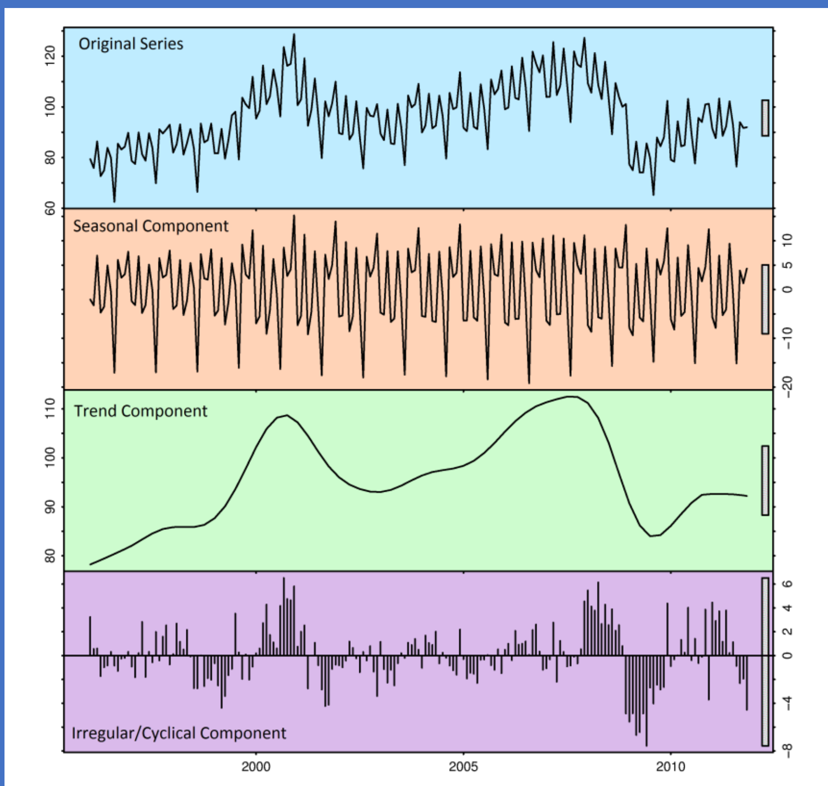
IÉSEG SCHOOL OF MANAGEMENT

EMPOWERING CHANGEMAKERS FOR A BETTER SOCIETY

FORECASTING INDIVIDUAL ASSIGNMENT



Submitted By:

PRINEET KAUR BHURJI

EXERCISE 1

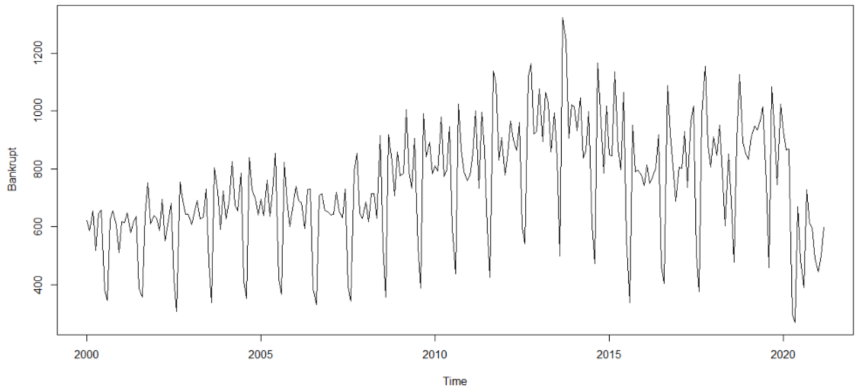
The purpose of this document is to describe the application of various Forecasting Models and the corresponding evaluation of them with an aim to come up with the expected Best Forecasts. For the first task, I used the "**Bankrupt**" dataset which contains the number of bankruptcies in Belgium (per month) for all economic activities, from January 2000 to March 2021. As a data scientist would need to analyse data to check for relevant trends and forecast the number of bankruptcies for Belgium until December 2022.

The Forecasting Process will consist of five main phases –

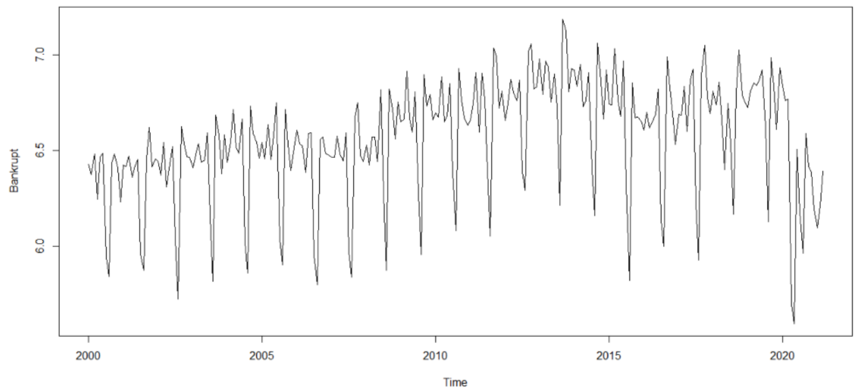
- *Exploratory Data Analysis (Understanding and Decomposing the Time Series Data)*
- *Model Fitting (experimenting using different Methods)*
- *Evaluation of each of the Models*
- *Comparing All Models (picking the Best for Final Forecast)*
- *Final Forecast*

EXPLORATORY DATA ANALYSIS

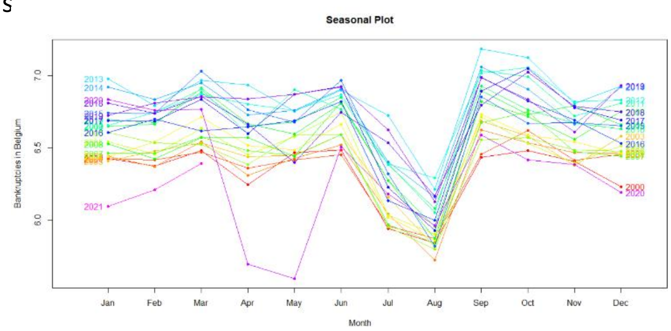
To explore the given data, I read the file and obtained the following graph by applying the plot() function.

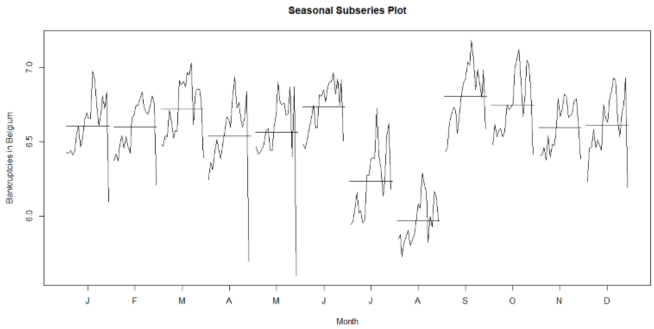


From the above plot we can observe that there is no stationarity, instead, there is a clear increasing trend and a strong seasonal pattern. It can also be seen that the seasonality variation increases slightly as the level of the series increases (with an exception at the end), therefore it was a good idea to log transform the time series data. Below plot shows the data behaviour after the log transformation was done using log() function–



After the application of log transformation in the previous step, the time series variations seemed to be more compact now. To further explore the seasonality component, I went ahead with plotting the Seasonal and Seasonal subseries graphs





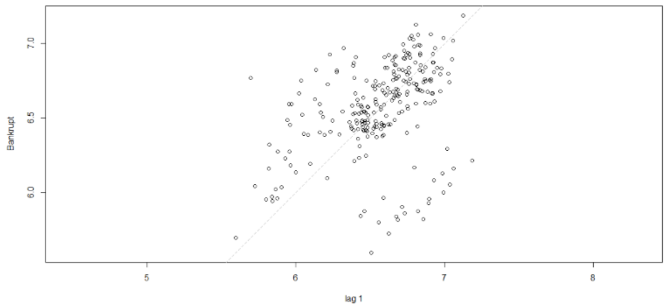
<u>Both the graphs above show that there is Seasonality as well as a Trend in the data under consideration</u>.

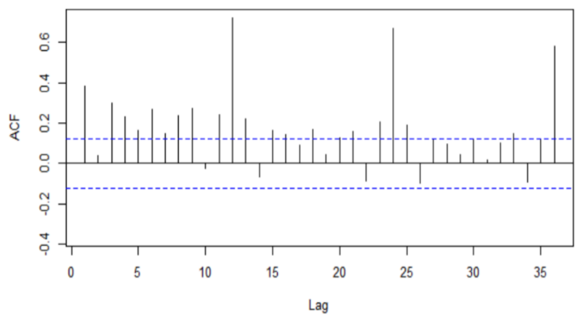
Further, the need was to check is there is any autocorrelation or partial autocorrelation if at all it exists.

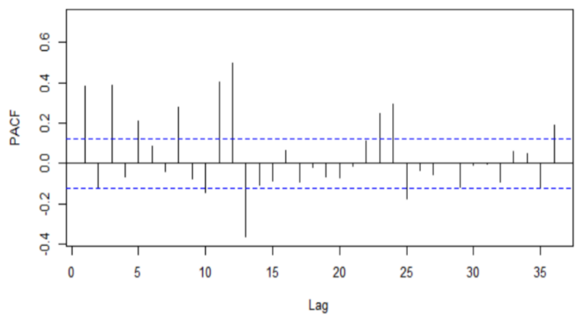
Let us understand what we mean by ACF and PACF first –

- **ACF** is an (complete) auto-correlation function which gives us values of autocorrelation of any series with its lagged values. In simple terms, it describes how well the present value of the series is related with its past values.
- **PACF** is a partial autocorrelation function. Thus, instead of finding correlations of present with lags like ACF, it finds correlation of the residuals with the next lag value hence 'partial' and not 'complete' as we remove already found variations before we find the next correlation.

The graphs below show some points of high autocorrelation values occurring at lags 12, 24 and 36 which are above of 0.60 and even some component of seasonality. From this information, one may wonder about the lack of white noise. But if we see the PACF graph we see that there is not much white noise component.





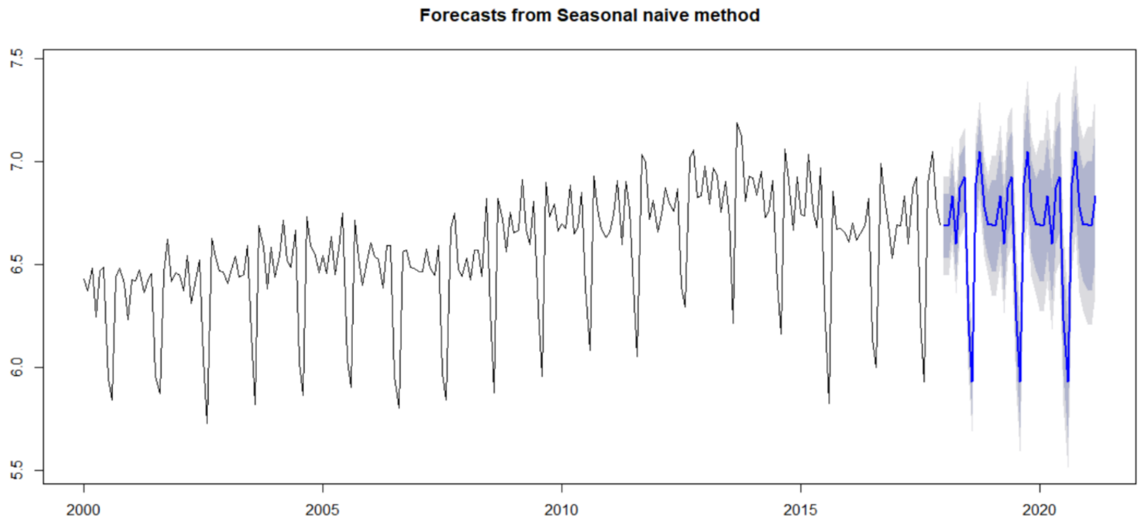


MODEL FITTING

To begin with the data was first splitted into the Train and Test sets. The Train set includes data from **January 2000 to December 2017** and the Test set includes data from **January 2018 to March 2021**.

Seasonal Naive Method

For the Forecasting purpose, I first chose the Seasonal Naive Method. Used to fit the Training set and evaluate the performance on both Train and Test sets. The results were as follows –

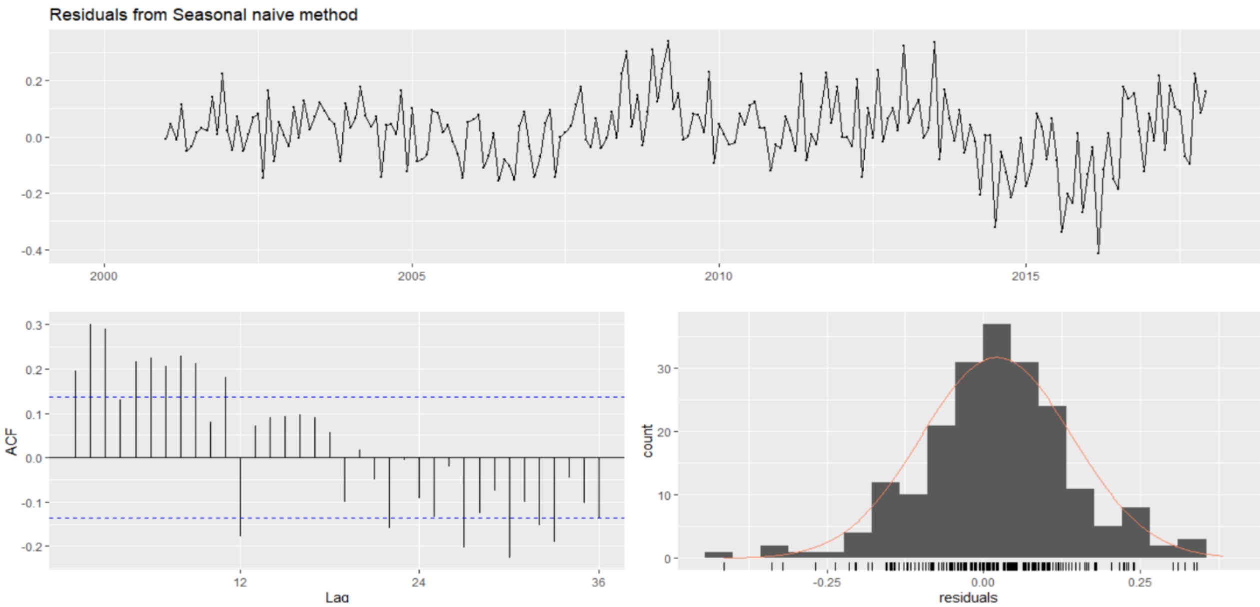


EVALUATION OF MODEL

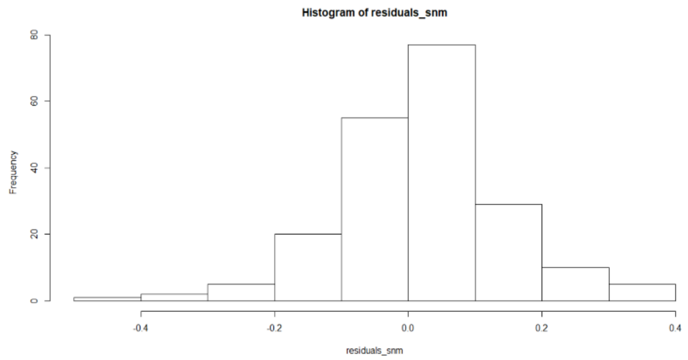
For evaluating the performance of the method used to build the models, I went ahead with checking the Residuals and applying the Ljung–Box test. Let us first understand what they signify –

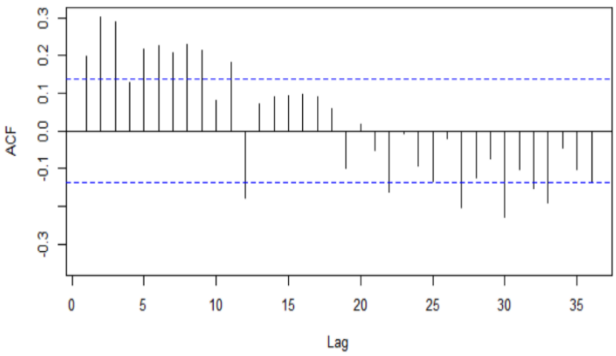
- **Residuals** in a time series model are what is left over after fitting a model. They show the difference between the observations and the corresponding fitted values. Residuals are useful in checking whether a model has adequately captured the information in the data and a good forecasting method will yield residuals with no or very less correlation.
- **Ljung–Box test** is a type of statistical test of whether any of a group of autocorrelations of a time series are different from zero. Instead of testing randomness at each distinct lag, it tests the "overall" randomness based on a number of lags and is therefore a portmanteau test.

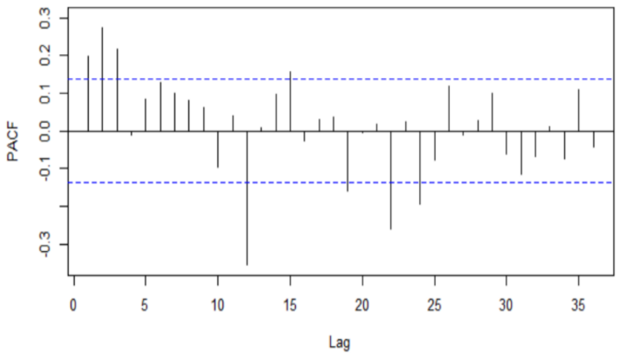
The below plots show the behaviour of the residuals and the accuracy of the model –



Went ahead with plotting the histogram of the residuals and the ACF and PACF plots to understand the autocorrelation and the partial autocorrelation if any.








```
> LjungBox(residuals_snm[-c(1:12),
 lags  statistic df      p-value
    1   8.013276  1 4.643566e-03
    2  26.888712  2 1.449407e-06
    3  44.466212  3 1.201425e-09
    4  48.008554  4 9.399166e-10
    5  57.999535  5 3.146161e-11
    6  68.809770  6 7.170931e-13
    7  77.893785  7 3.697043e-14
    8  89.175093  8 6.661338e-16
    9  98.920409  9 0.000000e+00
   10 100.324128 10 0.000000e+00
   11 107.441513 11 0.000000e+00
   12 114.387167 12 0.000000e+00
   13 115.553955 13 0.000000e+00
   14 117.374331 14 0.000000e+00
   15 119.313308 15 0.000000e+00
   16 121.379064 16 0.000000e+00
   17 123.239848 17 0.000000e+00
   18 123.986039 18 0.000000e+00
   19 126.257970 19 0.000000e+00
   20 126.332952 20 0.000000e+00
   21 126.907991 21 0.000000e+00
   22 132.842555 22 0.000000e+00
   23 132.852609 23 0.000000e+00
   24 134.860759 24 0.000000e+00
```

```
        Box-Ljung test

data:  residuals_snm
X-squared = 134.86, df = 24, p-value < 2.2e-16
```

From the above Residual Plots visualizations and the Ljung Box tests performed for this model, we can see that the p-values (parameters in the tests) are less than .05 which means that they are significant. So, we can reject the null hypothesis of white noise and there is still something in the residuals that is not been captured by the model.

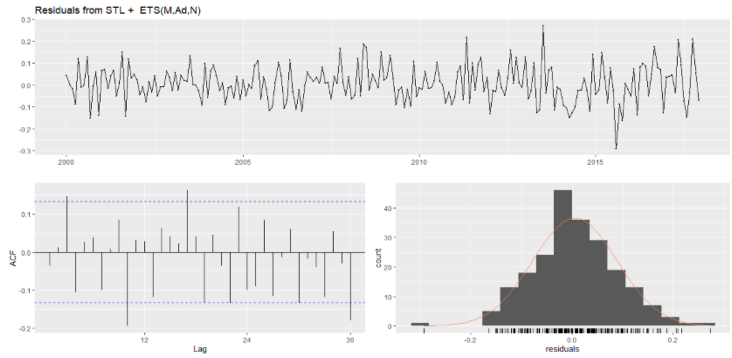
In terms of accuracy, the following results we obtained. These will be used further as reference for model selection purposes.

```
> accuracy(snm, test)[,c(2,3,4,5,6,7)]
                  RMSE        MAE        MPE     MAPE     MASE      ACF1
Training set 0.1222211 0.09289657  0.3073768 1.409821 1.000000 0.1967453
Test set     0.3637954 0.24789251 -2.0490865 3.960417 2.668479 0.6127023
```

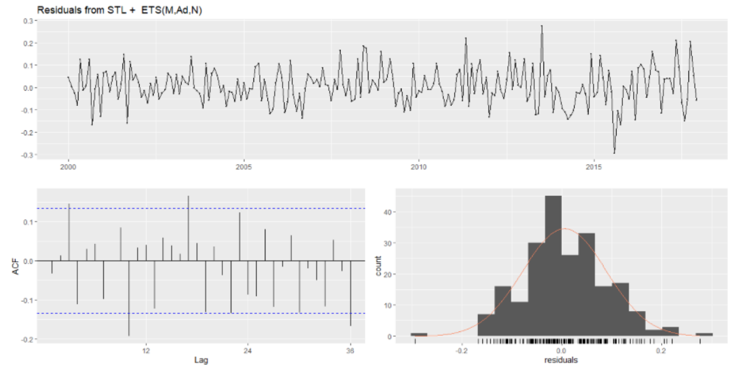
Seasonal Naive Method with STL Decomposition

Three different cases with varying t.window values were taken (while keeping s.window as “periodic”) along with rwdrift function. Residuals were checked for all cases and the results were as follows –

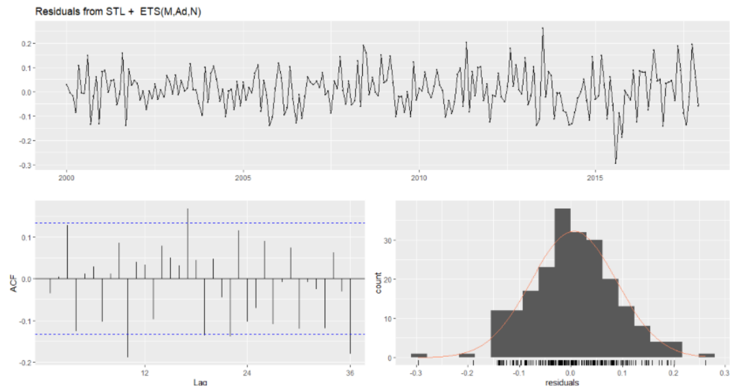
t.window = 12



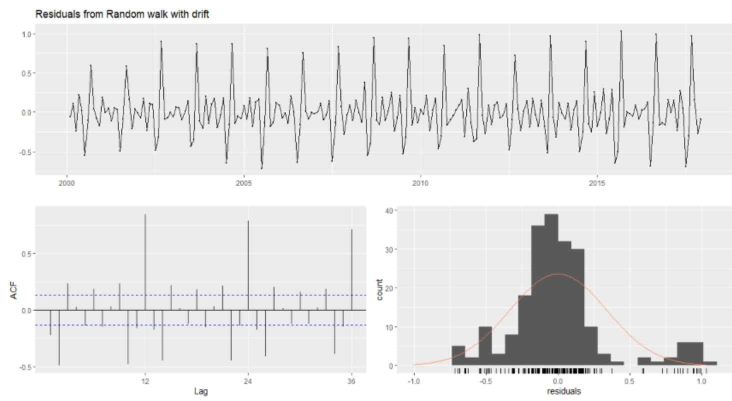
t.window = 6



t.window = 3

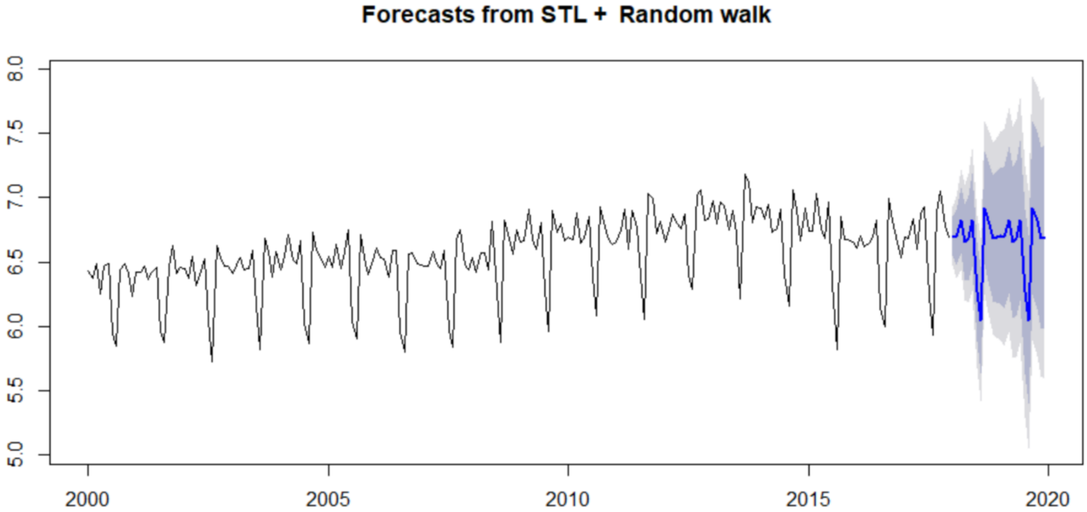


rwdrift



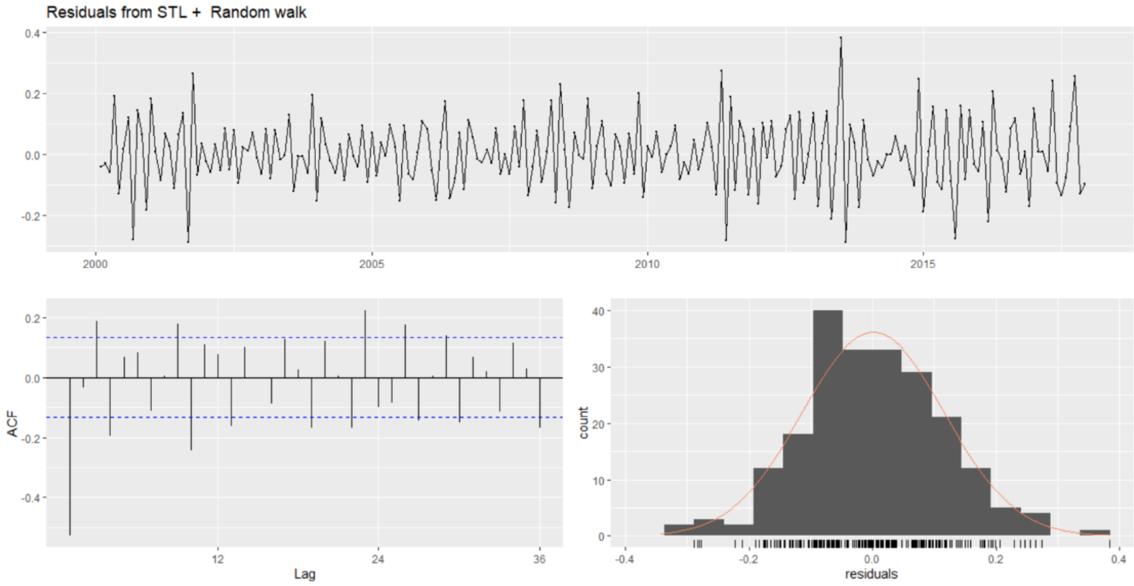
After analysing the above residual graphs, realized that the least residuals were seen in the case of t.window = 6. Also, the autocorrelation also seemed to be least. Thus, went ahead with choosing the same.

The seasonal naïve model with STL decomposition was fitted on the training set and the evaluation was done using both training and testing set. The Forecasting results were as follows –

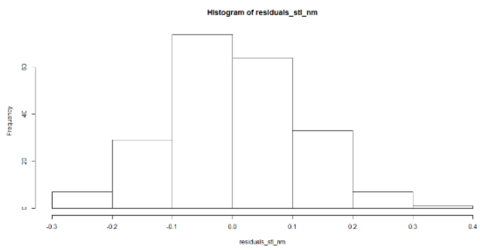


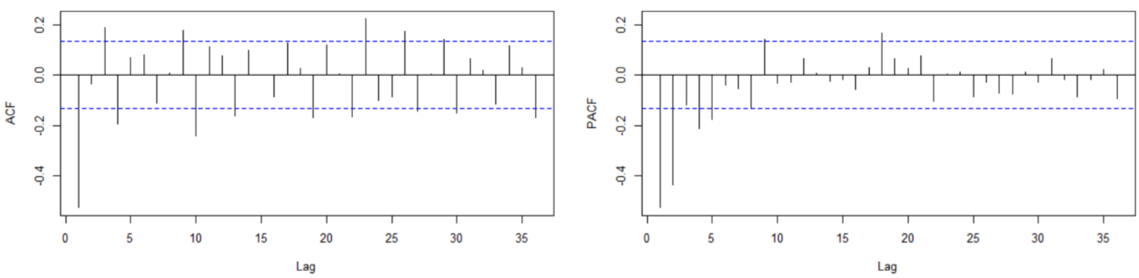
EVALUATION OF MODEL

The below plots show the behaviour of the residuals and the accuracy of the model.



Went ahead with plotting the histogram of the residuals and the ACF and PACF plots to understand the autocorrelation and the partial autocorrelation if any.





```
        Box-Ljung test

data:  residuals_stl_nm
X-squared = 153.63, df = 24, p-value < 2.2e-16
```

```
LjungBox(residuals_stl_nm[-1],
lags statistic df      p-value
   1  61.01598  1 5.662137e-15
   2  61.28023  2 4.929390e-14
   3  69.12171  3 6.550316e-15
   4  77.51116  4 5.551115e-16
   5  78.57749  5 1.665335e-15
   6  80.07257  6 3.441691e-15
   7  82.89794  7 3.552714e-15
   8  82.90751  8 1.265654e-14
   9  90.10455  9 1.554312e-15
  10 103.57013 10 0.000000e+00
  11 106.39788 11 0.000000e+00
  12 107.78682 12 0.000000e+00
  13 113.86515 13 0.000000e+00
  14 116.19262 14 0.000000e+00
  15 116.19268 15 0.000000e+00
  16 117.99014 16 0.000000e+00
  17 121.87945 17 0.000000e+00
  18 122.04871 18 0.000000e+00
  19 128.75021 19 0.000000e+00
  20 132.28214 20 0.000000e+00
  21 132.28552 21 0.000000e+00
  22 139.02660 22 0.000000e+00
  23 151.24299 23 0.000000e+00
  24 153.62827 24 0.000000e+00
```

From the above Residual Plots visualizations and the Ljung Box tests performed for this model, we can see that the p-values (parameters in the tests) are less than .05 which means that they are significant. So, we can reject the null hypothesis of white noise and there is still something in the residuals that is not been captured by the model.

In terms of accuracy, the following results we obtained. These will be used further as reference for model selection purposes –

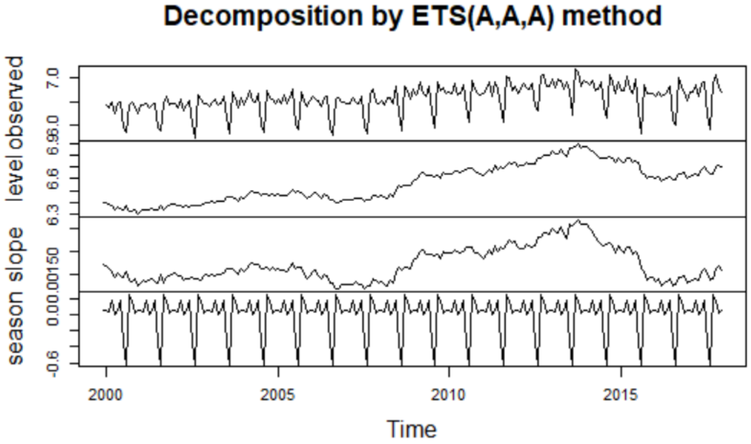
	RMSE	MAE	MPE	MAPE	MASE	ACF1
Training set	0.1143733	0.0904064	0.002379906	1.381122	0.9731942	-0.52902914
Test set	0.1519006	0.1211428	1.105732567	1.808763	1.3040611	0.06472759

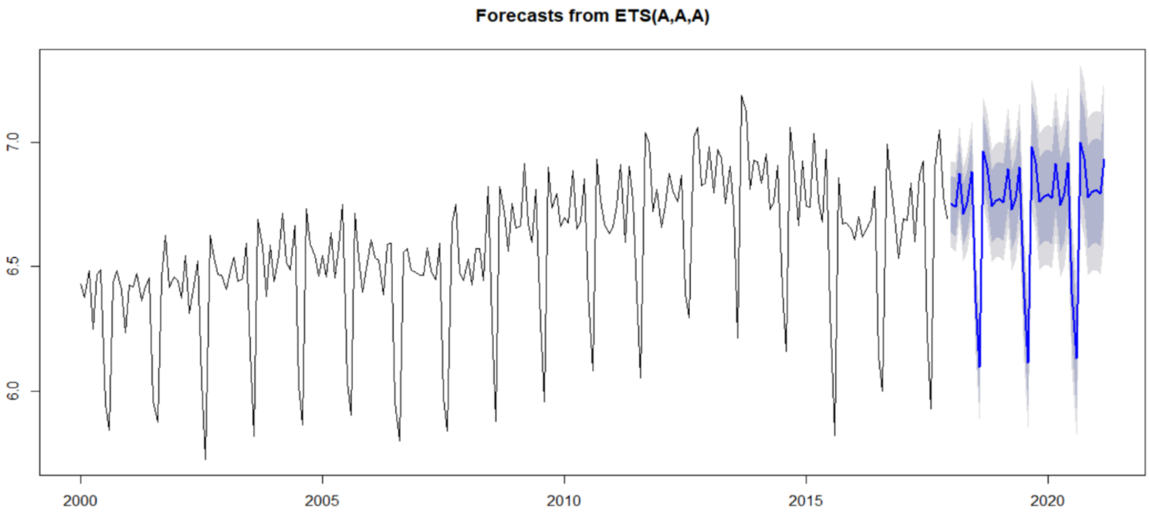
Exponential Smoothening Method (ETS)

For forecasting using the ETS method, I handpicked only 8 models that incorporate the seasonality and trend components from the 18 available models. The list of the models chosen includes –

- ETS1 (A, N, A)
- ETS2 (A, N, Ad)
- ETS3 (A, A, A)
- ETS4 (A, A, Ad)
- ETS5 (M, N, A)
- ETS6 (M, N, Ad)
- ETS7 (M, A, A)
- ETS8 (M, A, Ad)

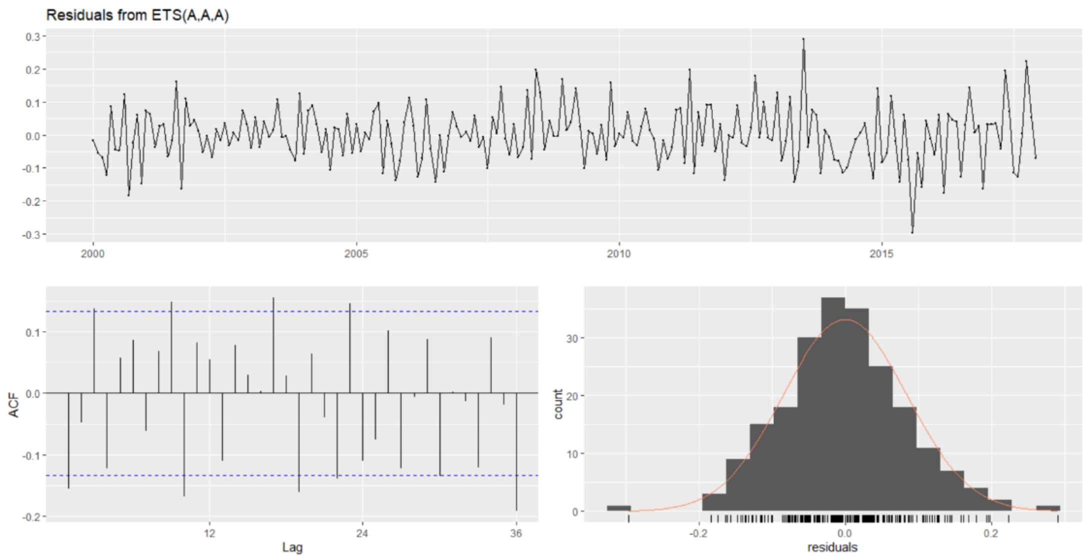
Checked the Residuals and Accuracy of each of the above models. (For all individual plots for these methods kindly refer to the code, for report purpose only included the results for the Auto ETS). As the results of all the models were quite similar, went ahead with applying the automated ETS method. It showed **ETS(A,A,A)** as the best performing method so went ahead with using the same for further analysis. The results of the auto ETS were as follows –



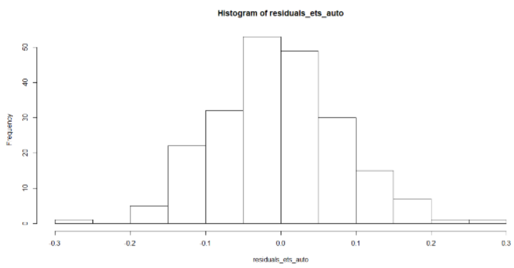


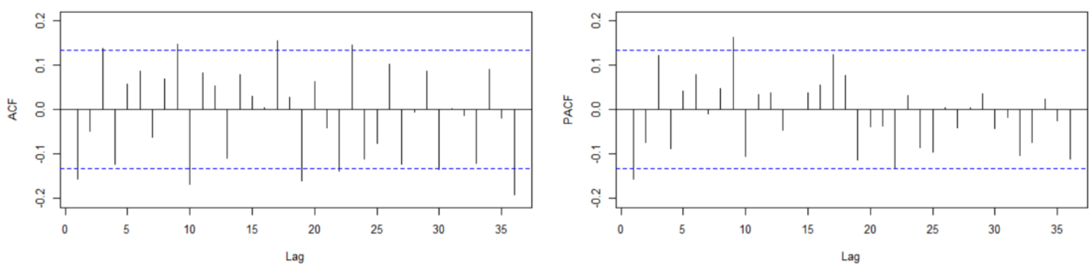
EVALUATION OF MODEL

For evaluating the performance of the method used to build the models, I went ahead with checking the Residuals and applying the Ljung–Box test. The below plots show the behaviour of the residuals and the accuracy of the model -



Went ahead with plotting the histogram of the residuals and the ACF and PACF plots to understand the autocorrelation and the partial autocorrelation if any.





```
        Box-Ljung test

data:  residuals_ets_auto
X-squared = 62.121, df = 24, p-value = 3.187e-05
```

```
LjungBox(residuals_ets_auto, la
lags statistic df       p-value
   1  5.330841  1 2.095127e-02
   2  5.826836  2 5.428986e-02
   3  9.991152  3 1.864150e-02
   4 13.290490  4 9.940269e-03
   5 14.034880  5 1.538936e-02
   6 15.717257  6 1.535477e-02
   7 16.585828  7 2.027116e-02
   8 17.660034  8 2.392479e-02
   9 22.690711  9 6.929596e-03
  10 29.143578 10 1.181061e-03
  11 30.706043 11 1.226322e-03
  12 31.374487 12 1.726785e-03
  13 34.180715 13 1.129769e-03
  14 35.620587 14 1.188920e-03
  15 35.833235 15 1.868387e-03
  16 35.836808 16 3.048243e-03
  17 41.557701 17 7.766388e-04
  18 41.740044 18 1.202162e-03
  19 47.908651 19 2.644342e-04
  20 48.897327 20 3.179206e-04
  21 49.280468 21 4.589753e-04
  22 53.952563 22 1.670369e-04
  23 59.136455 23 5.087817e-05
  24 62.120546 24 3.186914e-05
```

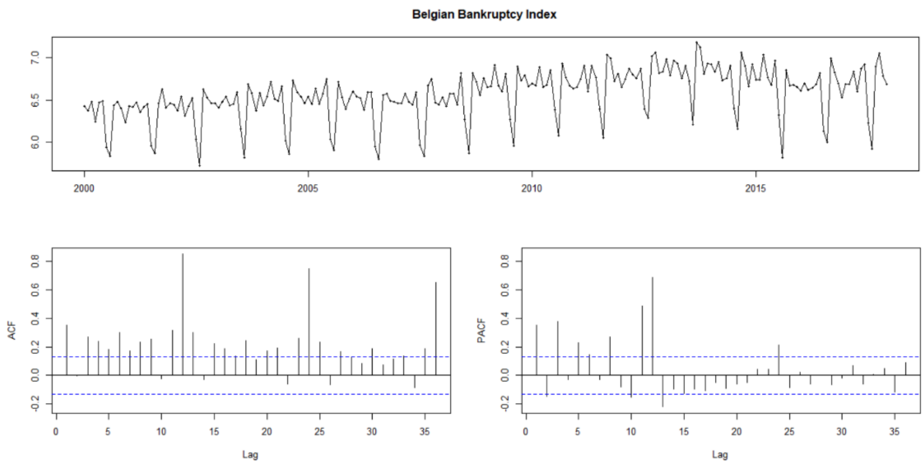
From the above Residual Plots visualizations and the Ljung Box tests performed for this model, we can see that the p-values (parameters in the tests) are less than .05 which means that they are significant. So, we can reject the null hypothesis of white noise and there is still something in the residuals that is not been captured by the model.

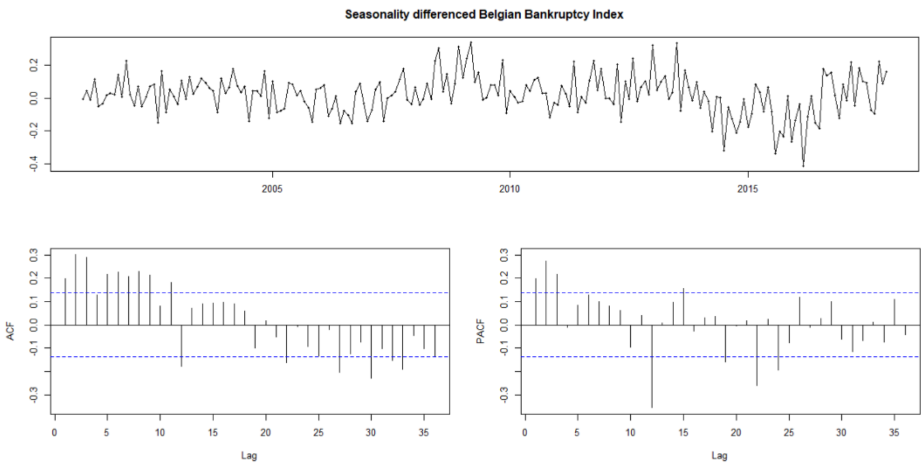
In terms of accuracy, the following results we obtained. These will be used further as reference for model selection purposes –

	RMSE	MAE	MPE	MAPE	MASE	ACF1
Training set	0.08448937	0.06532629	-0.02118966	0.997086	0.7032153	-0.1560134
Test set	0.36699750	0.23652551	-2.85637851	3.803810	2.5461168	0.6882844

<u>Seasonal ARIMA Method</u>

Before applying the auto ARIMA method, to add a component of stationarity and to help stabilize the mean, first computed the difference/variation in the observations and the output was plotted. The two graphs below show the graphs before and after the difference was computed –



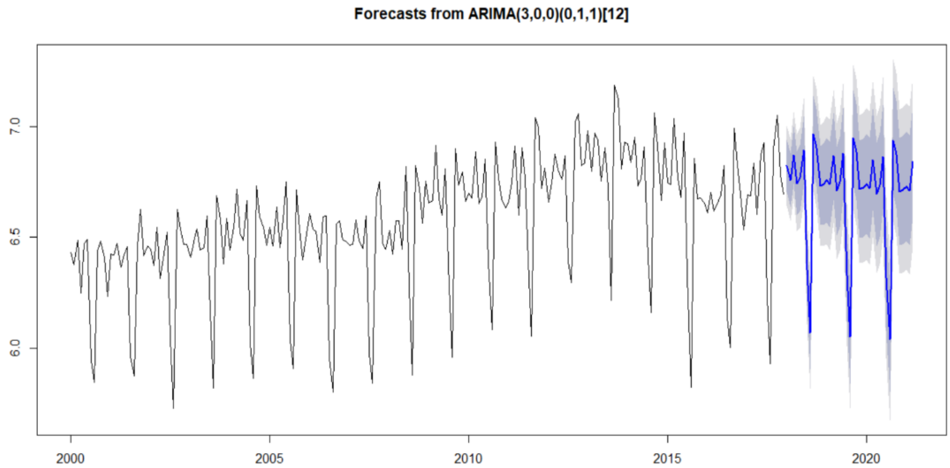


We can still observe some significant peaks in the PACF graph, in order to address this situation, some models were handpicked manually adjusting the elements in the trending and seasonal components (P, D, Q). The list of the models chosen includes –

- ARIMA (1, 0, 1) (2, 1, 1)[12]
- ARIMA (3, 0, 1) (0, 1, 1)[12]
- ARIMA (3, 0, 1) (3, 1, 1)[12]
- ARIMA (3, 1, 1) (3, 1, 1)[12]
- ARIMA (3, 0, 3) (3, 1, 1)[12]
- ARIMA (3, 0, 3) (3, 1, 3)[12]

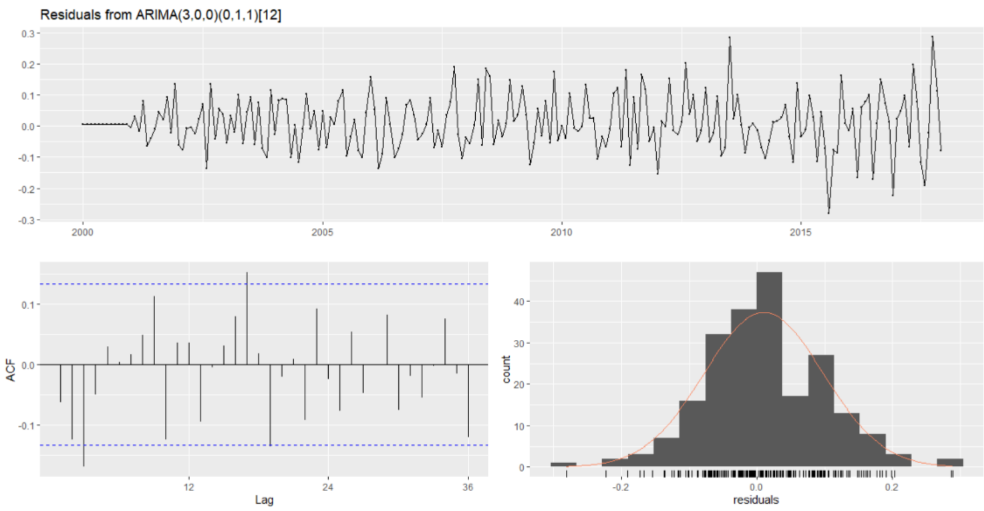
Checked the Residuals and Accuracy of each of the above models. (For all individual plots for these kindly refer to the code, for report purpose only included the results for the Auto ARIMA). As the results of all the models were not so great and to ensure that I choose the best model, went ahead with applying the automated ARIMA method. It showed **ARIMA(3,0,0)(0,1,1)[12]** as the best performing method so went ahead with using the same for further analysis. The results of the auto ARIMA were as follows –

```
ARIMA(3,0,0)(0,1,1)[12]

Coefficients:
         ar1     ar2     ar3     sma1
      0.2031  0.3051  0.4224  -0.8823
s.e.  0.0647  0.0621  0.0663   0.0832

sigma^2 estimated as 0.008388:  log likelihood=191.19
AIC=-372.37   AICc=-372.07   BIC=-355.78
```

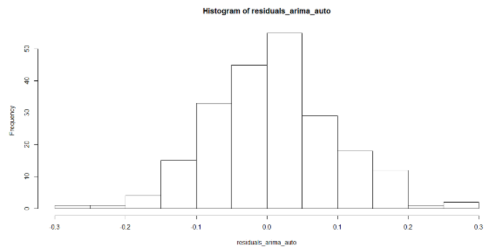


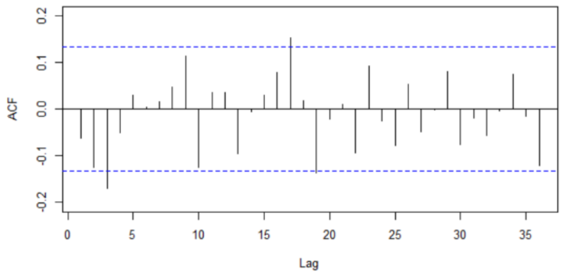
EVALUATION OF MODEL

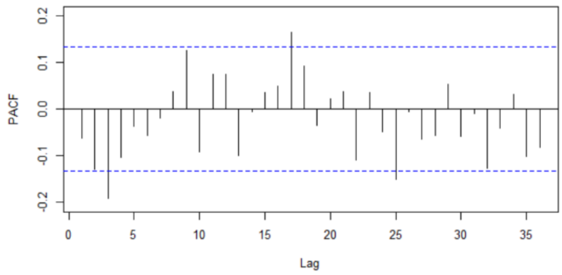
For evaluating the performance of the method used to build the models, I went ahead with checking the Residuals and applying the Ljung–Box test. The below plots show the behaviour of the residuals and the accuracy of the model -



Went ahead with plotting the histogram of the residuals and the ACF and PACF plots to understand the autocorrelation and the partial autocorrelation if any.








```
        Box-Ljung test

data:  residuals_arima_auto
X-squared = 37.378, df = 24, p-value = 0.0401
```

From the above Residual Plots visualizations and the Ljung Box tests performed for this model, we can see that the p-values (parameters in the tests) are less than .05 which means that they are significant. So, we can reject the null hypothesis of white noise and there is still something in the residuals that is not been captured by the model.

```
> LjungBox(residuals_arima_auto,
 lags  statistic df    p-value
    1  0.8688761  1 0.35126664
    2  4.2662580  2 0.11846603
    3 10.6604848  3 0.01371084
    4 11.2311411  4 0.02408552
    5 11.4266844  5 0.04354651
    6 11.4300093  6 0.07596156
    7 11.4921839  7 0.11854457
    8 12.0296619  8 0.14988534
    9 14.9615970  9 0.09199626
   10 18.4981166 10 0.04712034
   11 18.8073988 11 0.06463960
   12 19.1076709 12 0.08596296
   13 21.1965484 13 0.06913100
   14 21.2012219 14 0.09658603
   15 21.4267584 15 0.12373939
   16 22.9027475 16 0.11635913
   17 28.4540931 17 0.03990401
   18 28.5272651 18 0.05447154
   19 32.9525226 19 0.02434502
   20 33.0532189 20 0.03328978
   21 33.0756907 21 0.04538453
   22 35.1526952 22 0.03734495
   23 37.2349358 23 0.03071778
   24 37.3778769 24 0.04010404
```

In terms of accuracy, the following results we obtained –

```
                   RMSE        MAE        MPE     MAPE      MASE        ACF1
Training set 0.08812702 0.06757031  0.1318679 1.028642 0.7273714 -0.06298586
Test set     0.33700432 0.21993784 -2.2999930 3.529802 2.3675561  0.65957101
```

COMPARING ALL MODELS

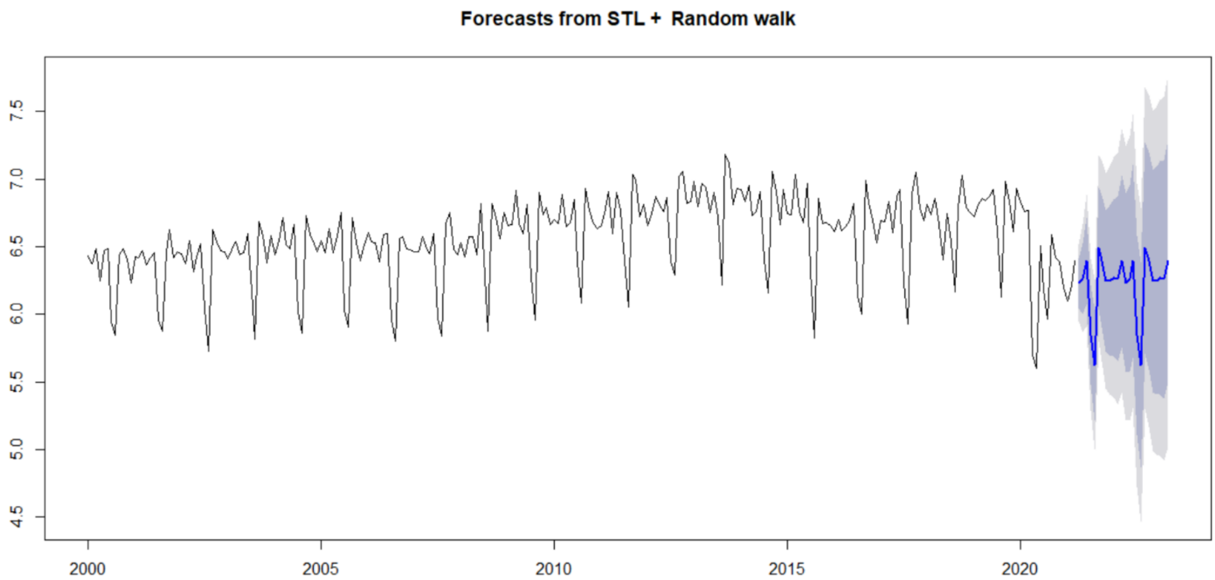
To choose which model performed better, let us compare the results of the final accuracy scores –

KPI	SNAIVE	SNAIVE STL	Auto ETS	Auto ARIMA
p-value	2.2E-16	2.2E-16	3.19E-05	0.0401
RMSE	0.3637954	**0.1519006**	0.3669975	0.33700432
MAE	0.24789251	0.1211428	0.23652551	0.21993784
MPE	-2.0490865	1.105732567	-2.85637851	-2.299993
MAPE	3.960417	1.808763	3.80381	3.529802
MASE	2.668479	1.3040611	2.5461168	2.3675561
ACF1	0.6127023	0.06472759	0.6882844	0.65957101

We can conclude that even though the models used did not perform well on the residuals test, but looking at the accuracy scores across various KPIs, we can say that the **Seasonal Naïve Method with STL Decomposition** performed overall the best. So, I went ahead with applying the same for coming up with the Final forecast.

FINAL FORECAST

To get the Final Forecast applied the evaluated Best Model in the previous step (i.e., Seasonal Naïve Method with STL) to the entire dataset and the results were as follows –



FINAL MODEL ACCURACY

RMSE	MAE	MPE	MAPE	MASE	ACF1
0.1432197	0.1025948	-0.0373248	1.5752022	0.8435015	-0.2731486

With the above scores for some of the important KPI’s like the MAPE, MAE and RMSE we can say that the Final Model accuracy is pretty good in terms of predictions and can be relied upon.

Thus, from the final forecast as seen from above graph, we can observe and conclude that the trend of Bankruptcies is downward in coming years and that the Bankruptcies are forecasted to decrease by December 2022.

EXERCISE 2

For this forecasting task, I used the "**Total Energy Monthly Data for the US**" since January 1973 to January 2021 (which is the latest data released by the **U.S. Energy Information Administration** on 27th April 2021).

Website Link - [Total Energy Monthly Data - U.S. Energy Information Administration (EIA)](http://www.eia.gov/totalenergy/data/monthly/)

Dataset Link - http://www.eia.gov/totalenergy/data/monthly/#electricity (Source: Table 7.6.)

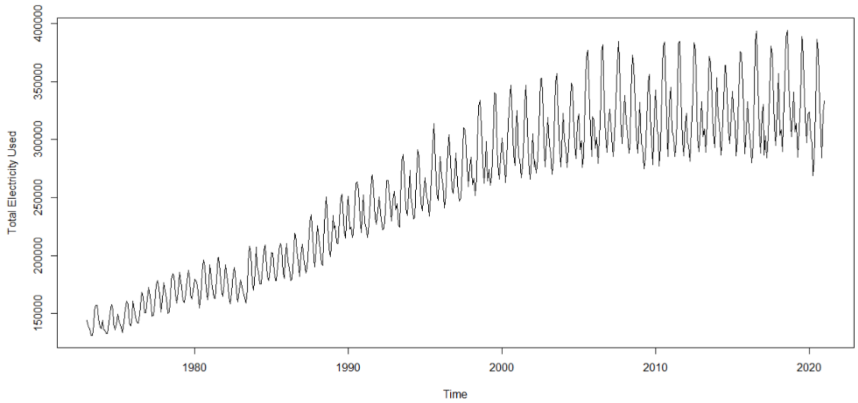
Further, a new sheet was created to just include the "**Month**" & "**Total Electricity Used**" columns and was saved as "**Data_Formatted**" in the main excel file.

The Forecasting Process will consist of five main phases –

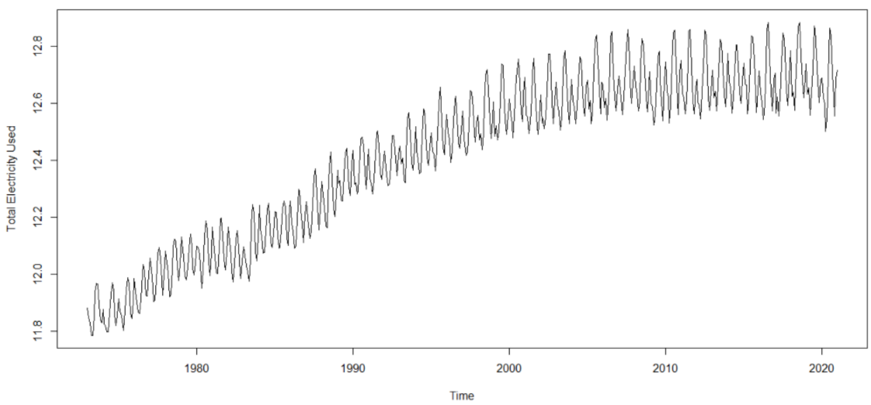
- *Exploratory Data Analysis (Understanding and Decomposing the Time Series Data)*
- *Model Fitting (experimenting using different Methods)*
- *Evaluation of each of the Models*
- *Comparing All Models (picking the Best for Final Forecast)*
- *Final Forecast*

EXPLORATORY DATA ANALYSIS

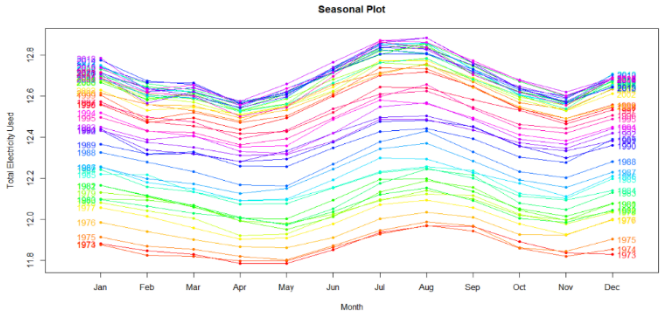
To explore the given data, I read the file and obtained the following graph by applying the plot() function.

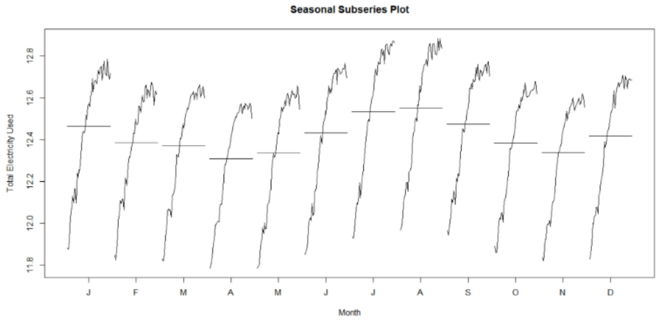


From the above plot we can observe that there is no stationarity, instead, there is a clear increasing trend and a strong seasonal pattern. It can also be seen that the seasonality variation increases slightly as the level of the series increases, therefore it was a good idea to log transform the time series data. The below plot shows the data behaviour after the log transformation was done using log() function –



After the application of log transformation in the previous step, the time series variations seemed to be more compact now. To further explore the seasonality component, I went ahead with plotting the Seasonal and Seasonal subseries graphs.



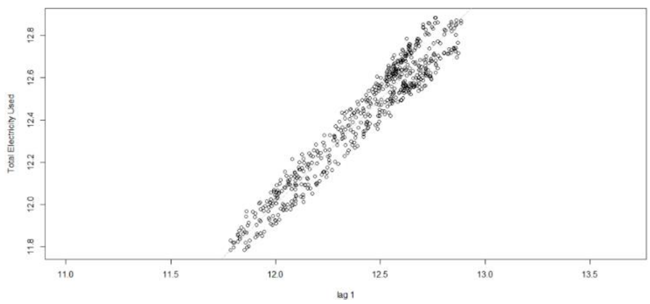


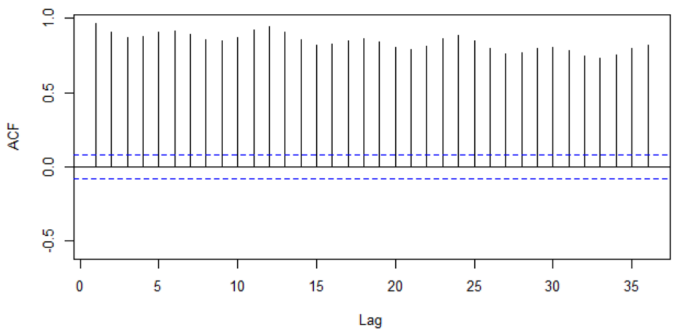
<u>Both the graphs above show that there is Seasonality as well as a Trend in the data under consideration</u>.

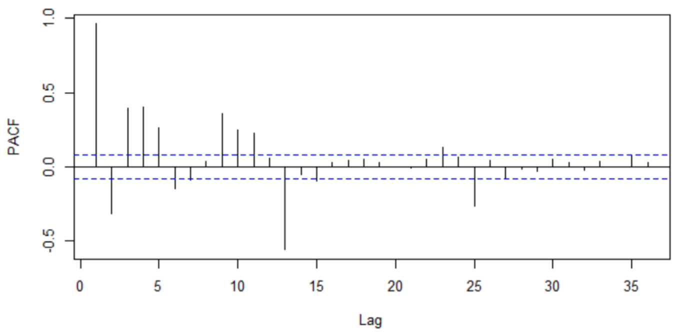
Further, the need was to check is there is any autocorrelation or partial autocorrelation if at all it exists. Let us understand what we mean by ACF and PACF first –

- **ACF** is an (complete) auto-correlation function which gives us values of autocorrelation of any series with its lagged values. In simple terms, it describes how well the present value of the series is related with its past values.
- **PACF** is a partial auto-correlation function. Thus, instead of finding correlations of present with lags like ACF, it finds correlation of the residuals with the next lag value hence 'partial' and not 'complete' as we remove already found variations before we find the next correlation.

The graphs below show that there is a strong autocorrelation with so many high values in the ACF graph and even a strong component of seasonality. From this information, one may wonder about the lack of white noise. But if we see the PACF graph we see that there is not much white noise component.





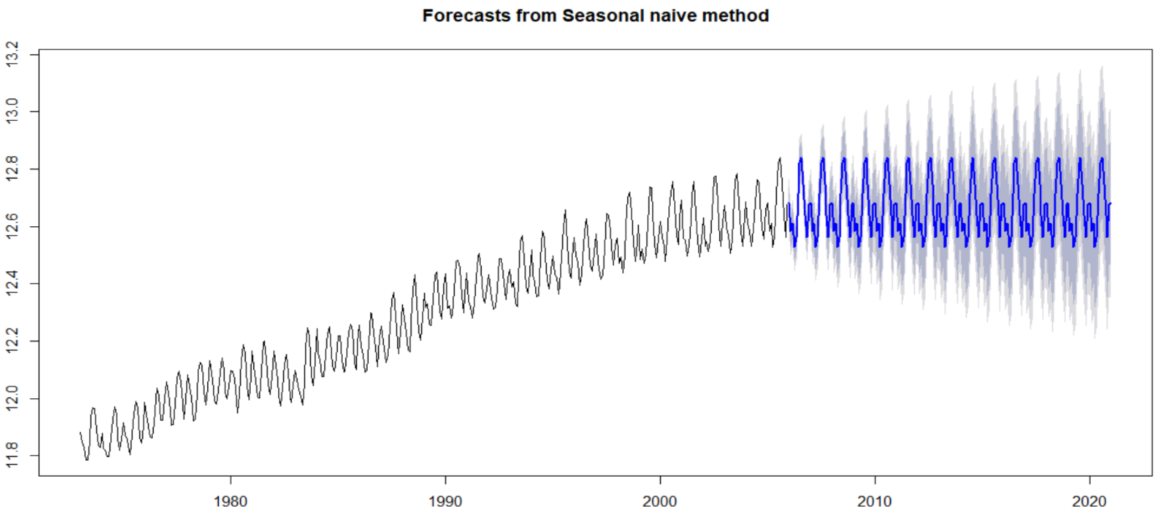


MODEL FITTING

To begin with the data was first splitted into the Train and Test sets. The Train set includes data from **January 1973 to December 2005** and the Test set includes data from **January 2006 to January 2021**.

Seasonal Naive Method

For the Forecasting purpose I first chose the Seasonal Naive Method. Used to fit the Training set and evaluate the performance on both Train & Test sets. The results were as follows -

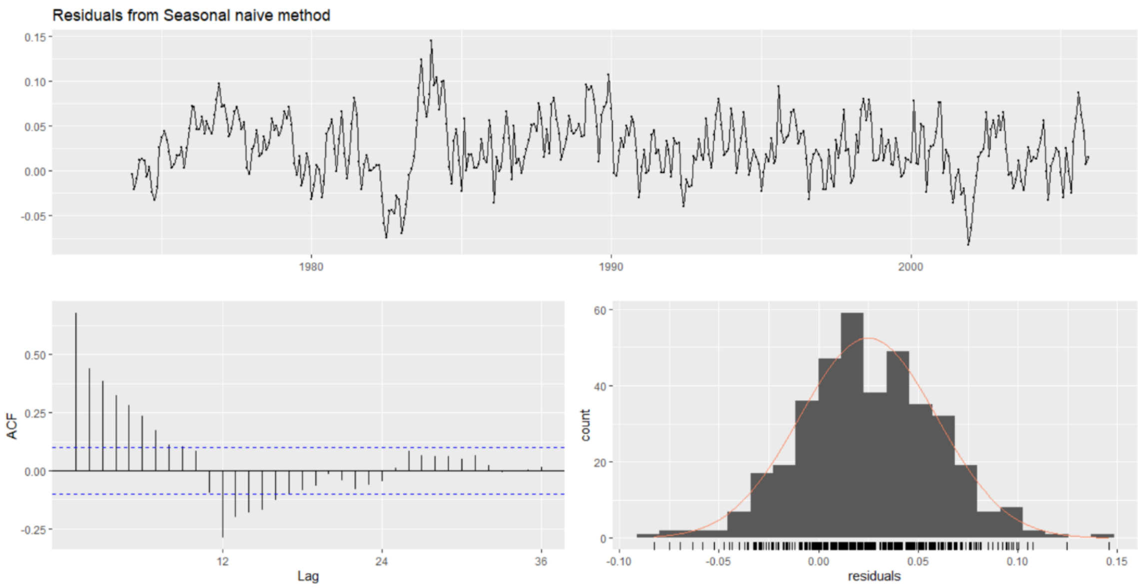


EVALUATION OF MODEL

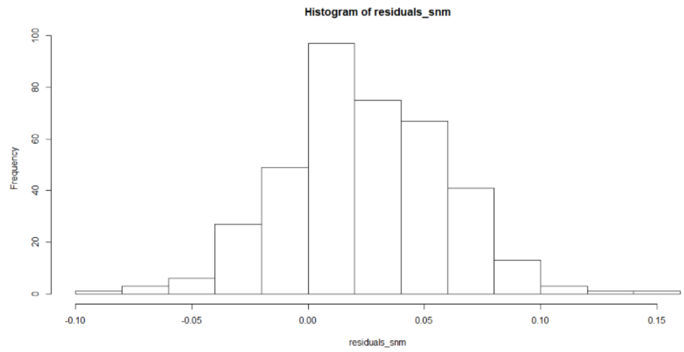
For evaluating the performance of the method used to build the models, I went ahead with checking the Residuals and applying the Ljung–Box test. Let us first understand what they signify –

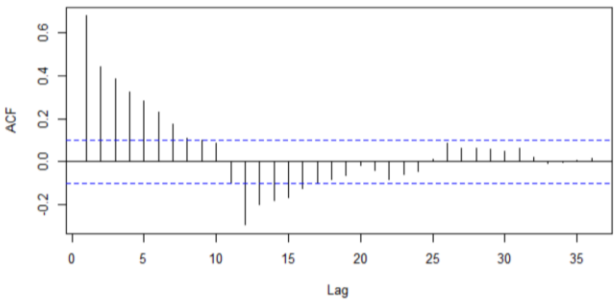
- **Residuals** in a time series model are what is left over after fitting a model. They show the difference between the observations and the corresponding fitted values. Residuals are useful in checking whether a model has adequately captured the information in the data and a good forecasting method will yield residuals with no or very less correlation.
- **Ljung–Box test** is a type of statistical test of whether any of a group of autocorrelations of a time series are different from zero. Instead of testing randomness at each distinct lag, it tests the "overall" randomness based on a number of lags and is therefore a portmanteau test.

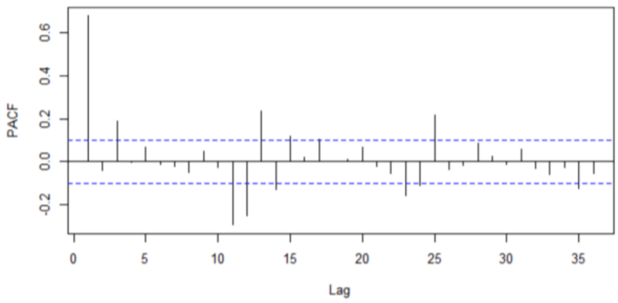
The below plots show the behaviour of the residuals and the accuracy of the model –



Went ahead with plotting the histogram of the residuals and the ACF and PACF plots to understand the autocorrelation and the partial autocorrelation if any.








```
LjungBox(residuals_snm[-c(
lags statistic df p-value
   1  178.6735  1        0
   2  253.7501  2        0
   3  311.2887  3        0
   4  352.2908  4        0
   5  383.0616  5        0
   6  404.4466  6        0
   7  416.2229  7        0
   8  421.0039  8        0
   9  425.1546  9        0
  10  428.0129 10        0
  11  431.6029 11        0
  12  465.3512 12        0
  13  481.4224 13        0
  14  494.6559 14        0
  15  506.0641 15        0
  16  512.4813 16        0
  17  516.4547 17        0
  18  519.2983 18        0
  19  521.0086 19        0
  20  521.1008 20        0
  21  521.8282 21        0
  22  524.4787 22        0
  23  525.9869 23        0
  24  526.8319 24        0
```

```
        Box-Ljung test

data:  residuals_snm
X-squared = 526.83, df = 24, p-value < 2.2e-16
```

From the above Residual Plots visualizations and the Ljung Box tests performed for this model, we can see that the p-values (parameters in the tests) are less than .05 which means that they are significant. So, we can reject the null hypothesis of white noise and there is still something in the residuals that is not been captured by the model.

In terms of accuracy, the following results we obtained. These will be used further as reference for model selection purposes.

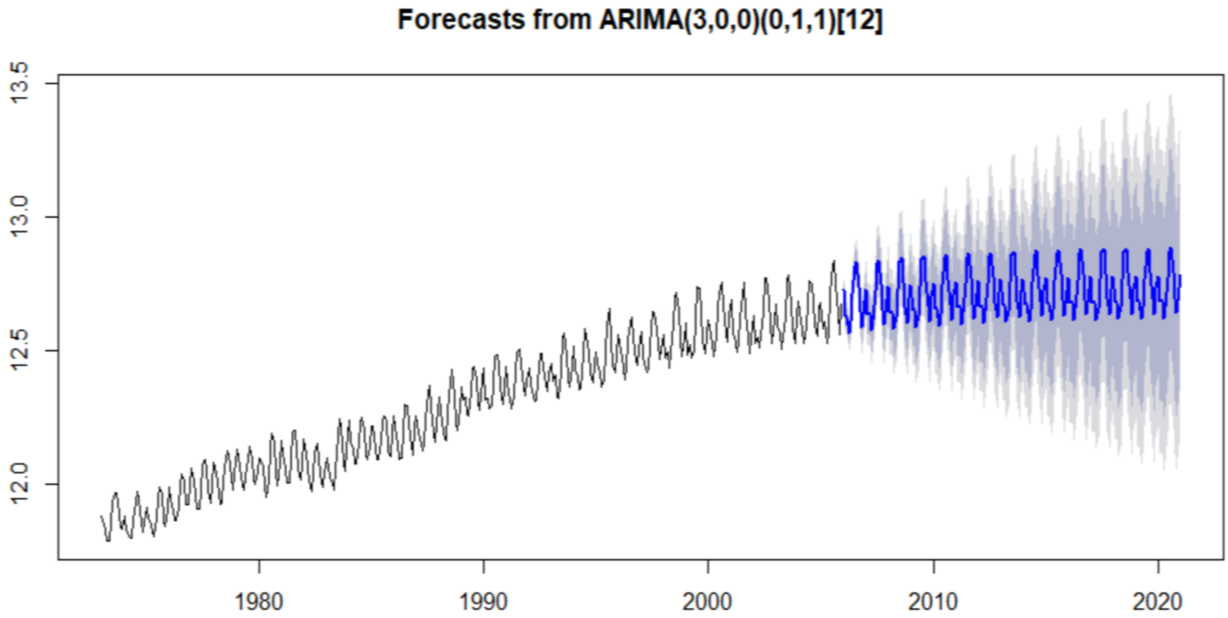
```
> accuracy(snm, test)[,c(2,3,4,5,6,7)]
                   RMSE        MAE      MPE      MAPE      MASE      ACF1
Training set 0.04237624 0.03372629 0.202401 0.2742641 1.0000000 0.6794698
Test set     0.03602182 0.02901740 0.131627 0.2287681 0.8603794 0.5167383
```

Seasonal ARIMA Method

Further, the auto ARIMA method was used to pick up the best model. The result of was **ARIMA (3, 0, 0) (0, 1, 1)[12]** with the following coefficients -

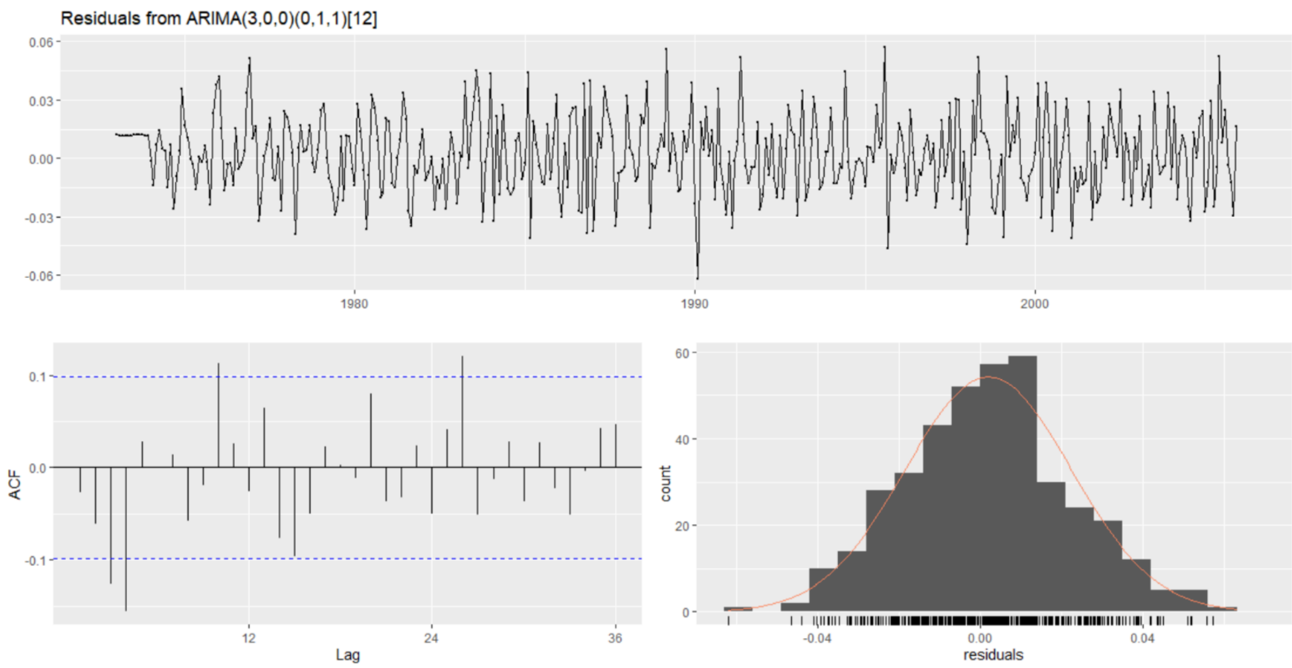
```
Coefficients:
         ar1      ar2     ar3     sma1
      0.8710  -0.1525  0.2644  -0.7255
s.e.  0.0499   0.0661  0.0494   0.0353

sigma^2 estimated as 0.0004389:  log likelihood=938.38
AIC=-1866.76   AICc=-1866.61   BIC=-1847.01
```

The Forecast from the auto ARIMA was as follows –

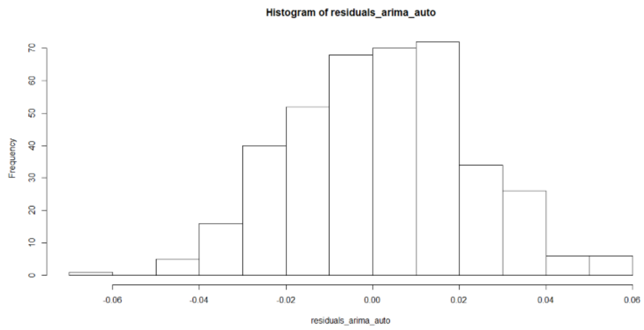


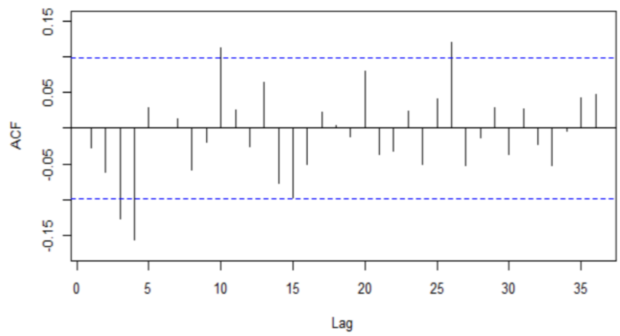
EVALUATION OF MODEL

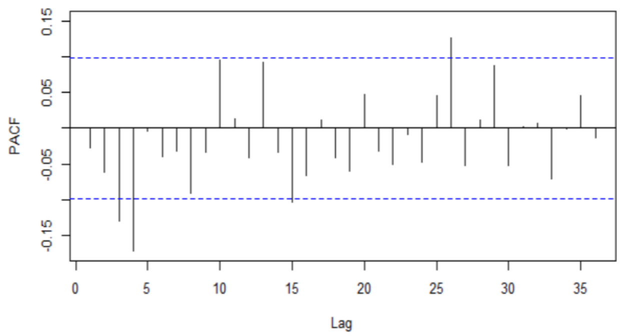
For evaluating the performance of the method used to build the models, I went ahead with checking the Residuals and applying the Ljung–Box test. The below plots show the behaviour of the residuals and the accuracy of the model –



Even went ahead with plotting the histogram of the residuals and the ACF and PACF plots to understand the autocorrelation and the partial autocorrelation if any.








```
        Box-Ljung test

data:  residuals_arima_auto
X-squared = 39.817, df = 24, p-value = 0.02237
```

lags	statistic	df	p-value
1	0.3018367	1	0.582733255
2	1.7942443	2	0.407741387
3	8.1291335	3	0.043416949
4	17.9365136	4	0.001269857
5	18.2568784	5	0.002641110
6	18.2568786	6	0.005621306
7	18.3278765	7	0.010575094
8	19.6783492	8	0.011623932
9	19.8389030	9	0.018933039
10	25.0341312	10	0.005281185
11	25.3064677	11	0.008219783
12	25.5787630	12	0.012306256
13	27.2970835	13	0.011325598
14	29.6849280	14	0.008431393
15	33.5284353	15	0.003964094
16	34.5856724	16	0.004526620
17	34.7856426	17	0.006637006
18	34.7898219	18	0.010044801
19	34.8479865	19	0.014567194
20	37.5008199	20	0.010184116
21	38.0751944	21	0.012629590
22	38.5217836	22	0.015999372
23	38.7569241	23	0.021093758
24	39.8171182	24	0.022373951

From the above Residual Plots visualizations and the Ljung Box tests performed for this model, we can see that the p-values (parameters in the tests) are less than .05 which means that they are significant. So, we can reject the null hypothesis of white noise and there is still something in the residuals that is not been captured by the model.

In terms of accuracy, the following results we obtained –

```
> accuracy(arimaf_auto, test)[,c(2,3,4,5,6,7)]
                   RMSE        MAE         MPE      MAPE      MASE        ACF1
Training set 0.02052207 0.01639897  0.01600656 0.1331904 0.4862369 -0.02750397
Test set     0.04662397 0.03880271 -0.29312656 0.3066585 1.1505183  0.64917552
```

COMPARING BOTH THE MODELS

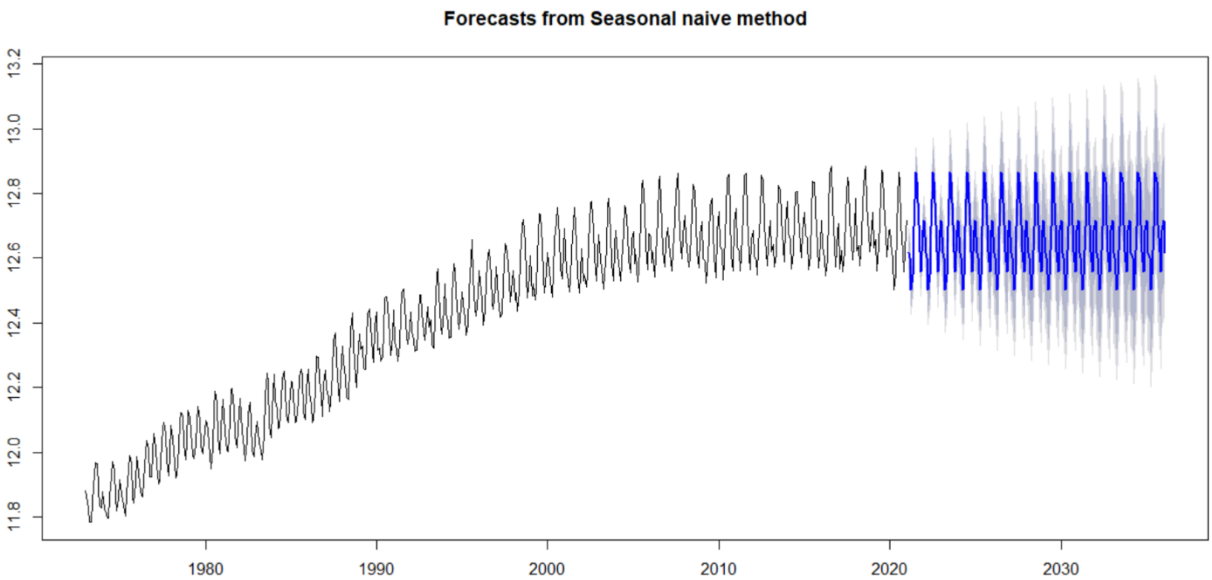
To choose which model performed better, let us compare the results of the final accuracy scores –

KPI	SNAIVE	ARIMA
p-value	2.2E-16	0.02237
RMSE	**0.03602182**	0.04662397
MAE	0.02901740	0.03880271
MPE	0.131627	-0.29312656
MAPE	0.2287681	0.3066585
MASE	0.8603794	1.1505183
ACF1	0.5167383	0.64917552

We can conclude that even though the models used did not perform really well on the residuals test, but looking at the accuracy scores across various KPIs, we can say that the **Seasonal Naïve Method** performed better than the Seasonal ARIMA (3, 0, 0) (0, 1, 1)[12].

FINAL FORECAST

To get the Final Forecast applied the evaluated Best Model (i.e. Seasonal Naïve Method) to the entire dataset and the results were as follows –



FINAL MODEL ACCURACY

RMSE	MAE	MPE	MAPE	MASE	ACF1
0.1432197	0.1025948	-0.0373248	1.5752022	0.8435015	-0.2731486

With the above scores for some of the important KPI's like the MAPE, MAE and RMSE we can say that the Final Model accuracy is pretty good in terms of predictions and can be relied upon.

Thus, from the final forecast as seen from above graph, we can observe and conclude that the trend of Total Electricity usage in the US will remain comparable to the immediate previous years.

REFERENCES

- What Is Time Series Forecasting? (machinelearningmastery.com)
- What is Time Series Forecasting? Everything you Need to Know | Springboard Blog
- Time Series Forecasting Theory | AR, MA, ARMA, ARIMA | Data Science - Bing video
- Time Series in R: Forecasting and Visualisation (robjhyndman.com)
- Time Series Forecasting Using R | Pluralsight
- Perform Time Series Analysis And Forecasting Using R Programming Language - Analytics India Magazine
- The Absolute Best Way to Measure Forecast Accuracy | Axsium (axsiumgroup.com)
- Ljung-Box Test: Definition + Example - Statology
- Total Energy Monthly Data - U.S. Energy Information Administration (EIA)
- Significance of ACF and PACF Plots In Time Series Analysis | by Jayesh Salvi | Towards Data Science
- Time series forecasting: from naive to ARIMA and beyond | by Mahbubul Alam | Towards Data Science
- Forecast KPIs: RMSE, MAE, MAPE & Bias | by Nicolas Vandeput | Towards Data Science
- 3.3 Residual diagnostics | Forecasting: Principles and Practice (2nd ed) (otexts.com)
- Ljung–Box test - Wikipedia
- Ljung-Box Q-test for residual autocorrelation - MATLAB lbqtest (mathworks.com)
- 7.7 Forecasting with ETS models | Forecasting: Principles and Practice (2nd ed) (otexts.com)