

ShovanBiswas-DATA24_Homework_03

Shovan Biswas

9/20/2020

Libraries

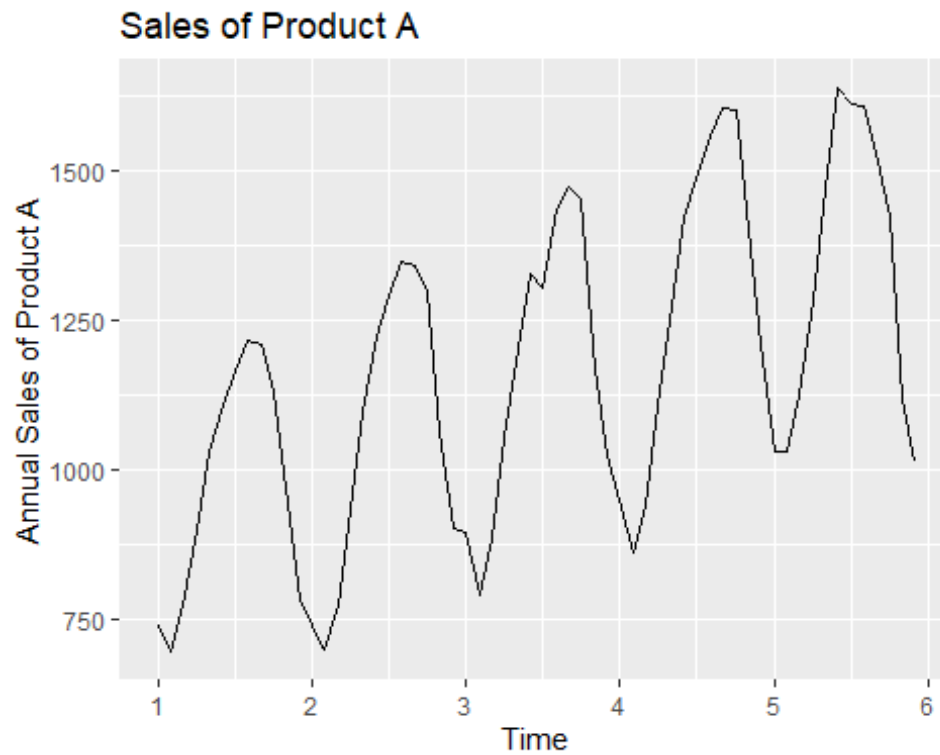
```
library(tidyverse)
library(knitr)
library(kableExtra)
library(fpp2)
library(gridExtra)
library(seasonal)
library(readxl)
library(forcats)
```

Exercise 6.2

The plastics data set consists of the monthly sales (in thousands) of product A for a plastics manufacturer for 5 years.

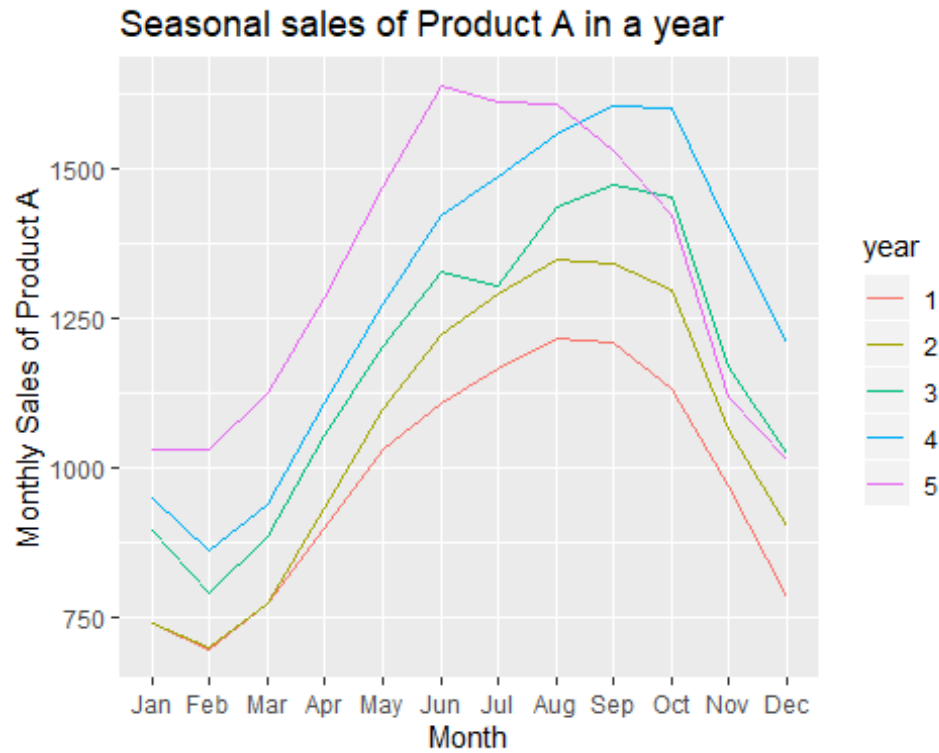
- Plot the time series of sales of product A. Can you identify seasonal fluctuations and/or a trend-cycle?

```
autoplot(plastics) + ggtitle("Sales of Product A") + ylab("Annual Sales of Product A")
```



The curve is trending upwards, with a seasonal frequency of 1 year. Refer `ggseasonplot` in earlier chapters.

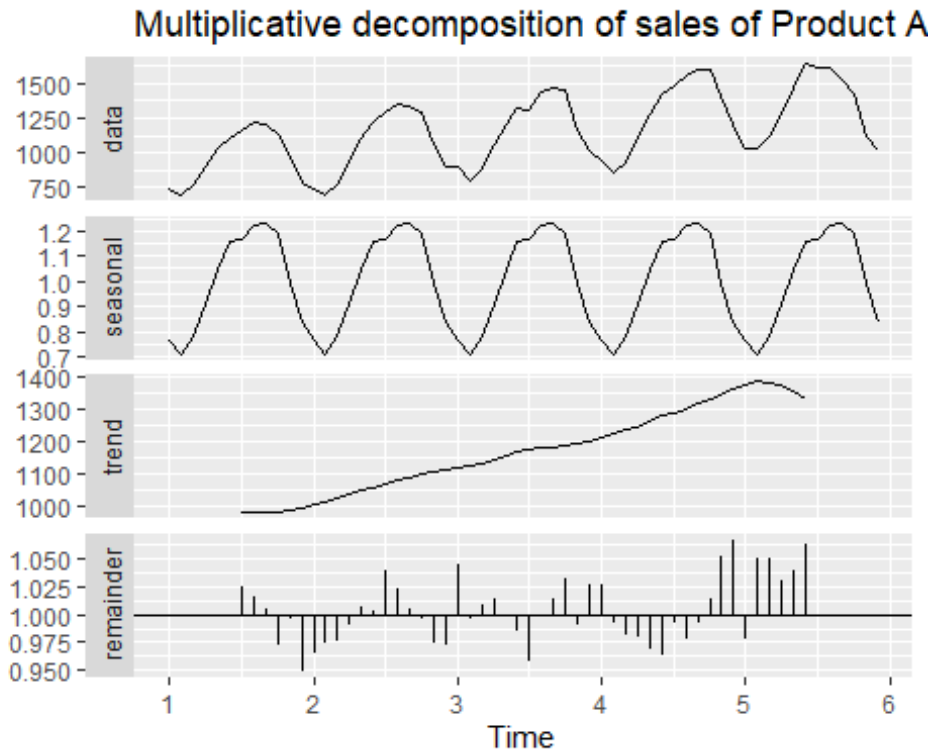
```
ggseasonplot(plastics) + ggtitle("Seasonal sales of Product A in a year") +  
ylab("Monthly Sales of Product A")
```



Sales starts trending up from Feb-Mar, attains a maxima from Jul-Oct, then keeps falling up to December.

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

```
decompose(plastics, type = "multiplicative") %>% autoplot() +  
ggtitle("Multiplicative decomposition of sales of Product A")
```



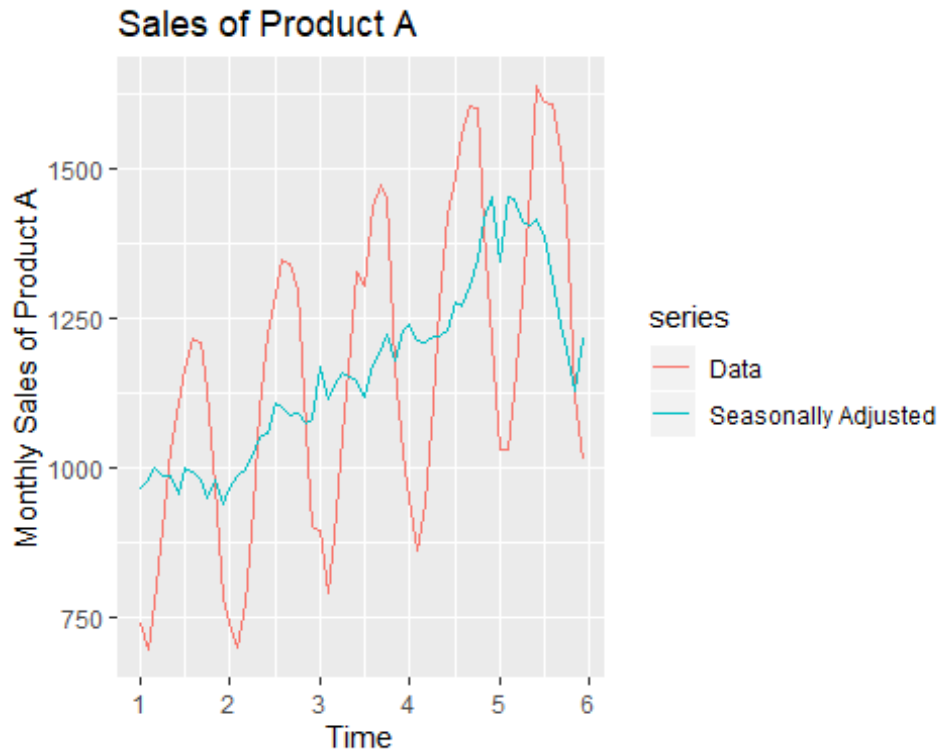
Although sales in all four graphs trend upwards, there is seasonality within a year.

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

Results of Classical Decomposition are consistent with observations from part a – upward trend, with a seasonal component.

d. Compute and plot the seasonally adjusted data.

```
autoplot(plastics, series = "Data") + autolayer(seasadj(plastics %>%
decompose(type = "multiplicative")), series = "Seasonally Adjusted") +
ggtitle("Sales of Product A") + ylab("Monthly Sales of Product A")
```



Upward trending, after 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?

I broke the task into three code chunks for better readability

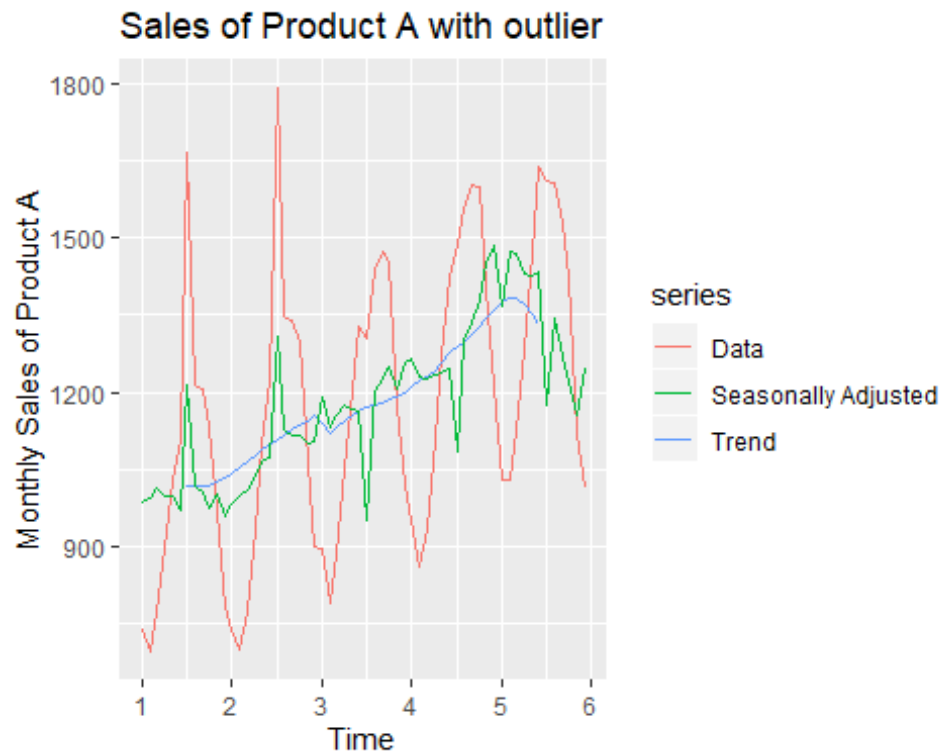
```
# Creating the outliers in July of two consecutive years.
outlier <- plastics
outlier[7] <- outlier[7] + 500
outlier[19] <- outlier[19] + 500
```

As opposed to earlier code chunk (d), I am storing **outlier %>% decompose(type = "multiplicative")** in a variable fit, because I'll use it twice.

```
fit <- outlier %>% decompose(type = "multiplicative")

# The actual graphing happens here.
autoplot(outlier, series = "Data") + autolayer(trendcycle(fit), series =
"trend") + autolayer(seasadj(fit), series = "Seasonally Adjusted") +
  ggtitle("Sales of Product A with outlier") + ylab("Monthly Sales of Product
A")

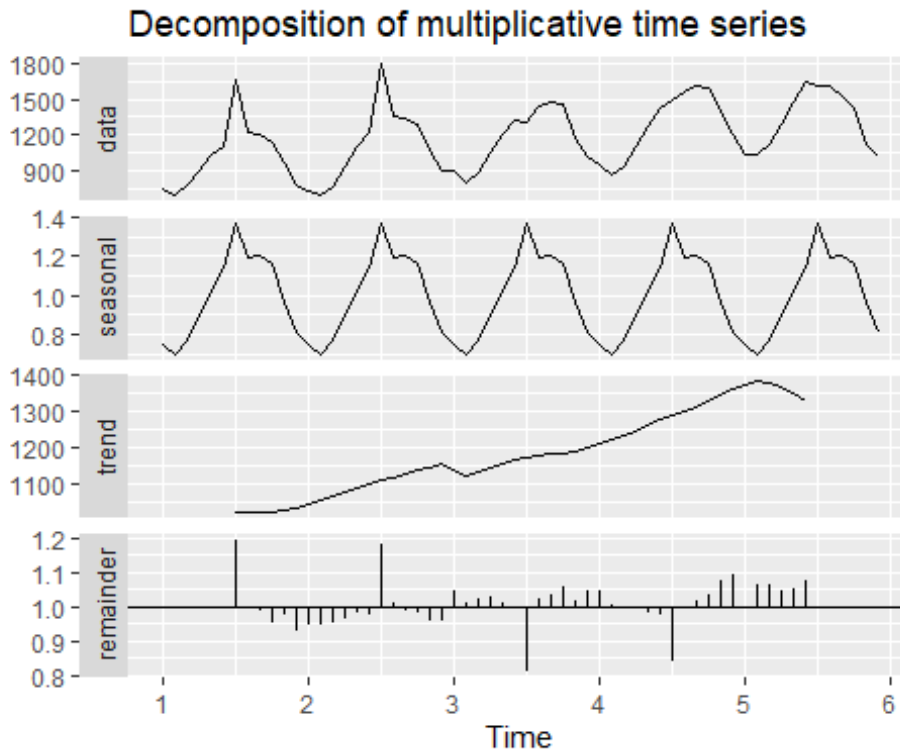
## Warning: Removed 12 rows containing missing values (geom_path).
```



I created two outliers in the 7th and 19th indices i.e. on the July of two consecutive years. So, the curve spiked up at those two points, everything remains the same – trends upwards and still seasonal.

The below autoplot() vindicates the same observation.

```
fit %>% autoplot() # used fit, to reduce
```



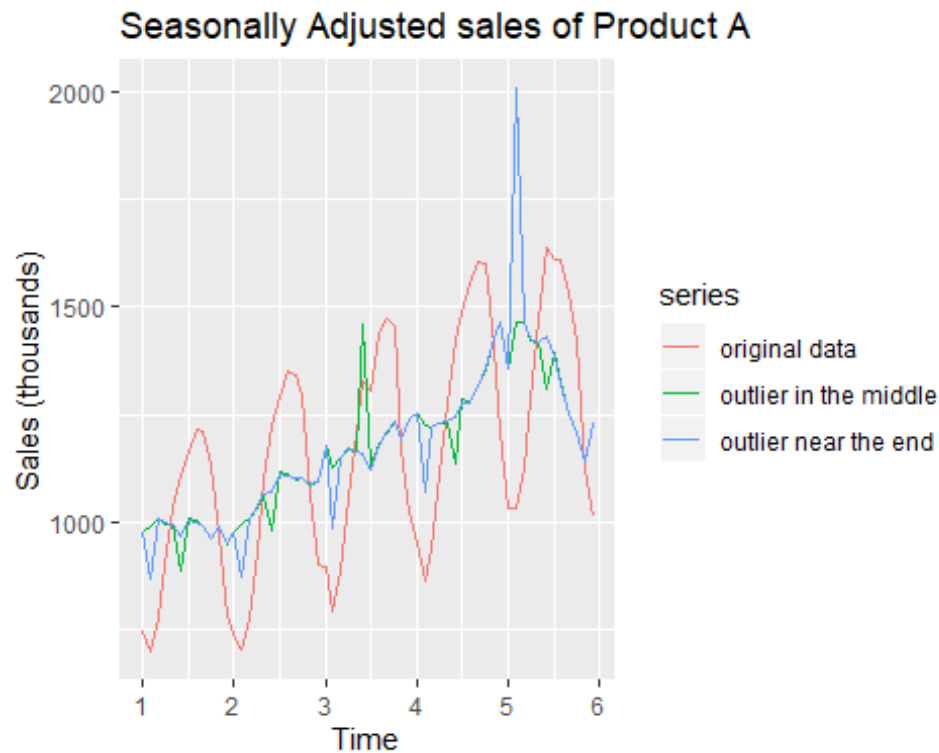
- f. Does it make any difference if the outlier is near the end rather than in the middle of the time series?

I broke the task into two code chunks for better readability

```
# Creating outlier in the middle
outlier_mid <- plastics
outlier_mid[30] <- outlier_mid[30] + 500

# Creating outlier in the End
outlier_end <- plastics
outlier_end[50] <- outlier_end[50] + 500

# The actual graphing happens here.
autoplot(plastics, series = 'original data') +
  autolayer(outlier_mid / decompose(outlier_mid, type =
    "multiplicative")$seasonal, series = 'outlier in the middle') +
  autolayer(outlier_end / decompose(outlier_end, type =
    "multiplicative")$seasonal, series = 'outlier near the end') +
  ylab("Sales (thousands)") + ggtitle("Seasonally Adjusted sales of Product A")
```

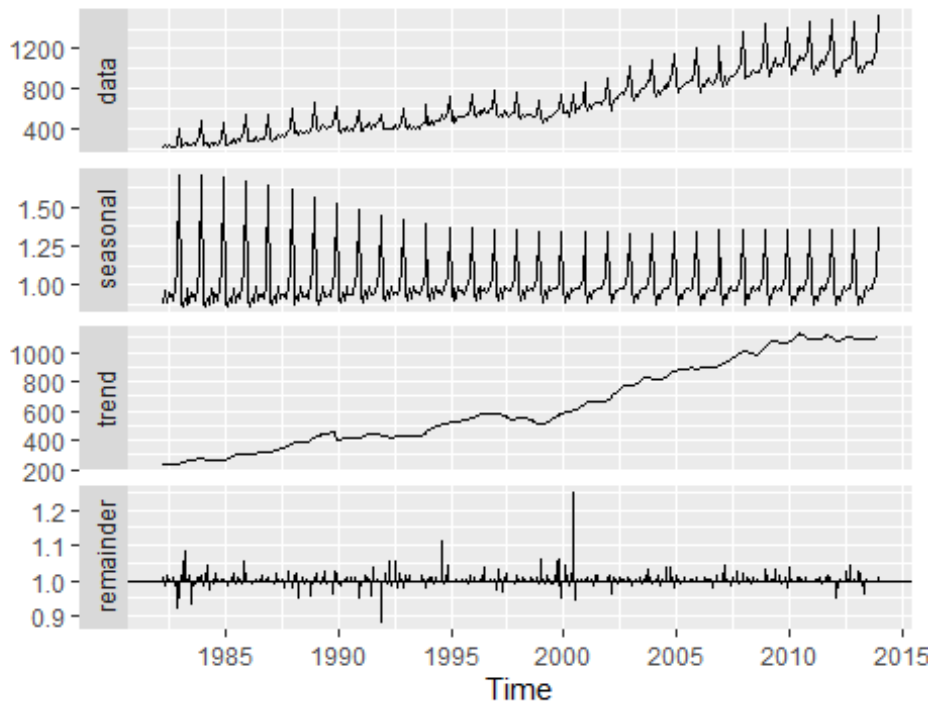


Exercise 6.3

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?

```
retail_data <- read_excel("retail.xlsx", skip = 1)
retail <- ts(retail_data[, "A3349397X"], frequency = 12, start = c(1982, 4))
x11_retail <- seas(retail, x11 = "")
autoplot(x11_retail) + ggtitle("X11 Decomposition of Retail Sales")
```


X11 Decomposition of Retail Sales



Seasonality with frequency of 1 year. Remainder has two major spikes, plus a few here and there, which probably indicate outliers.