeseries <- ts(tute1[,-1], start=1981, frequency=4)
#The [,-1] removes the first column which contains the quarters
View(mytimeseries)
```

2c) Construct the plots of each of the three series

```
#Construct time series plots of each of the three series
library(fpp2)
autoplot(mytimeseries, facets=TRUE)
```



What happens when you don’t include faces=TRUE?



Instead of drawing the 3 series into 3 different graphs, the 3 series are plotted in the same graph with the same y-axis. As a result, the scale in the y-axis are standardized across the 3 series unlike when plotted over 3 separate graphs. If the absolute variation is small, then the graph appears to be small which might give a wrong impression of the volatility in the series (for example, GDP). The graph of each series may overlap with another if the values lie within their range, resulting in the graph to be messy.

Homework 1 – Exercise 3

3a) Read the data into R

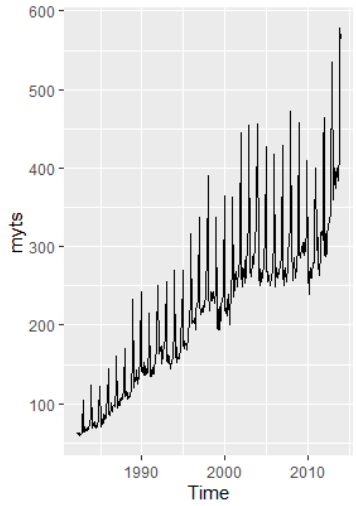
```
install.packages("readxl")
library(readxl)

#Load data
retaildata <- readxl::read_excel("retail.xlsx",skip=1)
#skip=1 because the Excel sheet has two header rows
```

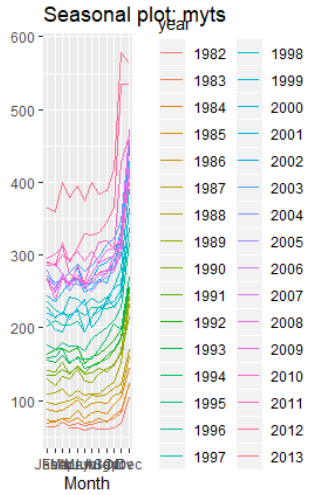
3b) Select one of the time series – A3349873A

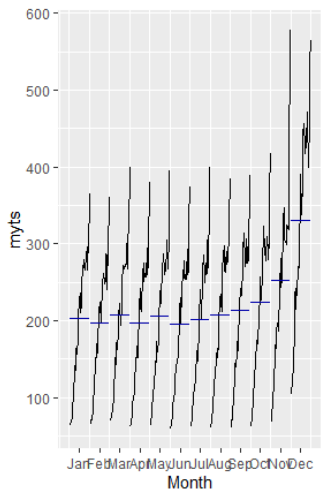
```
#Select a time series from one of the columns - A3349873A
myts <- ts(retaildata[,"A3349873A"],
           frequency=12, start=c(1982,4))
```

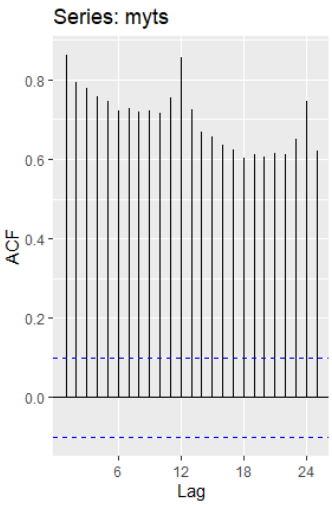
3c) Plot the graphs

```
autoplot(myts)
ggseasonplot(myts)
ggsubseriesplot(myts)
ggAcf(myts)
```









From the autoplot, we can see that there is a clear and increasing trend over the years, and a strong seasonal pattern that increases in size as level of the series increases. From the seasonal plots, there is a large jump in sales from November onwards across the years, and sales are beginning to jump from October onwards in the more recent 3 years (2011, 2012, 2013). While the average sales are

highest in December, they have also shown the greatest variation across the years as shown in the seasonal subseries plots. In the acf plot, the slow decrease in the ACF shows a clear trend as the sales are highly correlated with its lags while the ‘scalloped’ shape shows seasonality every 12 months.