Homework 6.2

By: Shaya Engelman, Julia Ferris, Amanda Fox, and Jean Jimenez

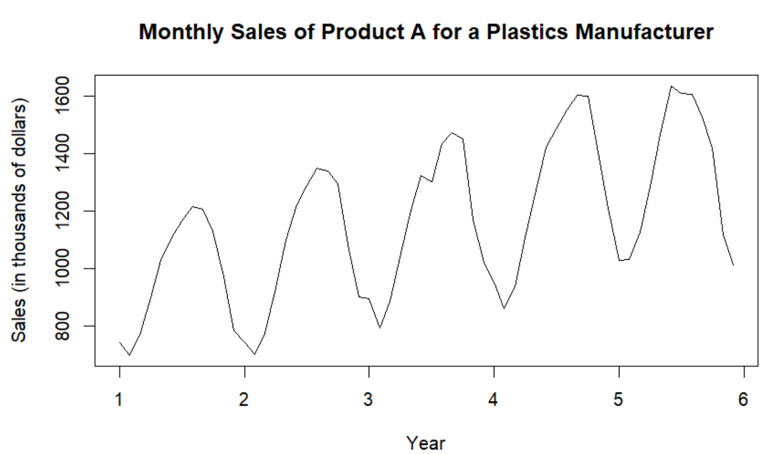
06/15/2024

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

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

   In the graph below, you can see the overall sales. There appears to be a clear cyclic pattern. It is very consistent and repeats yearly. The pattern seems to shift upward after each year, which indicates a possible increasing trend.

```
```{r}
library(fma)
library(ggplot2)
data(plastics)
plot(plastics, main = "Monthly Sales of Product A for a Plastics Manufacturer",
    xlab="Year", ylab="Sales (in thousands of dollars")
```
```



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

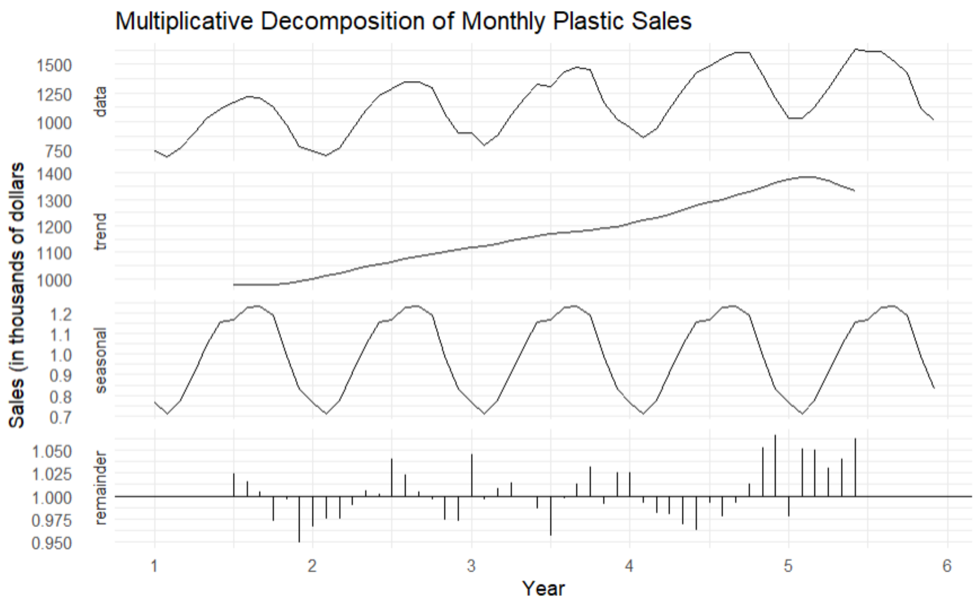
   The first section below shows the values of the indices for the trend-cycle and then the seasonal cycle. The values for the trend-cycle show an increase until right near the end, at which point they decrease slightly. The values for the seasonal cycle are exactly the same for each month. All the values for January are equal, all the values for February are equal, and so on. Different months have different values. In order, they increase until September. After September, they begin to decrease.

```{r}
decomposition <- decompose(plastics, type = "multiplicative")

trend_cycle <- decomposition$trend
seasonal <- decomposition$seasonal

trend_cycle
seasonal
```

```
> trend_cycle
        Jan       Feb       Mar       Apr       May       Jun       Jul       Aug       Sep       Oct       Nov       Dec
1        NA        NA        NA        NA        NA        NA  976.9583  977.0417  977.0833  978.4167  982.7083  990.4167
2 1000.4583 1011.2083 1022.2917 1034.7083 1045.5417 1054.4167 1065.7917 1076.1250 1084.6250 1094.3750 1103.8750 1112.5417
3 1117.3750 1121.5417 1130.6667 1142.7083 1153.5833 1163.0000 1170.3750 1175.5000 1180.5417 1185.0000 1190.1667 1197.0833
4 1208.7083 1221.2917 1231.7083 1243.2917 1259.1250 1276.5833 1287.6250 1298.0417 1313.0000 1328.1667 1343.5833 1360.6250
5 1374.7917 1382.2083 1381.2500 1370.5833 1351.2500 1331.2500        NA        NA        NA        NA        NA        NA
> seasonal
        Jan       Feb       Mar       Apr       May       Jun       Jul       Aug       Sep       Oct       Nov       Dec
1 0.7670466 0.7103357 0.7765294 0.9103112 1.0447386 1.1570026 1.1636317 1.2252952 1.2313635 1.1887444 0.9919176 0.8330834
2 0.7670466 0.7103357 0.7765294 0.9103112 1.0447386 1.1570026 1.1636317 1.2252952 1.2313635 1.1887444 0.9919176 0.8330834
3 0.7670466 0.7103357 0.7765294 0.9103112 1.0447386 1.1570026 1.1636317 1.2252952 1.2313635 1.1887444 0.9919176 0.8330834
4 0.7670466 0.7103357 0.7765294 0.9103112 1.0447386 1.1570026 1.1636317 1.2252952 1.2313635 1.1887444 0.9919176 0.8330834
5 0.7670466 0.7103357 0.7765294 0.9103112 1.0447386 1.1570026 1.1636317 1.2252952 1.2313635 1.1887444 0.9919176 0.8330834
```

   The second section below shows a graph of these indices along with the original data and a remainder component.

```{r}
autoplot(decomposition) +
  ggtitle("Multiplicative Decomposition of Monthly Plastic Sales") +
  xlab("Year") +
  ylab("Sales (in thousands of dollars)") +
  theme_minimal()
```



3. **Do the results support the graphical interpretation from part a?**

   Yes, the results support the graphical interpretation from part a. After multiplicative decomposition, it is clear to see the upward trend that was assumed in part a, and the consistent cyclical pattern is also shown in the results. The seasonal indices additionally show that the cycle repeats after every year.

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

   In this section, the seasonally adjusted data consists of all the data without the seasonal component. The seasonal component is divided out because it was originally calculated by using multiplicative decomposition. The first plot shows the original data, and the second plot shows the adjusted data. The two graphs together demonstrate the difference between original data and seasonally-adjusted data.

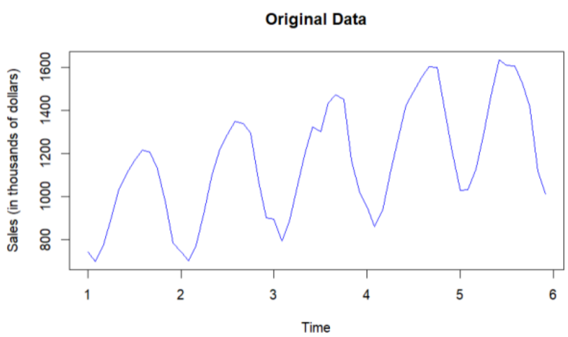
```
```{r}

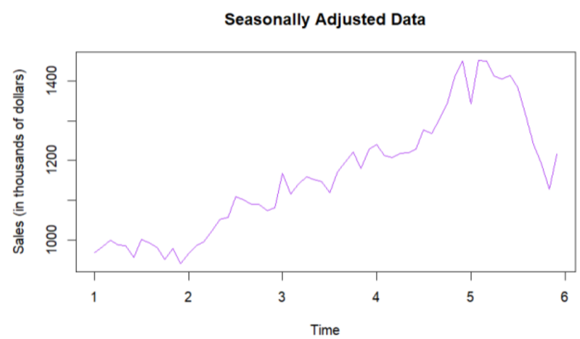
decomposed_data <- decompose(plastics, type = "multiplicative")


seasonal <- decomposed_data$seasonal

seasonally_adjusted <- plastics / seasonal
```
```

```
plot(plastics, main = "Original Data", ylab = "Sales (in thousands of dollars)", col =
    "blue")
plot(seasonally_adjusted, main = "Seasonally Adjusted Data",
    ylab = "Sales (in thousands of dollars)", col = "purple")
```





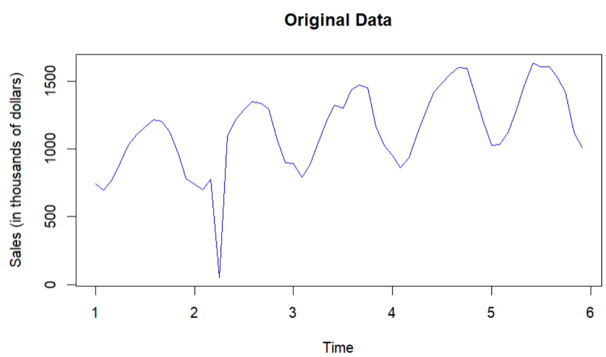
5. **Change one observation to be an outlier (e.g., add 500 to one observation), and recompute the seasonally adjusted data. What is the effect of the outlier?**

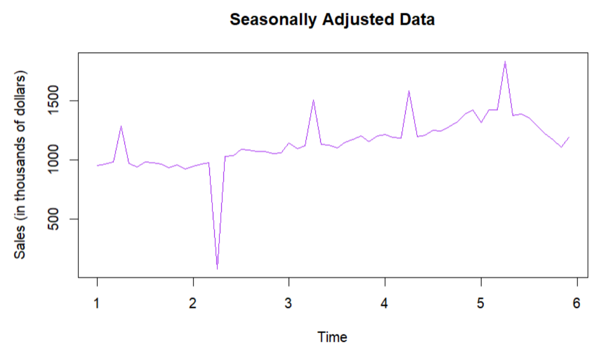
   The outlier of 50 was changed for the month of April in the second year of data collection. The results below were computed after this outlier was added to the data.

   The outlier changed the seasonally-adjusted data. The rest of the data appeared to have the same upward trend as the original data when looking at the scale on the y-axis. The outlier in the data was seen as a major dip in Sales. This dip is much different from the original seasonally-adjusted data. This example demonstrates the power of outliers and the need to identify them when analyzing data.

```{r}
plastics[16] <- 50
decomposed_data <- decompose(plastics, type = "multiplicative")


seasonal <- decomposed_data$seasonal
seasonally_adjusted <- plastics / seasonal
```

```
plot(plastics, main = "Original Data", ylab = "Sales (in thousands of dollars)", col =
    "blue")
plot(seasonally_adjusted, main = "Seasonally Adjusted Data", ylab = "Sales (in thousands
    of dollars)", col = "purple")
```





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

   To determine if the location of the outlier results in any difference, an outlier was added near the end of the time series and compared to the results of data with the other outlier.

   Based on the graph from below, the outlier still changed the graph, but since the outlier was at the end, it did not appear to have much of an effect on the cyclical pattern of the data as the other outlier did. The outlier from the previous question resulted in small peaks throughout the graph, while this outlier did not appear to result in any peaks. Instead, the graph shows an increasing trend and then a major dip where the outlier occurred.

```{r}
data(plastics)
plastics[58] <- 75
```

```
decomposed_data <- decompose(plastics, type = "multiplicative")
seasonal <- decomposed_data$seasonal
seasonally_adjusted <- plastics / seasonal

plot(plastics, main = "Original Data", ylab = "Sales (in thousands of dollars)", col =
    "blue")
plot(seasonally_adjusted, main = "Seasonally Adjusted Data", ylab = "Sales (in thousands
    of dollars)", col = "purple")
```