Data 624 - HW 1

Coco Donovan

2024-06-03

Loading Necessary Packages:

```
library(fpp2)
```

```
## Registered S3 method overwritten by 'quantmod':
##   method            from
##   as.zoo.data.frame zoo

## -- Attaching packages ---------------------------------------------- fpp2 2.5 --

## v ggplot2   3.4.4      v fma       2.5
## v forecast  8.22.0     v expsmooth 2.3

## Warning: package 'ggplot2' was built under R version 4.3.1

## Warning: package 'forecast' was built under R version 4.3.1

##
```

```
library(forecast)
library(ggplot2)
```

HA - 2.1

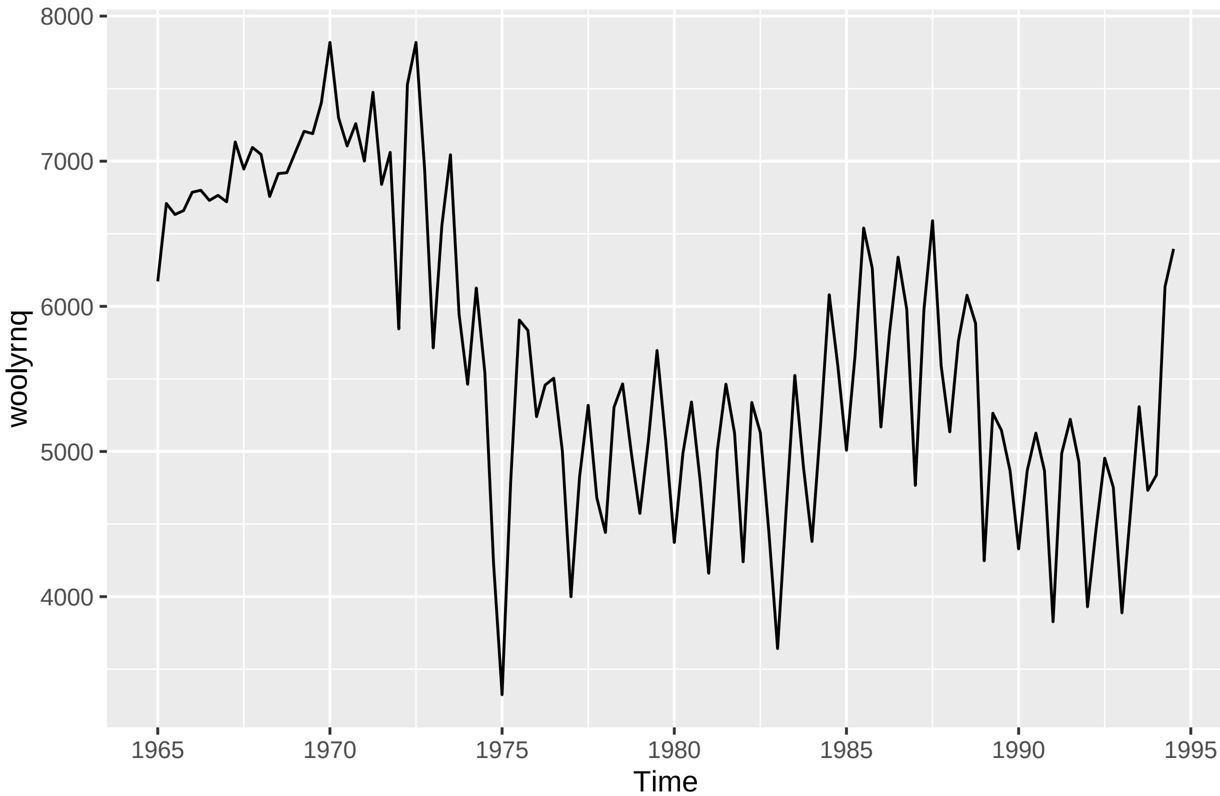
Use the help function to explore what the series gold, woolyrnq and gas represent.

Use autoplot() to plot each of these in separate plots. What is the frequency of each series? Hint: apply the frequency() function. Use which.max() to spot the outlier in the gold series. Which observation was it?

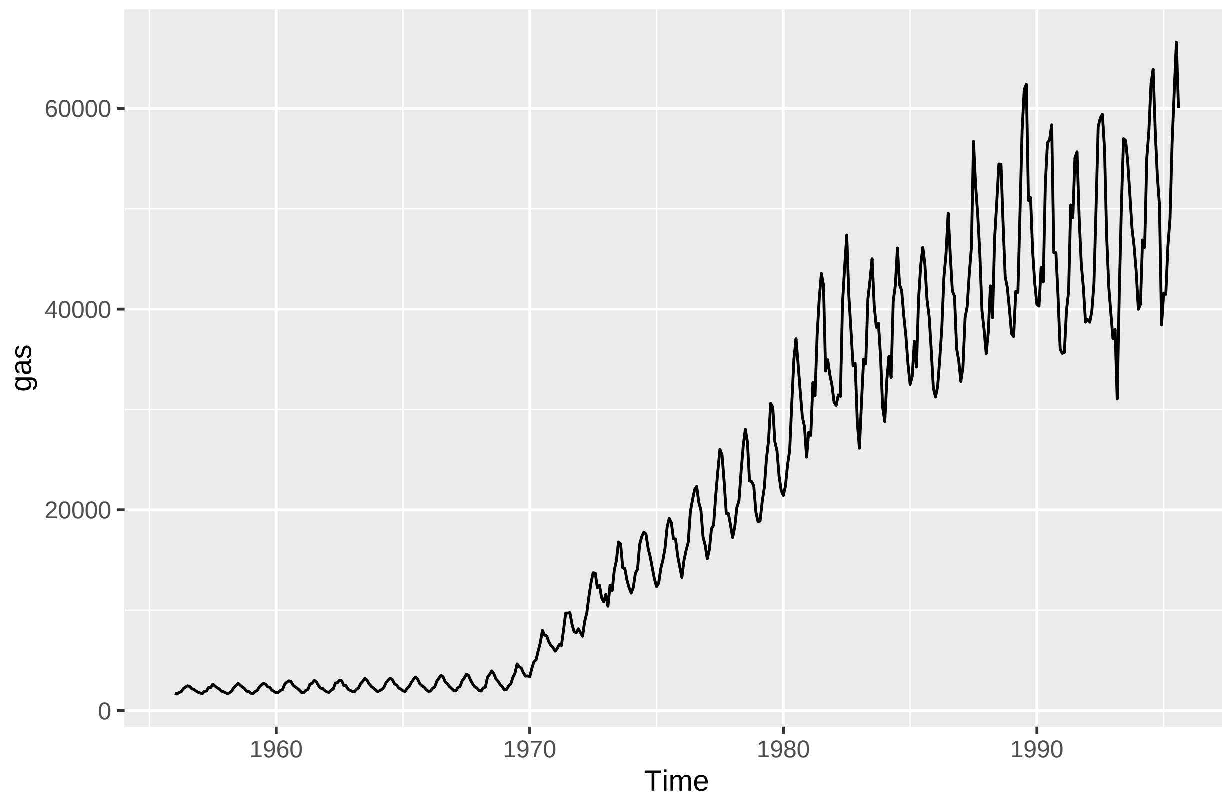
```
autoplot(gold)
```

autoplot()

```
autoplot(woolyrnq)
```




```
autoplot(gas)
```




```
frequency(gold)
```

frequency()

```
## [1] 1
```

```
frequency(woolyrnq)
```

```
## [1] 4
```

```
frequency(gas)
```

```
## [1] 12
```

The frequencies of the referenced time series are as follows:
- gold: 1
- woolyrnq: 4
- gas: 12

HA - 2.3

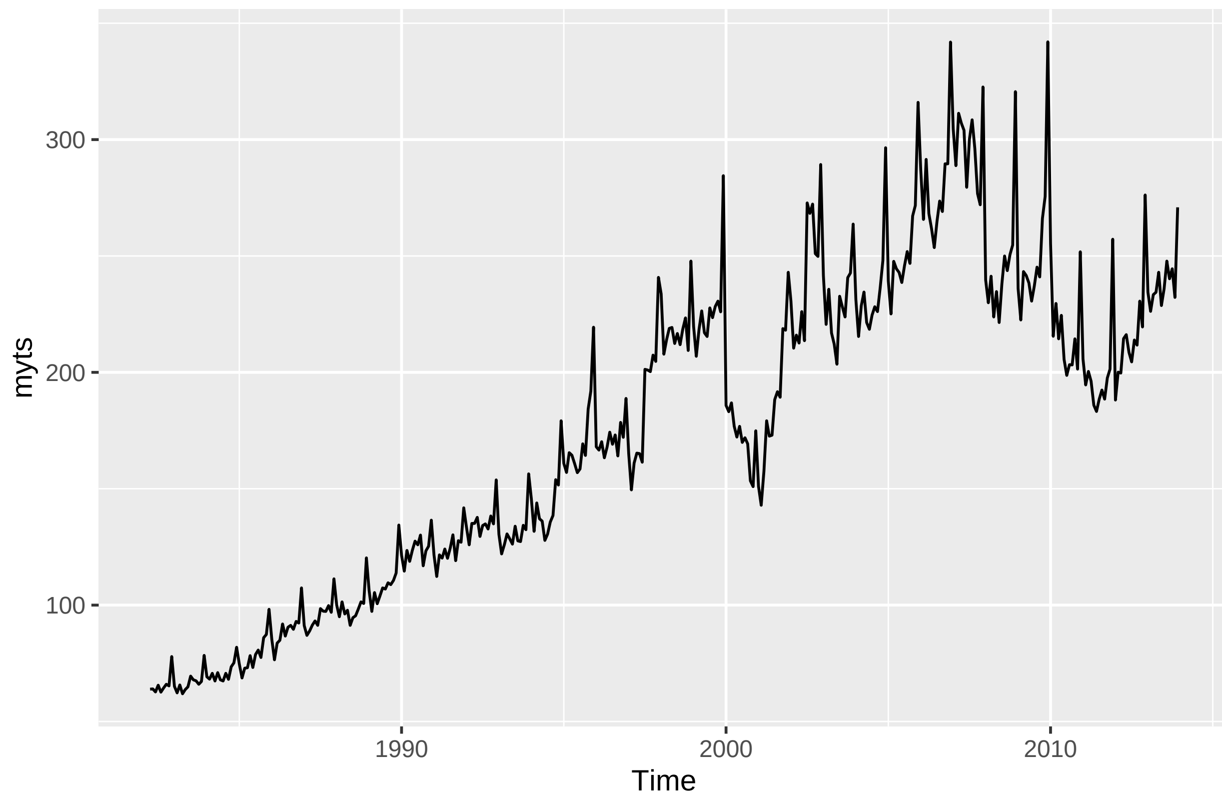
```
retaildata <- readxl::read_excel("retail.xlsx", skip=1)
```

a.

```
myts <- ts(retaildata[,"A3349338X"],
  frequency=12, start=c(1982,4))
```

b.

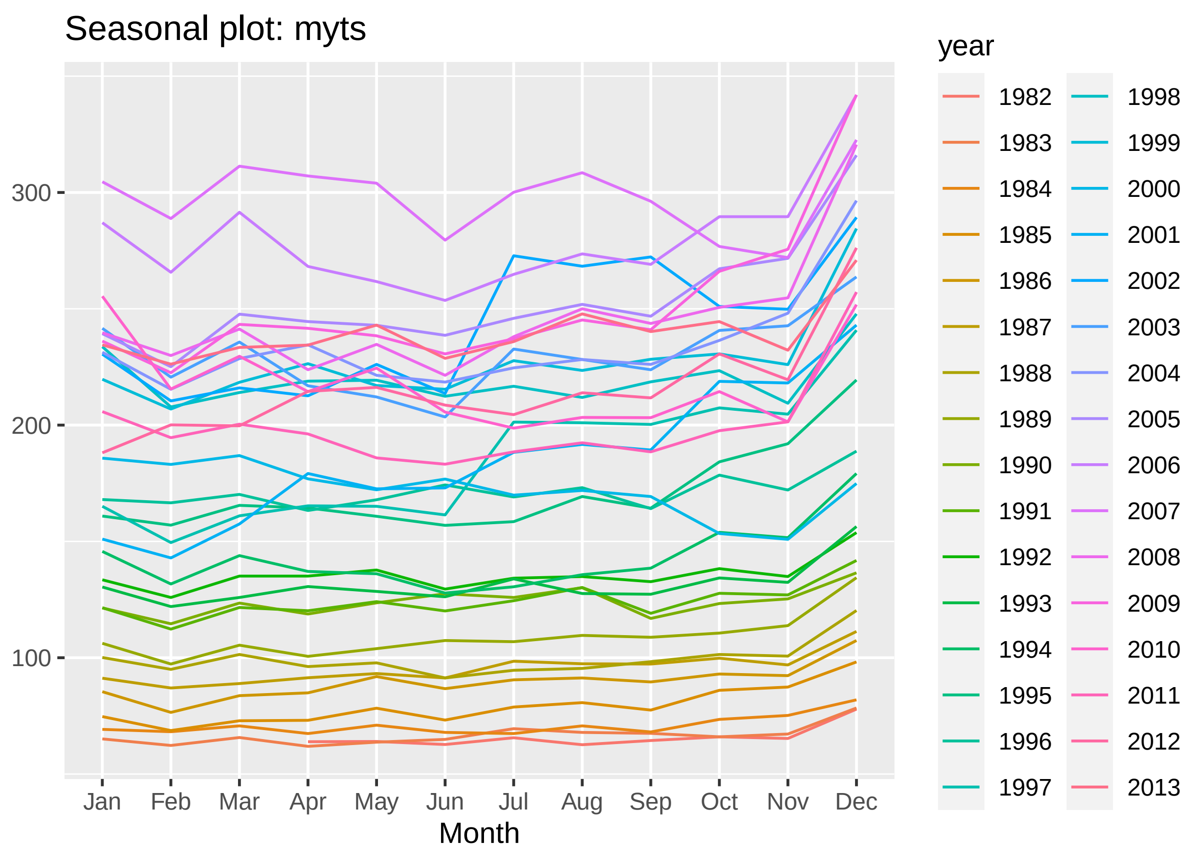
```
autoplot(myts)
```



c.

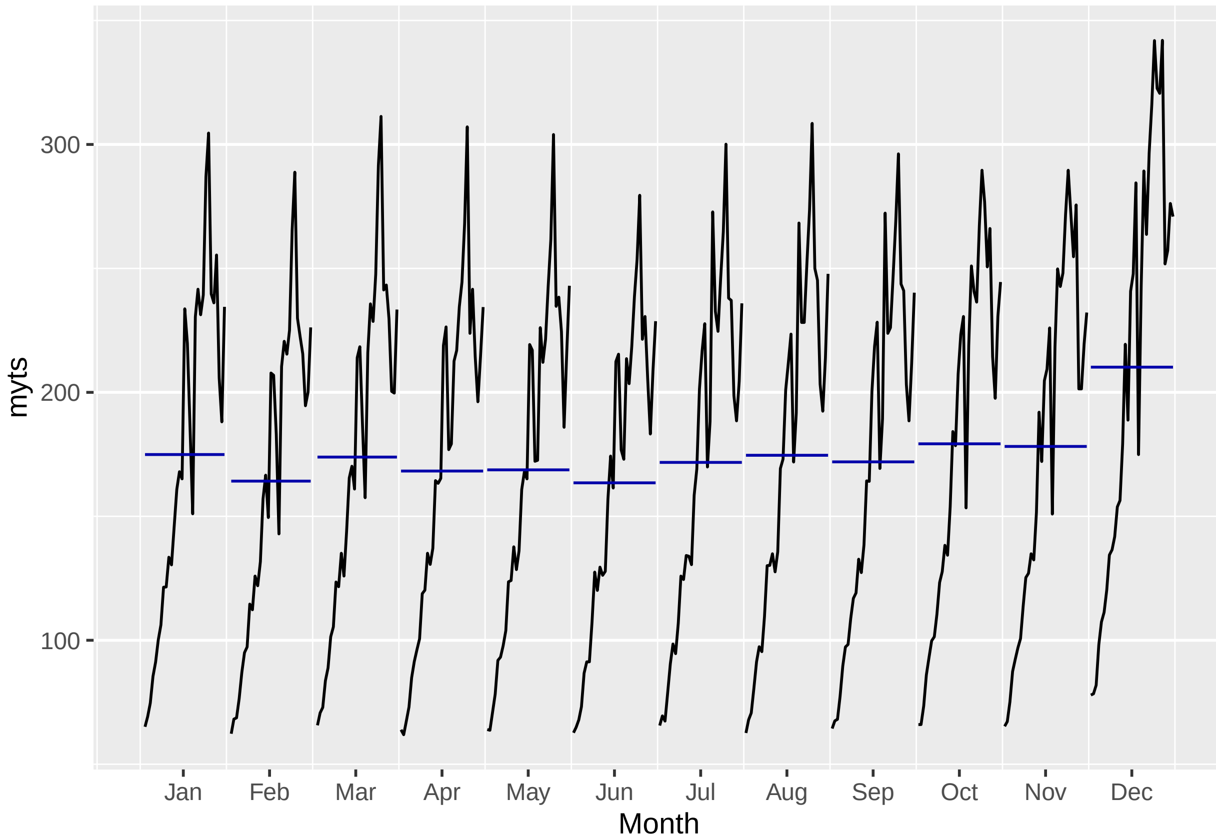
When looking at the autoplot(), it seems that the retail data has a general increasing trend since the beginning of data being recorded. From this plot and my untrained eyes, it is unclear whether there is a cyclical pattern present (although, I would lean towards saying there is no cyclical nature to this data).

```
ggseasonplot(myts)
```



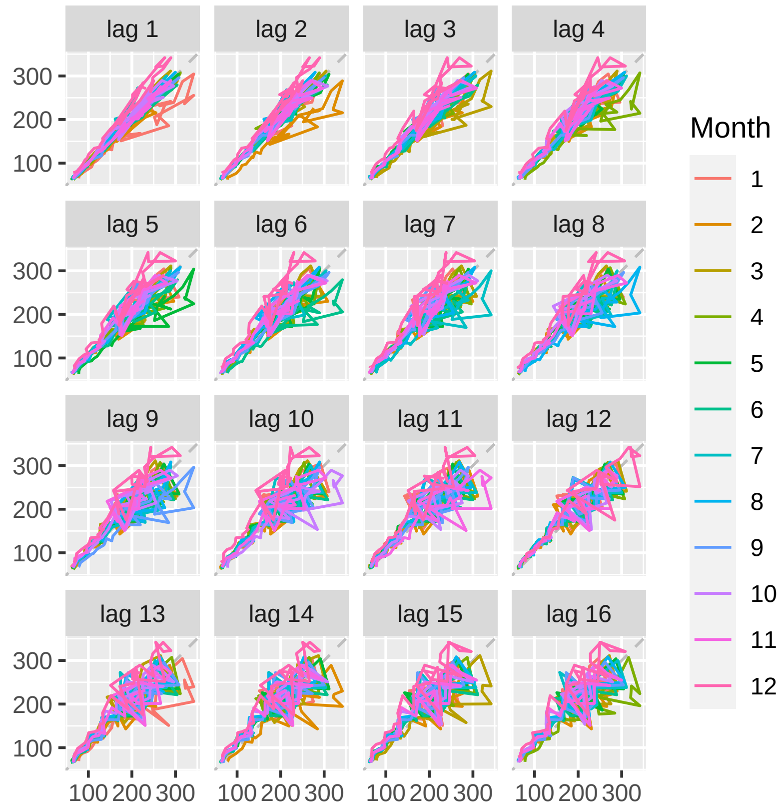
When looking at the seasonal plot, it is clear that there are some seasonality that occurs with A3349338X. For example, there is typically an increase that occurs from November to December, and a decrease from January to February.

```
ggsubseriesplot(myts)
```

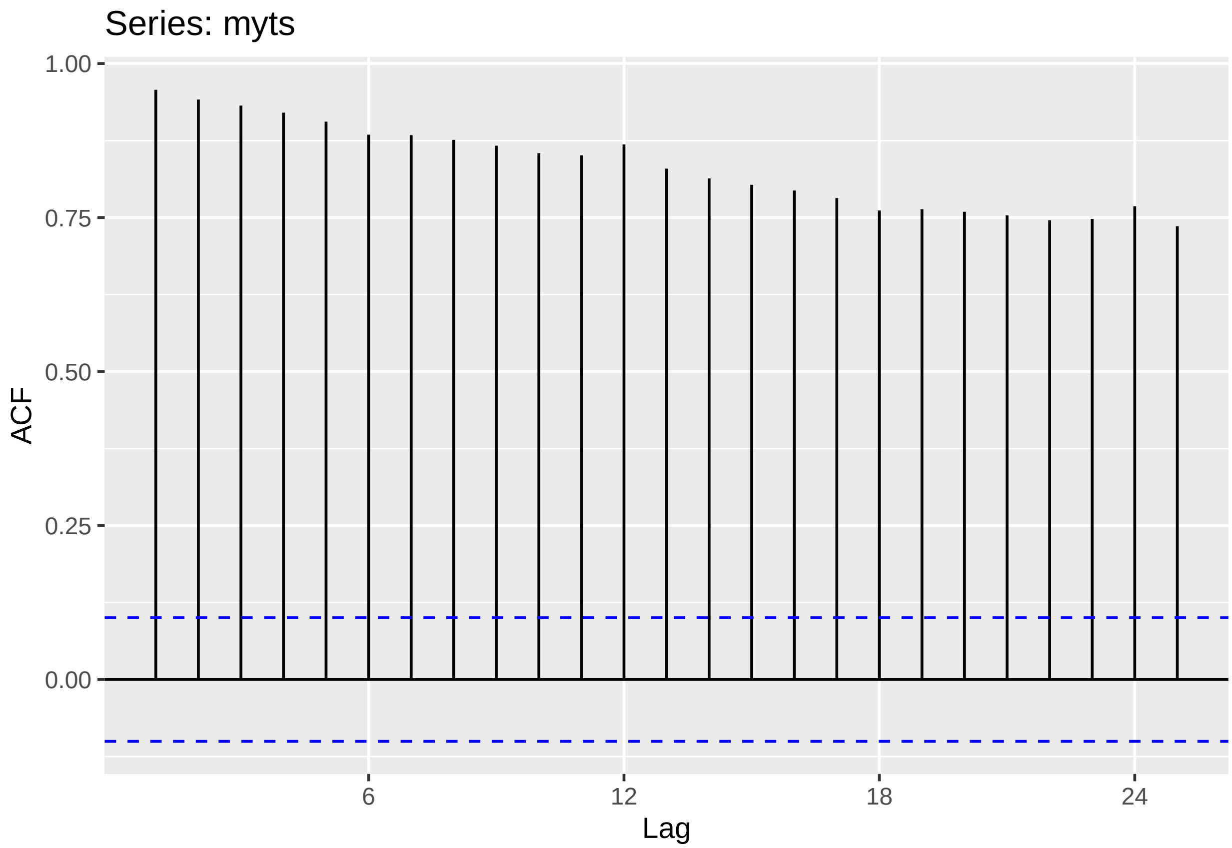


These observations concerning seasonality are confirmed when I look at the subseries plot. There are more subtle fluctuations month to month, and the first observations I presented are more pronounced.

```
gglagplot(myts)
```


lag 1
lag 2
lag 3
lag 4
lag 5
lag 6
lag 7
lag 8
lag 9
lag 10
lag 11
lag 12
lag 13
lag 14
lag 15
lag 16
300
200
100
100 200 300
Month
1
2
3
4
5
6
7
8
9
10
11
12

```
ggAcf(myts)
```

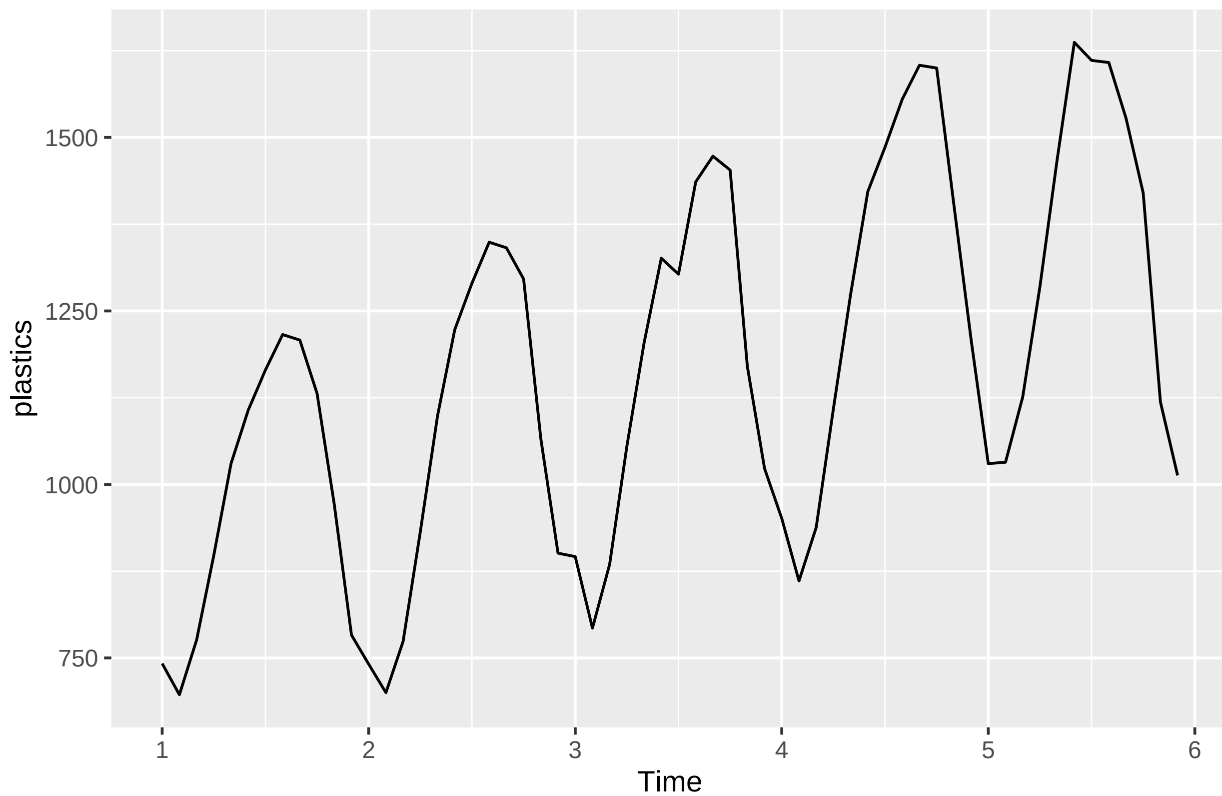


HA - 6.2

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

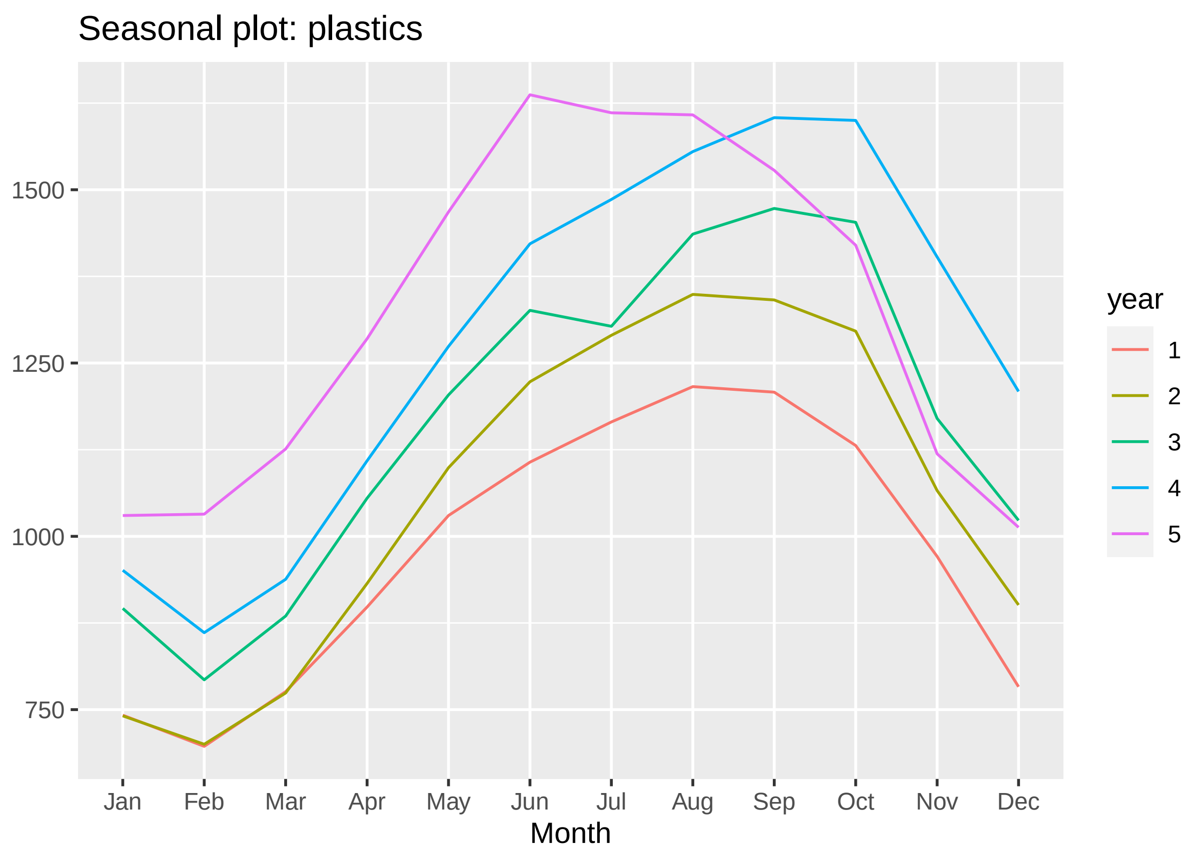
a. Plot the time series of sales of product A. Can you identify seasonal fluctuations and/or a trend-cycle?

```
autoplot(plastics)
```

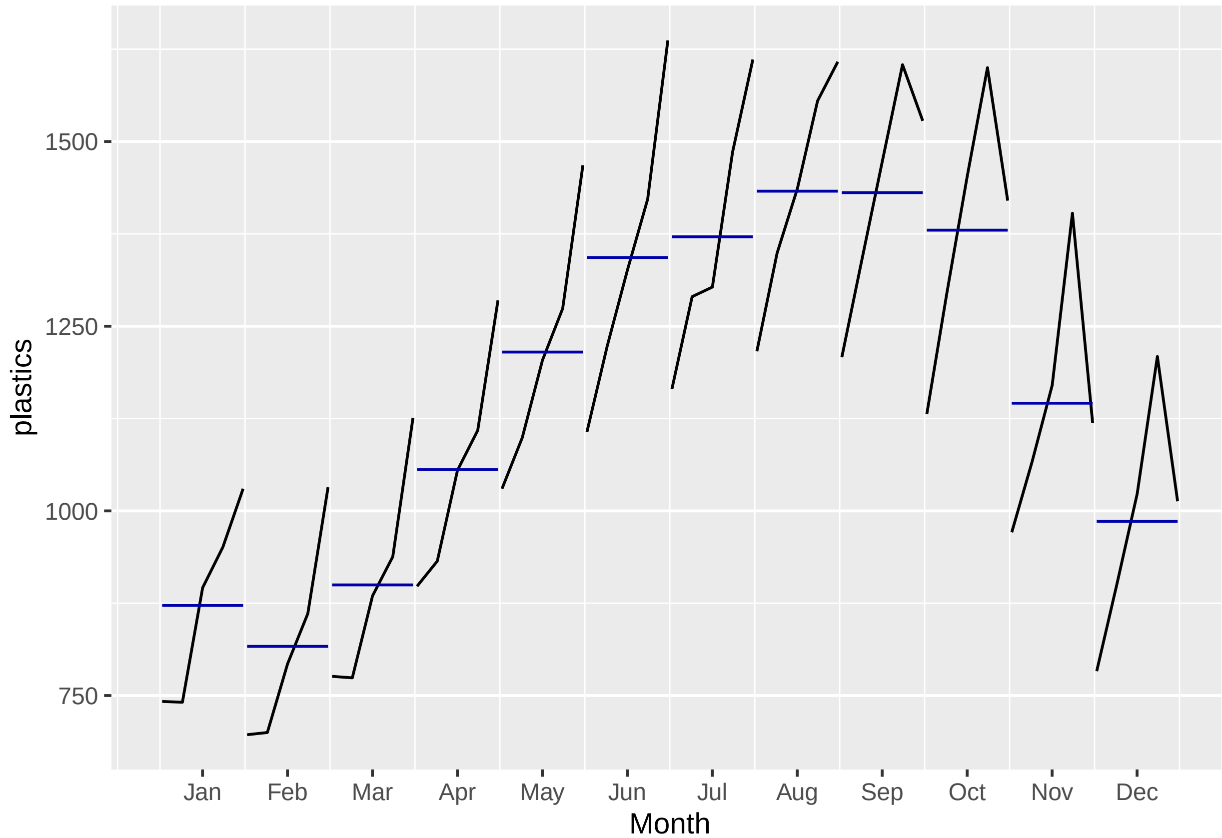


From the basic autoplot() of the plastics data, it seems that product A experiences significant seasonality in terms of sales figures. There seems to be steep increase from a little bit in the calendar year. This increase stops a little after the middle of the calendar year and is then followed by a steep decrease until a little bit into the next calendar year. In addition to a very transparent view of seasonality, there also seems to be an overall trend of increasing sales figures year over year (for the five years of data).

```
ggseasonplot(plastics)
```




```
ggsubseriesplot(plastics)
```

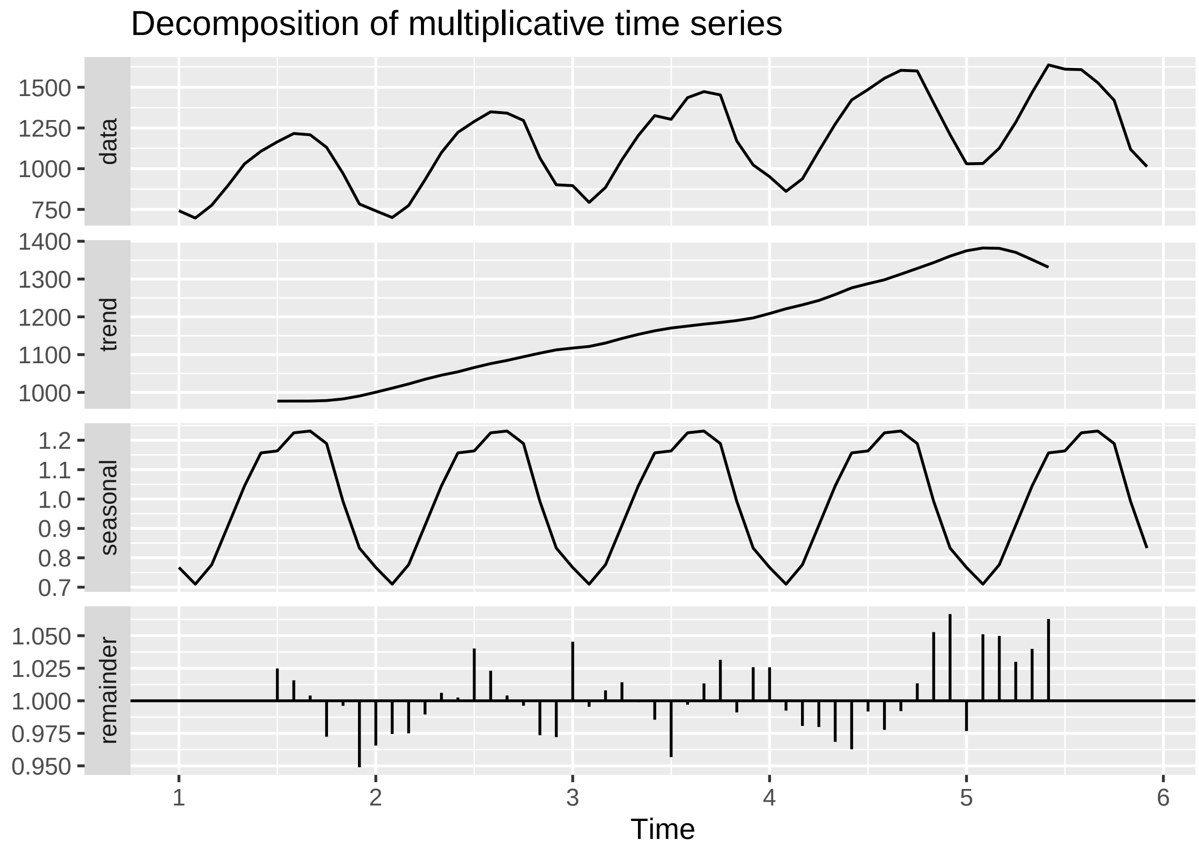


This observation about seasonality is confirmed when using the ggseasonplot() and the ggsubseriesplot() functions.

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

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

decomp %>% autoplot()
```



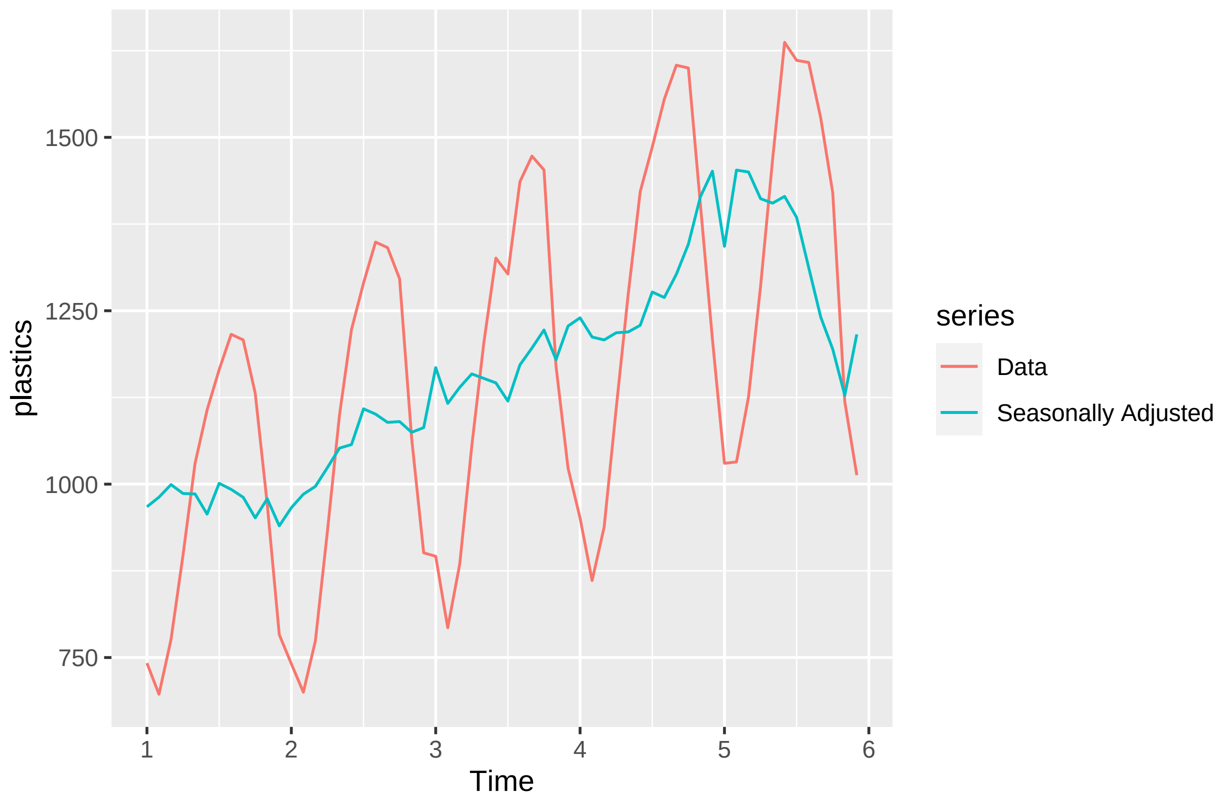
c. Do the results support the graphical interpretation from part a?

Answer:

Yes, the classical decomposition supports the answer from part a. The same seasonality that was discussed during part a is present in our seasonal plot and the same general increasing trend is displayed in the "trend" plot.

d. Compute and plot the seasonally adjusted data.

```
autoplot(plastics, series="Data") +
  autolayer(seasadj(decomp), series="Seasonally Adjusted")
```



e. 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?

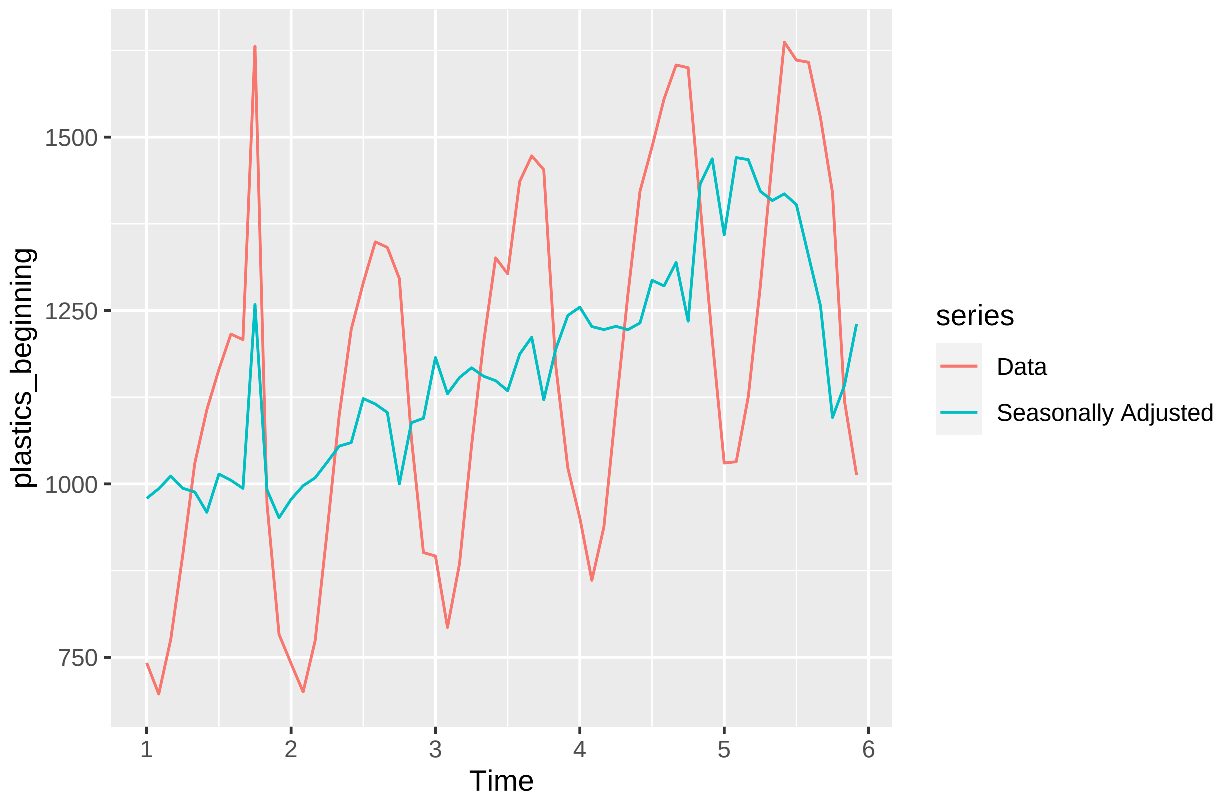
```
# Beginning

plastics_beginning = plastics

plastics_beginning[10] = plastics_beginning[10] + 500

decomp_w_beginning_outlier <- plastics_beginning %>% decompose(type="multiplicative")

autoplot(plastics_beginning, series="Data") +
  autolayer(seasadj(decomp_w_beginning_outlier), series="Seasonally Adjusted")
```



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

Middle:

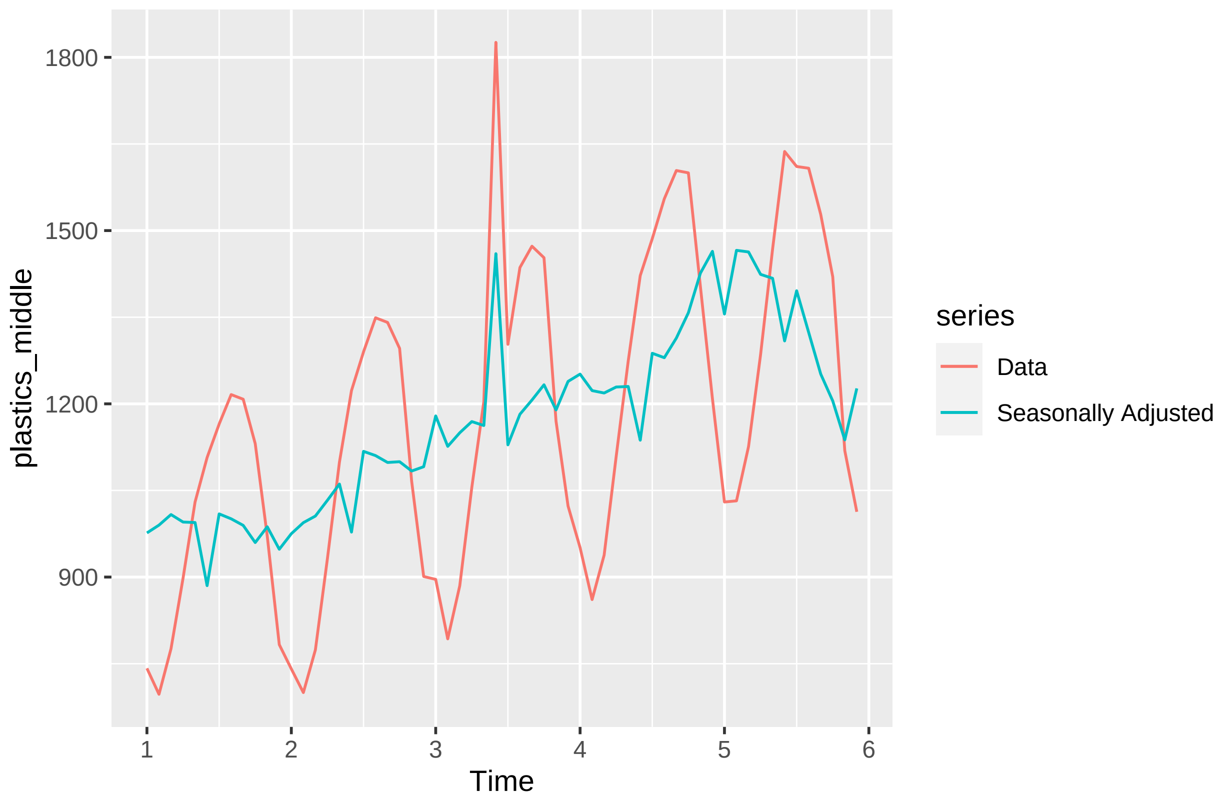
```
# Middle

plastics_middle = plastics

plastics_middle[30] = plastics_middle[30] + 500

decomp_w_middle_outlier <- plastics_middle %>% decompose(type="multiplicative")

autoplot(plastics_middle, series="Data") +
  autolayer(seasadj(decomp_w_middle_outlier), series="Seasonally Adjusted")
```



End:

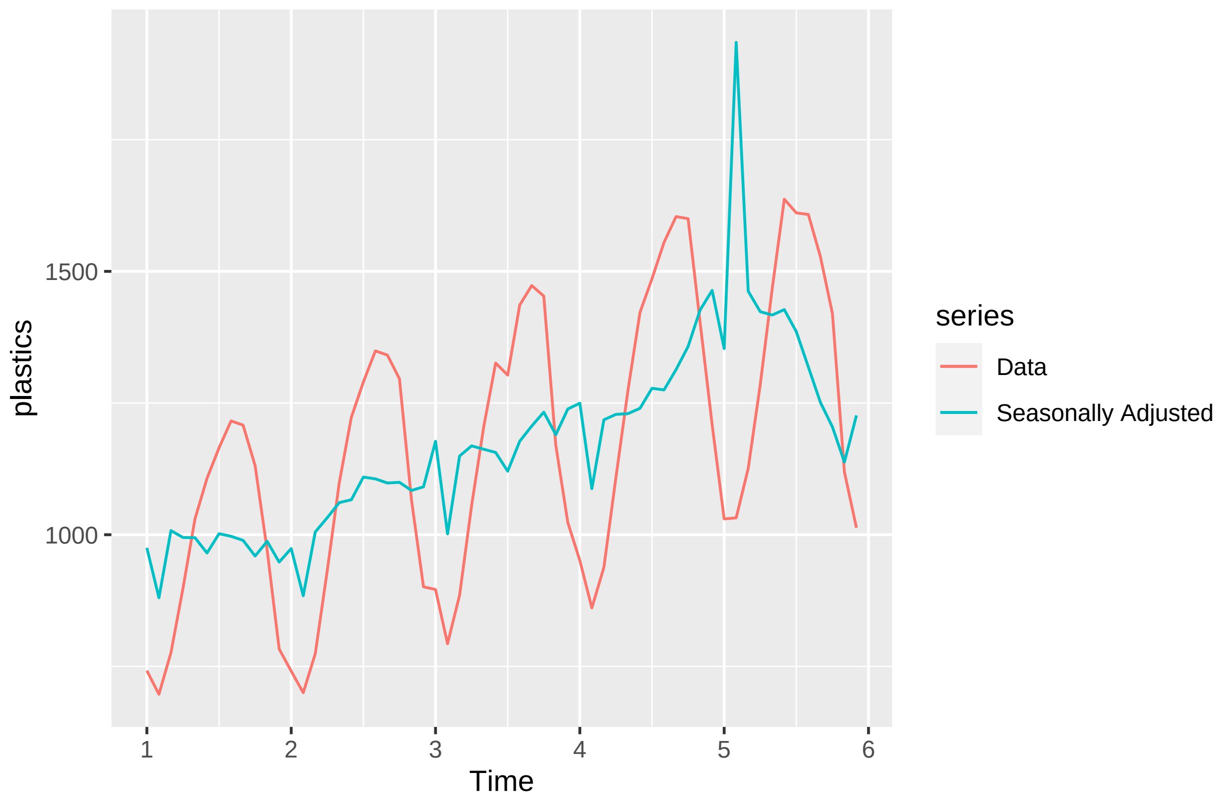
```
# End

plastics_end = plastics

plastics_end[50] = plastics_end[50] + 500

decomp_w_end_outlier <- plastics_end %>% decompose(type="multiplicative")

autoplot(plastics, series="Data") +
  autolayer(seasadj(decomp_w_end_outlier), series="Seasonally Adjusted")
```


plastics
1500
1000
1
2
3
4
5
6
Time
series
Data
Seasonally Adjusted