# Data624 - Homework3

Amit Kapoor

02/18/2021

## Contents

6.2 . . . . . 1
6.3 . . . . . 9

```
library(fpp2)
library(dplyr)
library(gridExtra)
library(seasonal)
```

### 6.2

**The plastics data set consists of the monthly sales (in thousands) of product A for a plastics manufacturer for five years.**

```
?plastics
```

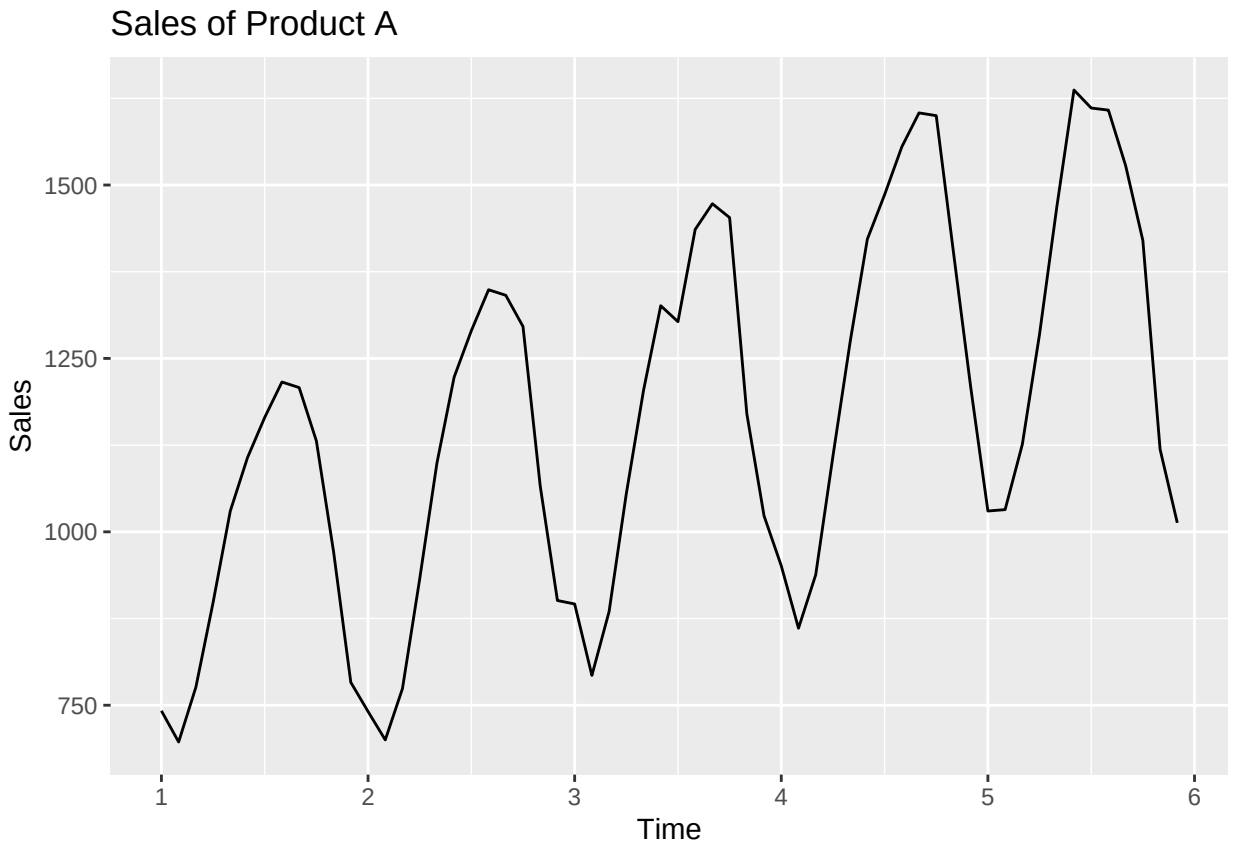
plastics - Monthly sales of product A for a plastics manufacturer.

```
glimpse(plastics)
```

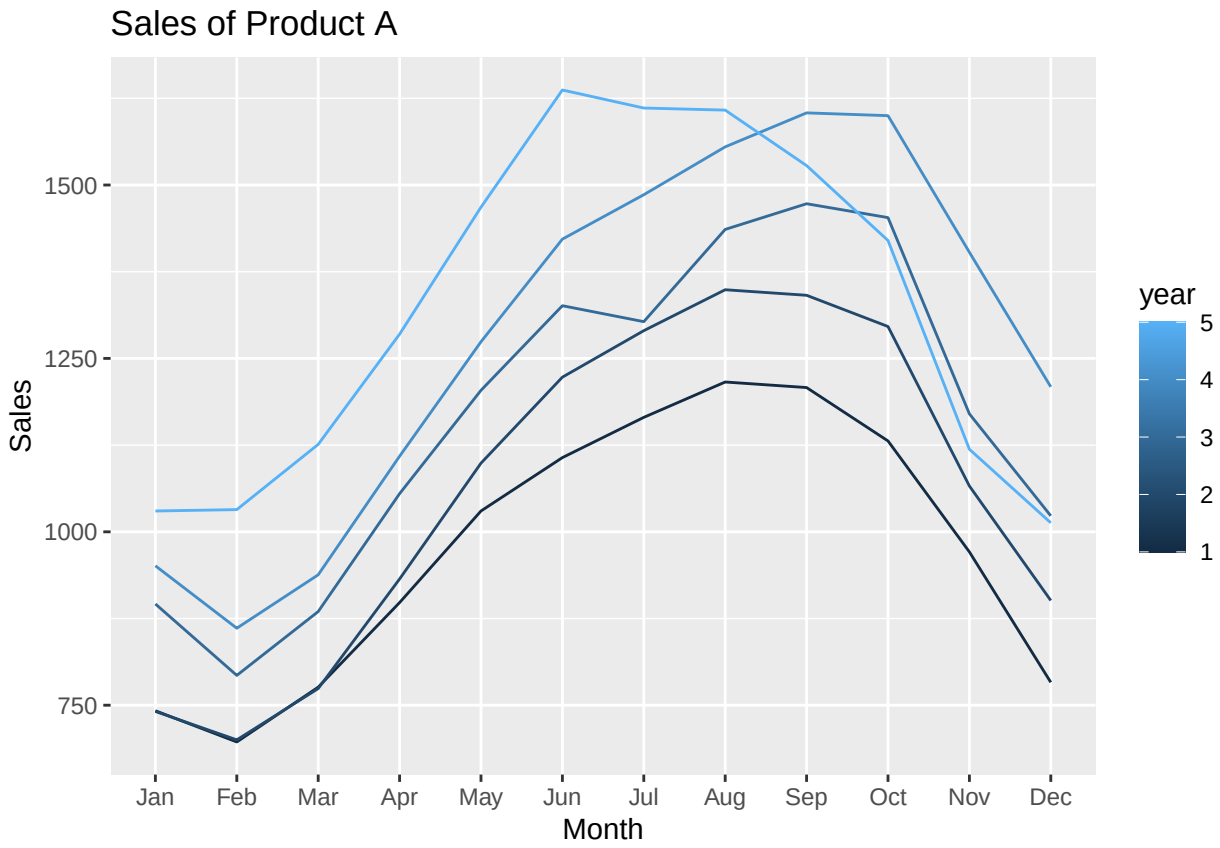
```
##  Time-Series [1:60] from 1 to 5.92: 742 697 776 898 1030 ...
```

- **Plot the time series of sales of product A. Can you identify seasonal fluctuations and/or a trend-cycle?**

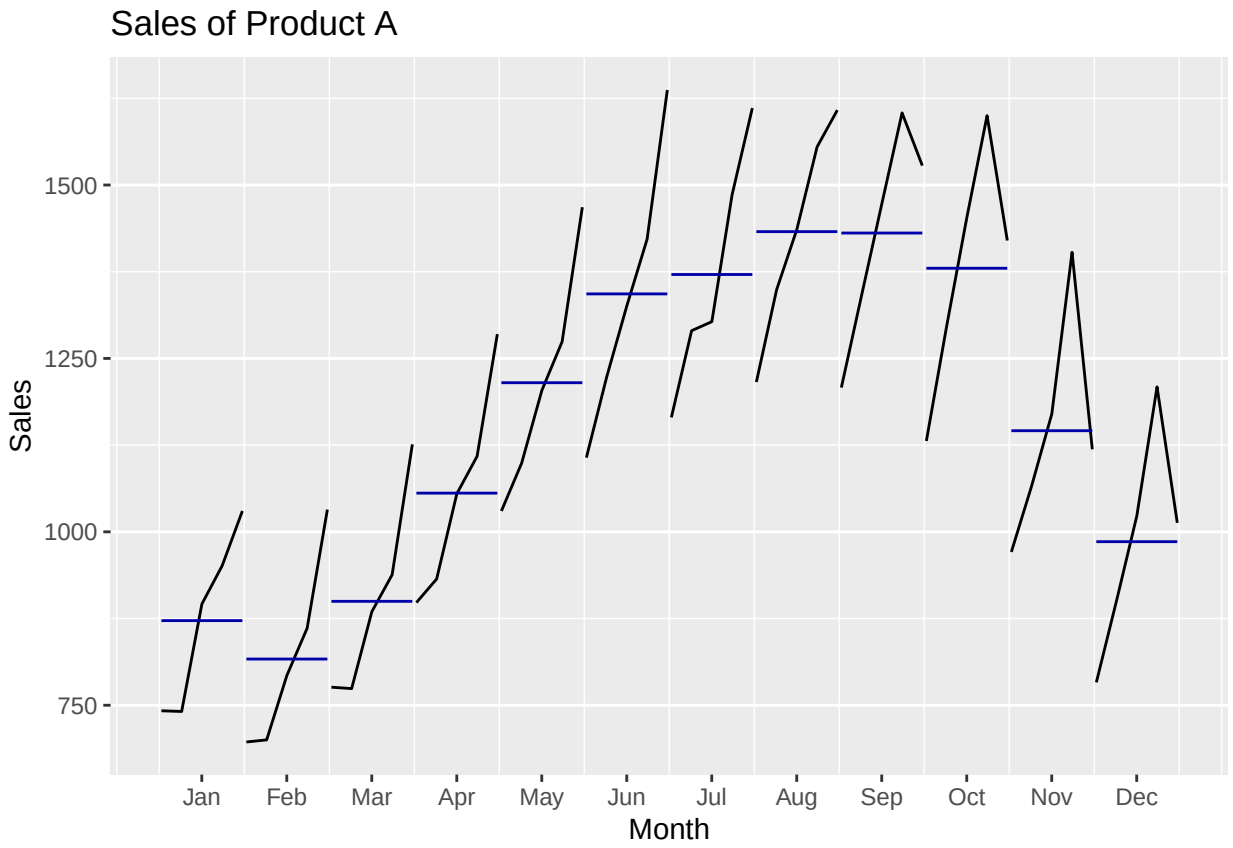
```
autoplot(plastics) +
  labs(title = "Sales of Product A", y="Sales")
```




```
ggseasonplot(plastics, continuous = T) +
  labs(title = "Sales of Product A", y="Sales")
```



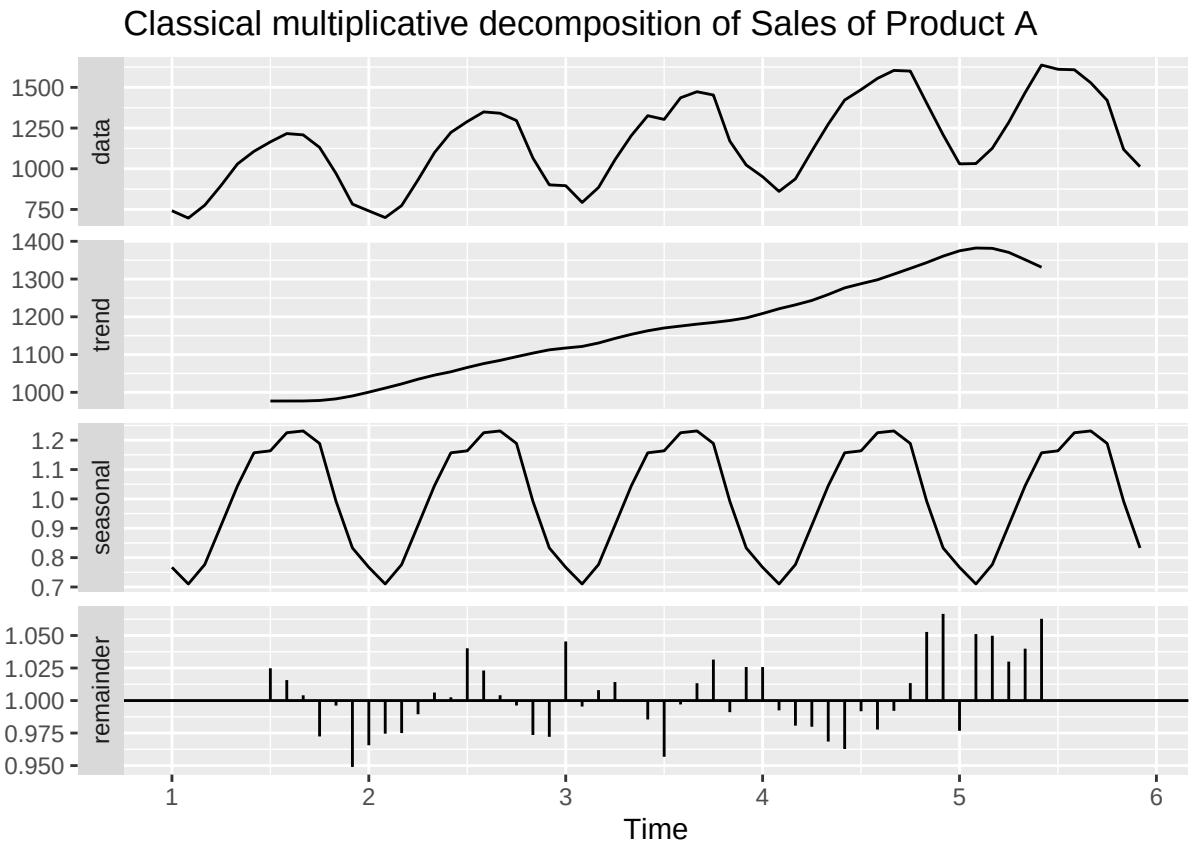

```
ggsubseriesplot(plastics) +
  labs(title = "Sales of Product A", y="Sales")
```




The plot above shows strong seasonality with a positive trend. There is no strong evidence of cyclic behavior. Also seasonal plot shows increase in sale that starts from Feb with peak in June and then declines..

- **Use a classical multiplicative decomposition to calculate the trend-cycle and seasonal indices.**

```
plastics %>% decompose(type="multiplicative") %>%
  autoplot() +
  ggtitle("Classical multiplicative decomposition of Sales of Product A")
```

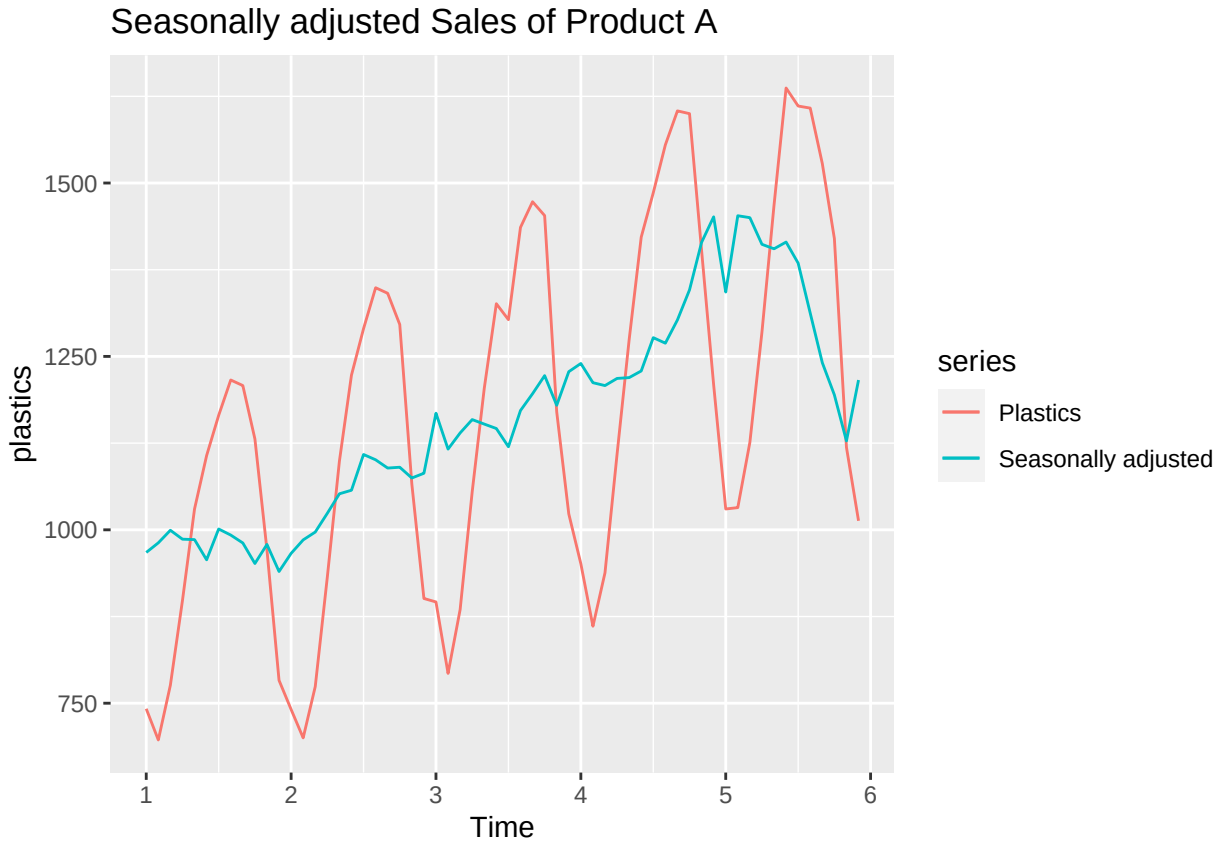



- **Do the results support the graphical interpretation from part a?**

The overall data shows the same upward trend as of part A. The seasonal plot here is not on the same X scale as in part A (which is in months) so a little difficult to interpret.

- **Compute and plot the seasonally adjusted data.**

```
plas_decom <- plastics %>% decompose(type="multiplicative")

autoplot(plastics, series = "Plastics") +
  autolayer(seasadj(plas_decom), series = "Seasonally adjusted") +
  ggtitle("Seasonally adjusted Sales of Product A")
```

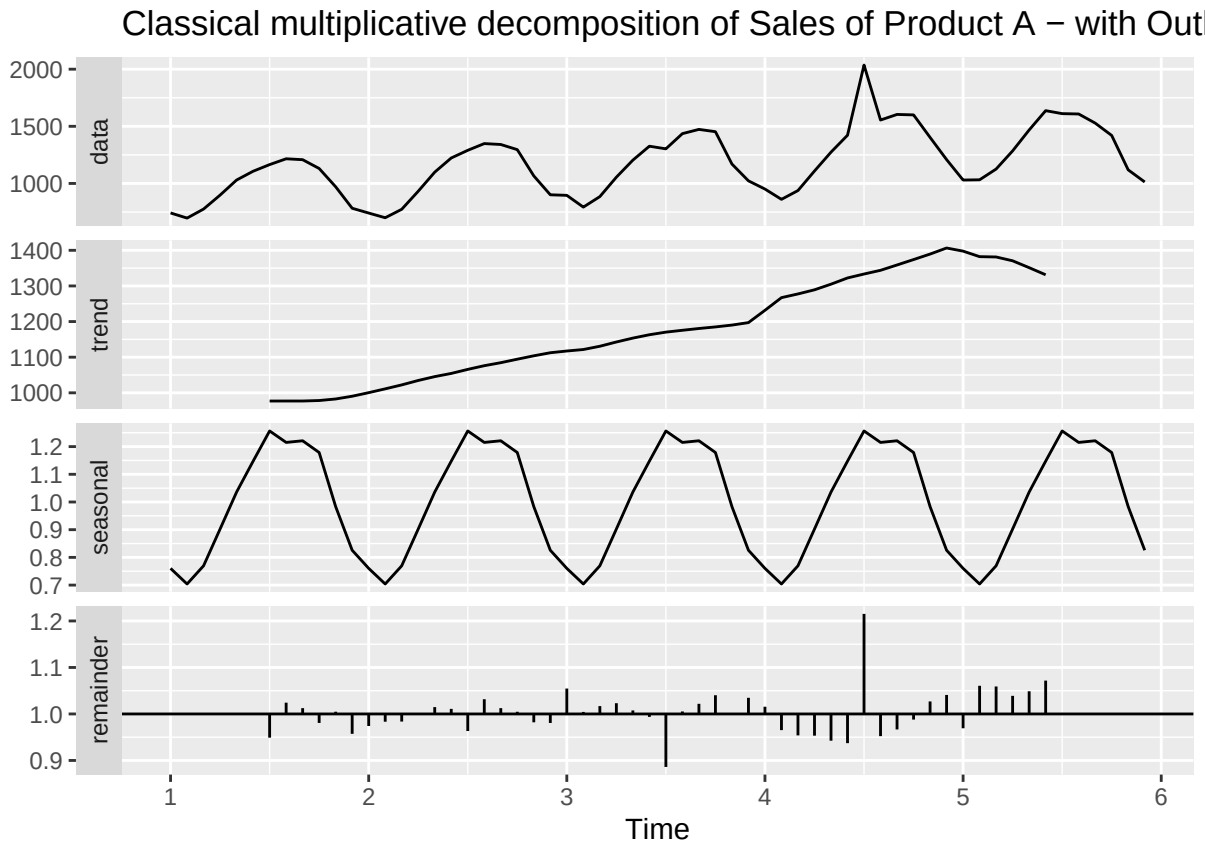



- **Change one observation to be an outlier (e.g., add 500 to one observation), and recompute the seasonally adjusted data. What is the effect of the outlier?**

```
plastics
```

```
##    Jan  Feb  Mar  Apr  May  Jun  Jul  Aug  Sep  Oct  Nov  Dec
## 1  742  697  776  898 1030 1107 1165 1216 1208 1131  971  783
## 2  741  700  774  932 1099 1223 1290 1349 1341 1296 1066  901
## 3  896  793  885 1055 1204 1326 1303 1436 1473 1453 1170 1023
## 4  951  861  938 1109 1274 1422 1486 1555 1604 1600 1403 1209
## 5 1030 1032 1126 1285 1468 1637 1611 1608 1528 1420 1119 1013
```

```
# Change one observation to be an outlier
plastics_outlier <- plastics
plastics_outlier[43] <- plastics_outlier[43] + 550

plastics_outlier %>% decompose(type="multiplicative") %>%
  autoplot() +
  ggtitle("Classical multiplicative decomposition of Sales of Product A - with Outlier")
```



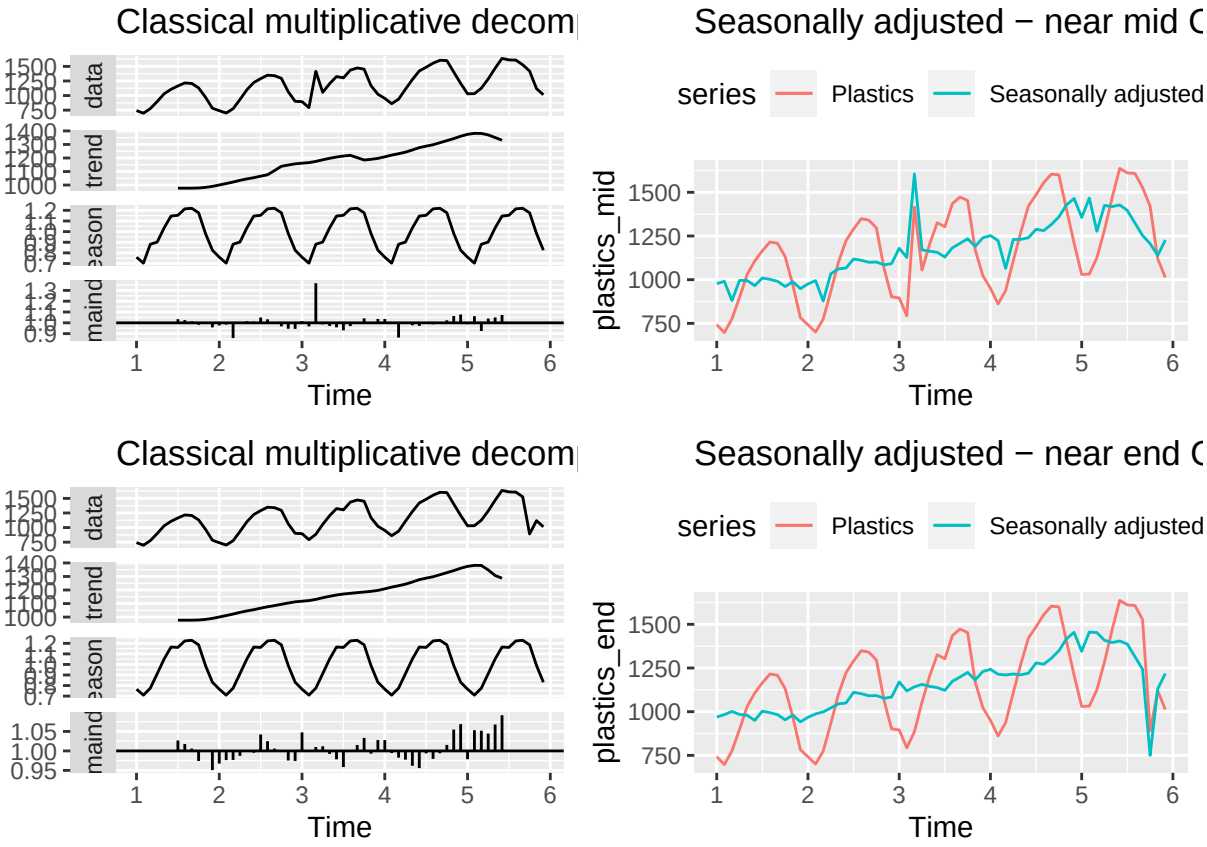

```
# decomposition
plas_out_decom <- plastics_outlier %>% decompose(type="multiplicative")

# plot data having outlier
autoplot(plastics_outlier, series = "Plastics") +
  autolayer(seasadj(plas_out_decom), series = "Seasonally adjusted") +
  ggtitle("Seasonally adjusted Sales of Product A - with Outlier")
```


We see outlier impacted the trend more compared to seasonality of time series data. Also we see a spike in seasonally adjusted data.

- **Does it make any difference if the outlier is near the end rather than in the middle of the time series?**

```
# middle outlier
plastics_mid <- plastics
plastics_mid[27] <- plastics_mid[27] + 530

m1 <- plastics_mid %>% decompose(type="multiplicative") %>%
  autoplot() +
  ggtitle("Classical multiplicative decomposition - near mid Outlier")

# mid decomposition
plas_mid_decom <- plastics_mid %>% decompose(type="multiplicative")

# plot data with mid outlier
m2 <- autoplot(plastics_mid, series = "Plastics") +
  autolayer(seasadj(plas_mid_decom), series = "Seasonally adjusted") +
  ggtitle("Seasonally adjusted - near mid Outlier") +
  theme(legend.position = "top")



# end outlier
plastics_end <- plastics
plastics_end[58] <- plastics_end[58] - 530
```

```
e1 <- plastics_end %>% decompose(type="multiplicative") %>%
  autoplot() +
  ggtitle("Classical multiplicative decomposition - near end Outlier")

# end decomposition
plas_end_decom <- plastics_end %>% decompose(type="multiplicative")

# plot data with end outlier
e2 <- autoplot(plastics_end, series = "Plastics") +
  autolayer(seasadj(plas_end_decom), series = "Seasonally adjusted") +
  ggtitle("Seasonally adjusted - near end Outlier") +
  theme(legend.position = "top")

grid.arrange(m1, m2, e1,e2, nrow=2, ncol=2)
```

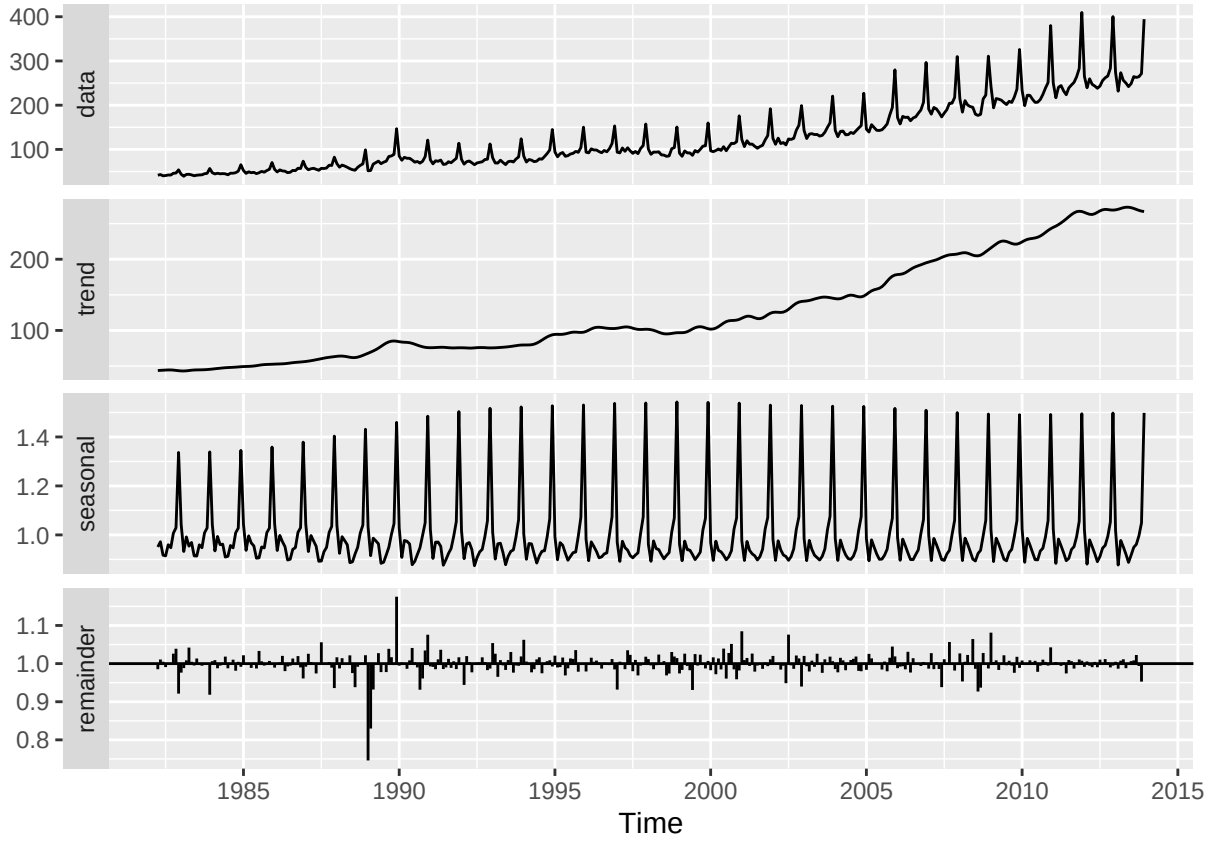



We have 2 outliers in above 2 sets of graphs. For the near middle outlier added 530 and for the near end outlier subtracted 530. It is clearly visible that no matter where the outlier added, it effects the time series trend. Almost no impact on seasonality though its variance differ based on outlier. We also notice spikes in seasonality adjusted plots based on outlier. For end outlier plots, remainder doesnt capture the outlier since classical decomposition doesn't include end points.

### 6.3

**Recall your retail time series data (from Exercise 3 in Section 2.10). Decompose the series using X11. Does it reveal any outliers, or unusual features that you had not noticed previously?**

```
retaildata <- readxl::read_excel("retail.xlsx", skip=1)
myts <- ts(retaildata[,"A3349627V"], frequency=12, start=c(1982,4))


myts %>% seas(x11="") %>% autoplot()
```




X-11 decomposition reveals trend and seasonality remain consistent but remainder does show spikes near the year 1990, confirming outlier. Trend seems increasing over the time. This retail data shows strong trend and seasonality.