CANDY PRODUCTION

- **INTRODUCTION**

Sweets, chocolates, and candy are universally enjoyed. In the US, there are holidays themed around giving candy! All this consumption first needs production. The dataset below shows monthly production of candy in the US. The industrial production index measures the real output of all relevant establishments located in the United States, regardless of their ownership, but not those located in U.S. territories.

https://fred.stlouisfed.org/series/IPG3113N

We will proceed towards our forecasting as follows :

Import Data

Plot and Inference

Central Tendency

Decomposition

Naïve Method

Simple Moving Averages

Simple Smoothing

Holt-Winters

Accuracy Summary

Conclusion

- # Introduction

Sweets, chocolates, and candy are universally enjoyed. In the US, there are holidays themed around giving candy! All this consumption first needs production. The dataset below shows monthly production of candy in the US. The industrial production index measures the real output of all relevant establishments located in the United States, regardless of their ownership, but not those located in U.S. territories.

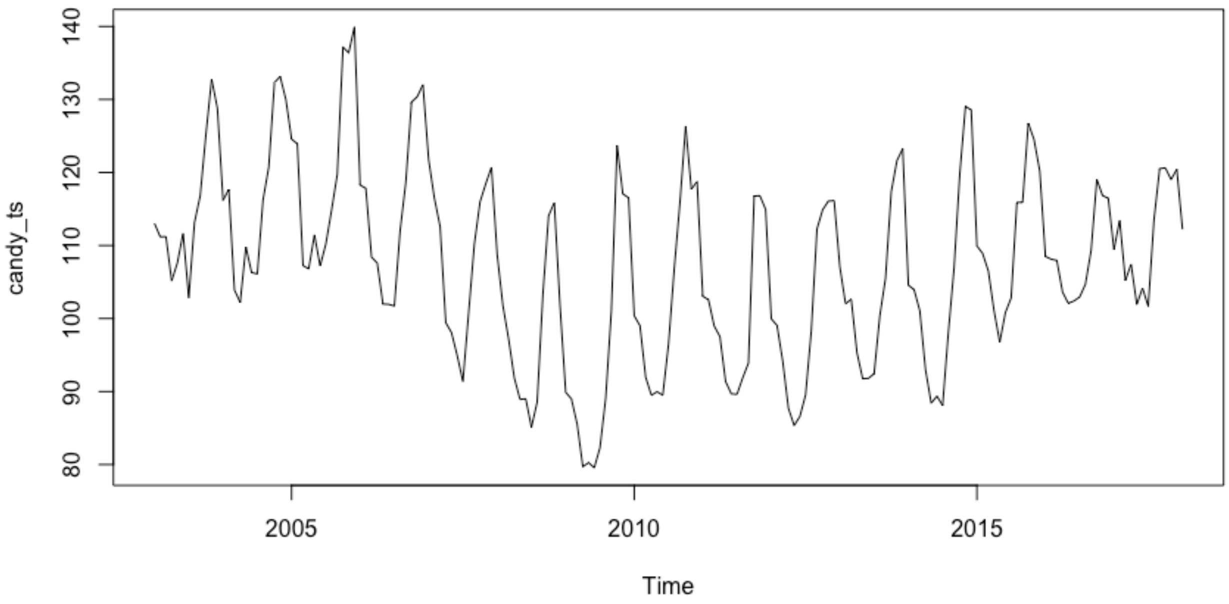
https://fred.stlouisfed.org/series/IPG3113N

- # Import Data

```
Console /Rutgers Doc/Semester 2/BF/Midterm/WorkingDirectory/
> #### Import Data
> library(readr)
> IPG3113N_Spring18<-read_csv("/Rutgers Doc/Semester 2/BF/Midterm/IPG3113N_Spring18.csv")
Parsed with column specification:
cols(
  DATE = col_character(),
  IPG3113N = col_double()
)
> #### Convert into Time Series
> candy_ts<-ts(IPG3113N_Spring18$IPG3113N,frequency = 12,start = c(2003,1))
> #### Plot Time Series
> plot(candy_ts)
> 
```

- ## [Plot and Inference]

- A time series plot.



- Summerising observations of the times series plot

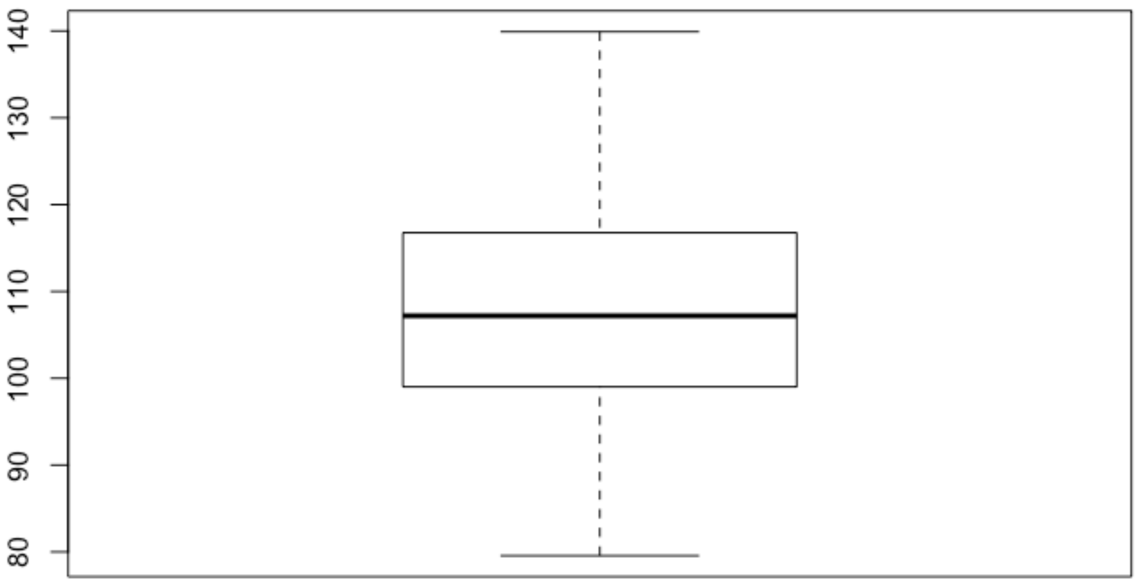
From the above time series, we observe that the data follows a seasonal pattern with peaks and troughs with regular interval of time.

- ## Central Tendency

- Finding out the min, max, mean, median, 1st and 3rd Quartile values of the times series.

```
Console /Rutgers Doc/Semester 2/BF/Midterm/WorkingDirectory/
> #### Central Tendency
> ## min, max, mean, median, 1 st and 3 rd Quartile values
> summary(candy_ts)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  79.57   99.02  107.19  107.45  116.76  139.92
>
```

- Ploting the box plot.



- Summarising observation about the time series from the summary stats and box plot.

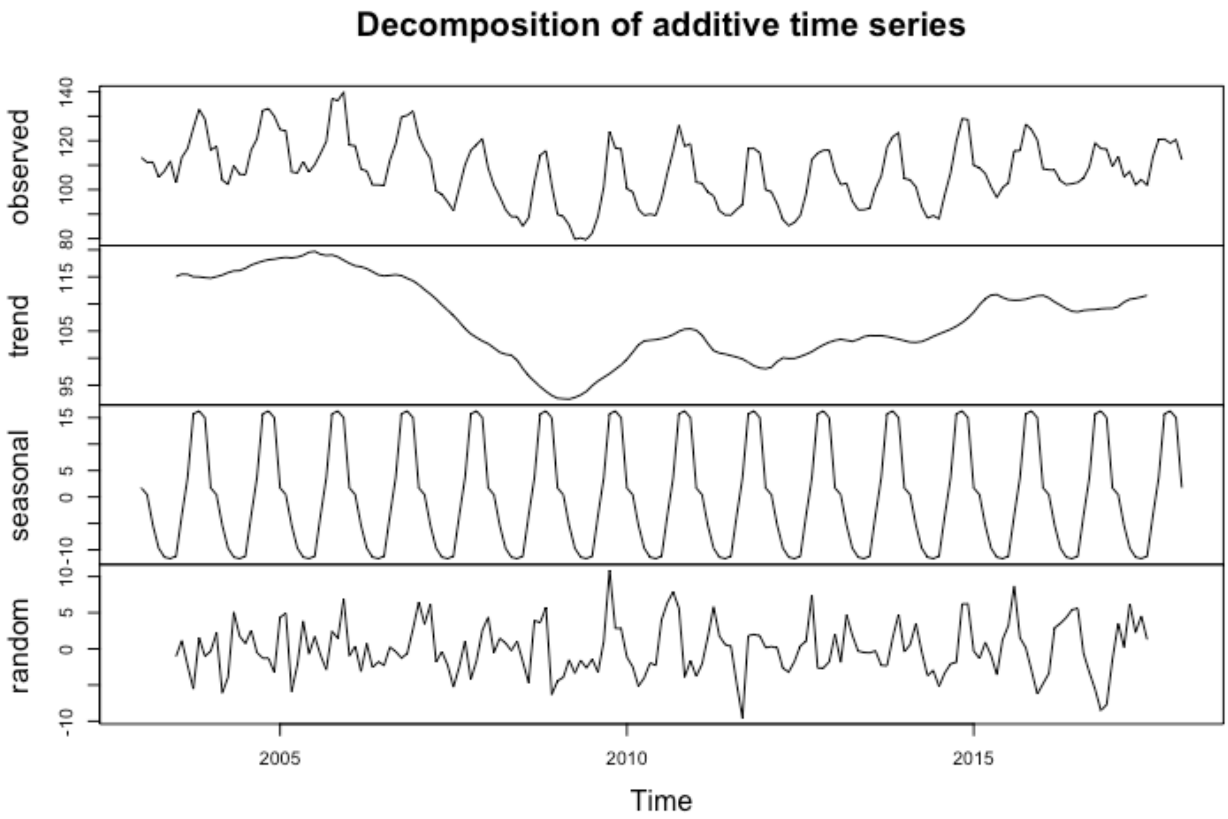
From the above box plot, we observe that the production of candies is high at the end of the year and decreases gradually.

(1st Quartile + 3rd Quartile)/2 = (99.02 + 116.76)/2 = 107.89 ~ (Similar to) Median

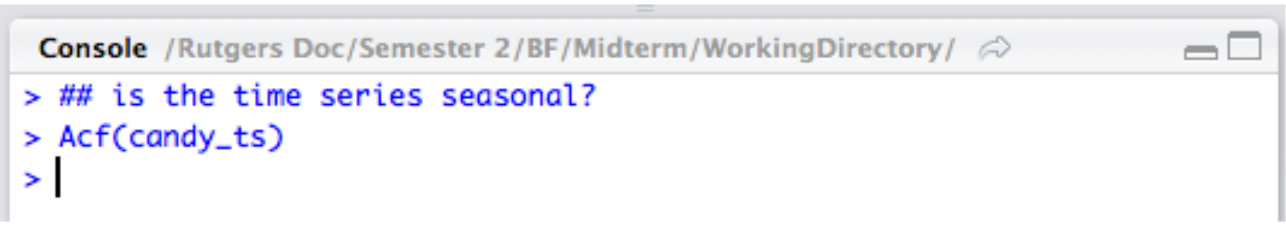
- ## Decomposition

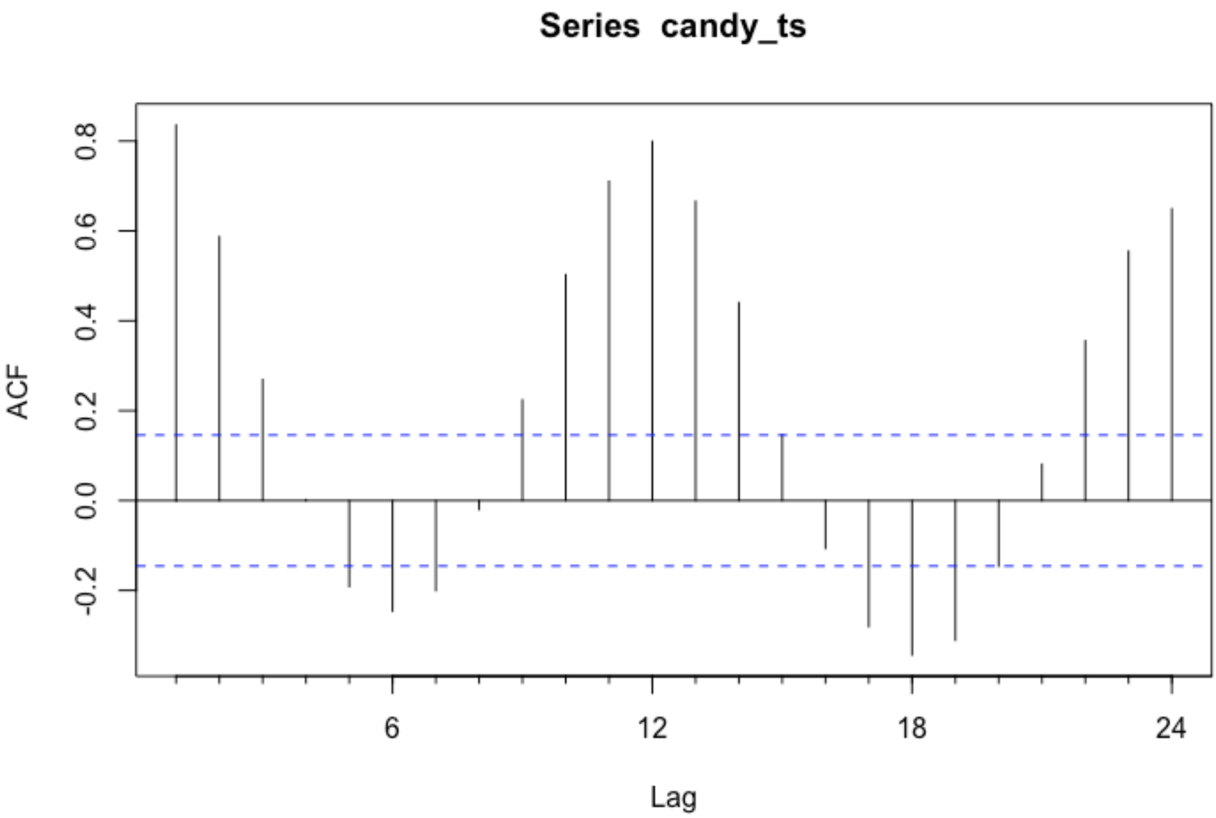
- Plot the decomposition of the time series.

```
Console  /Rutgers Doc/Semester 2/BF/Midterm/WorkingDirectory/
> #### Decompose Candy Time Series
> decomp<-decompose(candy_ts)
> #### Plot Decomposition
> plot(decomp)
>
```



- Finding out whether the times series seasonal or not.





From the above ACF we can observe that our time series is Seasonal.

- Finding out whether the decomposition additive or multiplicative.

Console /Rutgers Doc/Semester 2/BF/Midterm/WorkingDirectory/

```
$type
[1] "additive"
```

- If decomposition is seasonal, let's find the values of the seasonal monthly indices.

Console /Rutgers Doc/Semester 2/BF/Midterm/WorkingDirectory/

```
$seasonal
            Jan        Feb        Mar        Apr        May        Jun        Jul        Aug        Sep
2003   1.7367141  0.4089563 -5.3388684 -9.6736722 -11.3775282 -11.6560576 -11.1830346 -3.4903600  3.6323090
2004   1.7367141  0.4089563 -5.3388684 -9.6736722 -11.3775282 -11.6560576 -11.1830346 -3.4903600  3.6323090
2005   1.7367141  0.4089563 -5.3388684 -9.6736722 -11.3775282 -11.6560576 -11.1830346 -3.4903600  3.6323090
2006   1.7367141  0.4089563 -5.3388684 -9.6736722 -11.3775282 -11.6560576 -11.1830346 -3.4903600  3.6323090
2007   1.7367141  0.4089563 -5.3388684 -9.6736722 -11.3775282 -11.6560576 -11.1830346 -3.4903600  3.6323090
2008   1.7367141  0.4089563 -5.3388684 -9.6736722 -11.3775282 -11.6560576 -11.1830346 -3.4903600  3.6323090
2009   1.7367141  0.4089563 -5.3388684 -9.6736722 -11.3775282 -11.6560576 -11.1830346 -3.4903600  3.6323090
2010   1.7367141  0.4089563 -5.3388684 -9.6736722 -11.3775282 -11.6560576 -11.1830346 -3.4903600  3.6323090
2011   1.7367141  0.4089563 -5.3388684 -9.6736722 -11.3775282 -11.6560576 -11.1830346 -3.4903600  3.6323090
2012   1.7367141  0.4089563 -5.3388684 -9.6736722 -11.3775282 -11.6560576 -11.1830346 -3.4903600  3.6323090
2013   1.7367141  0.4089563 -5.3388684 -9.6736722 -11.3775282 -11.6560576 -11.1830346 -3.4903600  3.6323090
2014   1.7367141  0.4089563 -5.3388684 -9.6736722 -11.3775282 -11.6560576 -11.1830346 -3.4903600  3.6323090
2015   1.7367141  0.4089563 -5.3388684 -9.6736722 -11.3775282 -11.6560576 -11.1830346 -3.4903600  3.6323090
2016   1.7367141  0.4089563 -5.3388684 -9.6736722 -11.3775282 -11.6560576 -11.1830346 -3.4903600  3.6323090
2017   1.7367141  0.4089563 -5.3388684 -9.6736722 -11.3775282 -11.6560576 -11.1830346 -3.4903600  3.6323090
2018   1.7367141
            Oct        Nov        Dec
2003  15.6952043 16.2695507 14.9767867
2004  15.6952043 16.2695507 14.9767867
2005  15.6952043 16.2695507 14.9767867
2006  15.6952043 16.2695507 14.9767867
2007  15.6952043 16.2695507 14.9767867
2008  15.6952043 16.2695507 14.9767867
2009  15.6952043 16.2695507 14.9767867
2010  15.6952043 16.2695507 14.9767867
2011  15.6952043 16.2695507 14.9767867
2012  15.6952043 16.2695507 14.9767867
2013  15.6952043 16.2695507 14.9767867
2014  15.6952043 16.2695507 14.9767867
2015  15.6952043 16.2695507 14.9767867
2016  15.6952043 16.2695507 14.9767867
2017  15.6952043 16.2695507 14.9767867
2018
```

- Finding out which month is the value of time series high and for which month is it low.

From the above plots, we observe that the time series was high for the month of November i.e 16.2695 and was low for the month of June i.e. -11.65605

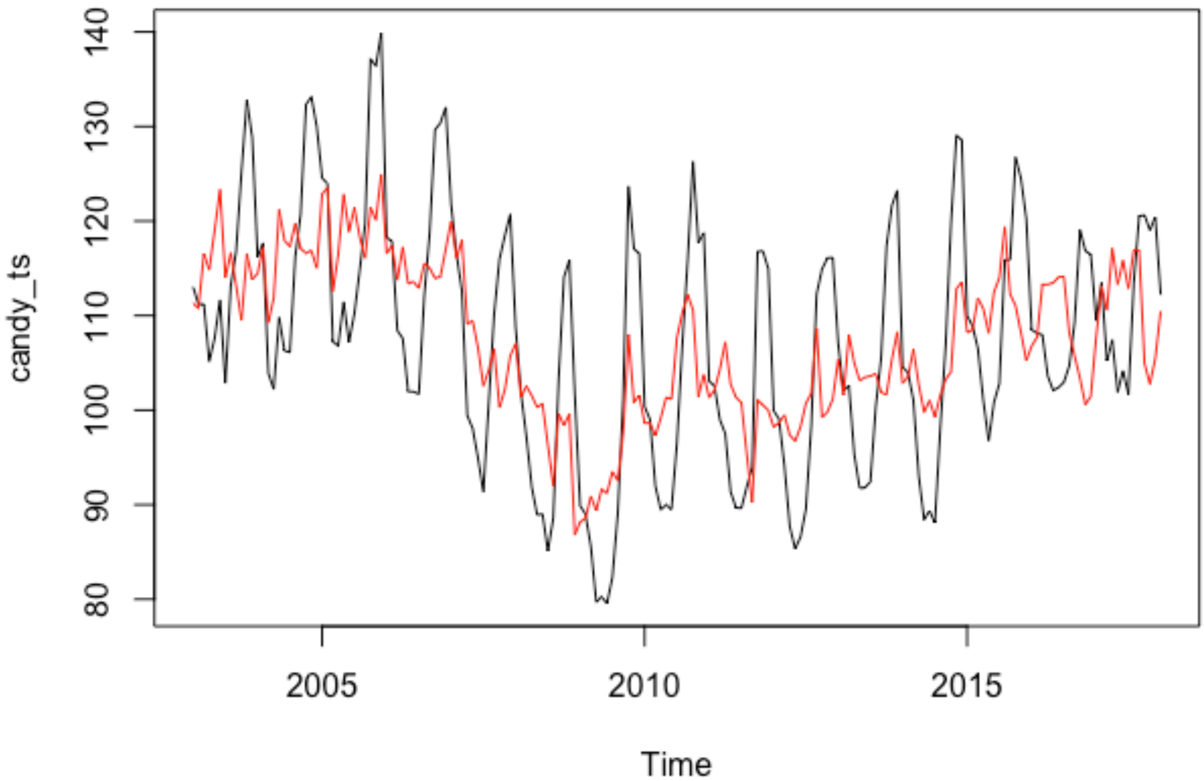
- Finding out the reason behind the value being high in those months and low in those months.

Since the period of year end has festivals like Halloween, Thanks Giving, Christmas and New Year, we predict that the candies are sold the most in this period than rest of the year which ultimately causes rise in the production of Candies.

- Showing the plot for time series adjusted for seasonality. Overlaying this with the line for actual time series. Finding out does seasonality have big fluctuations to the value of time series.

Console /Rutgers Doc/Semester 2/BF/Midterm/WorkingDirectory/

```
> #### Seasonal Adjusted & Overlay with Actual Time Series
> seas_adj<-seasadj(decomp)
> plot(candy_ts)
> lines(seas_adj,col='red')
>
```

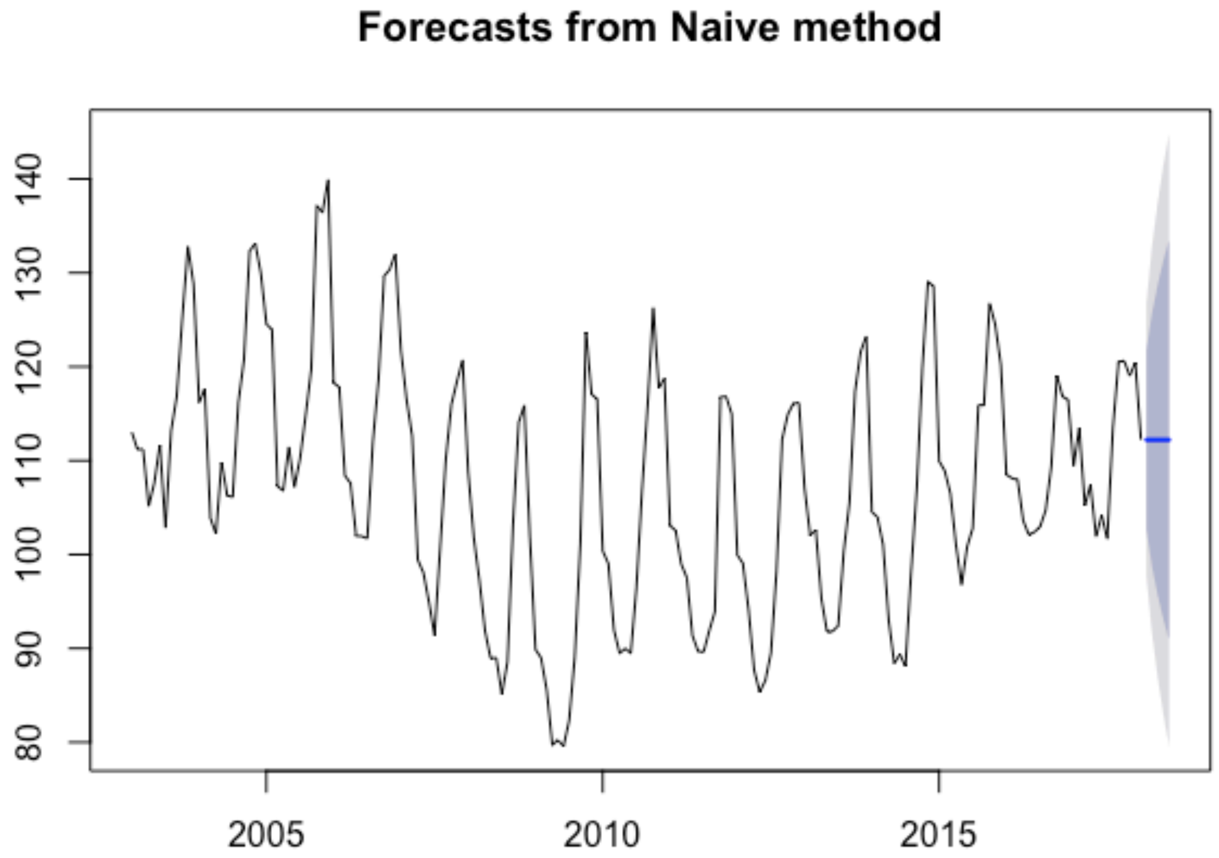


In the above plot, we observe that there is fluctuations in the peaks and troughs, making the plot highly seasonal.

- **Naïve Method**

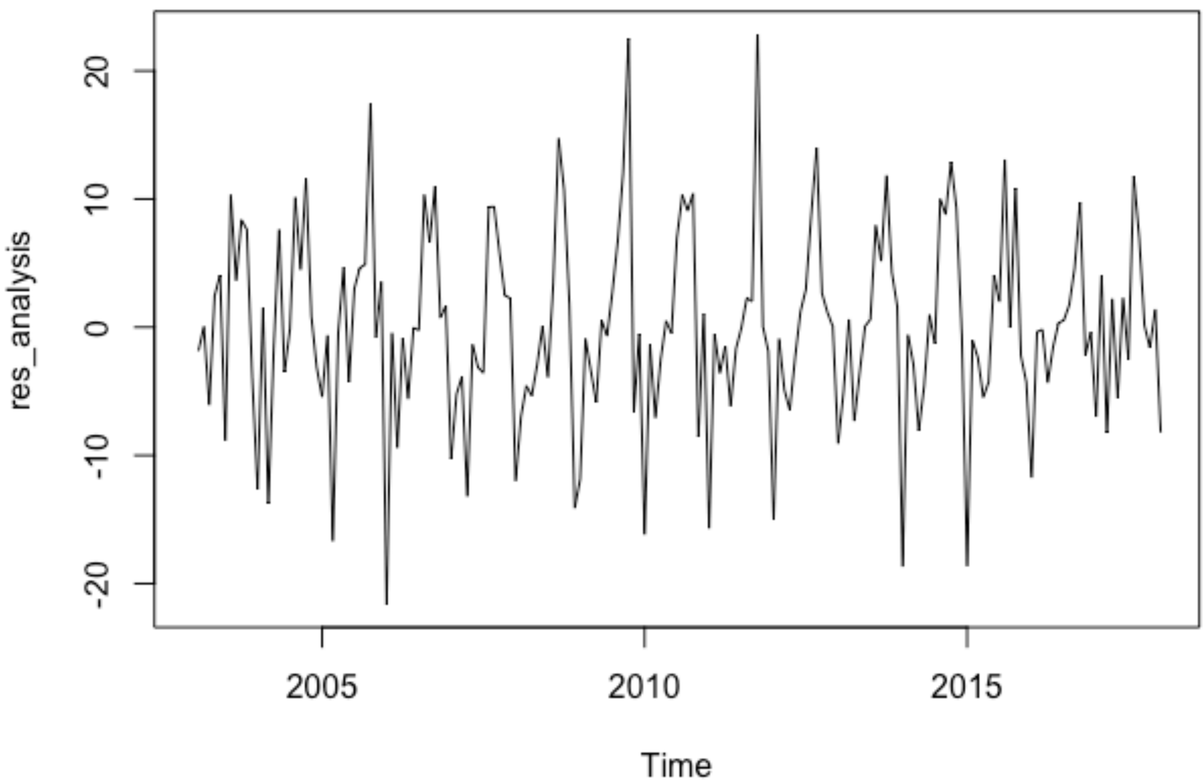
- Output

```
Console /Rutgers Doc/Semester 2/BF/Midterm/WorkingDirectory/
> #### Naive Forecast
> naive_fore<-naive(candy_ts,5)
> #### Plot Naive Forecast
> plot(naive_fore)
>
```



- Performing Residual Analysis for this technique.
  - Plot of residuals.

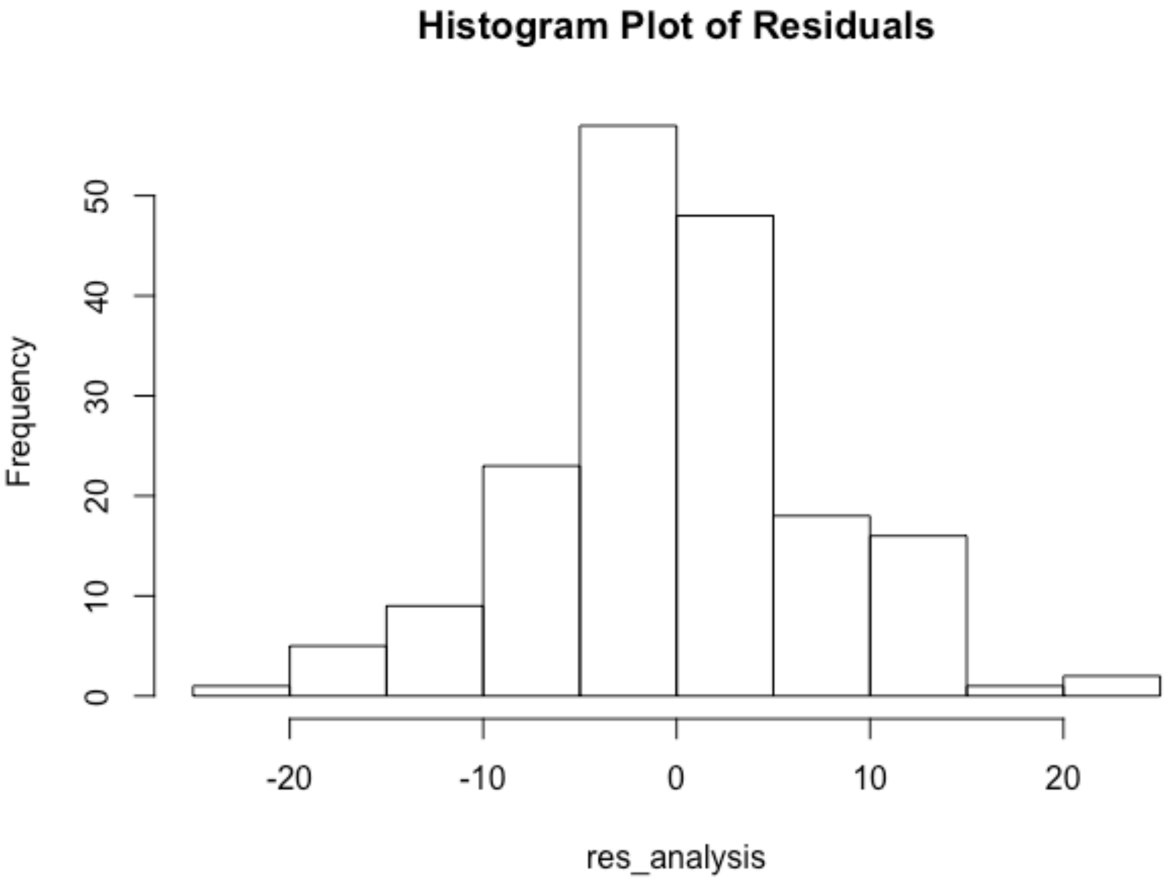
```
Console /Rutgers Doc/Semester 2/BF/Midterm/WorkingDirectory/
> #### Residual Analysis
> res_analysis<-residuals(naive_fore)
> #### Plot Residual Analysis
> plot(res_analysis)
>
```



We see that the residuals in the above plot has high variance

  - Histogram plot of residuals.

```
Console /Rutgers Doc/Semester 2/BF/Midterm/WorkingDirectory/
> #### Plot Histogram of Residual Analysis
> hist_res_analysis<-hist(res_analysis,breaks = 10,main = "Histogram Plot of Residuals")
>
```

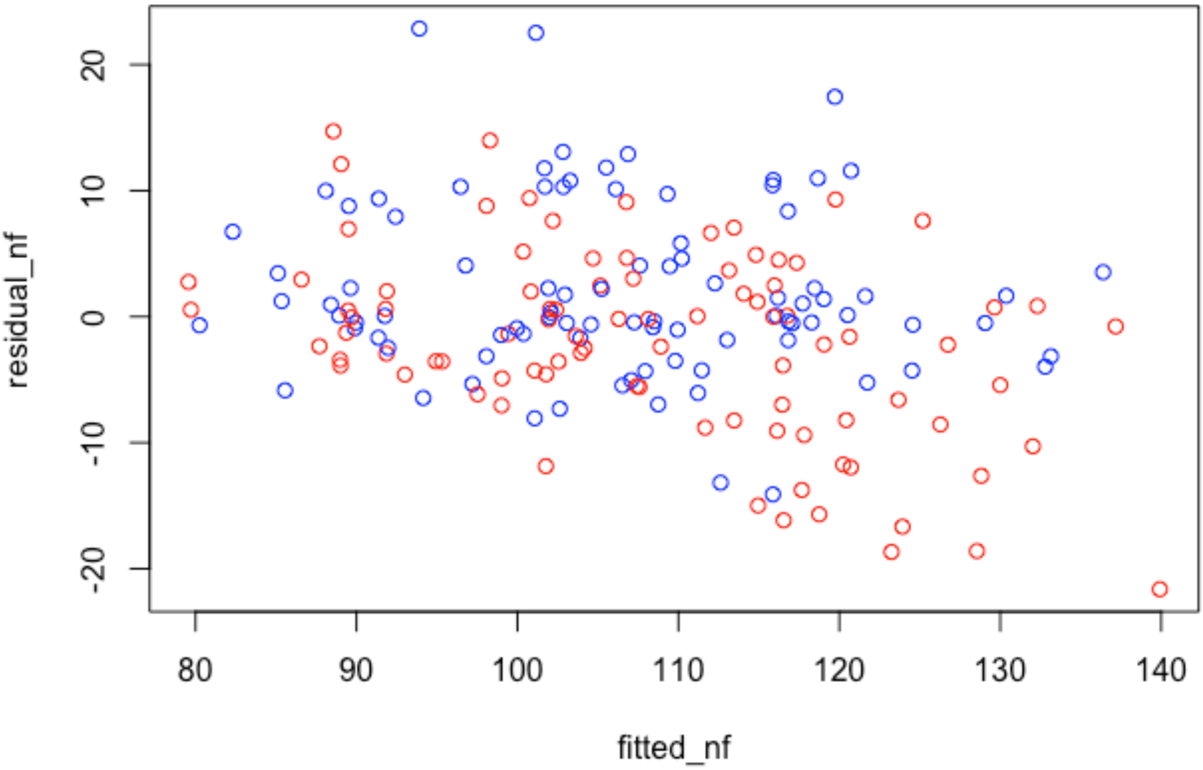


The histogram suggests that the residuals are normally distributed.

- Plot of fitted values vs. residuals.

Console /Rutgers Doc/Semester 2/BF/Midterm/WorkingDirectory/

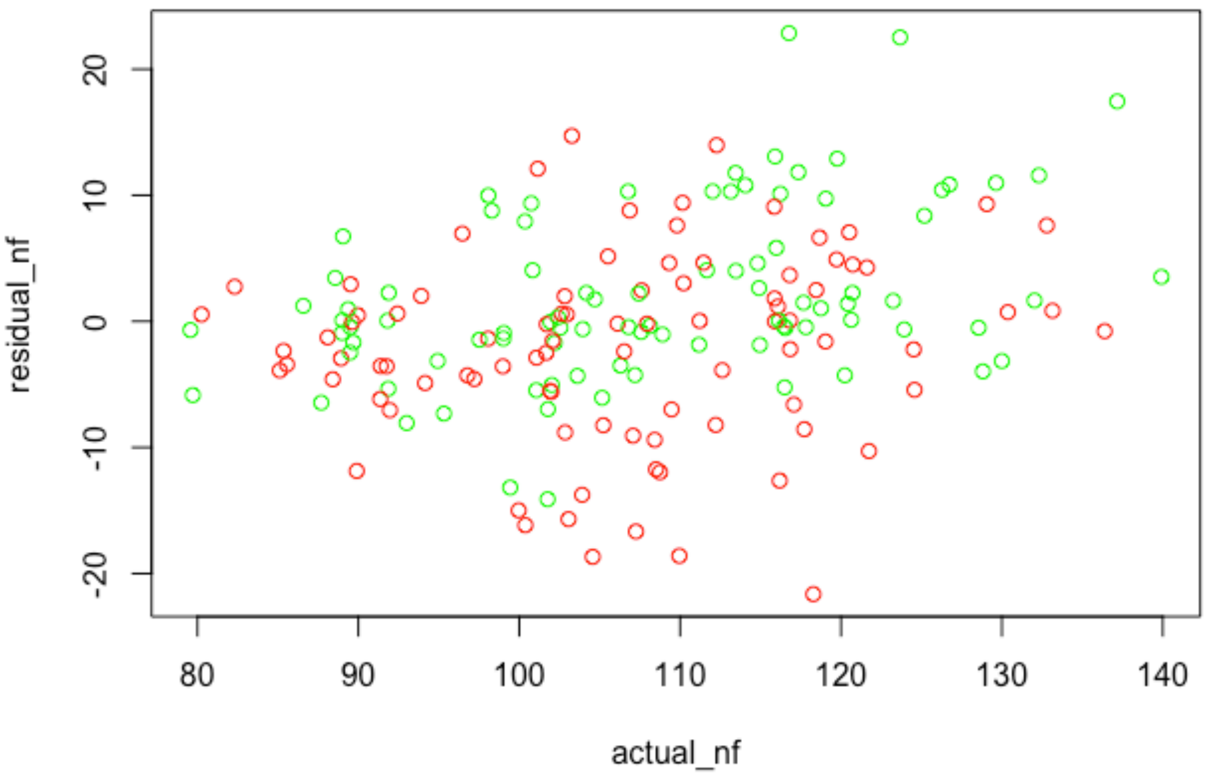
```
> #### Plot of Fitted values vs Residuals
> fitted_nf<-naive_fore$fitted
> residual_nf<-naive_fore$residuals
> plot(residual_nf~fitted_nf)
> plot(residual_nf~fitted_nf,col=c("red","blue"))
>
```



The scatter plot says that the residuals and fitted values are randomly distributed.

- Plot of actual values vs. residuals.

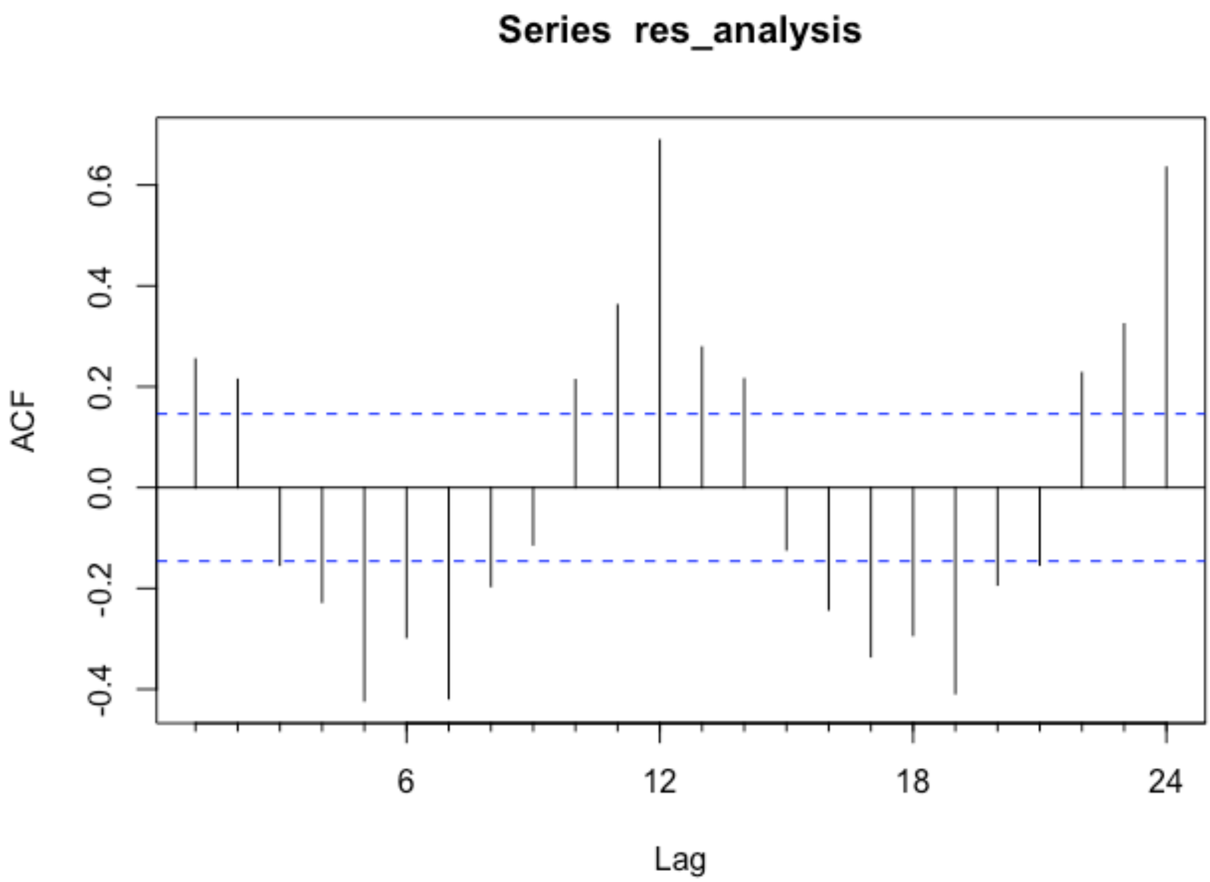
```
Console /Rutgers Doc/Semester 2/BF/Midterm/WorkingDirectory/
> #### Plot of Actual values vs Residuals
> actual_nf<-naive_fore$x
> plot(residual_nf~actual_nf)
> plot(residual_nf~actual_nf,col=c("red","green"))
>
```



The scatter plot says that the residuals and actual values are randomly distributed

- An ACF plot of the residuals.

```
Console /Rutgers Doc/Semester 2/BF/Midterm/WorkingDirectory/
> #### Acf of Residuals
> Acf(res_analysis)
>
```



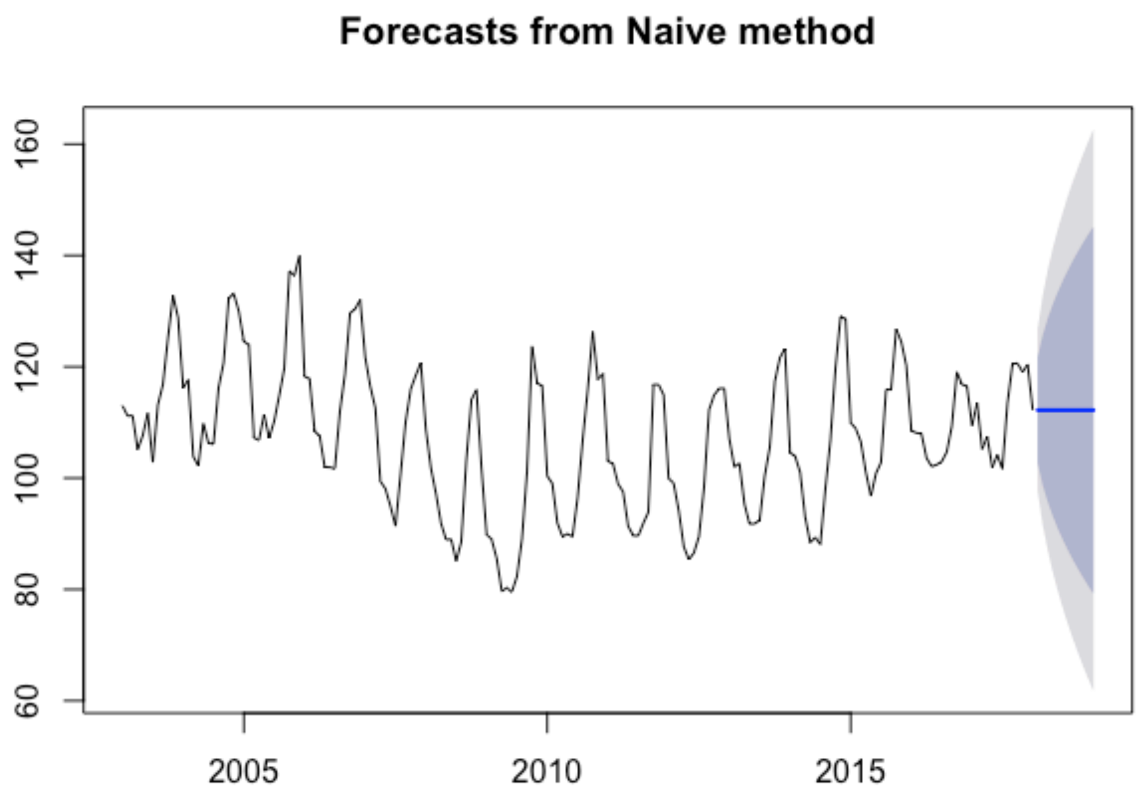
The ACF suggests that the production is high at the end of the year and less in the mid.

- Printing the 5 measures of accuracy for this forecasting technique

```
Console /Rutgers Doc/Semester 2/BF/Midterm/WorkingDirectory/
> #### Accuracy of Naive Forecast
> accuracy(naive_fore)
                       ME     RMSE      MAE        MPE     MAPE      MASE
Training set -0.004547778 7.422458 5.470242 -0.2333585 5.057813 0.9020712
                  ACF1
Training set 0.2547176
>
```

- Forecast
  - Time series value for next year.

```
Console /Rutgers Doc/Semester 2/BF/Midterm/WorkingDirectory/
> #### Forecast for the Next Year
> fore<-forecast(naive_fore)
> fore
         Point Forecast     Lo 80    Hi 80    Lo 95    Hi 95
Feb 2018       112.2117 102.69944 121.7240 97.66395 126.7595
Mar 2018       112.2117  98.75933 125.6641 91.63807 132.7853
Apr 2018       112.2117  95.73598 128.6874 87.01426 137.4091
May 2018       112.2117  93.18717 131.2362 83.11620 141.3072
Jun 2018       112.2117  90.94163 133.4818 79.68194 144.7415
Jul 2018       112.2117  88.91151 135.5119 76.57713 147.8463
Aug 2018       112.2117  87.04462 137.3788 73.72197 150.7014
Sep 2018       112.2117  85.30696 139.1164 71.06445 153.3590
Oct 2018       112.2117  83.67491 140.7485 68.56845 155.8550
Nov 2018       112.2117  82.13128 142.2921 66.20767 158.2157
Dec 2018       112.2117  80.66309 143.7603 63.96227 160.4611
Jan 2019       112.2117  79.26025 145.1631 61.81681 162.6066
>
```



- Summarising this forecasting technique

  - How good is the accuracy?

The RMSE for the model is high i.e. 7.4224 which says that there is huge difference in expected and actual values.

  - What does it predict the value of time series will be in one year?

The model predicts the exact same value for the next year. The value for the month Feb 2018 is 112.2117 and for the month Feb 2019 is 112.2117 which is exactly the same.

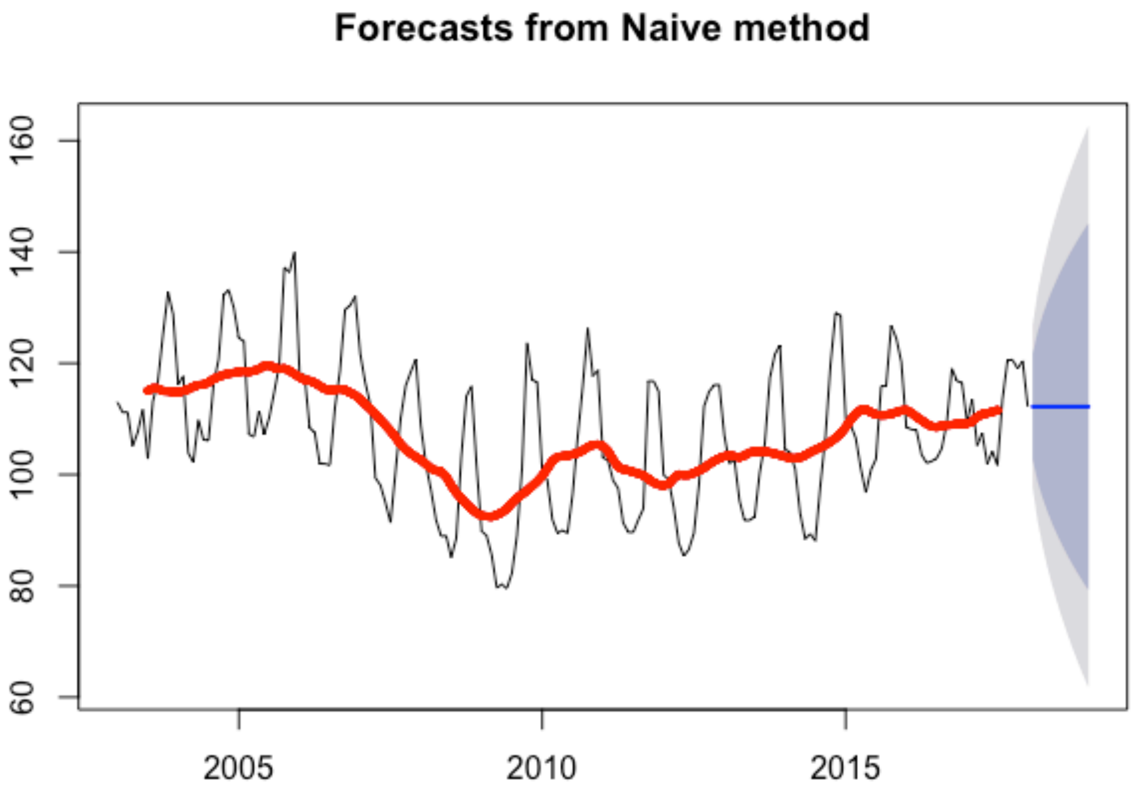
  - Other observation

We observe a minimal seasonality in the plot.

- ## Simple Moving Averages

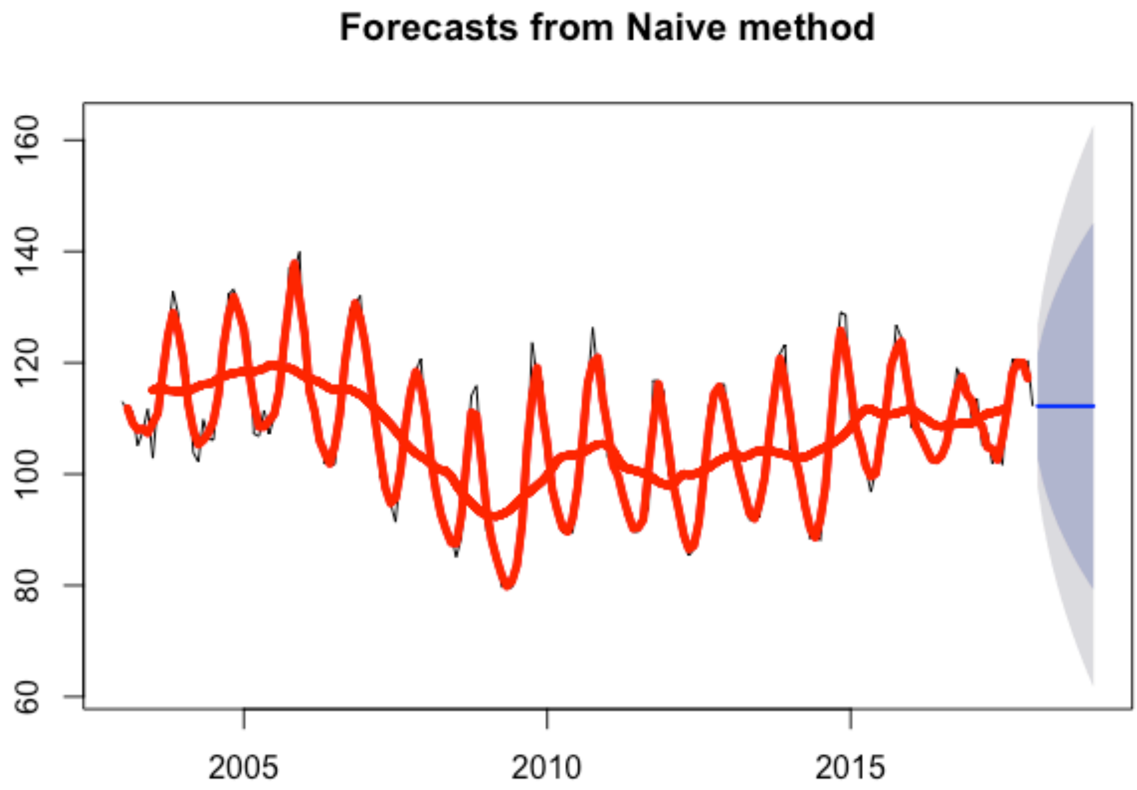
- Plot the graph for time series.

```
Console /Rutgers Doc/Semester 2/BF/Midterm/WorkingDirectory/
> #### Ploting the graph for Time Series
> ma<-ma(candy_ts,order = 12)
> lines(ma, col='Red', lwd=4)
>
```

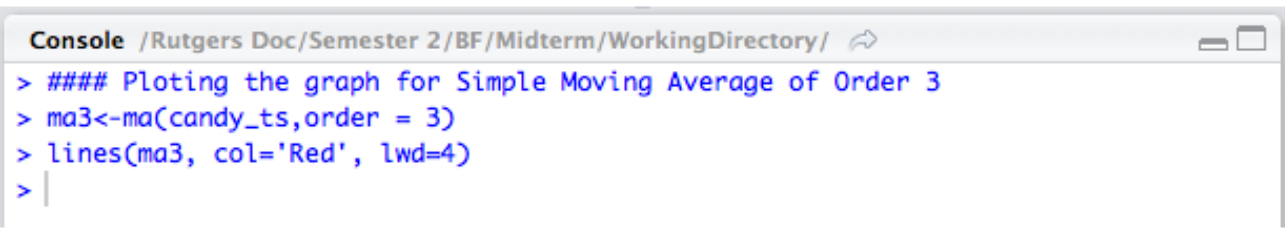


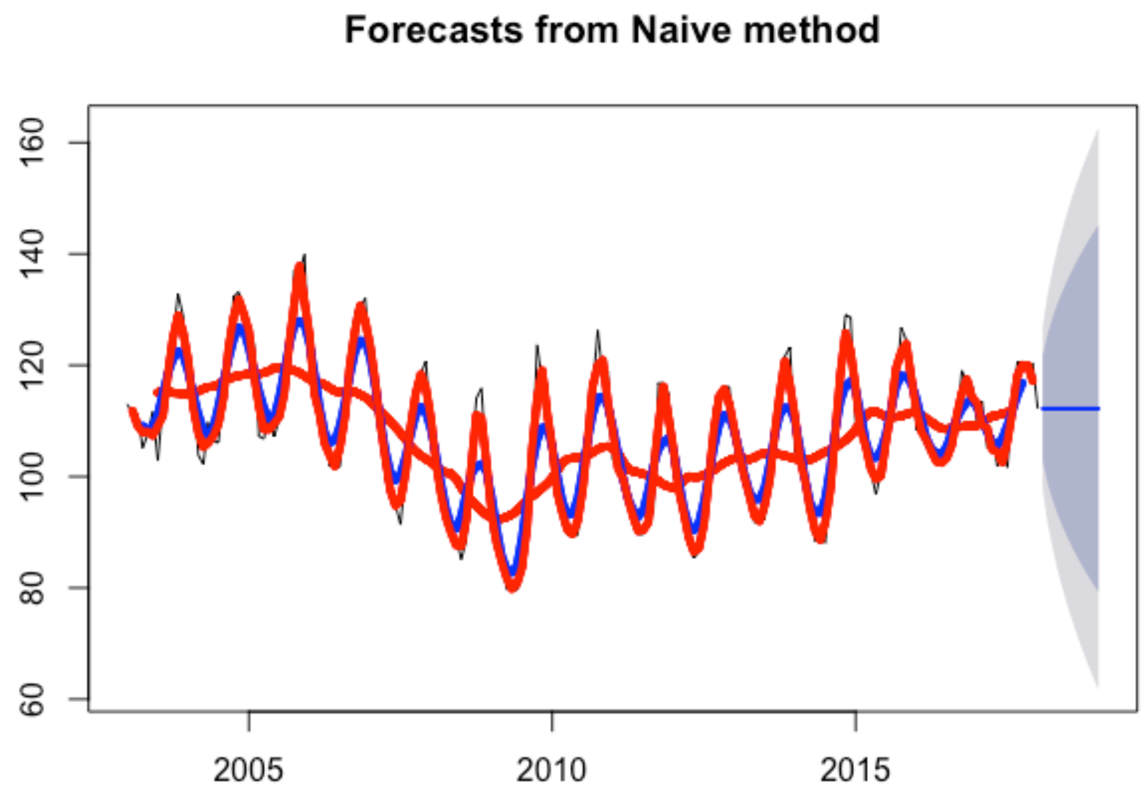
- Showing the Simple Moving average of order 3 on the plot above in Red

```
Console /Rutgers Doc/Semester 2/BF/Midterm/WorkingDirectory/
> #### Ploting the graph for Simple Moving Average of Order 3
> ma3<-ma(candy_ts,order = 3)
> lines(ma3, col='Red', lwd=4)
>
```



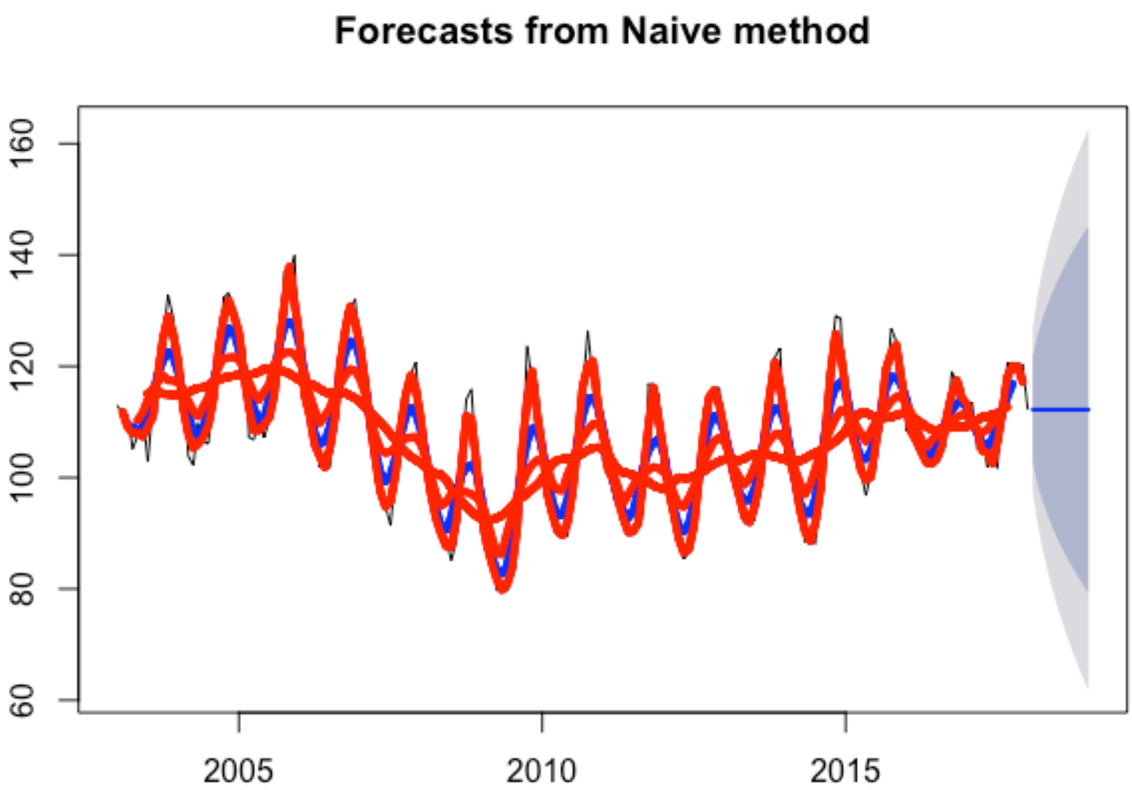
- Showing the Simple Moving average of order 6 on the plot above in Blue





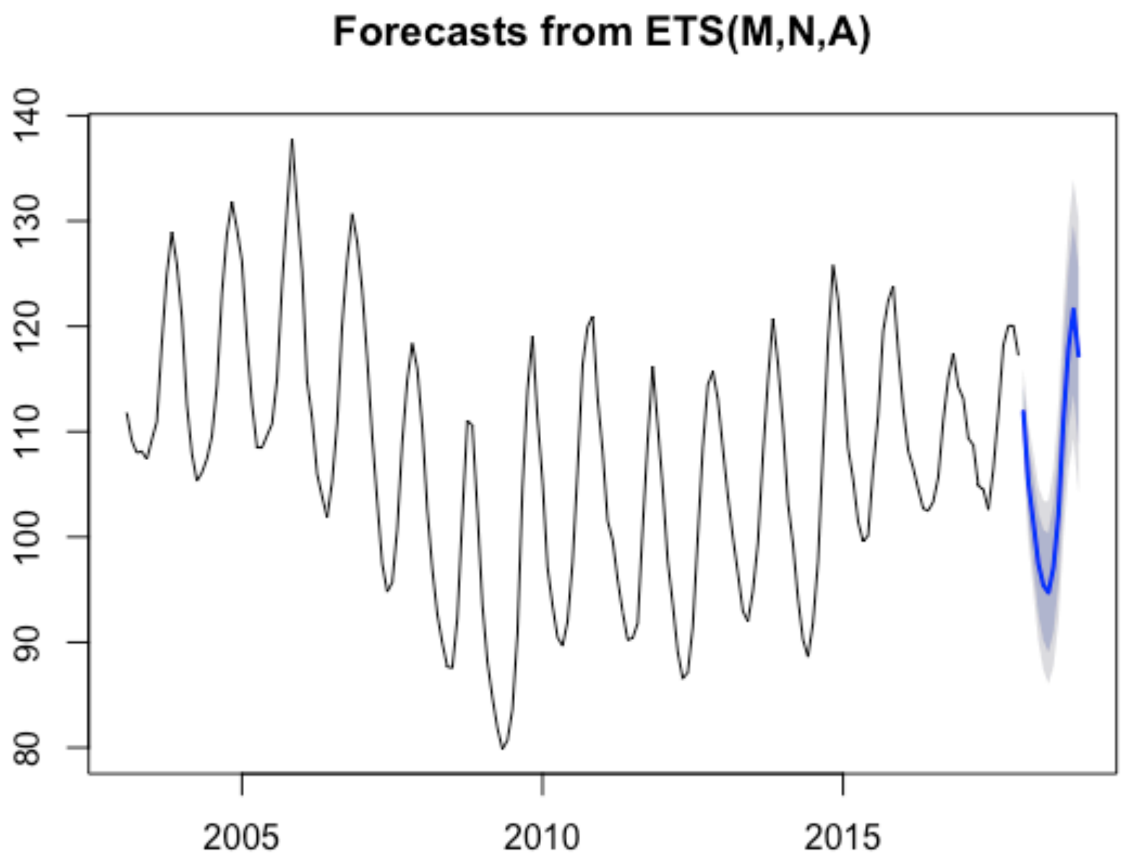
- Showing the Simple Moving average of order 9 on the plot above in Green

```
Console /Rutgers Doc/Semester 2/BF/Midterm/WorkingDirectory/
> #### Ploting the graph for Simple Moving Average of Order 9
> ma9<-ma(candy_ts,order = 9)
> lines(ma9, col='Red', lwd=4)
>
```



- Showing the forecast of next 12 months using one of the simple average order that we feel works best for time series

```
Console /Rutgers Doc/Semester 2/BF/Midterm/WorkingDirectory/
> #### Forecasting for next one year which gives optimum order
> simple_forecast <- forecast(ma3, 12)
```



- Observations of the plot as the moving average order goes up.

  The error increases as the order is increased.

- ## Simple Smoothing

    - Performing a simple smoothing forecast for next 12 months for the time series.

```
Console /Rutgers Doc/Semester 2/BF/Midterm/WorkingDirectory/
> #### Simple Smoothing Forecast for Next 12 months
> sim_smooth<-ses(candy_ts,h=12)
> sim_smooth
         Point Forecast     Lo 80    Hi 80    Lo 95    Hi 95
Feb 2018       112.2125 102.72633 121.6987 97.70464 126.7204
Mar 2018       112.2125  98.79769 125.6274 91.69631 132.7287
Apr 2018       112.2125  95.78305 128.6420 87.08582 137.3392
May 2018       112.2125  93.24156 131.1835 83.19894 141.2261
Jun 2018       112.2125  91.00245 133.4226 79.77452 144.6505
Jul 2018       112.2125  88.97813 135.4469 76.67859 147.7465
Aug 2018       112.2125  87.11657 137.3085 73.83158 150.5935
Sep 2018       112.2125  85.38387 139.0412 71.18164 153.2434
Oct 2018       112.2125  83.75648 140.6686 68.69276 155.7323
Nov 2018       112.2125  82.21725 142.2078 66.33871 158.0863
Dec 2018       112.2125  80.75324 143.6718 64.09971 160.3253
Jan 2019       112.2125  79.35440 145.0706 61.96037 162.4647
>
```

```
Console /Rutgers Doc/Semester 2/BF/Midterm/WorkingDirectory/
> summary(sim_smooth)

Forecast method: Simple exponential smoothing

Model Information:
Simple exponential smoothing

Call:
 ses(y = candy_ts, h = 12)

  Smoothing parameters:
    alpha = 0.9999

  Initial states:
    l = 113.0288

  sigma:  7.4021

     AIC     AICc      BIC
1671.567 1671.703 1681.163

Error measures:
                       ME     RMSE      MAE        MPE     MAPE      MASE      ACF1
Training set -0.004510014 7.402115 5.440284 -0.2320932 5.030116 0.8971309 0.2548242

Forecasts:
         Point Forecast     Lo 80    Hi 80    Lo 95    Hi 95
Feb 2018       112.2125 102.72633 121.6987 97.70464 126.7204
Mar 2018       112.2125  98.79769 125.6274 91.69631 132.7287
Apr 2018       112.2125  95.78305 128.6420 87.08582 137.3392
May 2018       112.2125  93.24156 131.1835 83.19894 141.2261
Jun 2018       112.2125  91.00245 133.4226 79.77452 144.6505
Jul 2018       112.2125  88.97813 135.4469 76.67859 147.7465
Aug 2018       112.2125  87.11657 137.3085 73.83158 150.5935
Sep 2018       112.2125  85.38387 139.0412 71.18164 153.2434
Oct 2018       112.2125  83.75648 140.6686 68.69276 155.7323
Nov 2018       112.2125  82.21725 142.2078 66.33871 158.0863
Dec 2018       112.2125  80.75324 143.6718 64.09971 160.3253
Jan 2019       112.2125  79.35440 145.0706 61.96037 162.4647
>
```

- What is the value of alpha? What does that value signify?

The value of Alpha is 0.9999 which signifies the optimum level of smoothing parameter.

- What is the value of initial state?

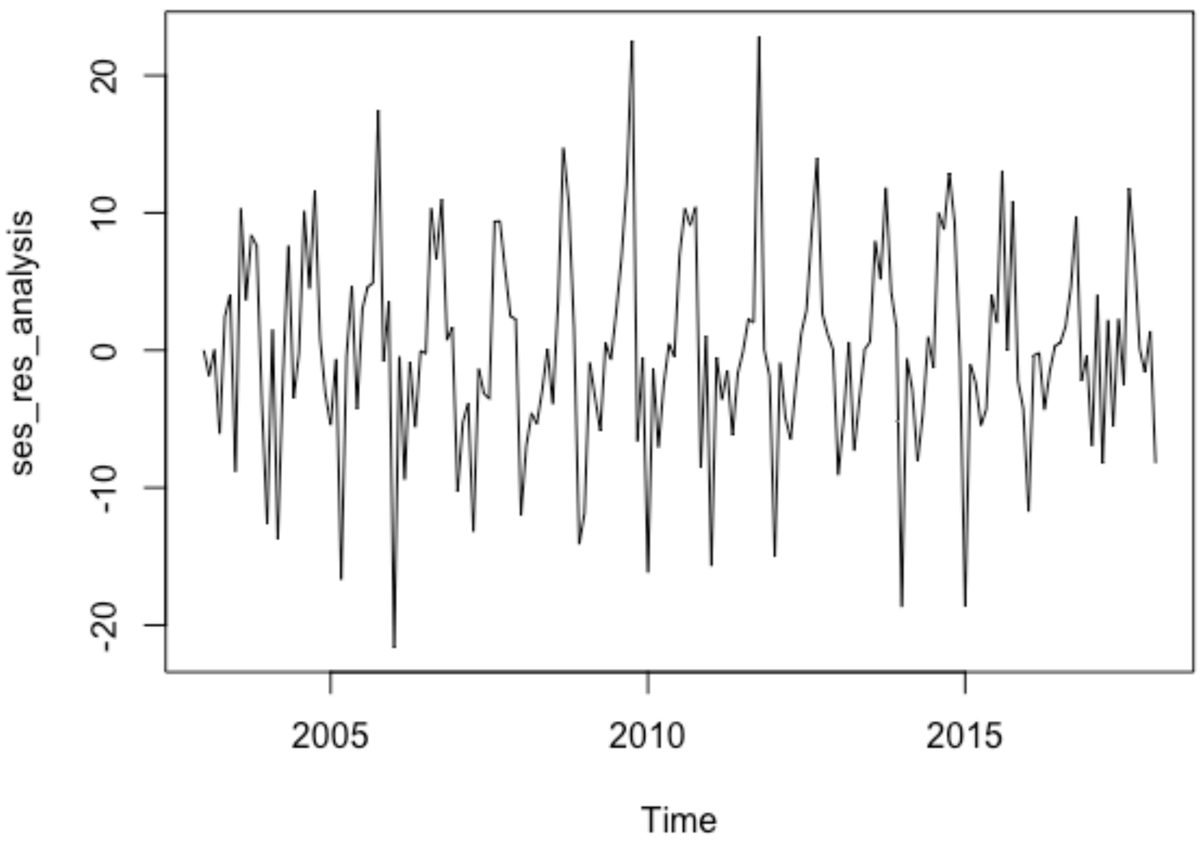
The initial state is 113.0288

- What is the value of sigma? What does the sigma signify?

The value of Sigma is 7.4021 which signifies that the variation of the residuals and is high around the mean.

- Performing Residual Analysis for this technique.
    - Plot of residuals.

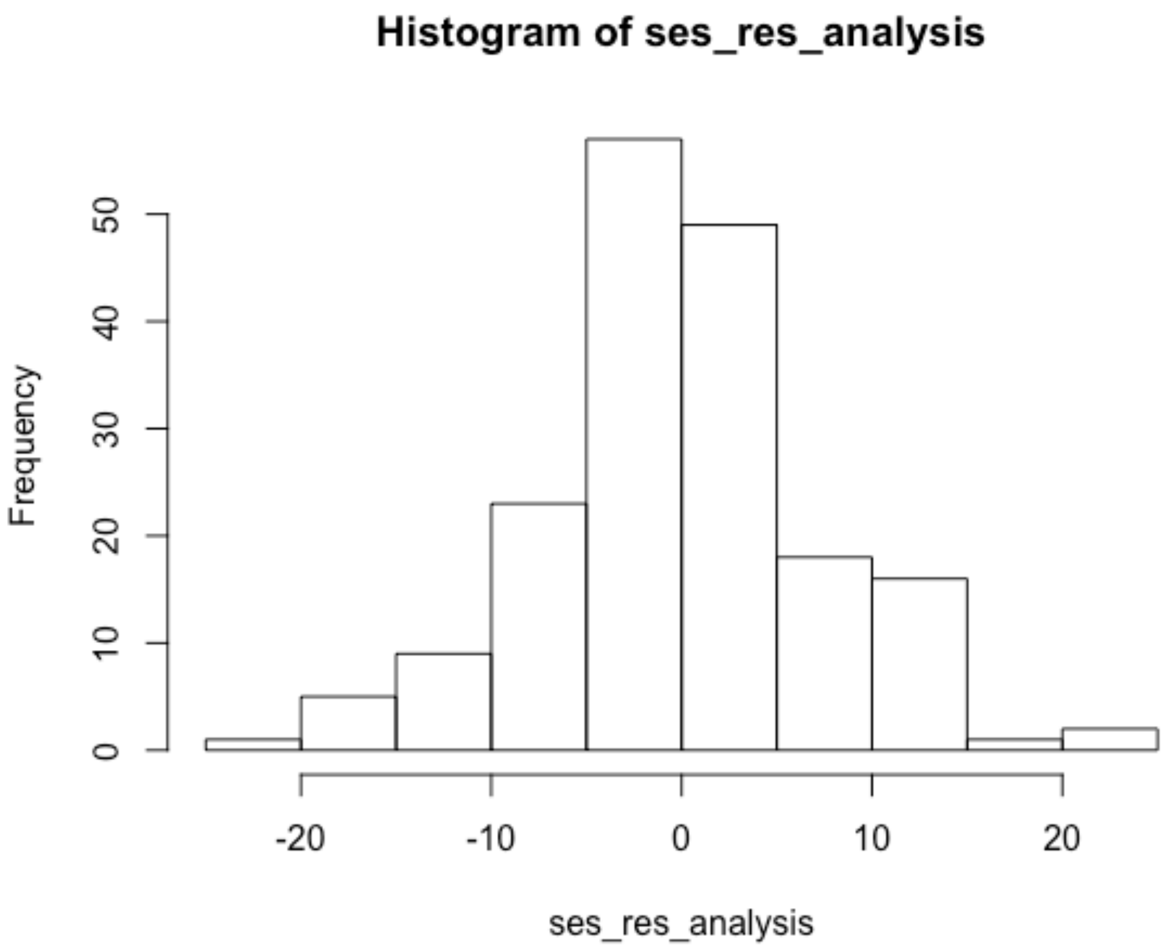
```
Console /Rutgers Doc/Semester 2/BF/Midterm/WorkingDirectory/
> #### Residual Analysis for Simple Smoothing
> ses_res_analysis<-residuals(sim_smooth,h=12)
> #### Plot for Residual Analysis for Simple Smoothing
> plot(ses_res_analysis)
>
```



We observe from the above plot that the residuals have high variability.

- Histogram plot of residuals.

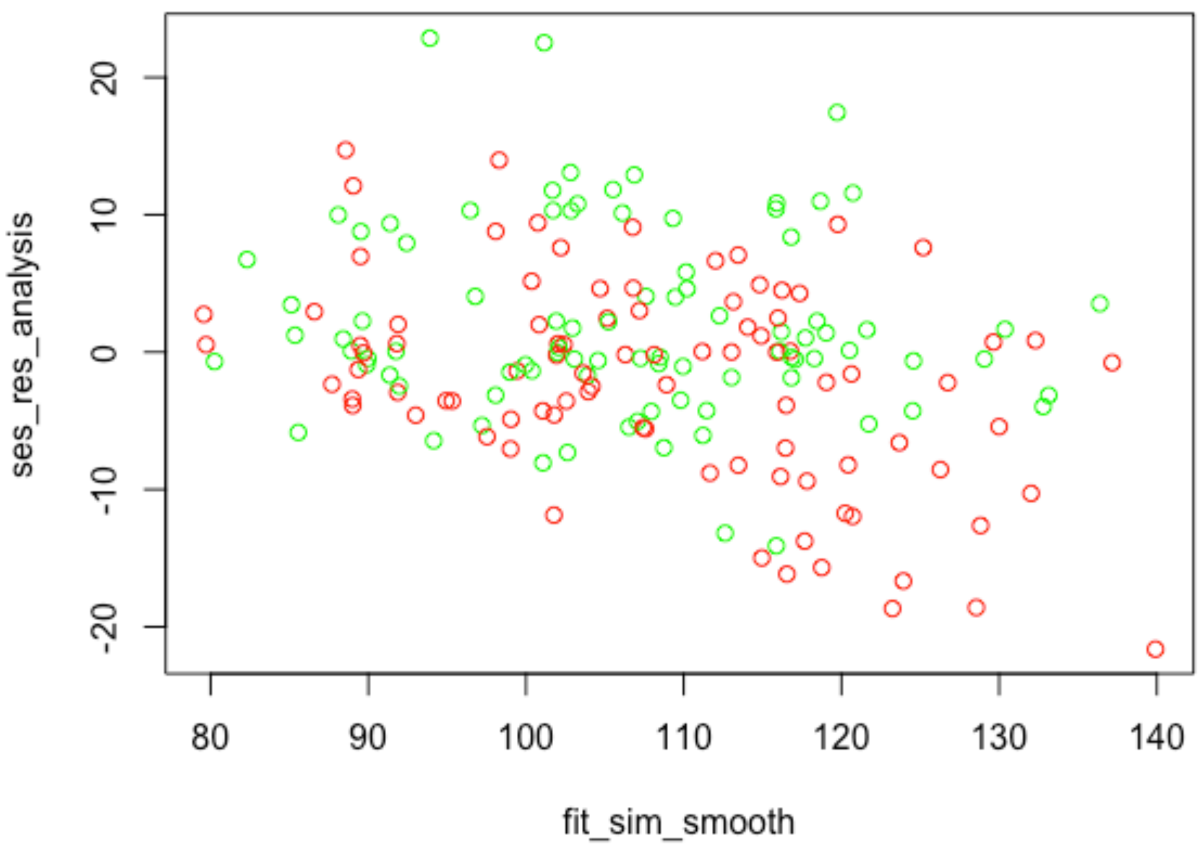
```
Console /Rutgers Doc/Semester 2/BF/Midterm/WorkingDirectory/
> #### Histogram Plot of Residuals
> hist_ses_res_analysis<-hist(ses_res_analysis)
>
```



The plot suggests that the residuals are evenly distributed.

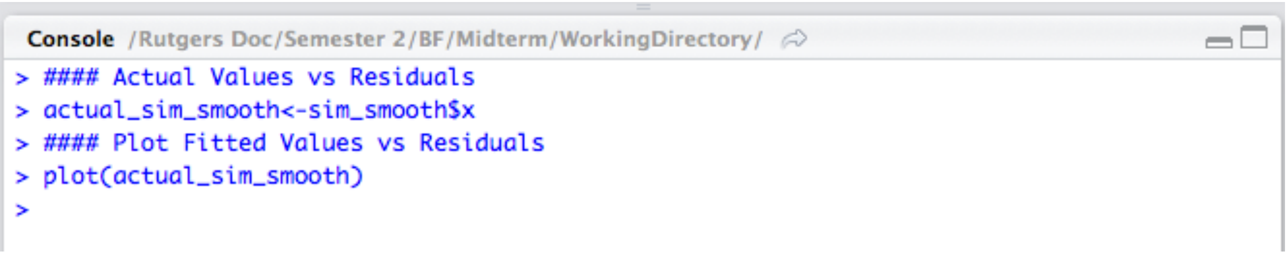
- Plot of fitted values vs. residuals.

```
Console /Rutgers Doc/Semester 2/BF/Midterm/WorkingDirectory/
> #### Fitted Values vs Residuals
> fit_sim_smooth<-sim_smooth$fitted
> #### Plot Fitted Values vs Residuals
> plot(fit_sim_smooth)
>
```



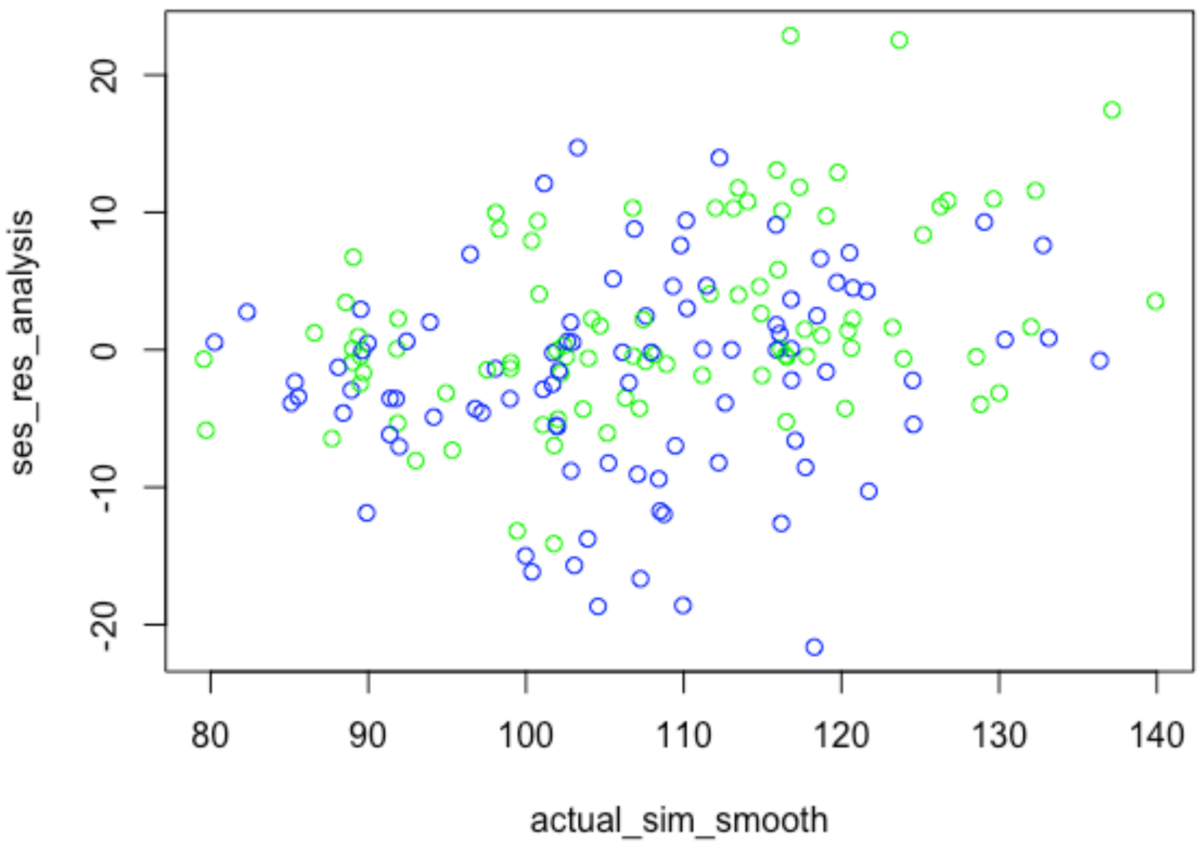
The above plot says that the residuals anydfitted values are randomly distributed.

- Plot of actual values vs. residuals.



Console /Rutgers Doc/Semester 2/BF/Midterm/WorkingDirectory/

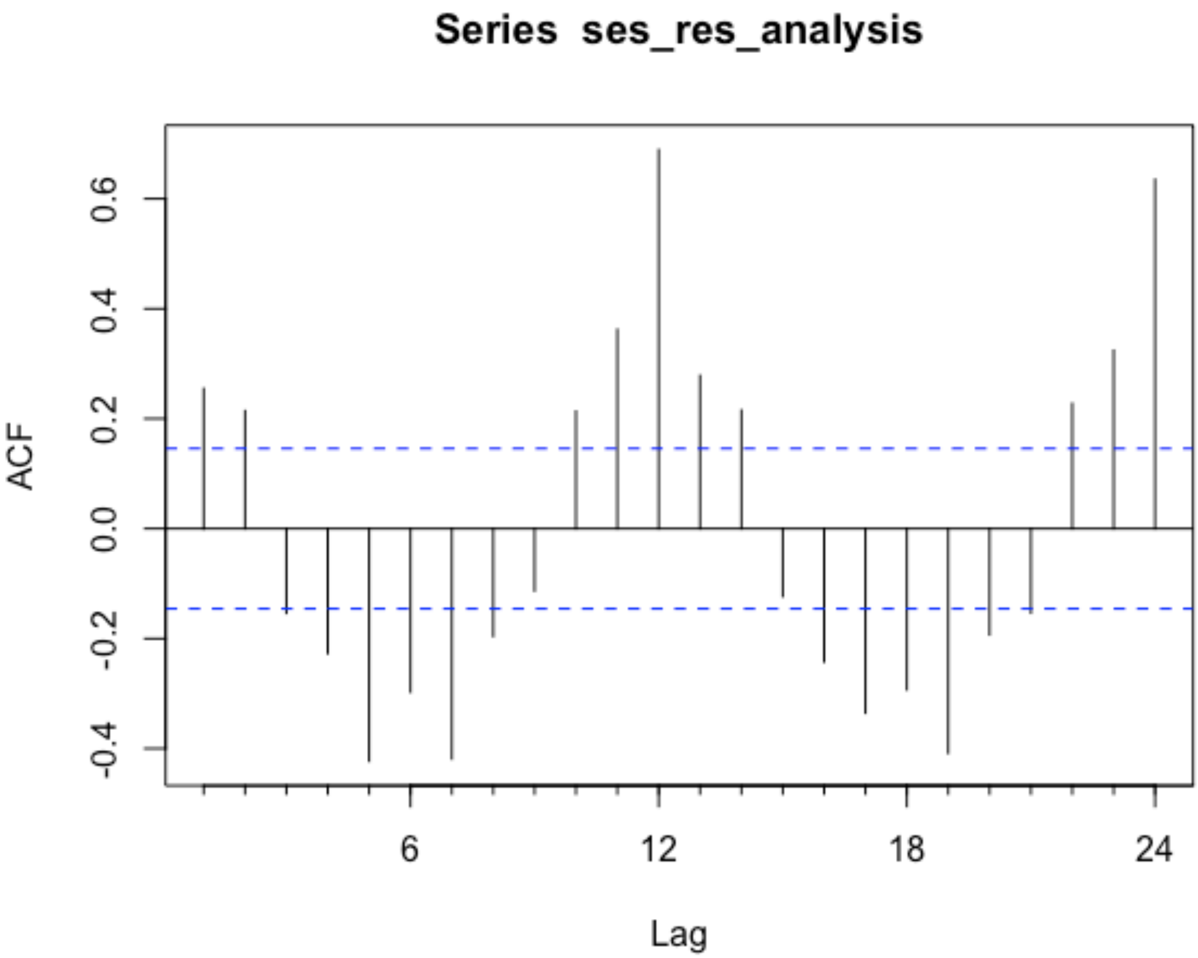
```
> #### Actual Values vs Residuals
> actual_sim_smooth<-sim_smooth$x
> #### Plot Fitted Values vs Residuals
> plot(actual_sim_smooth)
>
```



The above plot says that the actual values and the residuals are randomly distributed.

- An ACF plot of the residuals.

```
Console /Rutgers Doc/Semester 2/BF/Midterm/WorkingDirectory/
> #### ACF of the Residuals
> Acf(ses_res_analysis)
>
```



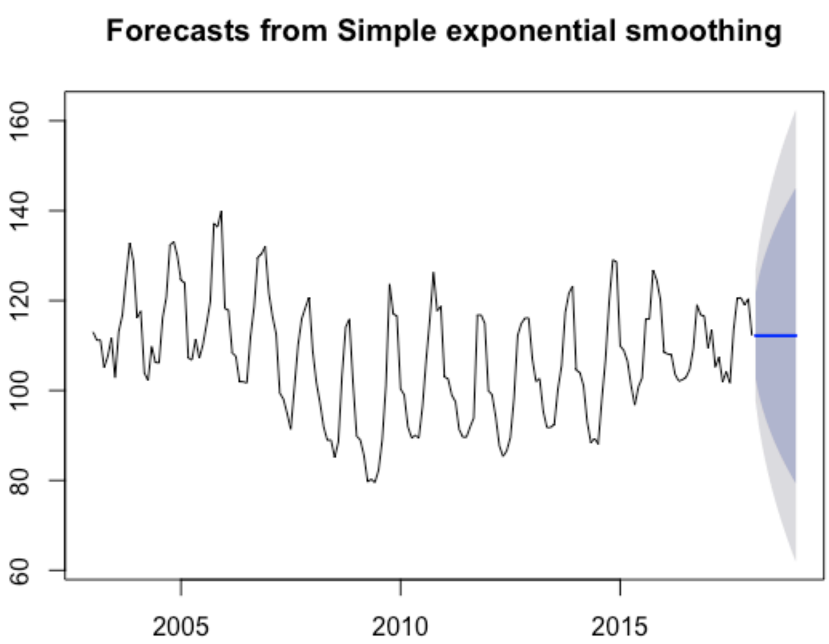
We observe that the values are at the peak in the end of the year and lags in the mid of the year

- Printing the 5 measures of accuracy for this forecasting technique

```
Console /Rutgers Doc/Semester 2/BF/Midterm/WorkingDirectory/
> accuracy(sim_smooth)
                       ME     RMSE      MAE        MPE     MAPE      MASE      ACF1
Training set -0.004510014 7.402115 5.440284 -0.2320932 5.030116 0.8971309 0.2548242
>
```

- Forecast
  - Time series value for next year. Showing table and plot

```
Console /Rutgers Doc/Semester 2/BF/Midterm/WorkingDirectory/
> forecast(sim_smooth,h=12)
         Point Forecast     Lo 80    Hi 80    Lo 95    Hi 95
Feb 2018       112.2125 102.72633 121.6987 97.70464 126.7204
Mar 2018       112.2125  98.79769 125.6274 91.69631 132.7287
Apr 2018       112.2125  95.78305 128.6420 87.08582 137.3392
May 2018       112.2125  93.24156 131.1835 83.19894 141.2261
Jun 2018       112.2125  91.00245 133.4226 79.77452 144.6505
Jul 2018       112.2125  88.97813 135.4469 76.67859 147.7465
Aug 2018       112.2125  87.11657 137.3085 73.83158 150.5935
Sep 2018       112.2125  85.38387 139.0412 71.18164 153.2434
Oct 2018       112.2125  83.75648 140.6686 68.69276 155.7323
Nov 2018       112.2125  82.21725 142.2078 66.33871 158.0863
Dec 2018       112.2125  80.75324 143.6718 64.09971 160.3253
Jan 2019       112.2125  79.35440 145.0706 61.96037 162.4647
>
```



- Summarising this forecasting technique
  - How good is the accuracy?

The MSE for the this model is 0.00002034 which is 0. Thus we say the accuracy is high.

  - What does it predict the value of time series will be in one year?

The value of the time series in the Feb 2019 will be 112.125

  - Other observation

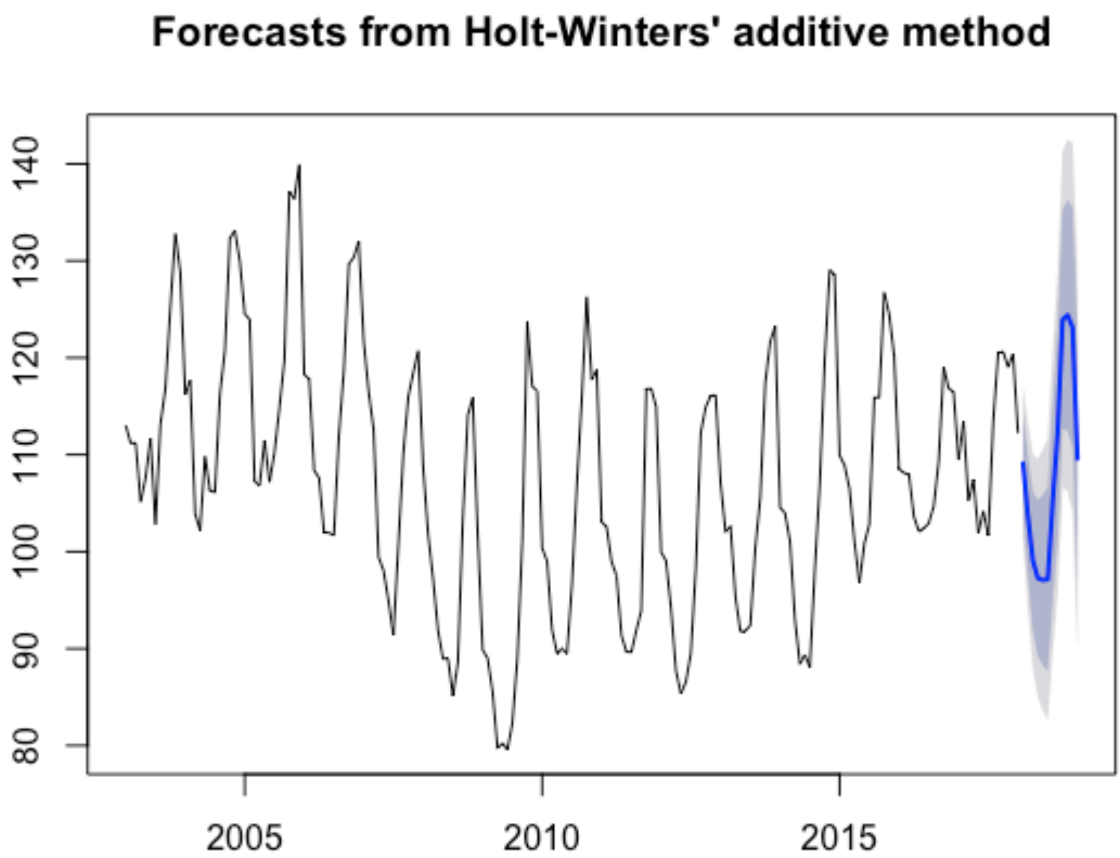
There is minimal seasonality in the plot.

- **Holt Winters**

- Performing Holt-Winters forecast for next 12 months for the time series.

```
Console /Rutgers Doc/Semester 2/BF/Midterm/WorkingDirectory/
> #### Holt Winters
> hw<-hw(candy_ts, h = 12)
> summary(hw)

Forecast method: Holt-Winters' additive method

Model Information:
Holt-Winters' additive method

Call:
 hw(y = candy_ts, h = 12)

  Smoothing parameters:
    alpha = 0.6989
    beta  = 1e-04
    gamma = 1e-04

  Initial states:
    l = 120.1657
    b = -0.078
    s=14.9725 16.1779 15.6251 3.989 -3.2511 -11.3932
           -11.5416 -11.4268 -9.6966 -5.1522 0.1811 1.5157

  sigma:  3.9859

     AIC     AICc      BIC
1475.491 1479.246 1529.866

Error measures:
                     ME     RMSE      MAE         MPE     MAPE     MASE       ACF1
Training set 0.02447927 3.985931 2.990258 -0.04846231 2.764973 0.493109 0.04635702

Forecasts:
         Point Forecast     Lo 80    Hi 80     Lo 95    Hi 95
Feb 2018      109.10428 103.99610 114.2125 101.29199 116.9166
Mar 2018      103.69354  97.46127 109.9258  94.16210 113.2250
Apr 2018       99.07185  91.88904 106.2547  88.08668 110.0570
May 2018       97.26444  89.24272 105.2862  84.99627 109.5326
Jun 2018       97.07188  88.29082 105.8529  83.64241 110.5013
Jul 2018       97.14224  87.66228 106.6222  82.64389 111.6406
Aug 2018      105.20684  95.07590 115.3378  89.71290 120.7008
Sep 2018      112.36941 101.62669 123.1121  95.93984 128.7990
Oct 2018      123.92753 112.60588 135.2492 106.61256 141.2425
Nov 2018      124.40309 112.53055 136.2756 106.24561 142.5606
Dec 2018      123.12053 110.72140 135.5197 104.15770 142.0834
Jan 2019      109.58617  96.68178 122.4906  89.85061 129.3217
>
```

```
Console /Rutgers Doc/Semester 2/BF/Midterm/WorkingDirectory/
> #### Plot HoltWinters
> plot(hw)
>
```



- What is the value of alpha? What does that value signify?

  The value of alpha is 0.6058406. The value signifies that the mean smoothing parameter has to be adjusted(added) to the level of the seasonal component of the time series.

- What is the value of beta? What does that value signify?

  The value of beta is 0. The value signifies that there is no need to smooth the trend parameter of the time series.

- What is the value of gamma? What does that value signify?

  The value of gamma is 0.6033215. The value signifies that there has to be some smoothing with the seasonal component(by adding the value).

- What is the value of initial states for the level, trend and seasonality? What do these values signify?

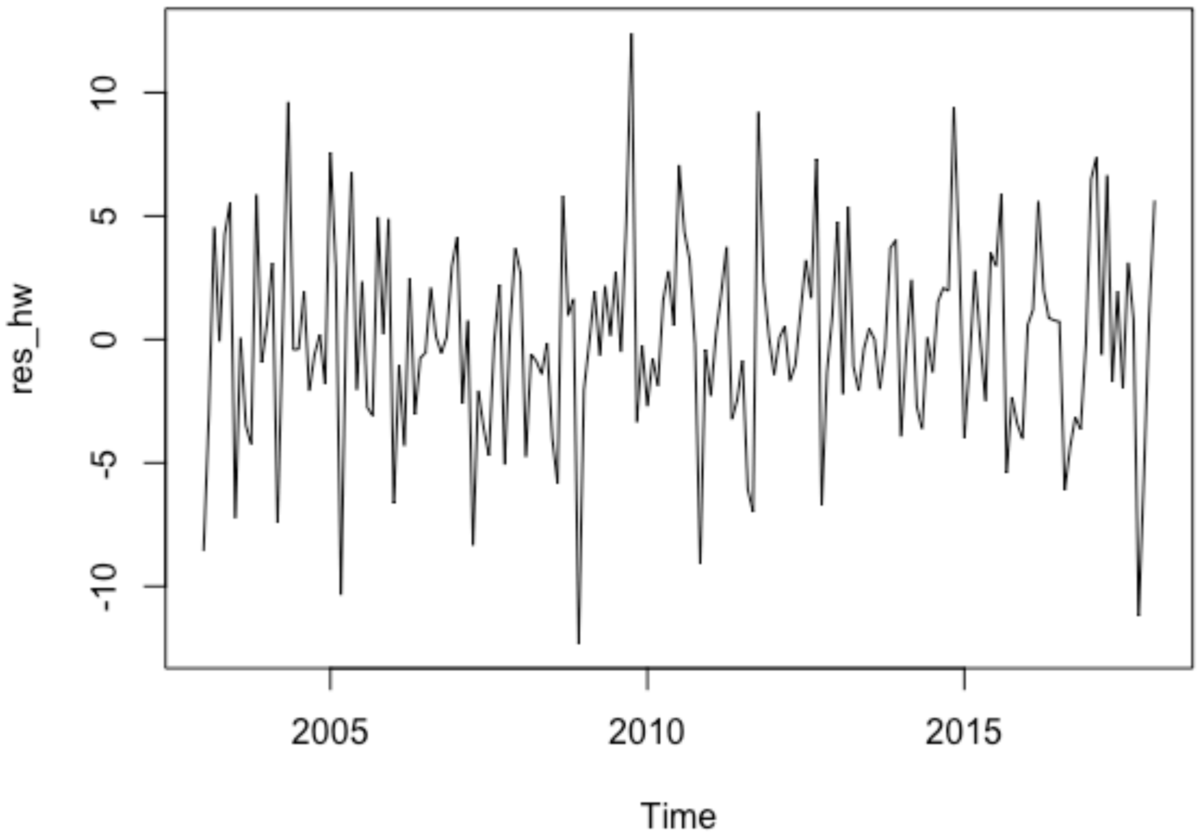
```
Console /Rutgers Doc/Semester 2/BF/Midterm/WorkingDirectory/

 Initial states:
   l = 120.1657
   b = -0.078
   s=14.9725 16.1779 15.6251 3.989 -3.2511 -11.3932
          -11.5416 -11.4268 -9.6966 -5.1522 0.1811 1.5157
```

- What is the value of sigma? What does the sigma signify?

The value of Sigma is 3.9859. The value signifies the standard deviation of the residuals.

- Performing Residual Analysis for this technique.
  - Plot of residuals.

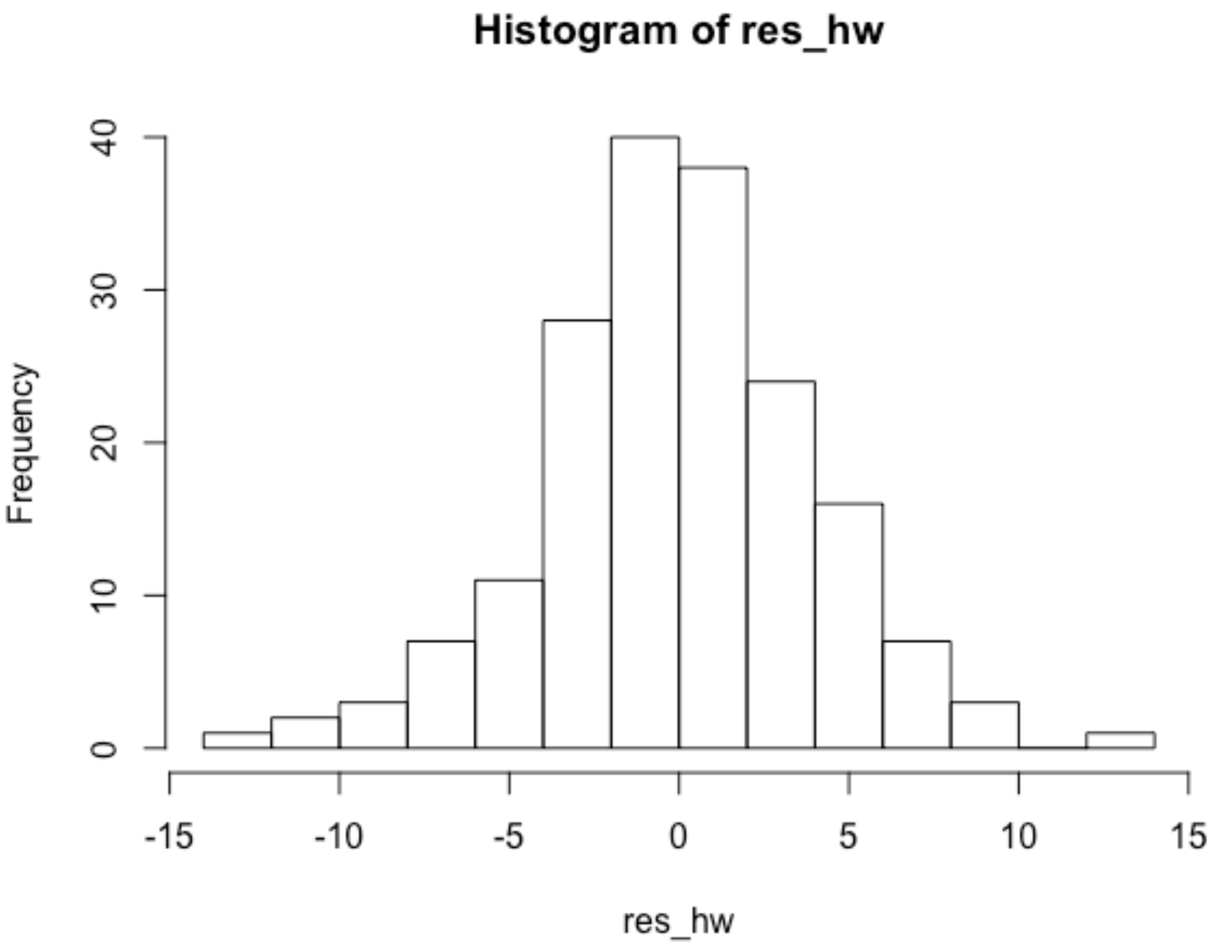
```
Console /Rutgers Doc/Semester 2/BF/Midterm/WorkingDirectory/
> #### Residuals
> res_hw<-residuals(hw)
> #### Plot HoltWinters
> plot(res_hw)
>
```



Residuals have high variance and are randomly distributed.

- Histogram plot of residuals.

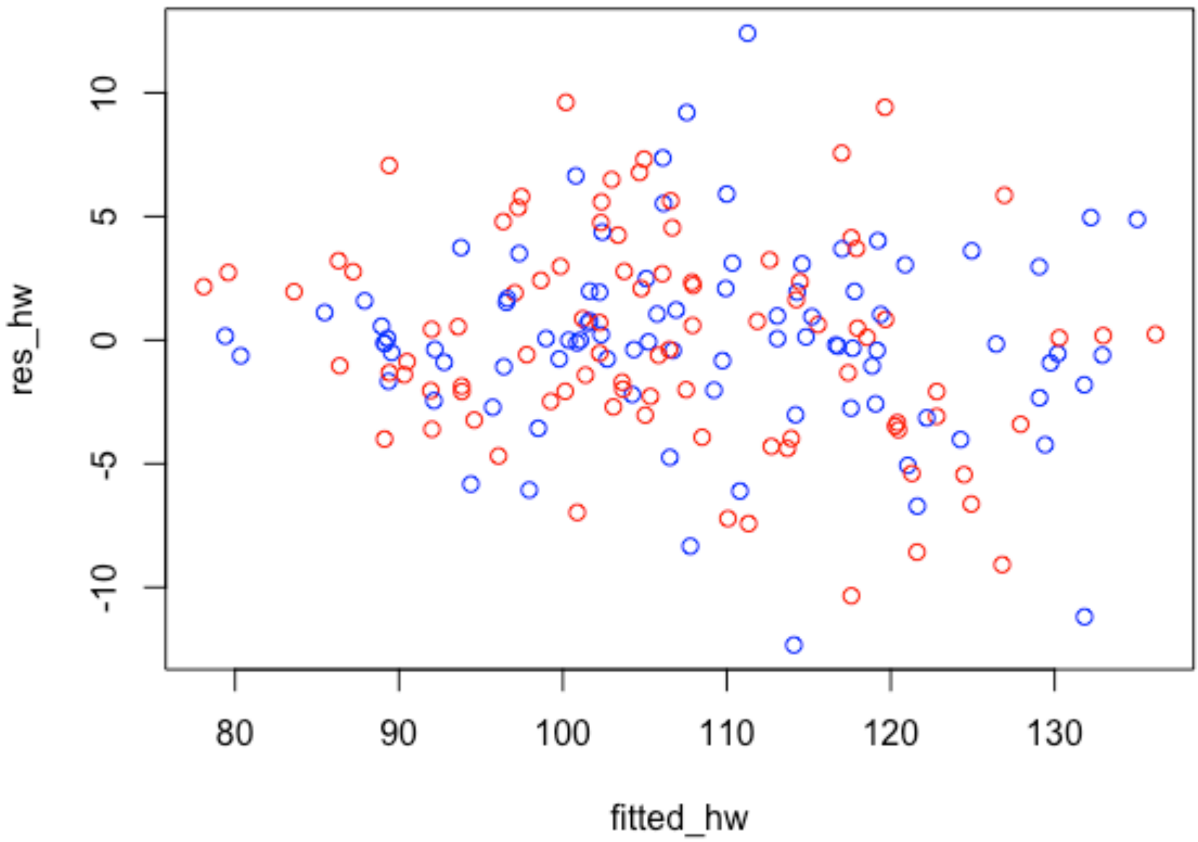
```
Console /Rutgers Doc/Semester 2/BF/Midterm/WorkingDirectory/
> #### Plot Histogram HoltWinters
> hist(res_hw)
>
```



Histogram says that the residuals are symmetrically distributed.

- Plot of fitted values vs. residuals.

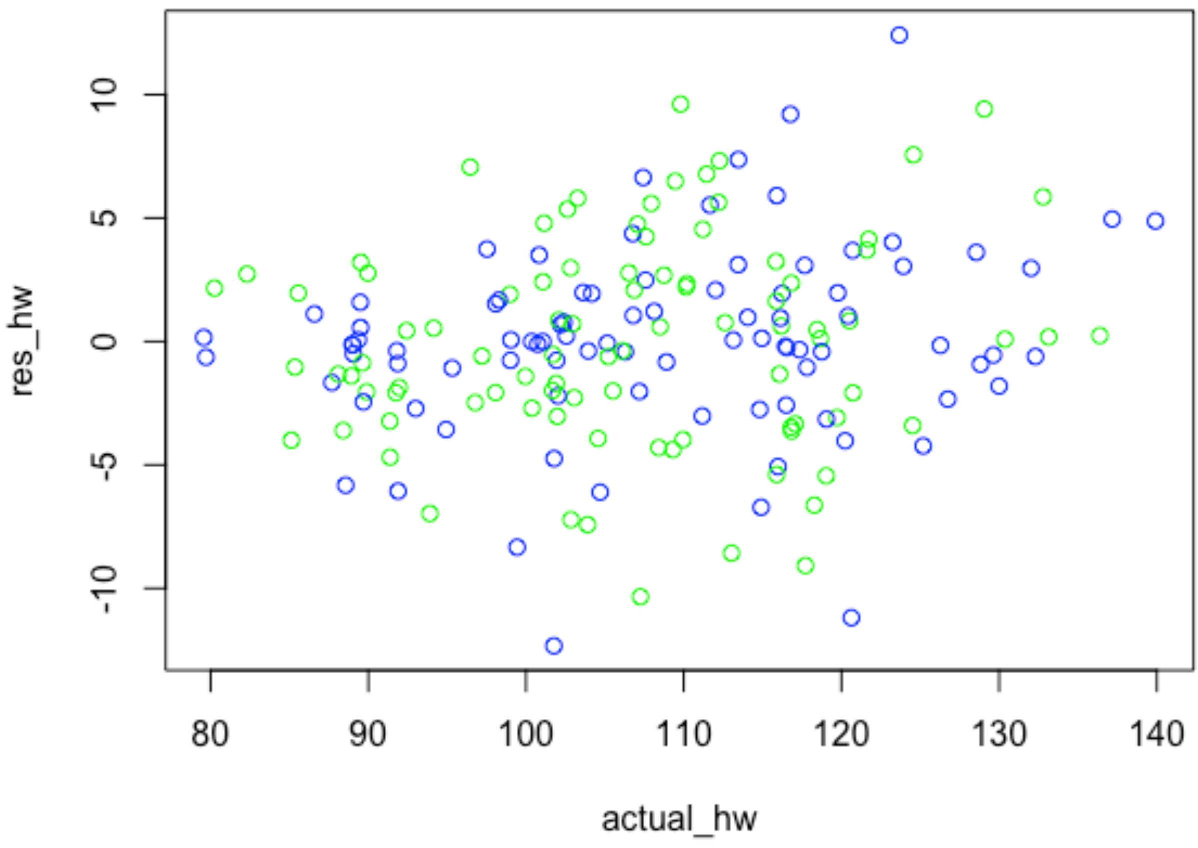
```
Console /Rutgers Doc/Semester 2/BF/Midterm/WorkingDirectory/
> #### Fitted HoltWinters
> fitted_hw<-hw$fitted
> #### Plot Fitted Holt vs Residual
> plot(fitted_hw,res_hw,col=c("red","blue"))
>
```



The plot for residual and the fitted values are randomly distributed.

- Plot of actual values vs. residuals.

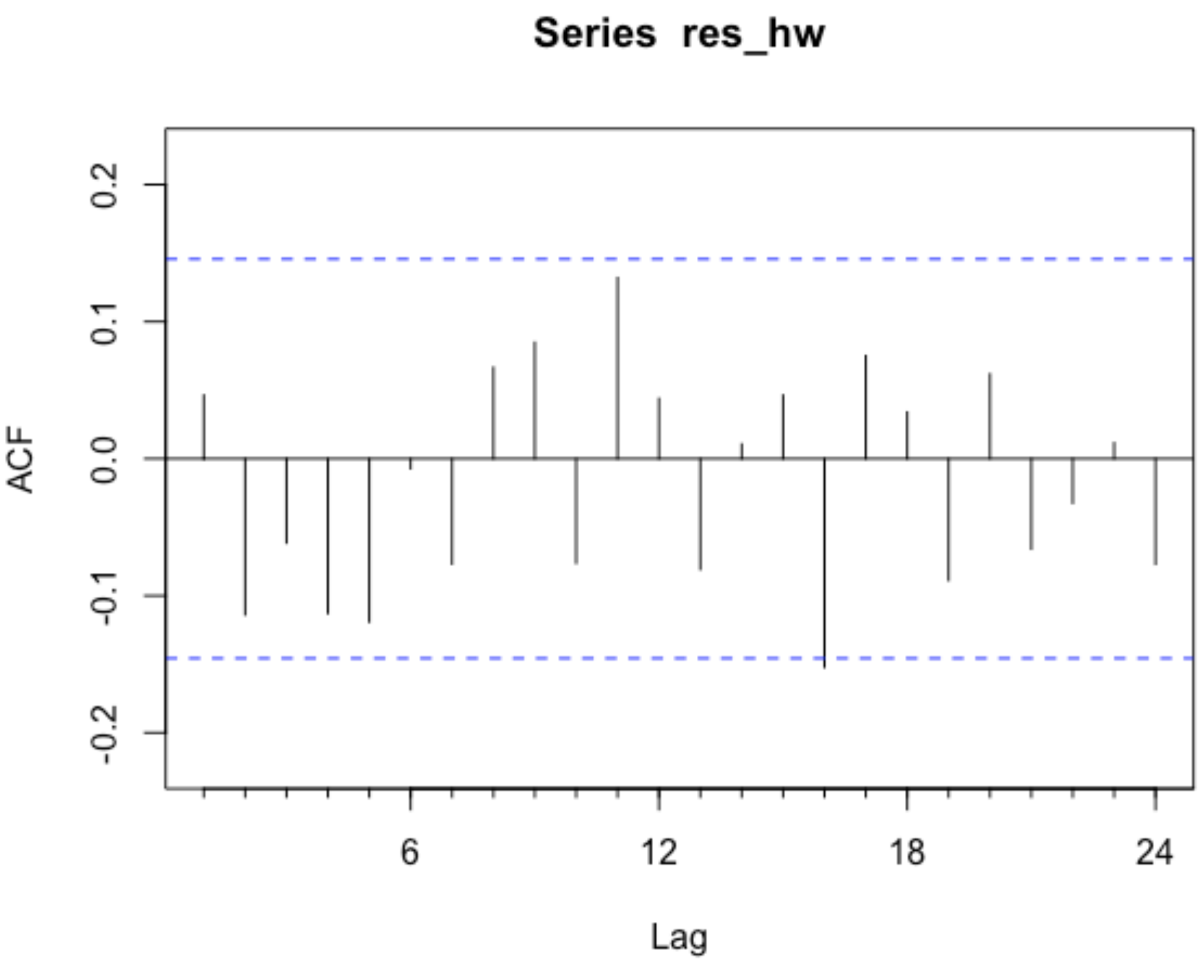
```
Console /Rutgers Doc/Semester 2/BF/Midterm/WorkingDirectory/
> #### Actual HoltWinters
> actual_hw<-hw$x
> #### Plot Fitted Holt vs Residual
> plot(actual_hw,res_hw,col=c("green","blue"))
>
```



The plot for residual and the actually values are randomly distributed.

- An ACF plot of the residuals.

```
Console /Rutgers Doc/Semester 2/BF/Midterm/WorkingDirectory/
> #### Plot ACF of Residuals
> Acf(res_hw)
>
```



The above plot shows that the ACF of residuals is random.

- Printing the 5 measures of accuracy for this forecasting technique

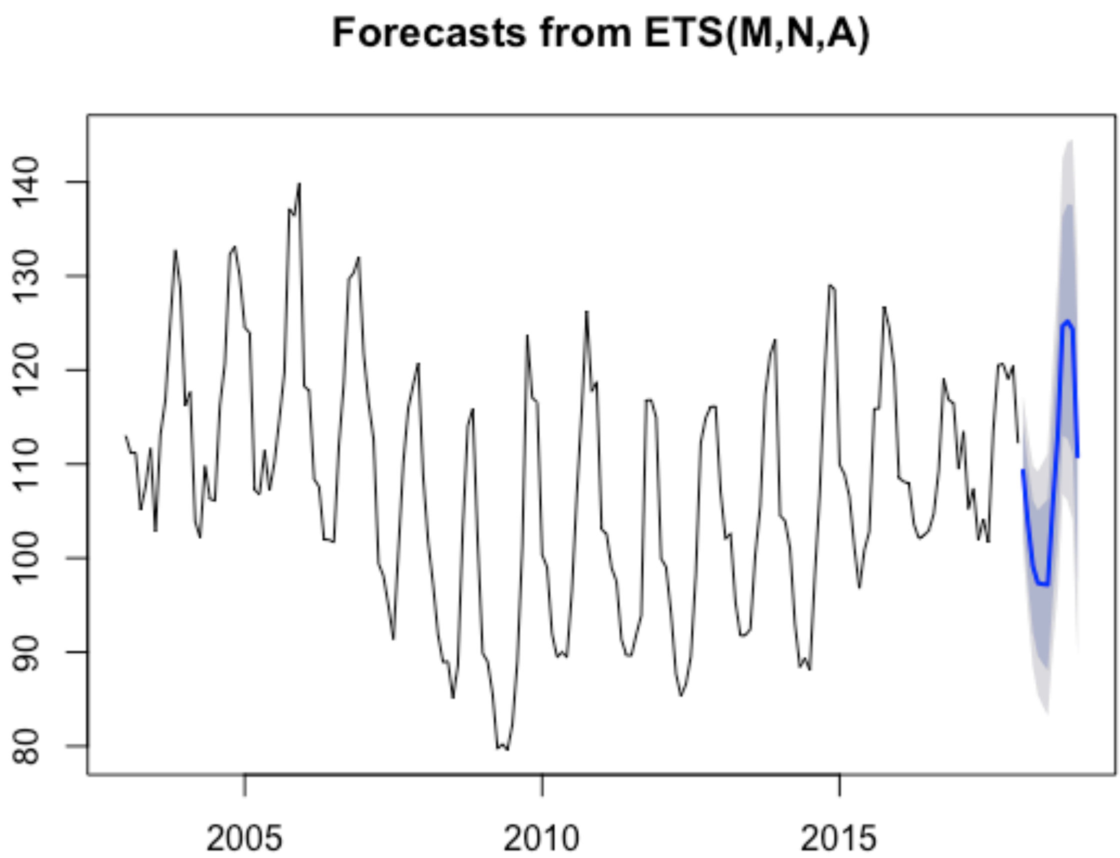
```
Console /Rutgers Doc/Semester 2/BF/Midterm/WorkingDirectory/
> #### Five Measures of Accuracy
> accuracy(hw)
                     ME     RMSE      MAE         MPE     MAPE     MASE       ACF1
Training set 0.02447927 3.985931 2.990258 -0.04846231 2.764973 0.493109 0.04635702
> 
```

- Forecast
  - Time series value for next year. Showing table and plot

```
Console /Rutgers Doc/Semester 2/BF/Midterm/WorkingDirectory/
> #### Forecast Candy Sales
> fore_hw<-forecast(candy_ts,h=12)
> fore_hw
         Point Forecast     Lo 80    Hi 80     Lo 95    Hi 95
Feb 2018      109.28137 104.21953 114.3432 101.53994 117.0228
Mar 2018      103.74316  97.61639 109.8699  94.37307 113.1133
Apr 2018       99.16507  92.19556 106.1346  88.50612 109.8240
May 2018       97.34774  89.61964 105.0758  85.52863 109.1668
Jun 2018       97.22752  88.79154 105.6635  84.32580 110.1292
Jul 2018       97.17709  88.08769 106.2665  83.27605 111.0781
Aug 2018      105.56456  95.67847 115.4507  90.44509 120.6840
Sep 2018      112.91012 102.19973 123.6205  96.52999 129.2903
Oct 2018      124.67594 113.00564 136.3462 106.82776 142.5241
Nov 2018      125.18774 112.72384 137.6516 106.12585 144.2496
Dec 2018      124.34403 111.15523 137.5328 104.17350 144.5146
Jan 2019      110.82761  97.19179 124.4634  89.97343 131.6818
> 
```

Console /Rutgers Doc/Semester 2/BF/Midterm/WorkingDirectory/

```
> #### Plot Forecast
> plot(fore_hw)
> 
```



- Summarising this forecasting technique
  - How good is the accuracy?

    Since the MSE for Holt Winters is 0.03511 is low, the Model is good for forecasting.

  - What does it predict the value of time series will be in one year?

    The value of time series in Feb 2019 will be 110.26

  - Other observation

    From the Holt Winters Forecasting technique, we observe that the model has low error and is also smooth for forecasting.

- **Accuracy Summary**

- Showing a table of all the forecast method above with their accuracy measures.

	ME	MSE	RMSE	MAE	MPE	MAPE
Naive	-0.0045477	0.00002068	7.422458	5.4702	-0.2333	5.05781
Simple Smoothing	-0.0045100	0.00002034	7.402115	5.4402	-0.2320	5.030116
Holt-Winters	-0.1873801	0.03511130	4.410365	3.3496	-0.2713	3.124352

- Separately defining each forecast method and why it is useful. Showing the best and worst forecast method for each of the accuracy measures.

Naive Forecasting : Estimating technique in which the last period's actuals are used as this period's forecast, without adjusting them or attempting to establish casual factors.

Simple Smoothing Forecasting : The model gives weight to more recent data and smooths the forecast.

Holt Winters Forecasting : This method is used for exponential smoothing to make short term forecast by using 'additive' or 'multiplicative' model with secular increasing or decreasing trend or seasonality

Model	ME	MSE	RMSE	MAE	MPE	MAPE
Best	SES	SES	HW	HIW	SES	SES
Worst	HW	HW	Naive	Naïve	HW	Naive

- ## Conclusion

    - Summarising my analysis of time series value over the time-period.

From the time series of two years data we observe that the sales values are peak in the month of November and decreases over the year and follows a pattern of seasonality. on the analysis and forecast above, do you think the value of the time series will increase, decrease or stay flat over the next year? How about next 2 years?

The pattern for the time series data is seasonal and the forecast for the next two years shows that the the production is high at the end of the year and decreases along the year.

    - Rank forecasting methods that best forecast for this time series based on historical values.

After considering the various factors of accuracy, we conclude that the best models for forecasting are as follows:-

1. Holt Winters
2. Simple Exponential Smoothing
3. Naive Forecasting