Live Unit 11 Assignment

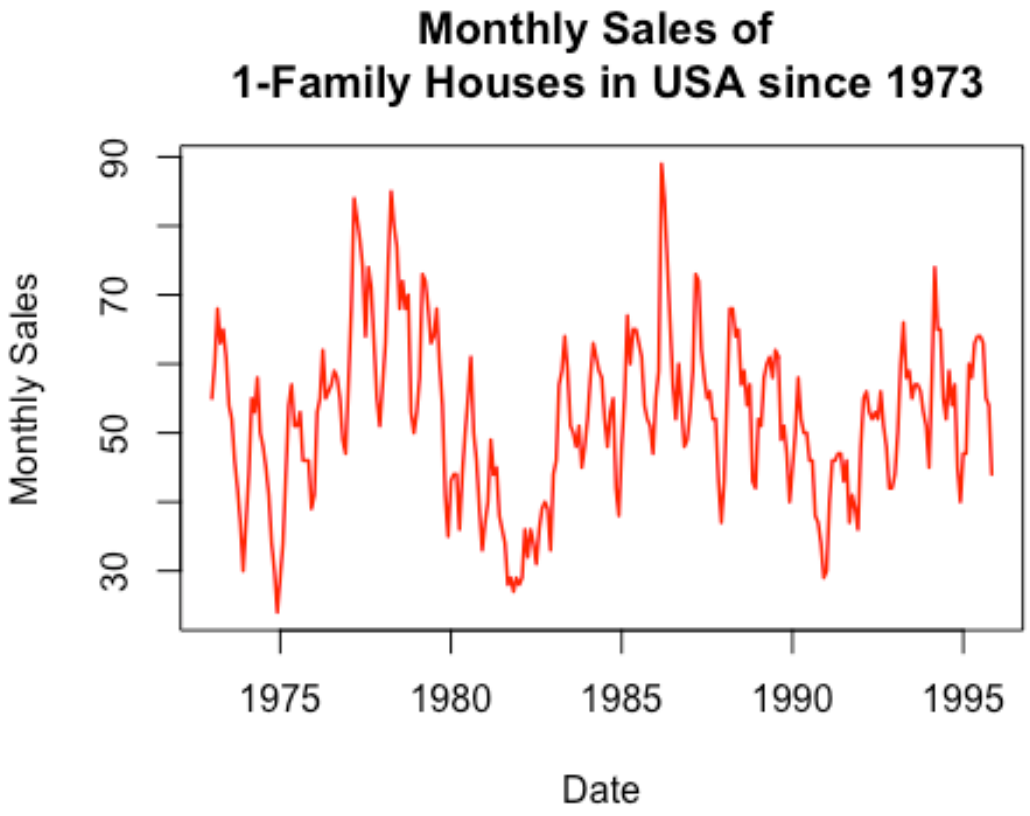
Angela Horacek, James Gage, Trace Smith

11/22/2016

- The data for this assignment is loaded from the 'Forecasting: principles and practice' (i.e. fpp) library. The source for the data 'hsales' which is the monthly sales of new one-family houses sold in the USA since 1973 is from: "Makridakis, Wheelwright and Hyndman (1998) Forecasting: methods and applications, John Wiley & Sons: New York. Chapter 3."

a.) Plot the time series. Can you identify seasonal fluctuations and/or a trend? - There appears to be a seasonality to the data. The data seem to be oscillating up and down corresponding to the calender year.In addition, a trend in the form of a sinusodal wave appears with a trough approximately every 9 years.

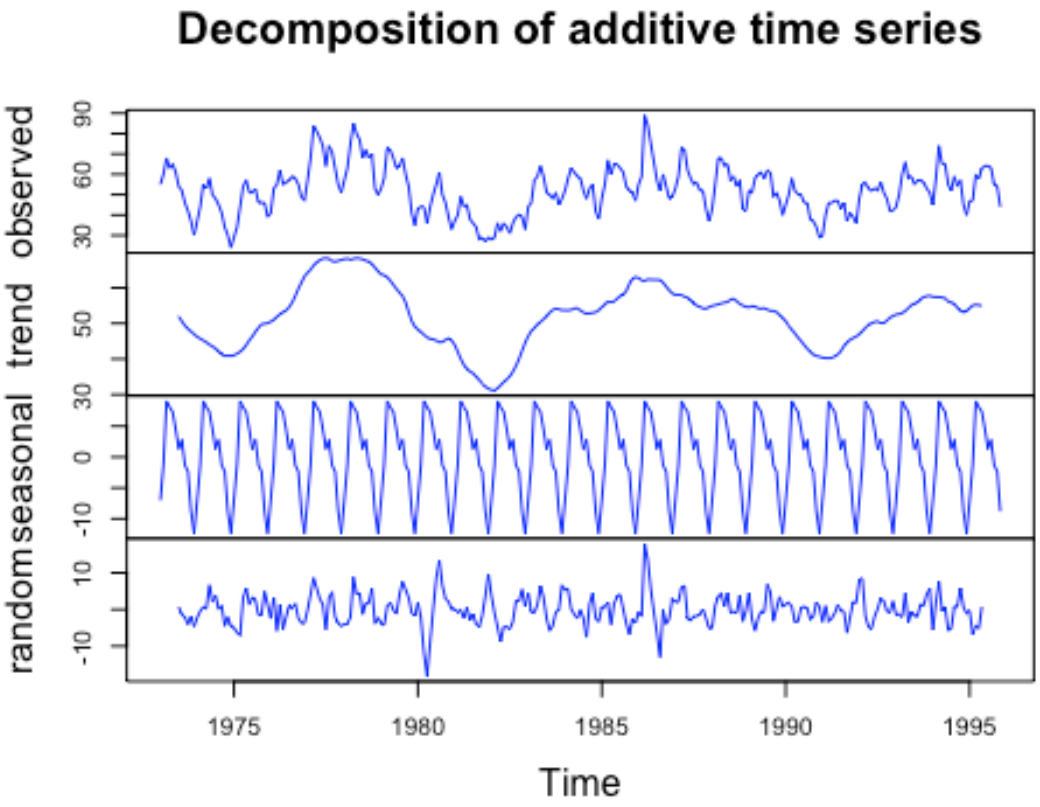
```
#plot the monthly sales of homes:
plot(hsales, col="red", lwd=1.5, ylab = "Monthly Sales", xlab="Date",
    main="Monthly Sales of \n 1-Family Houses in USA since 1973")
```



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

```
#Decompose a time series into seasonal, trend and irregular
#components using moving averages.
data <- decompose(hsales)

# Plot the estimated trend, seasonal, and irregular components
plot(data, col="blue")
```



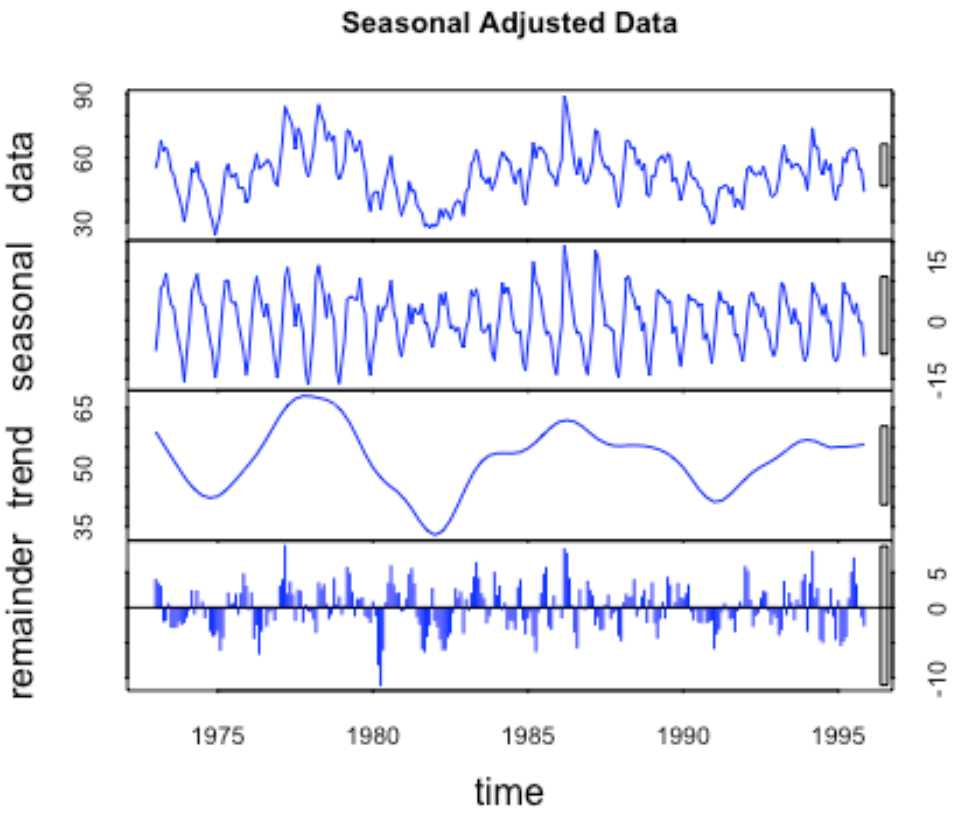
c.) Do the results support the graphical interpretation from part (a)?

- The results from the decomposition graphics indicate there is not only a seasonal component but also a trend component bolstering the interpretation of the original graph.

d) Compute and plot the seasonally adjusted data.

- Decompose a time series into seasonal, trend and irregular components using:

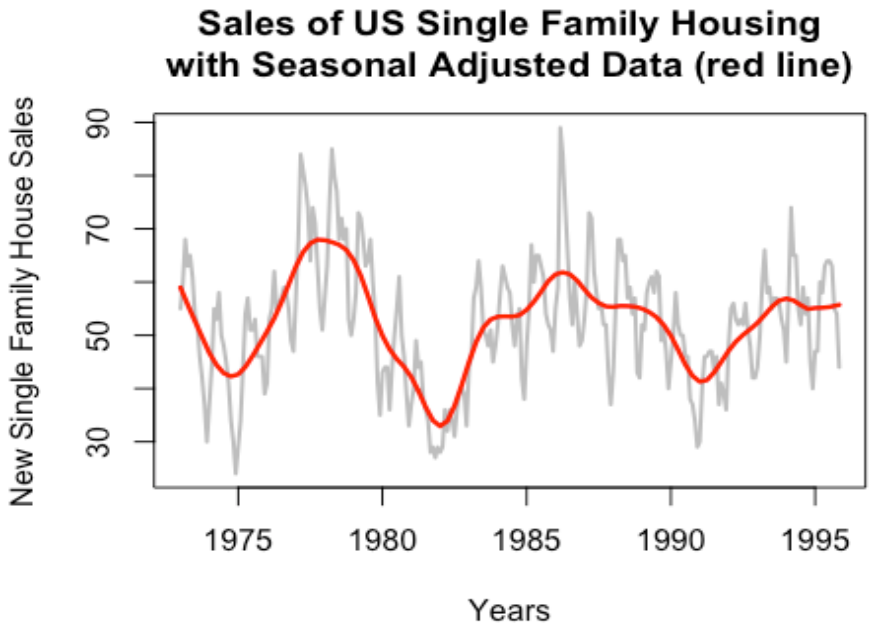
```
#Returns seasonally adjusted data constructed by removing the seasonal component.
fit <- stl(hsales, s.window=5)
plot(fit, main="Seasonal Adjusted Data",col="blue")
```



- Plot of hsales with the seasonal adjusted data (red line):

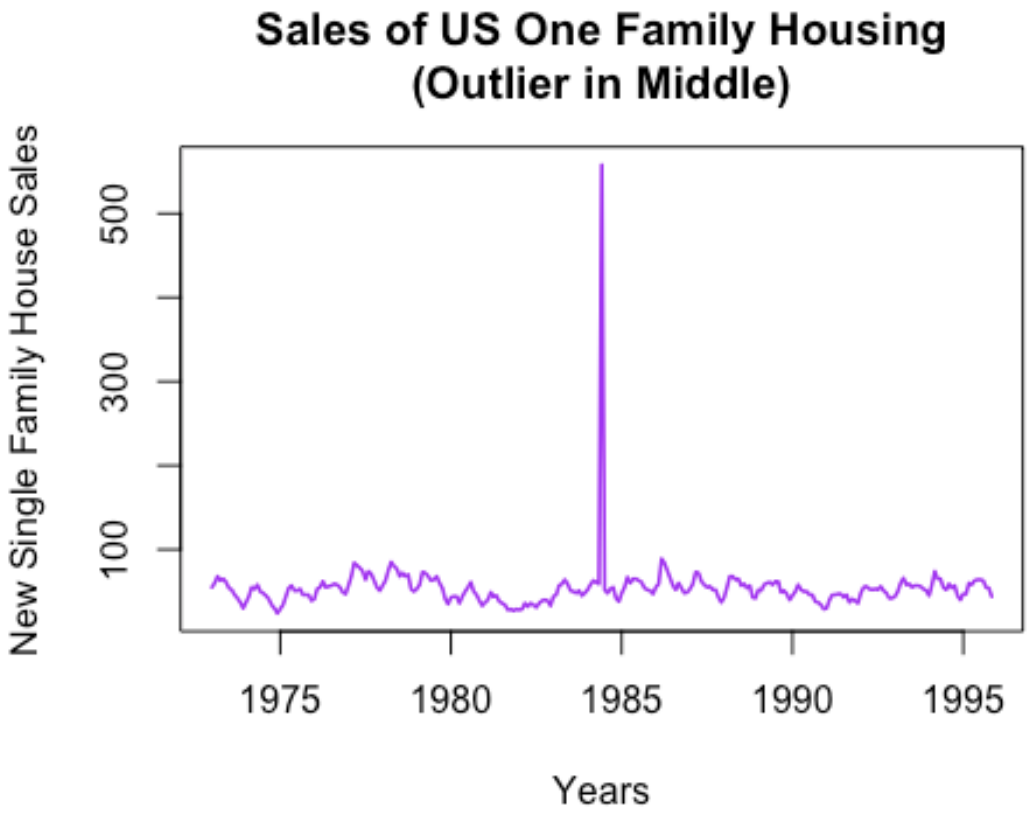
```
plot(hsales, col="gray",lwd=2, main="Sales of US Single Family Housing\nwith
Seasonal Adjusted Data (red line)", ylab="New Single Family House Sales",
xlab="Years")

lines(fit$time.series[,2],col="red",ylab="Trend",lwd=2.5)
```



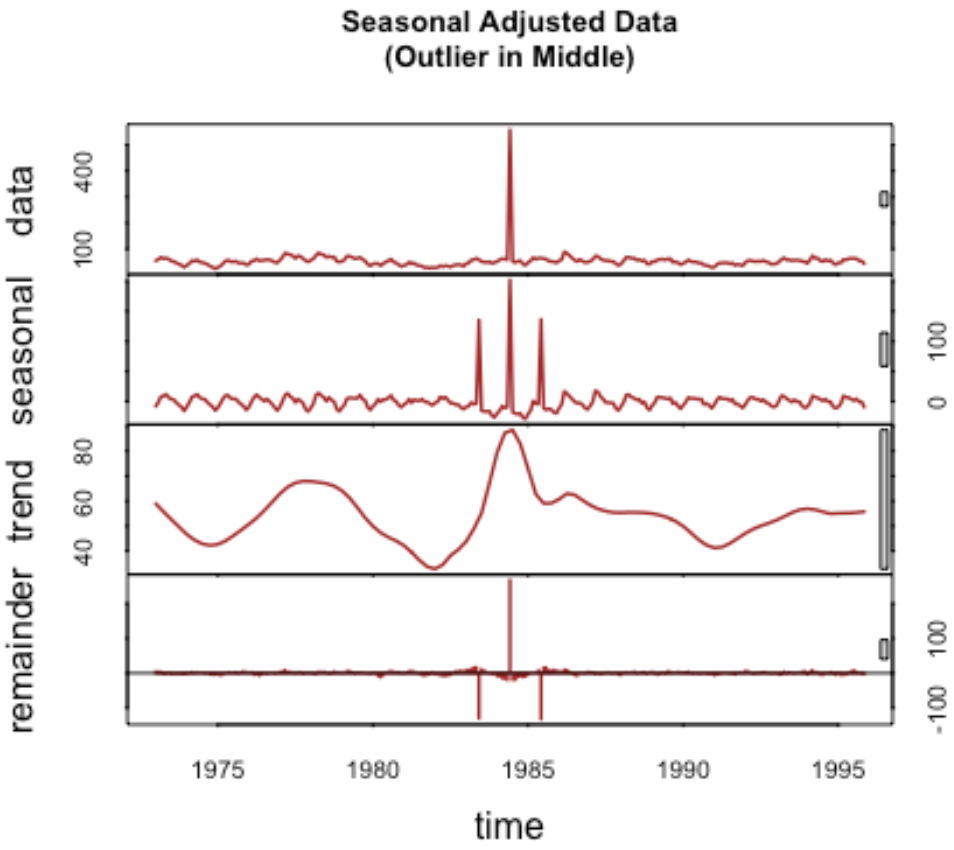
e) Change one observation to be an outlier (e.g., add 500 to one observation), and recompute the seasonally adjusted data. What is the effect of the outlier?

```
#start year=1973, frequency= 12 months
hsales2 <-
ts(c(hsales[1:137],hsales[138]+500,hsales[139:275]),start=c(1973,1),frequency=12)
plot(hsales2, col="purple",lwd=1.5, main="Sales of US One Family Housing\n(Outlier
in Middle)",
     ylab="New Single Family House Sales", xlab="Years")
```



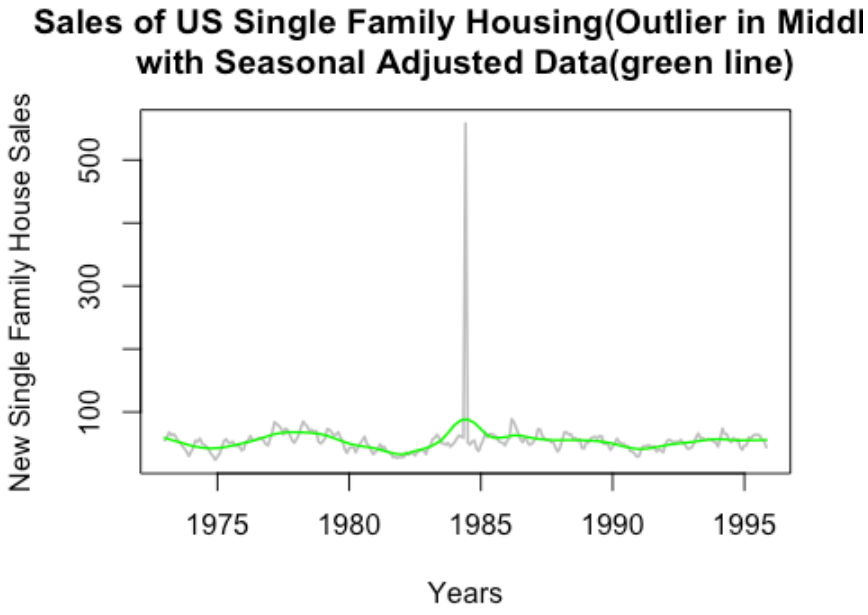
- Recompute the seasonally adjusted data.

```
 fit2 <- stl(hsales2, s.window=5)
plot(fit2, col="brown",lwd=1.5, main="Seasonal Adjusted Data\n(Outlier in Middle)")
```



- Plot of hsales2 with the seasonal adjusted data (green line)

```
plot(hsales2, col="gray", lwd=1.5,
     main="Sales of US Single Family Housing(Outlier in Middle)\nwith Seasonal
Adjusted Data(green line)",
     ylab="New Single Family House Sales", xlab="Years")
lines(fit2$time.series[,2],col="green",lwd=1.5, ylab="Trend")
```



- The outlier observation in the middle of the data increases the magnitude of the the seasonally adjusted data not only at that particular time but also on both sides of that time. Similarly, the trend line also peaks at that year pulling the trend curve up around that time distorting the sinusodial wave pattern.

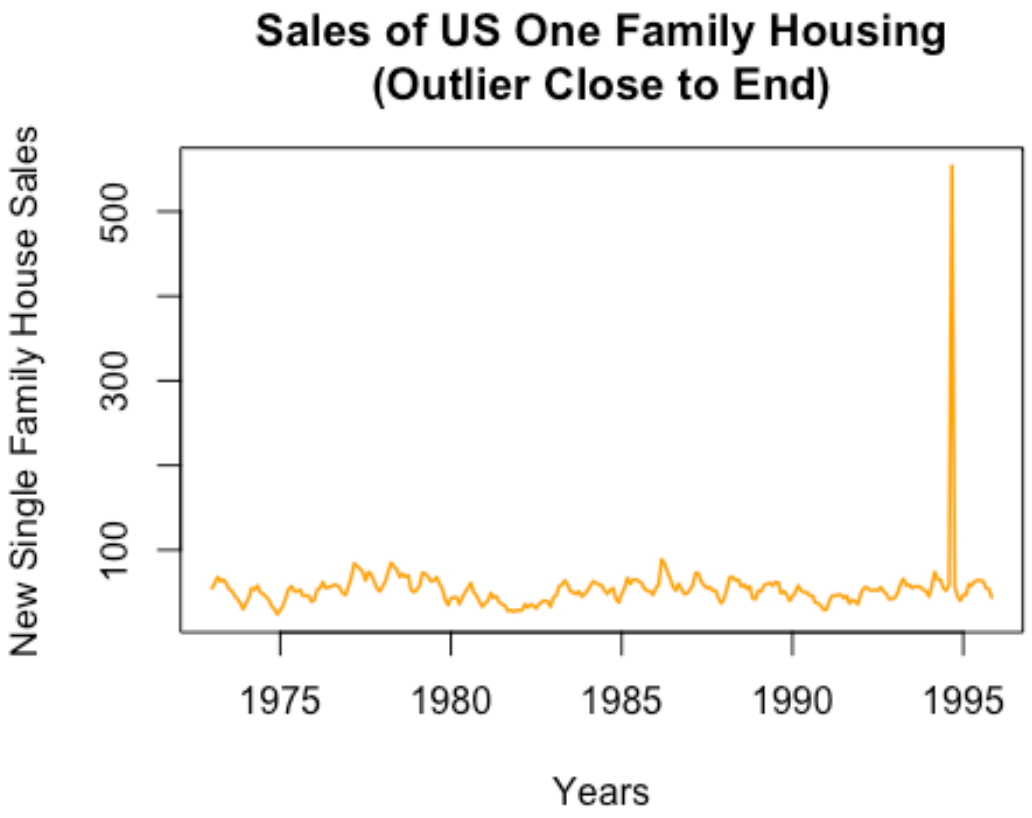
f) Does it make any difference if the outlier is near the end rather than in the middle of the time series?

- F-1 Step 1: Create an outlier near the end:

```
hsales4 <- ts(c(hsales[1:260],hsales[261]+500,
hsales[262:275]),start=c(1973,1),frequency=12)
```

- Plot the time series:

```
plot(hsales4, lwd=1.5, col="orange",main="Sales of US One Family Housing\n(Outlier
Close to End)",
ylab="New Single Family House Sales", xlab="Years")
```

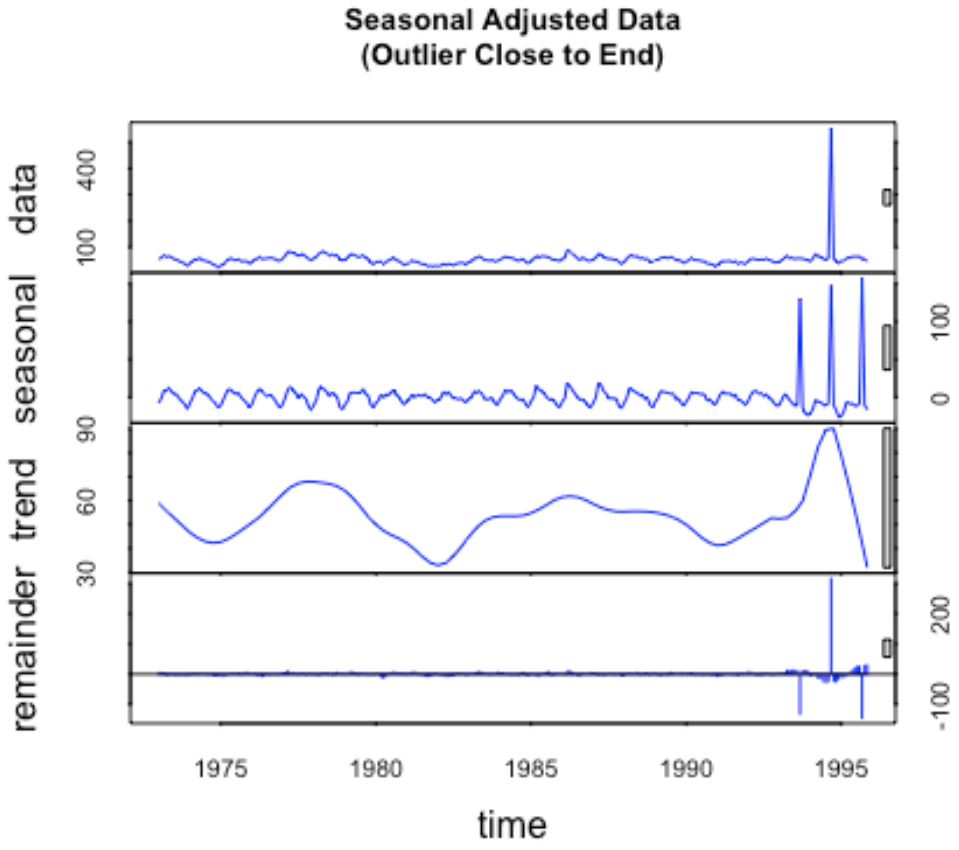


- The outlier observation near the end of the data increases the magnitude of the the seasonally adjusted data not only at that particular time but also on both sides of that time; This is the same effect of the outlier in the middle except at a different time. Similarly, the trend line also peaks at that year pulling the trend curve up around that time distorting the sinusodial wave pattern.

g) Now use STL to decompose the series.

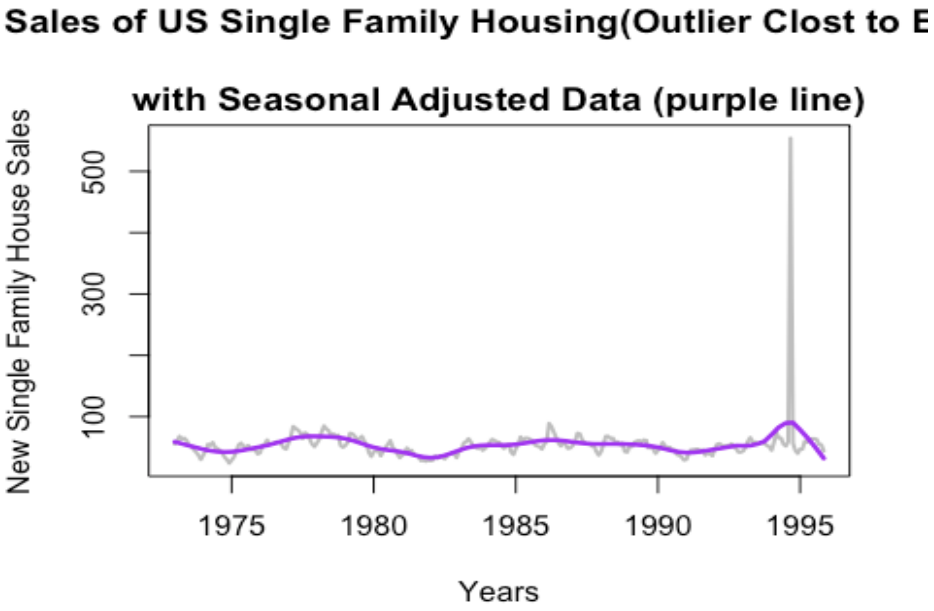
- Step G-1: Plot Seasonal Adjusted Data with Outlier at End:

```
fit4 <- stl(hsales4, s.window=5)
plot(fit4, col="blue",main="Seasonal Adjusted Data\n(Outlier Close to End)")
```



- Plot of hsales3 with the seasonal adjusted data(purple line), fit3

```
plot(hsales4, col="gray",lwd=2,main="Sales of US Single Family Housing(Outlier
Clost to End)
\nwith Seasonal Adjusted Data (purple line)",
ylab="New Single Family House Sales", xlab="Years")
lines(fit4$time.series[,2],col="purple",ylab="Trend",lwd=2.5)
```


Sales of US Single Family Housing(Outlier Clost to E
with Seasonal Adjusted Data (purple line)
New Single Family House Sales
100
300
500
1975
1980
1985
1990
1995
Years