Business Forecasting

Midterm Exam: Forecasting Candy Production in the USA

Sakshi Kalani
181007162

Business Forecasting Mid-Term Exam

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

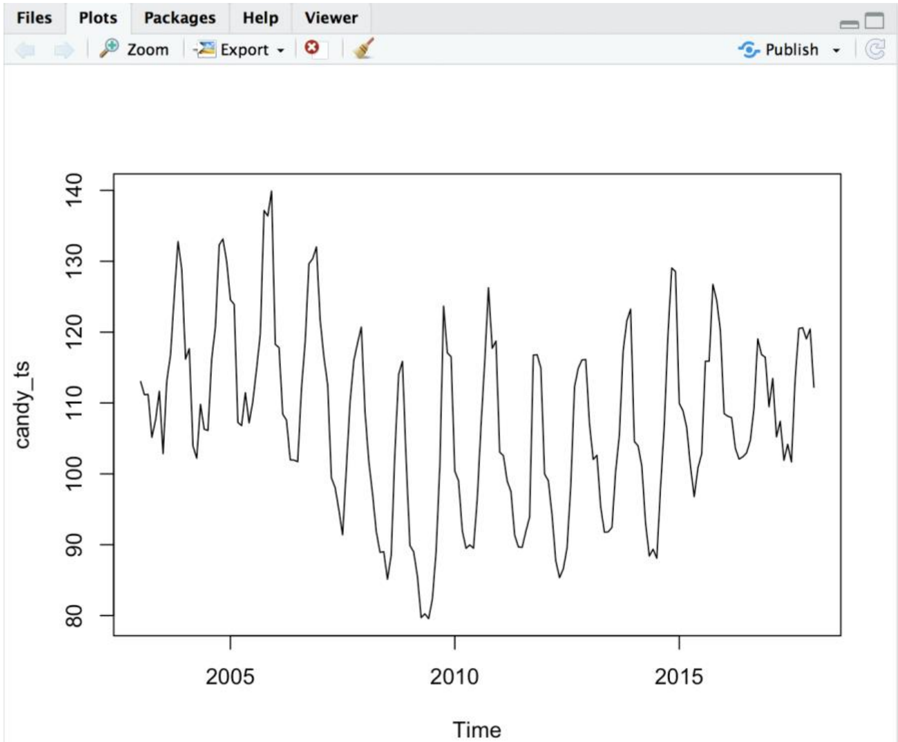
Please do the following steps once the csv file is on your desktop.

1. `library(readr)`
2. `IPG3113N_Spring18 <- read_csv("C:/Users/rrparikh/Desktop/RU/Business Forecasting/Mid-Term/IPG3113N_Spring18.csv")`
3. `candy_ts <- ts(IPG3113N_Spring18$IPG3113N,frequency = 12,start=c(2003,1))`
4. `plot(candy_ts)`

```
> install.packages("readr")
also installing the dependencies ‘pkgconfig’, ‘hms’, ‘BH’

trying URL 'https://cran.rstudio.com/bin/macosx/el-capitan/contrib/3.4/pkgconfig_2.0.1.tgz'
Content type 'application/x-gzip' length 15810 bytes (15 KB)
==================================================
downloaded 15 KB

trying URL 'https://cran.rstudio.com/bin/macosx/el-capitan/contrib/3.4/hms_0.4.2.tgz'
Content type 'application/x-gzip' length 32328 bytes (31 KB)
==================================================
downloaded 31 KB

trying URL 'https://cran.rstudio.com/bin/macosx/el-capitan/contrib/3.4/BH_1.66.0-1.tgz'
Content type 'application/x-gzip' length 10959784 bytes (10.5 MB)
==================================================
downloaded 10.5 MB

trying URL 'https://cran.rstudio.com/bin/macosx/el-capitan/contrib/3.4/readr_1.1.1.tgz'
Content type 'application/x-gzip' length 1967561 bytes (1.9 MB)
==================================================
downloaded 1.9 MB


The downloaded binary packages are in
	/var/folders/6f/cbmlt0qs4f7_q6v4lwk8yv400000gn/T//RtmpVloKAz/downloaded_packages
> library(readr)
> candy_data<-read_csv("/Users/sakshikalani/Desktop/Business Forecasting/Midterm/IPG3113N_Spring18.csv")
Parsed with column specification:
cols(
  DATE = col_character(),
  IPG3113N = col_double()
)
> candy_data<-read.csv("/Users/sakshikalani/Desktop/Business Forecasting/Midterm/IPG3113N_Spring18.csv")
> candy_ts <- ts(candy_data$IPG3113N, frequency= 12, start=c(2003,1))
> plot(candy_ts)
>
```

Plot and Inference

- Show a time series plot.
- Please summaries your observations of the times series plot



The time series plot does not show any significant trend however we can see seasonality as there is a repetitive with regularity making it a significant pattern.

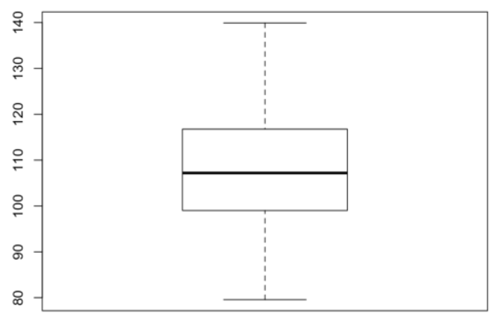
Central Tendency

- What are the min, max, mean, median, 1st and 3rd Quartile values of the times series?

```
> summary(candy_ts)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  79.57   99.02  107.19  107.45  116.76  139.92
>
```

- Show the box plot.

```
> boxplot(candy_ts)
>
```

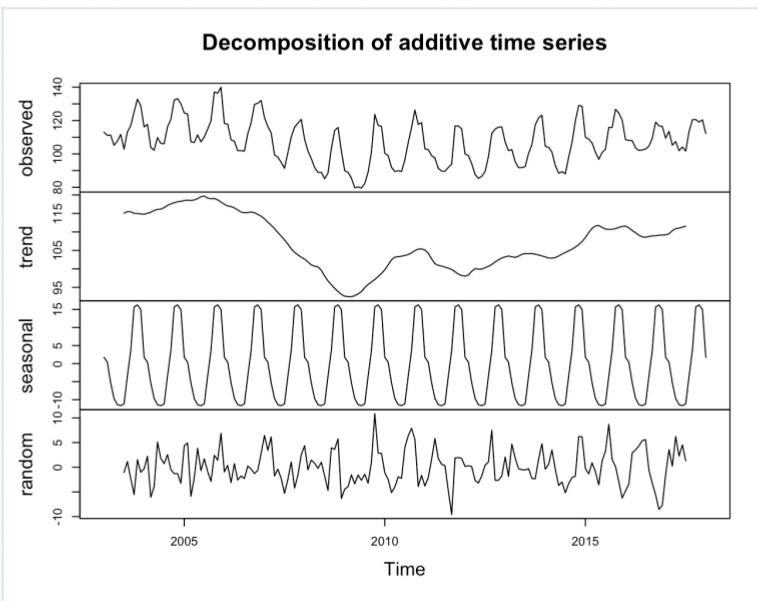


- Can you summarize your observation about the time series from the summary stats and box plot?
  The median of the time series is 107.19. Range of first quartile is from 80 to 99.02. Second ranges from 99.03 to 108, third ranges from 110 to 116.76 while fourth ranges from 116.77 to 140. The data range is approximately 60. Also, there are no outliers in the box plot.

Decomposition

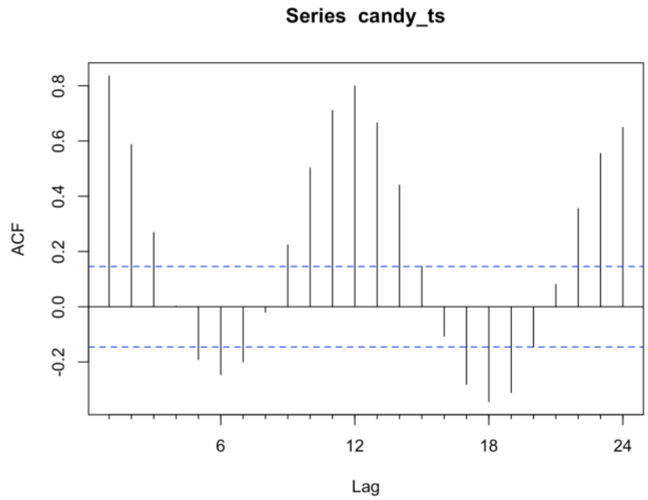
- Plot the decomposition of the time series.

```
> plot(decomp)
>
```



- Is the times series seasonal?

```
> Acf(candy_ts)
```



The Acf plot shows an oscillation indicative of seasonality. The peaks occur at lags of 12 months. Yes, the time series is seasonal.

- Is the decomposition additive or multiplicative?

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

- If seasonal, what are the values of the seasonal monthly indices?

```
> decomp
```

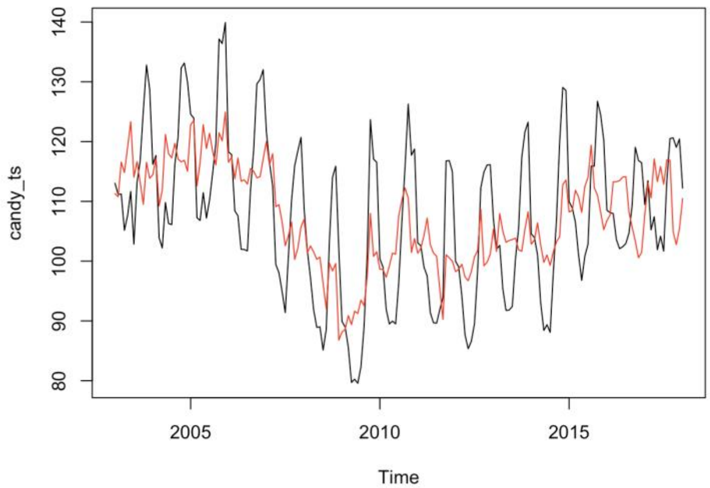
```
$seasonal
            Jan        Feb        Mar        Apr        May        Jun        Jul
2003  1.7367141  0.4089563 -5.3388684 -9.6736722 -11.3775282 -11.6560576 -11.1830346
2004  1.7367141  0.4089563 -5.3388684 -9.6736722 -11.3775282 -11.6560576 -11.1830346
2005  1.7367141  0.4089563 -5.3388684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

- For which month is the value of time series high and for which month is it low?
  Value of time series is low for the month of June and high for the month of November.

- Can you think of the reason behind the value being high in those months and low in those months?
  The possibility of values being high in November could be because of Thanksgiving and Halloween falling in and around the same month. The possibility of values being low in June could be possibly because of there are no major holidays/events.

- Show the plot for time series adjusted for seasonality. Overlay this with the line for actual time series? Does seasonality have big fluctuations to the value of time series?

```
> temp_sesAdjust<- seasadj(decomp)
>
>
> plot(candy_ts)
>
> lines(temp_sesAdjust,col="red")
>
```

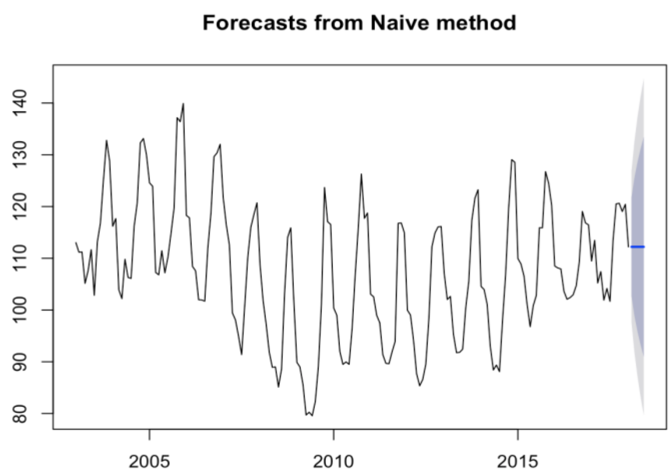


No, seasonality adjustment does not have big fluctuations to value of original time series.

Naïve Method

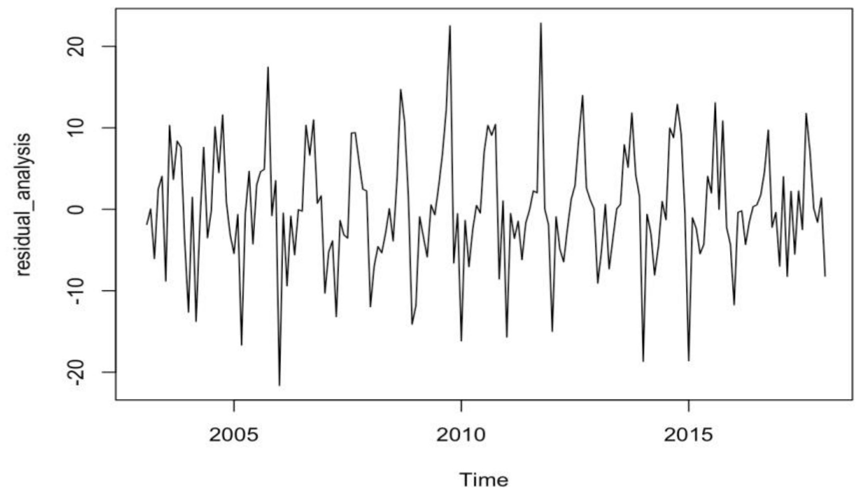
- Output

```
> naive_forecast <- naive(candy_ts,5)
> plot(naive_forecast)
>
```



- Perform Residual Analysis for this technique.
  - Do a plot of residuals. What does the plot indicate?

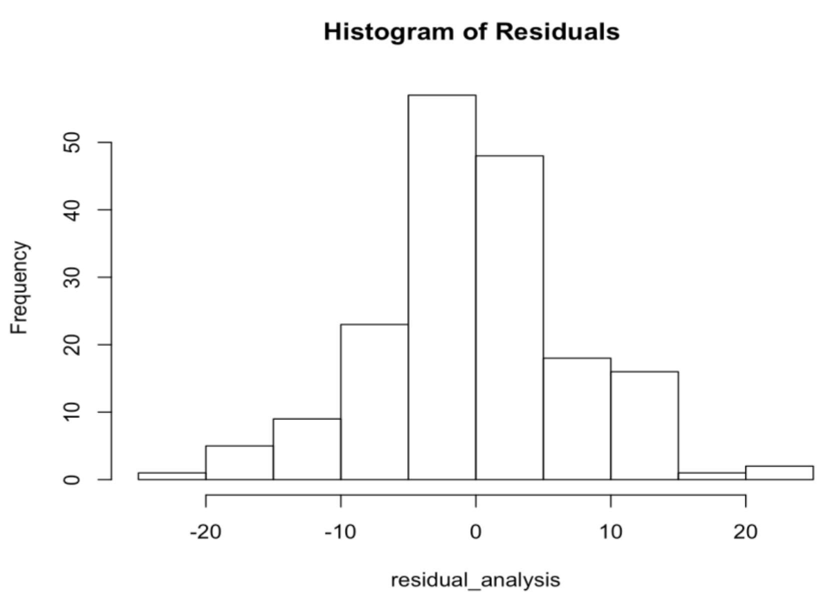
```
> residual_analysis <- residuals(naive_forecast)
> plot(residual_analysis)
```



The time plot of the residuals shows that the variation of the residuals stays much the same across the historical data and therefore the residual variance can be treated as constant.

- Do a Histogram plot of residuals. What does the plot indicate?

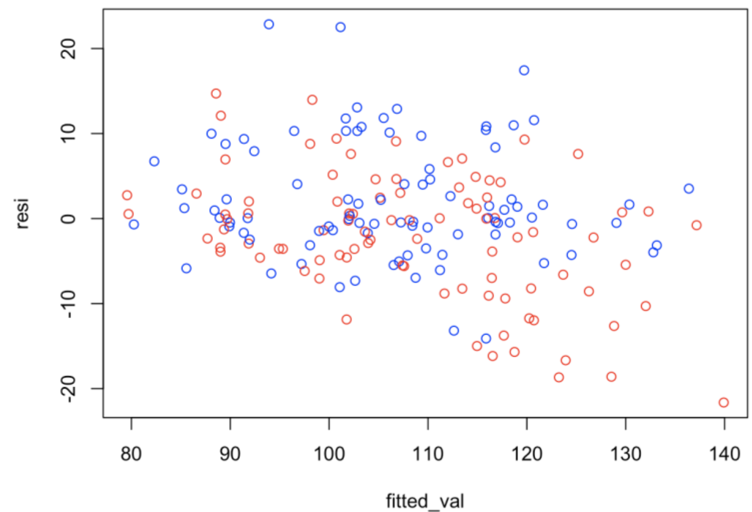
```
h<-hist(residual_analysis, breaks=10,main="Histogram of Residuals ")
```



The plot of histogram shows that it is skewed to the right. The fit of the distribution seems to be normal. The peak is around the mean which is zero.

- Do a plot of fitted values vs. residuals. What does the plot indicate?

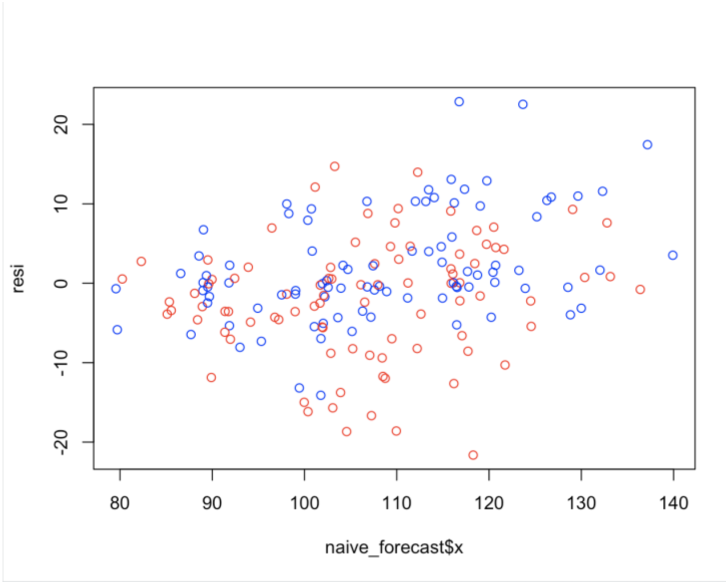
```
> fitted_val<- naive_forecast$fitted
> resi<-naive_forecast$residuals
> plot(resi ~ fitted_val)
> plot(resi ~ fitted_val, col=c("red","blue"))
>
```



The scatter plot of residuals and fitted values shows that there is negative correlation between the variables.

- Do a plot of actual values vs. residuals. What does the plot indicate?

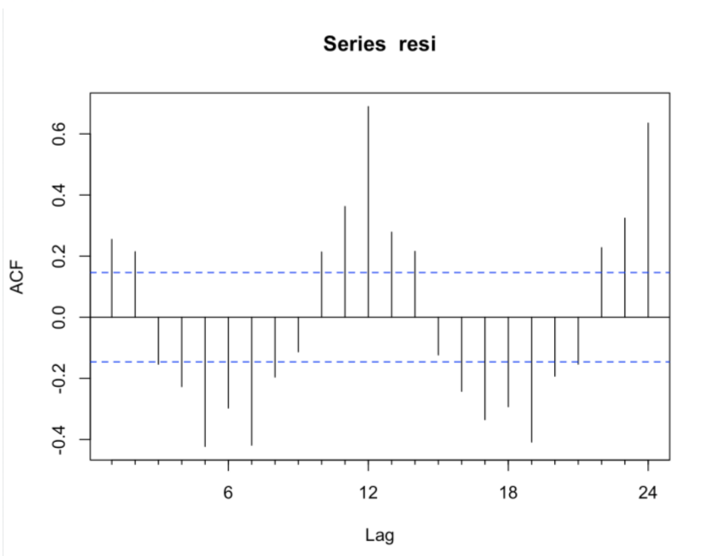
```
> plot(naive_forecast$x,resi, col=c("red","blue"))
>
```



The scatter plot of residuals and actual values shows that there is positive correlation between the variables.

- ○ Do an ACF plot of the residuals? What does this plot indicate?

```
> Acf(resi)
>
```



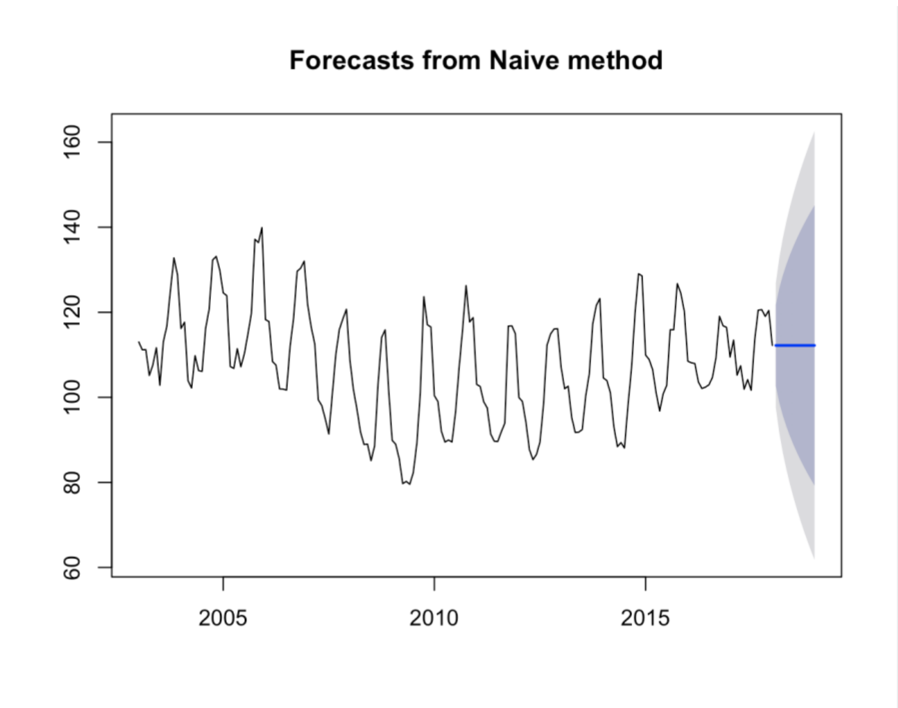
The Acf plot shows a high degree of correlation between residuals which is not suggested.

- Print the 5 measures of accuracy for this forecasting technique

```
> accuracy(naive_forecast)
                       ME     RMSE      MAE        MPE     MAPE      MASE      ACF1
Training set -0.004547778 7.422458 5.470242 -0.2333585 5.057813 0.9020712 0.2547176
>
```

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

```
> forecast(naive_forecast, h=12)
         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
> plot(forecast(naive_forecast, h=12))
```



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

RMSE is 7.4224. The RMSE being good, we would expect the accuracy of the model to be good too. We can see that the model accuracy isn't very good since there is a difference between the actual values and predicted values.

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

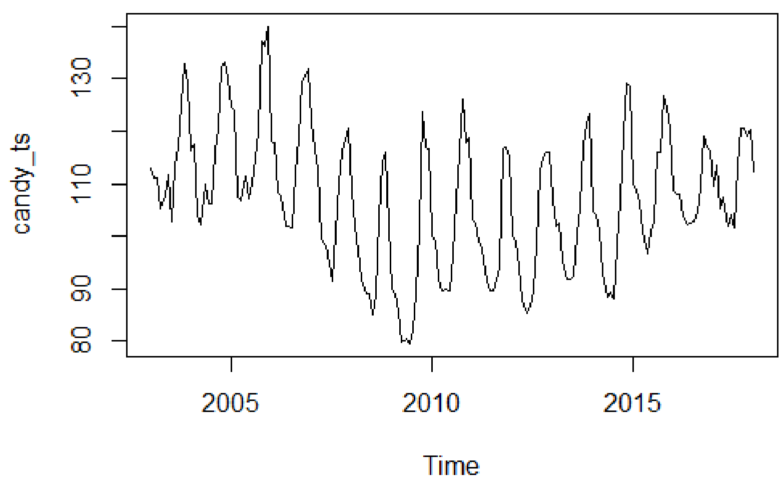
  The value will be the same after one year. If the value for June is 112.2117 for this year, it'll be the same for the next year too.

- Other observation

  The model performs average for the given dataset in comparison with other models.

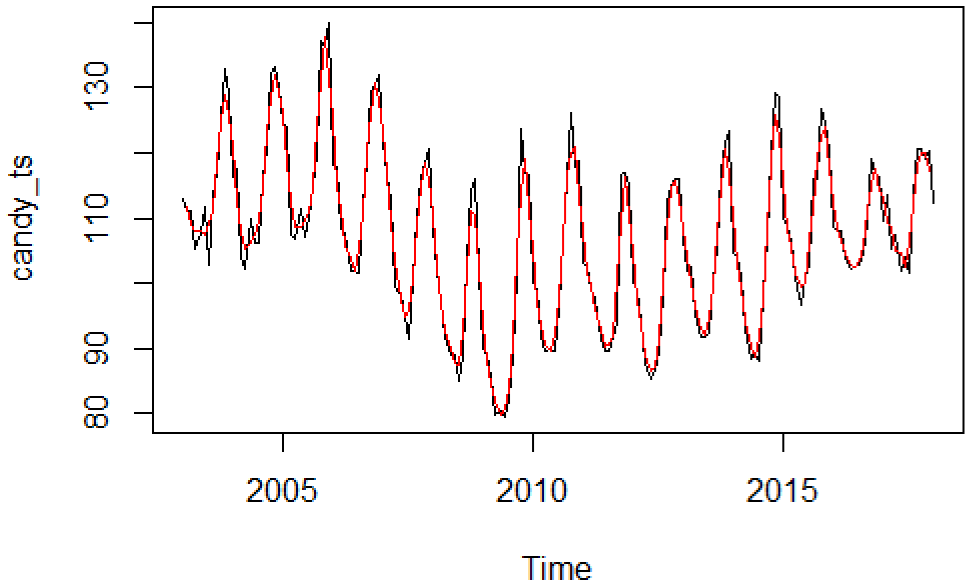
Simple Moving Averages

- Plot the graph for time series.



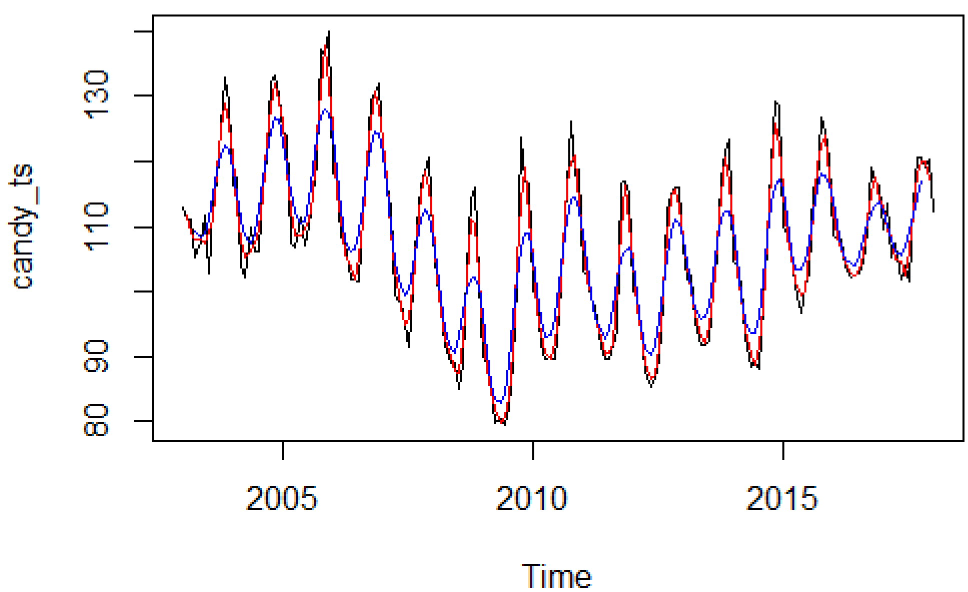
- Show the Simple Moving average of order 3 on the plot above in Red

```
> MA3_forecast <- ma(candy_ts,order=3)
> lines(MA3_forecast,col="Red",lwd=3)
> 
```



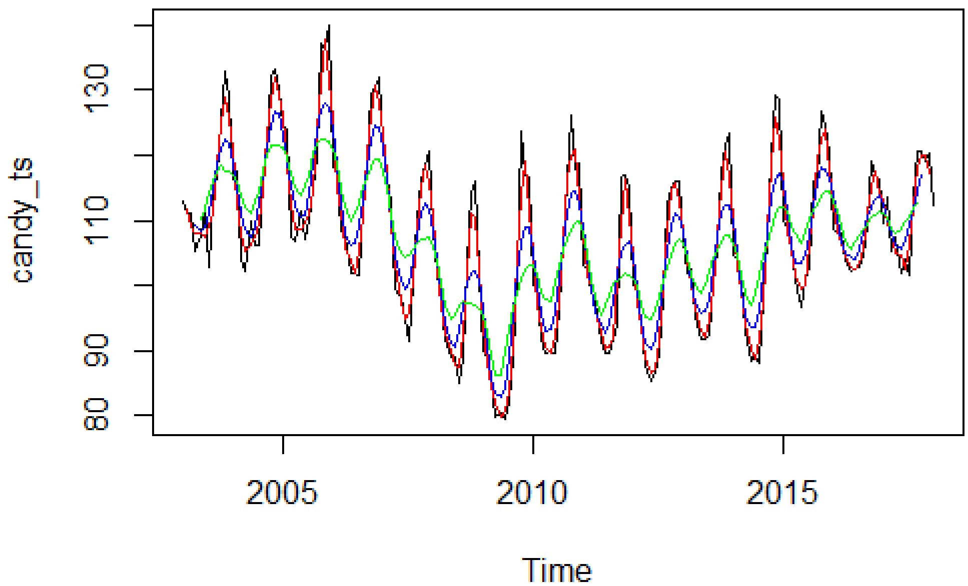
- Show the Simple Moving average of order 6 on the plot above in Blue

```
> MA6_forecast <- ma(candy_ts,order=6)
> lines(MA6_forecast,col="Blue",lwd=3)
>
```



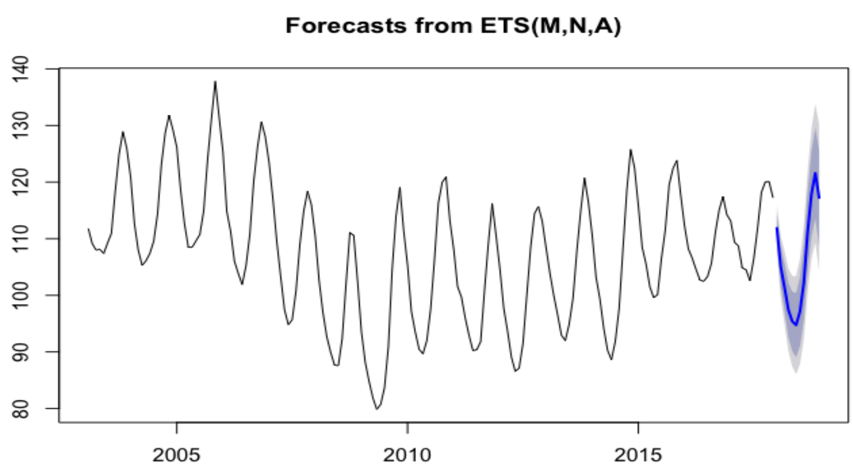
- Show the Simple Moving average of order 9 on the plot above in Green

```
> MA9_forecast <- ma(candy_ts,order=9)
> lines(MA9_forecast,col="Green",lwd=3)
> 
```



- (Bonus) show the forecast of next 12 months using one of the simple average order that you feel works best for time series

  Using 3 point moving average, we get -



- What are your observations of the plot as the moving average order goes up?

  The order of the moving average determines the smoothness of the trend-cycle estimate. In general, a larger order means a smoother curve. The trend is smoother than the original data and captures the main movement of the time series without all the minor fluctuations. The error increases and the mean narrows as we increase the order.

Simple Smoothing

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

```
> ses(candy_ts, 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
```

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

    The value of alpha is 0.99. Alpha signifies estimate of the level at the current time point. Values of alpha that are close to 1 mean that more weight is placed on the most recent observations when making forecasts of future values.

  - What is the value of initial state?

    The value of initial states (l) is 113.028.

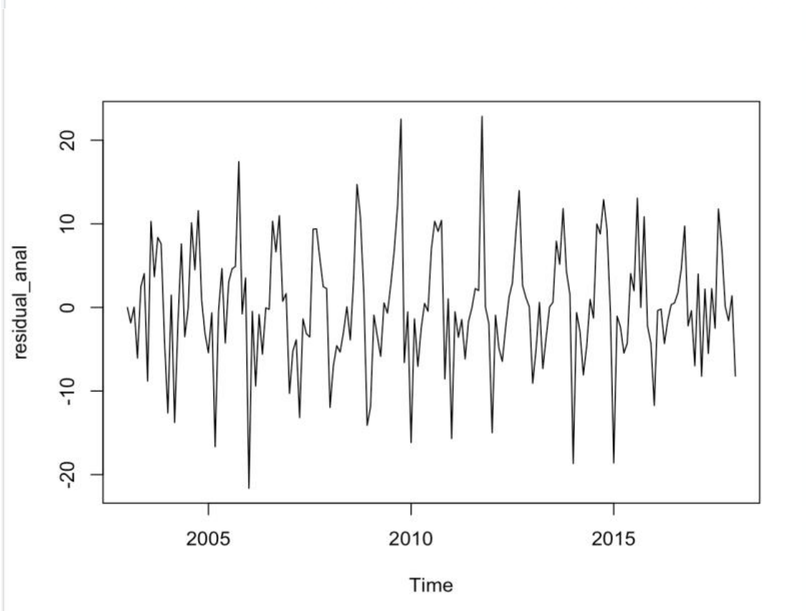
  - What is the value of sigma? What does the sigma signify?

    The value of sigma is 7.4. The Sigma value should approach zero for a model to be a good fit.

```
> summary(ses(candy_ts, h=12))

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
```

- Perform Residual Analysis for this technique.
  - Do a plot of residuals. What does the plot indicate?

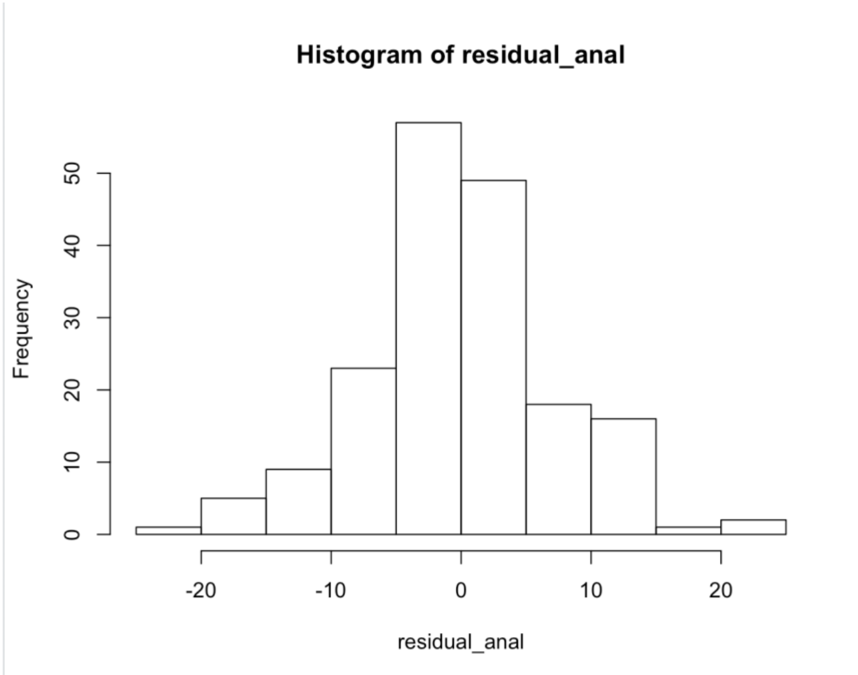
```
> ses_forecast<-ses(candy_ts, h=12)
> residual_anal<-residuals(ses_forecast)
> plot(residual_anal)
>
```



The time plot of the residuals shows that the variation of the residuals stays much the same across the historical data and therefore the residual variance can be treated as constant.

  - Do a Histogram plot of residuals. What does the plot indicate?

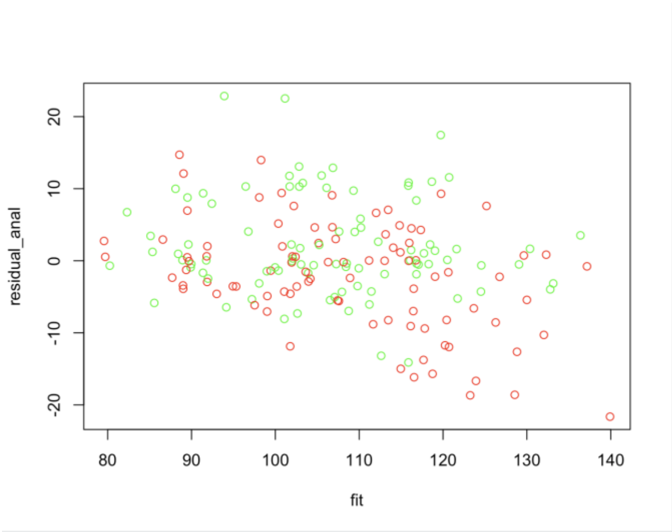
```
> h<-hist(residual_anal)
```



The plot of histogram shows that it is skewed to the right. The fit of the distribution seems to be normal. The peak is around the mean which is zero.

- Do a plot of fitted values vs. residuals. What does the plot indicate?

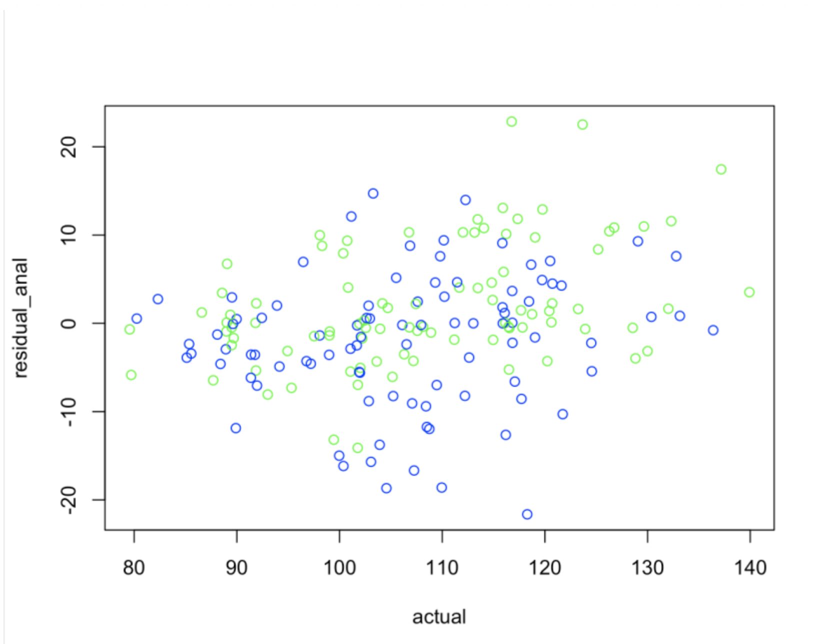
```
> attributes(ses_forecast)
$names
 [1] "model"     "mean"      "level"     "x"         "upper"     "lower"     "fitted"    "method"
 [9] "series"    "residuals"

$class
[1] "forecast"

> fit<- ses_forecast$fitted
> plot(fit,residual_anal,col=c("Red","Green"))
>
```



The scatter plot of residuals and fitted values shows that there is negative correlation between the variables.

- Do a plot of actual values vs. residuals. What does the plot indicate?

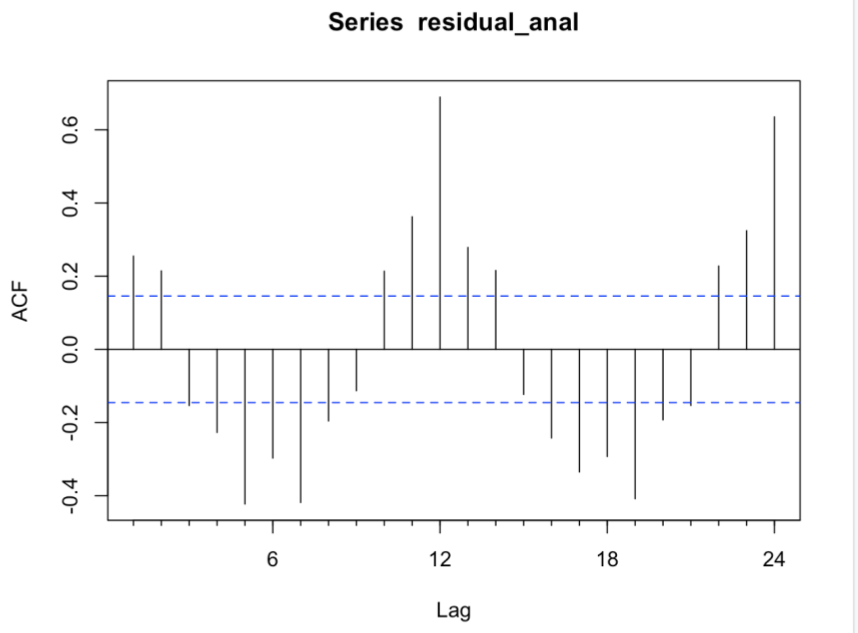
```
> actual<-ses_forecast$x
> plot(actual,residual_anal,col=c("Blue","Green"))
>
```



The scatter plot of residuals and actual values shows that there is positive correlation between the variables.

- Do an ACF plot of the residuals? What does this plot indicate?

```
> Acf(residual_anal)
>
```



The Acf plot shows a high degree of correlation between residuals which is not suggested.

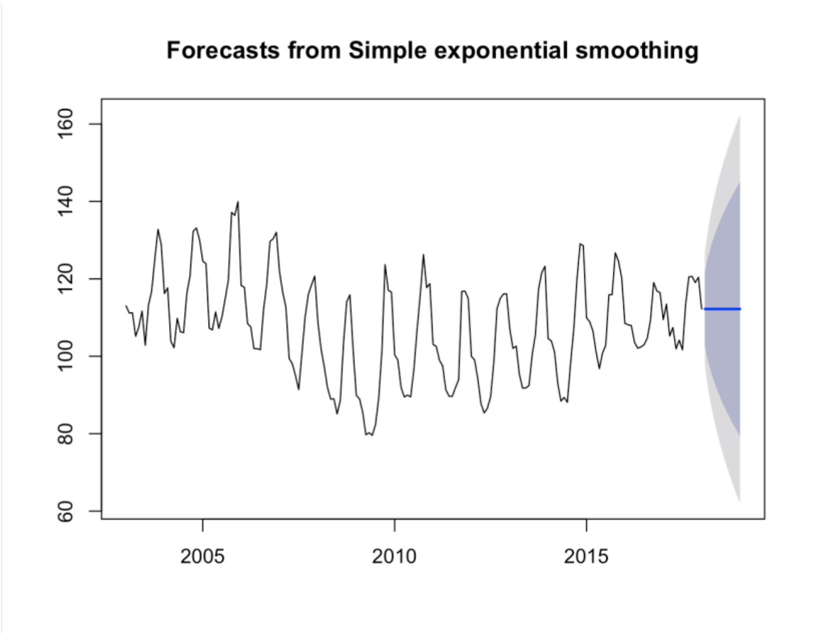
- Print the 5 measures of accuracy for this forecasting technique

```
> accuracy(ses_forecast)
                       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. Show table and plot

```
> forecast(ses_forecast,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
>
```

```
> plot(forecast(ses_forecast,h=12))
>
```

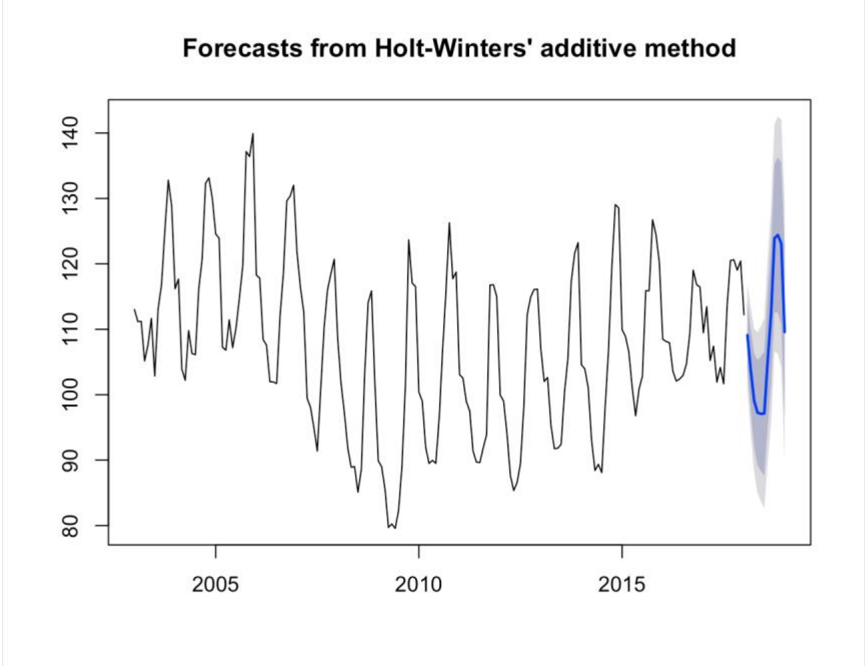


- Summarize this forecasting technique
    - How good is the accuracy?
      The MSE for Simple Smoothing is close to 0 which implies the accuracy is very good
    - What does it predict the value of time series will be in one year?
      Looking at the point forecast value in the 'summary', we can see that every month has the same value, i.e. 112.2125
    - Other observation
      It performs better than Naïve Method.

Holt-Winters

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

```
> holt<-hw(candy_ts,12)
> plot(holt)
> summary(holt)

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
>
```



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

  The value of alpha is 0.69. This is high, telling us that both the estimate of the current value of the level, and of the slope b of the trend component, are based mostly upon very recent observations in the time series

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

  The value of beta is 1e-04. This is high, telling us that both the estimate of the current value of the level, and of the slope b of the trend component, are based mostly upon very recent observations in the time series

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

  The value of gamma is 1e-04. The value is high indicating that the estimate of the seasonal component at the current time point is just based upon very recent observations.

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

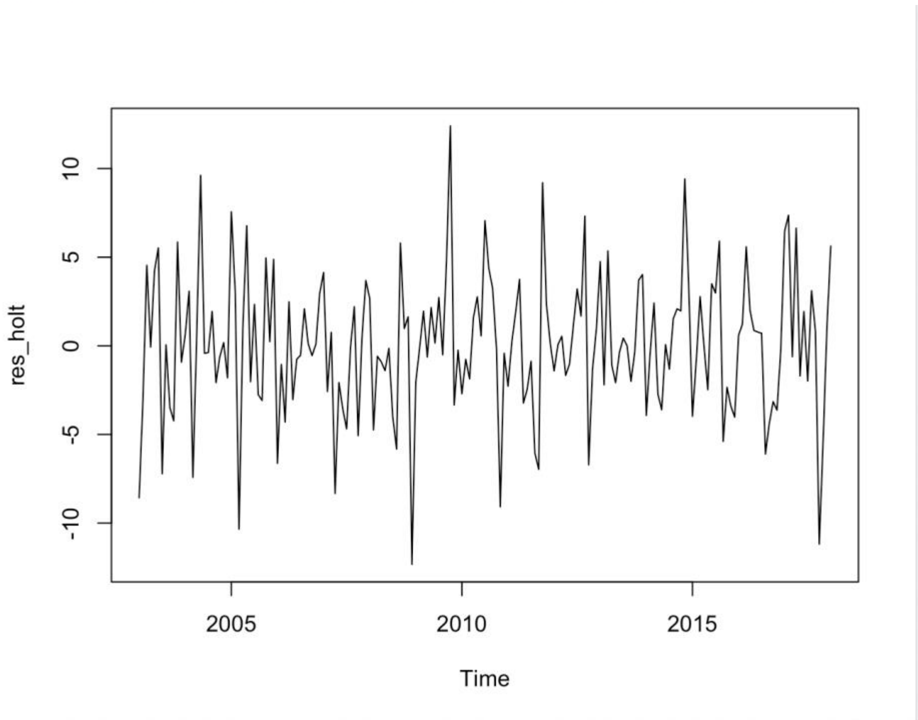
  The value of level (l) is 120.16, trend (b) is -0.078, seasonality (s) is 14.97.

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

The value of sigma is 3.98. The Sigma value should approach zero for a model to be a good fit.

- Perform Residual Analysis for this technique.
  - Do a plot of residuals. What does the plot indicate?

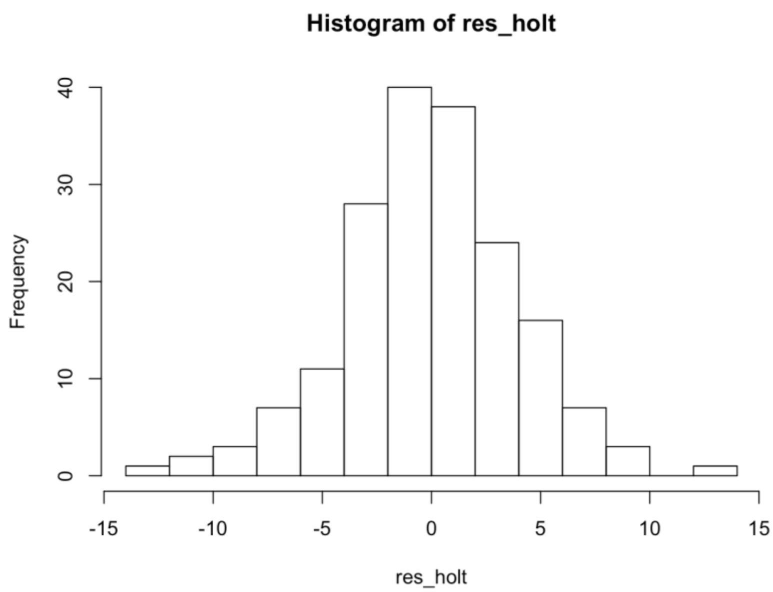
```
> res_holt<-residuals(holt)
> plot(res_holt)
>
```



The plot is random and shows high variation.

  - Do a Histogram plot of residuals. What does the plot indicate?

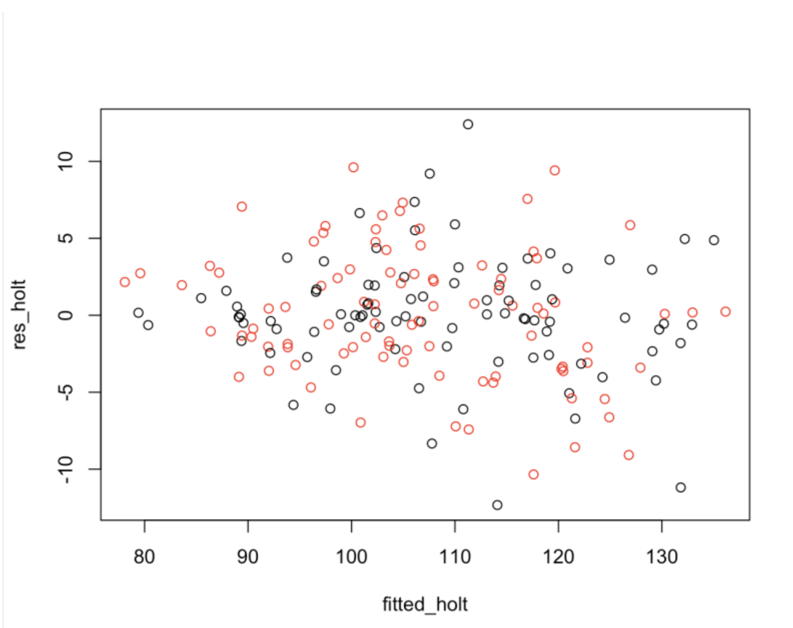
```
h<-hist(res_holt)
```



We can see an outlier between 10 to 15. The histogram is right skewed. The mean is around 0.

- Do a plot of fitted values vs. residuals. What does the plot indicate?

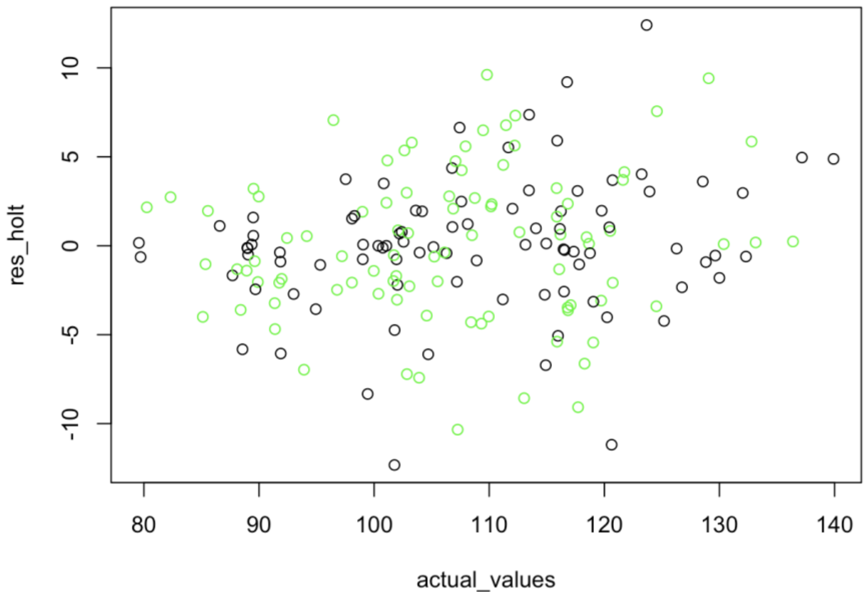
```
> fitted_holt<-holt$fitted
> plot(fitted_holt,res_holt,col=c("Red","Black"))
>
```



The plot seems quite vague to decipher any suitable information.

- Do a plot of actual values vs. residuals. What does the plot indicate?

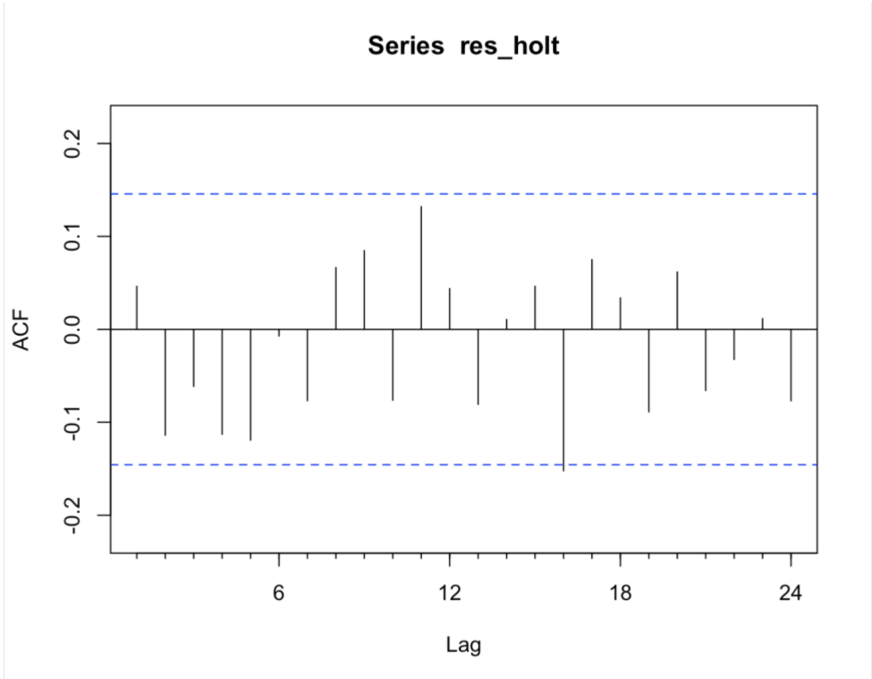
```
> actual_values<-holt$x
> plot(actual_values,res_holt,col=c("Green","Black"))
>
```



The scatter plot of residuals and actual values appears to be somewhat linear.

- Do an ACF plot of the residuals? What does this plot indicate?

```
Acf(res_holt)
```



We can see that there is no correlation here as the lines are well under the confidence band.

- Print the 5 measures of accuracy for this forecasting technique

```
> accuracy(holt)
                     ME     RMSE      MAE         MPE     MAPE     MASE       ACF1
Training set 0.02447927 3.985931 2.990258 -0.04846231 2.764973 0.493109 0.04635702
>
```

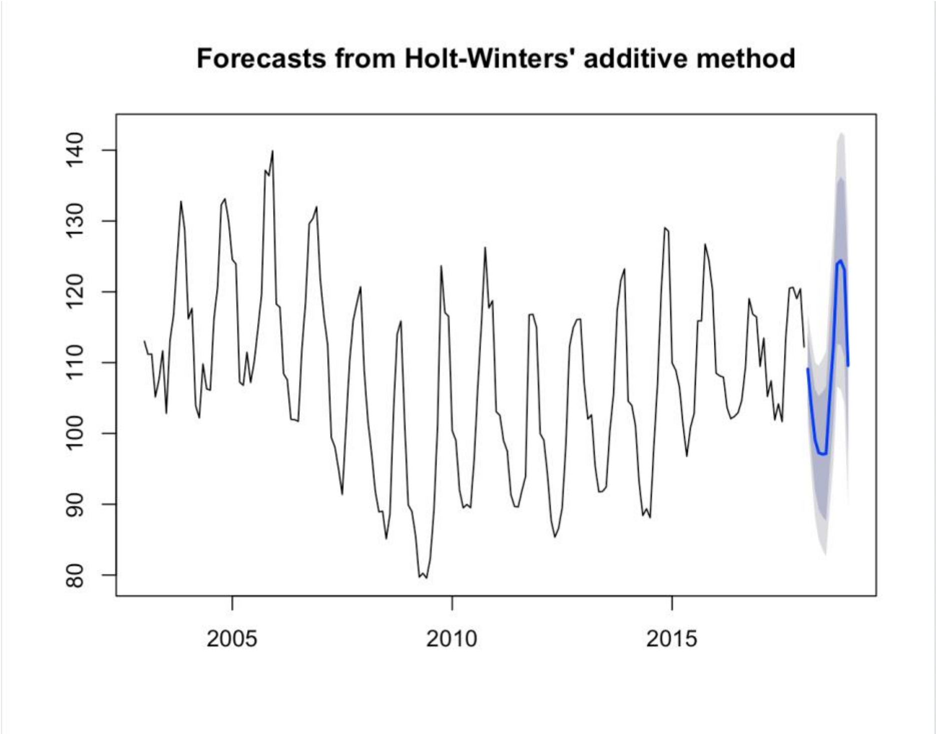
- Forecast
  - Time series value for next year. Show table and plot

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

>

```
> plot(forecast(holt))
>
```



- Summarize this forecasting technique.
  - How good is the accuracy?
    MSE for Holt Winters is close to 0, the model is good.

  - What does it predict the value of time series will be in one year?
    Since Acf shows that there isn't any relation, we can pass this of as randomness in the predictions. So the values are completely random.

  - Other observation
    Holt winters performs better than other two models.

Accuracy Summary

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

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
Naïve	-0.0045	7.422	5.4702	-0.233	5.057	0.902	0.2547
Simple Smoothing	-0.004	7.402	5.44	-0.232	5.03	0.897	0.254
Holt Winters	0.024	3.98	2.99	-0.0484	2.764	0.4931	0.0463

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

  Naïve – The model in which last period's data is used as this period's forecast without adjustment of any condition. Used mainly for comparison to other models.

  Simple Exponential Smoothing – This model smooths the data using the exponential window function and is used to assign exponentially decreasing weights over time. More useful when recent observations need to be given more weightage than past observations.

  Holt-Winters – Used to capture seasonality. Consists of three main components - level, trend and seasonality. Mainly used when we want the model to be fast as it is incremental and saves time. Also, the three components help us understand how the data is split**.**

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
BEST	Holt Winters	Naïve	Holt Winters	Simple Smoothing	Holt Winters	Holt Winters	Holt Winters
WORST	Simple Smoothing	Holt Winters	Naïve	Holt Winters	Naïve	Naïve	Naïve

Conclusion

- Summarize your analysis of time series value over the time-period.

  Looking at the time series, it seems that the production is maximum during the months of November and December each year. This could possibly be due to festivals like Halloween, Thanksgiving and Christmas. The production is lowest in the months of May, June indicating a dip in the sales of candies. There is ardent seasonality and general trend. The increase and decrease of production of candies will continue as per forecasts.

- Based on your 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?

  Based on the forecast, the production of candies may dip during the months of May, June in the year 2018 and it will be maximum in the months of November and December. Similar trend can be seen for consecutive year.

```
> plot(forecast(holt))
> forecast(candy_ts,h=24)
         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
Feb 2019      109.28137  95.13287 123.4299  87.64310 130.9196
Mar 2019      103.74316  89.17528 118.3111  81.46350 126.0228
Apr 2019       99.16507  84.21889 114.1113  76.30685 122.0233
May 2019       97.34774  82.02891 112.6666  73.91962 120.7759
Jun 2019       97.22752  81.53557 112.9195  73.22875 121.2263
Jul 2019       97.17709  81.12037 113.2338  72.62046 121.7337
Aug 2019      105.56456  89.03988 122.0892  80.29224 130.8369
Sep 2019      112.91012  95.87566 129.9446  86.85816 138.9621
Oct 2019      124.67594 107.01849 142.3334  97.67121 151.6807
Nov 2019      125.18774 106.99248 143.3830  97.36049 153.0150
Dec 2019      124.34403 105.64132 143.0467  95.74071 152.9473
Jan 2020      110.82761  91.80364 129.8516  81.73295 139.9223
> 
```

- Rank forecasting methods that best forecast for this time series based on historical values.
  1. Holt-Winters RMSE: 3.98
  2. Simple Exponential Smoothing RMSE: 7.402
  3. Naïve RMSE: 7.422