Homework - CUNY DATA 624 Sum I

Group 1: Daniel Sullivan, Jeff Parks, Lwin Shwe, Matthew Katz

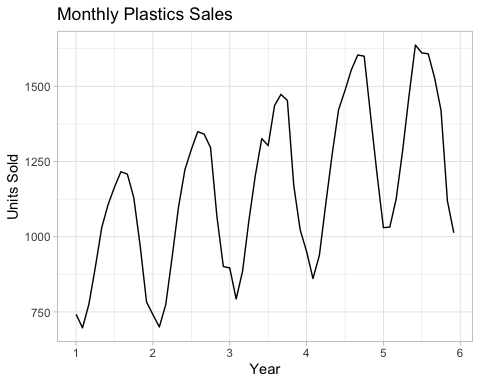
## HA Exercise 6.2

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?**

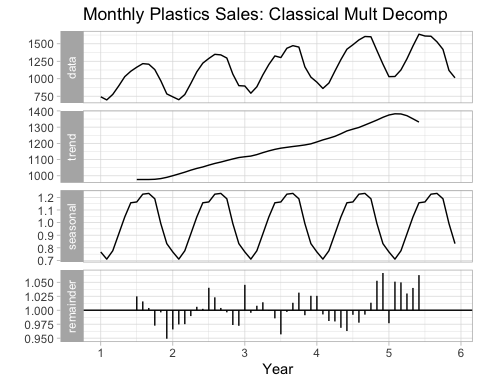
Before applying any calculations, we can observe a rising trend-cycle over the five years, with an annual seasonality where sales peak towards mid-year and then drop off rapidly.

autoplot(plastics) +  
 ggtitle('Monthly Plastics Sales') +  
 ylab('Units Sold') +  
 xlab('Year')



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

plastics %>% decompose(type='multiplicative') %>%  
 autoplot() +   
 xlab('Year') +   
 ggtitle("Monthly Plastics Sales: Classical Mult Decomp")

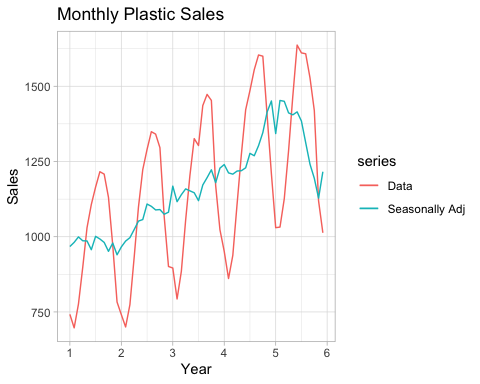


1. **Do the results support the graphical interpretation from part a?**

The results of the Classical Decomposition do seem to support the graphical interpretation. The trend-cycle decomposition shows a clear, increasing trend throughout the five years, and the seasonal decomposition shows a clear, repetitive annual cycle with a mid-year sales peak.

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

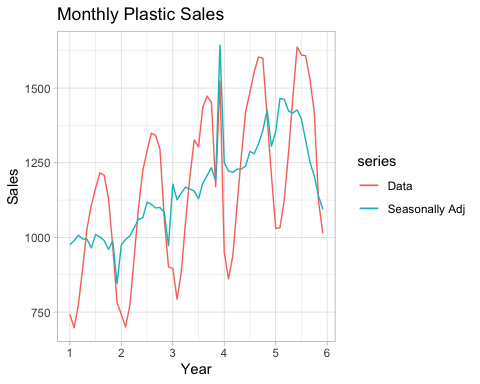
plastics\_decomp <- plastics %>% decompose(type='multiplicative')  
  
plastics\_decomp$seasonal\_adj <- plastics\_decomp$x / plastics\_decomp$seasonal  
  
autoplot(plastics\_decomp$x, series='Data') +   
 autolayer(plastics\_decomp$seasonal\_adj, series='Seasonally Adj') +  
 ggtitle('Monthly Plastic Sales') +   
 xlab('Year') +  
 ylab('Sales')



1. **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 addition of an outlier near the middle of the dataset produces a spike in both the non-decomposed data and the seasonally adjusted data. However the overall trend-cycle and seasonality seem to return to their established patterns immediately.

plastics\_out <- plastics  
plastics\_out[36] <- plastics\_out[36] + 500  
  
plastics\_out\_decomp <- plastics\_out %>% decompose(type='multiplicative')  
plastics\_out\_decomp$seasonal\_adj <- plastics\_out\_decomp$x / plastics\_out\_decomp$seasonal  
  
autoplot(plastics\_out\_decomp$x, series='Data') +   
 autolayer(plastics\_out\_decomp$seasonal\_adj, series='Seasonally Adj') +  
 ggtitle('Monthly Plastic Sales') +   
 xlab('Year') +  
 ylab('Sales')



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

The addition of an outlier towards the end of the dataset produces a spike in both the non-decomposed data and the seasonally adjusted data. However the overall trend-cycle and seasonality seem to return to their established patterns immediately - it does not seem to make much difference where this new outlier is placed.

plastics\_out2 <- plastics  
plastics\_out2[58] <- plastics\_out[58] + 500  
  
plastics\_out2\_decomp <- plastics\_out2 %>% decompose(type='multiplicative')  
plastics\_out2\_decomp$seasonal\_adj <- plastics\_out2\_decomp$x / plastics\_out2\_decomp$seasonal  
  
autoplot(plastics\_out2\_decomp$x, series='Data') +   
 autolayer(plastics\_out2\_decomp$seasonal\_adj, series='Seasonally Adj') +  
 ggtitle('Monthly Plastic Sales') +   
 xlab('Year') +  
 ylab('Sales')

