2022

Forecasting Individual Assignment

STUDENT – NITHESH RAMANNA

Contents

TASK 1 2
DATA EXPLORATION 2
BOX-COX TEST 5
SEASONAL NAIVE METHOD 5
STL DECOMPOSITION 7
EXPONENTIAL SMOOTHING 9
SEASONAL ARIMA METHOD 12
SELECTION OF BEST MODEL 14
INFERENCE FROM THE FORECAST 15
TASK 2 16
DATA EXPLORATION 16
BOX-COX TEST 17
STL DECOMPOSITION 17
EXPONENTIAL SMOOTING METHOD 19
ARIMA MODELS 19
FORECAST USING THE BEST MODEL 21
REFERENCES 22

TASK 1

For task one, the given dataset Airpass_BE is used. The data set Airpass_BE contains international intra-EU air passenger transport by Belgium and EU partner countries, from January 2003 to October 2021.

The time series is split into 3, from January 2003 to December 2017 as train set, from January 2018 to February 2020 as test and rest of the observations are used for later reference.

DATA EXPLORATION

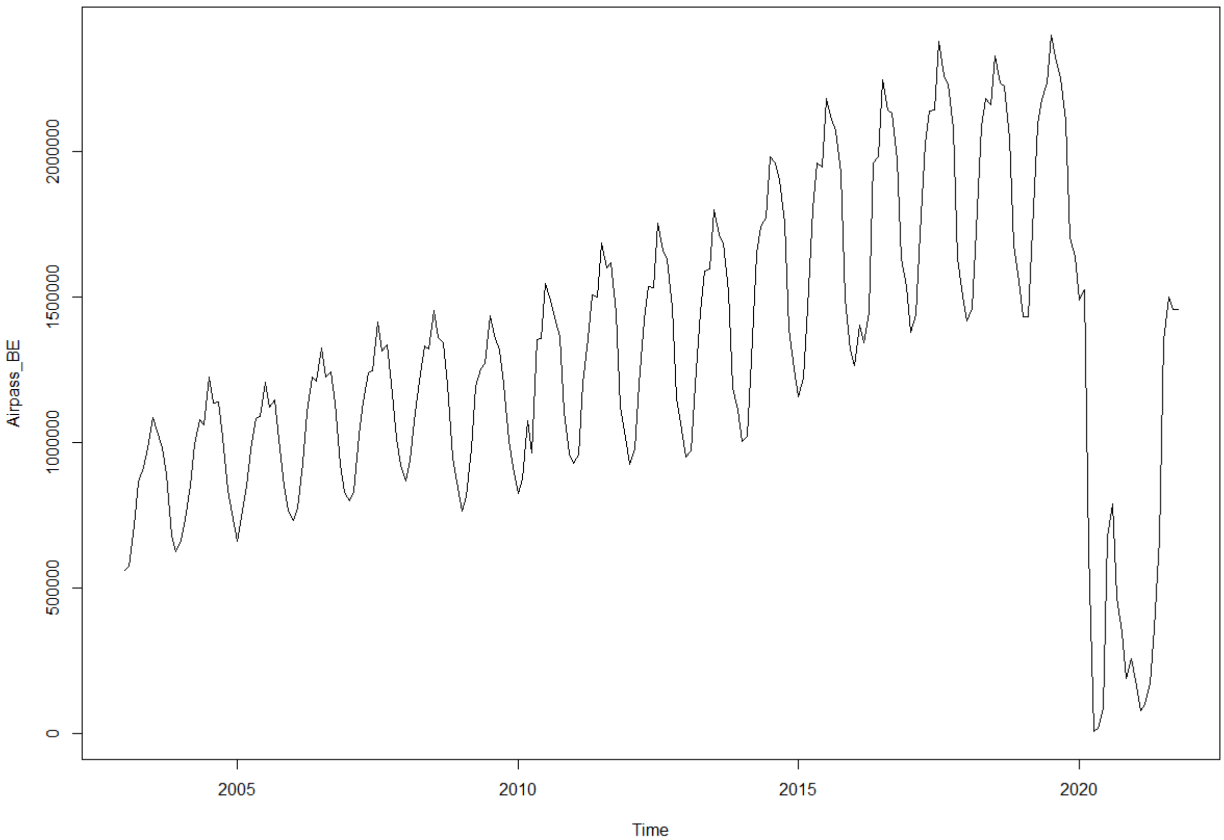


Figure 1

From the figure 1, we can clearly see that an increasing trend, but there is sudden down fall in the time series somewhere in the year 2020.So Sub-setting the data till Feb 2020 and plotting again as we are using the that part of the data for the future reference.

From figure 2, again we can see that there is an increasing trend. It can also be seen that the seasonality variation increases slightly as the level of the series increases.

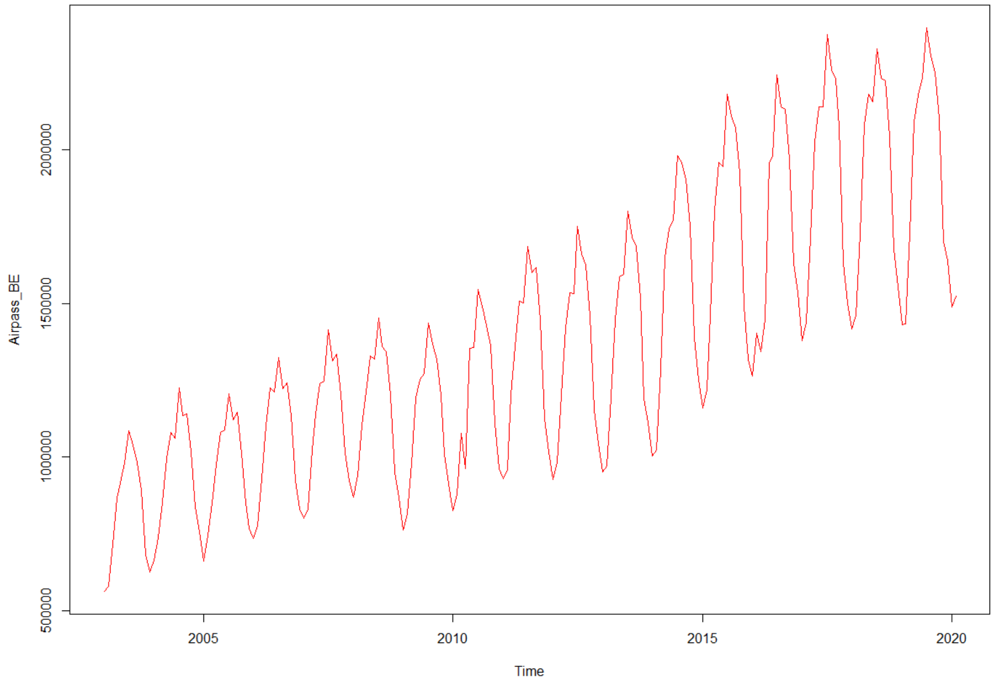


Figure 2

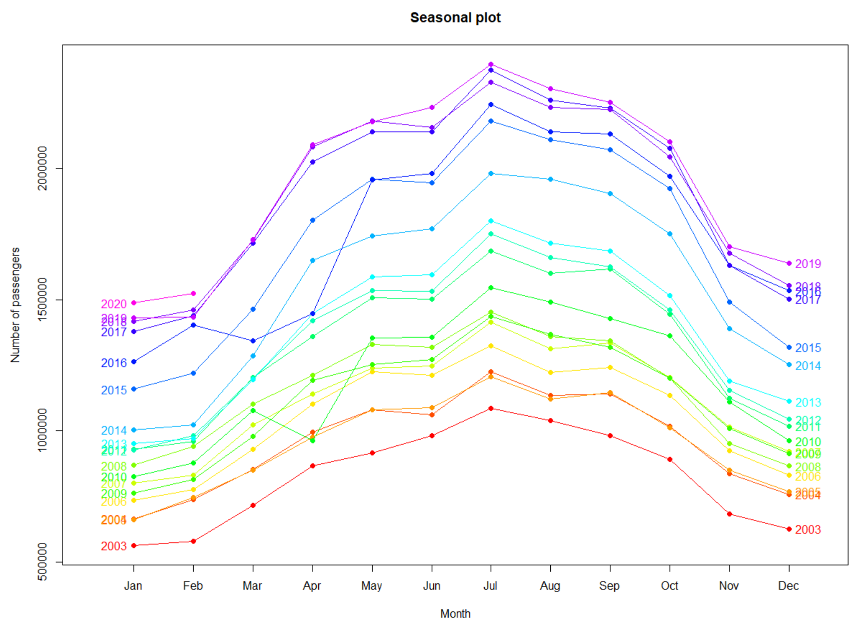


Figure 3

Figure 3 is the seasonal plot. Data plotted against the individual "seasons" in which the data were observed. For the given data, a "season" is a month. From the figure we can see all the curves look similar over each season i.e., the seasonal pattern can be seen clearly. We can also see that number of passengers over each season are increasing year after year.

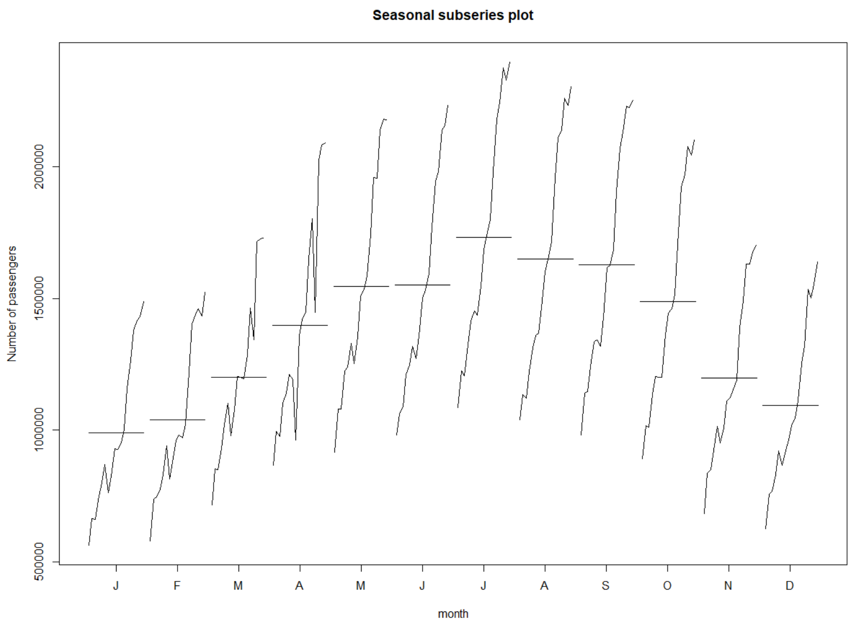


Figure 4

Figure 4 is the subseries plot, even with this plot we can see the increasing trend i.e., the number of passengers is increasing year after year.

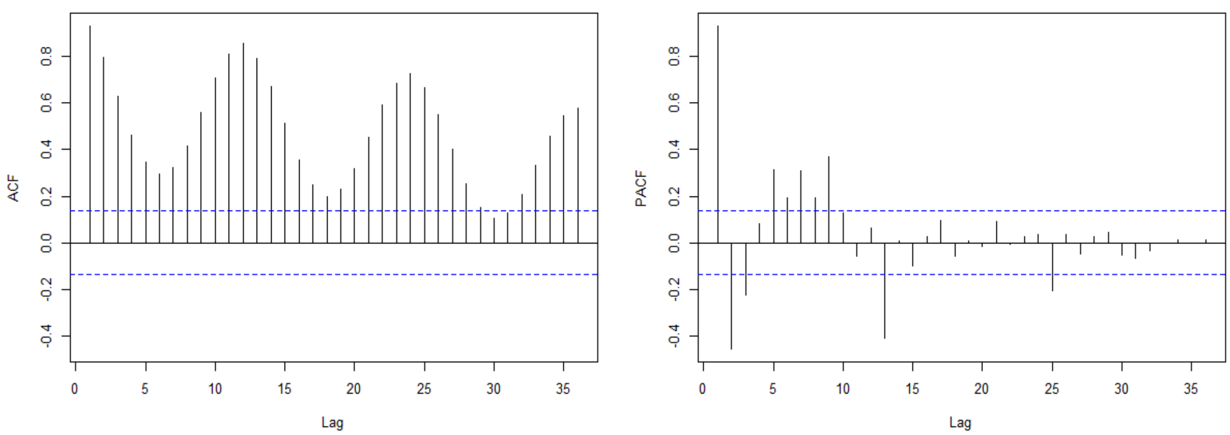


Figure 5

Figure 5 is the ACF and PACF plots.

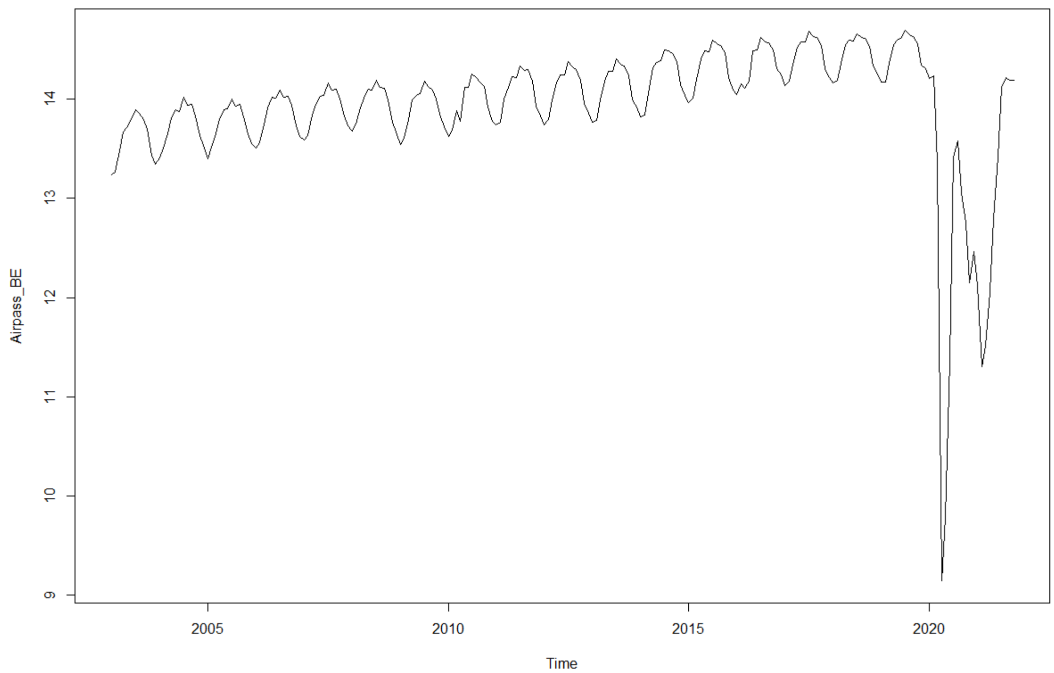
ACF is an (complete) auto-correlation function which gives us values of autocorrelation of any series with its lagged values. In simple terms, it describes how well the present value of the series is related with its past values.

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.

Figure 5 shows that there is a strong autocorrelation because we can see many high values in the ACF graph and even a strong component of seasonality. By looking at the PACF plot we can say that there is not much noise component.

BOX-COX TEST

As we saw that the increase in seasonality variation over the years from the time series plot in Figure 1, I feel like to take the log transformation is the good idea. The below figure shows the log transformed time series where we can see there is a less variation in seasonality.



To get it confirmed Box-Cox lambda test is conducted. The test gave the 0.01461 as optimal lambda value. Since the optimal lambda value is close to 0, I am using lambda value as 0 in the rest of the modeling process.

SEASONAL NAIVE METHOD

Create forecasts using the seasonal naive method. Check the residual diagnostics (including the Ljung-Box test) and the forecast accuracy (on the test set).

Seasonal Naive Method

The train set is used to fit the model and evaluation of performance of the model is conducted on both train and test set. The figure 6 shown below is the forecasts from the seasonal naive method.

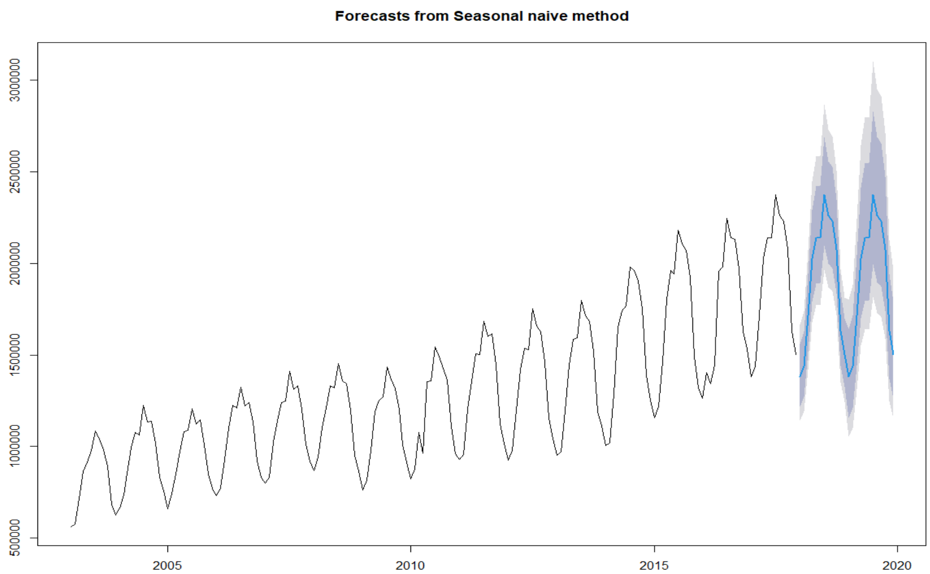


Figure 6

Ljung-Box test: The Ljung–Box test (named for Greta M. Ljung and George E. P. Box) 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 number of lags and is therefore a portmanteau test.

Residuals: Residuals are useful in checking whether a model has adequately captured the information in the data. A good forecasting method will yield residuals with no correlation and zero mean.

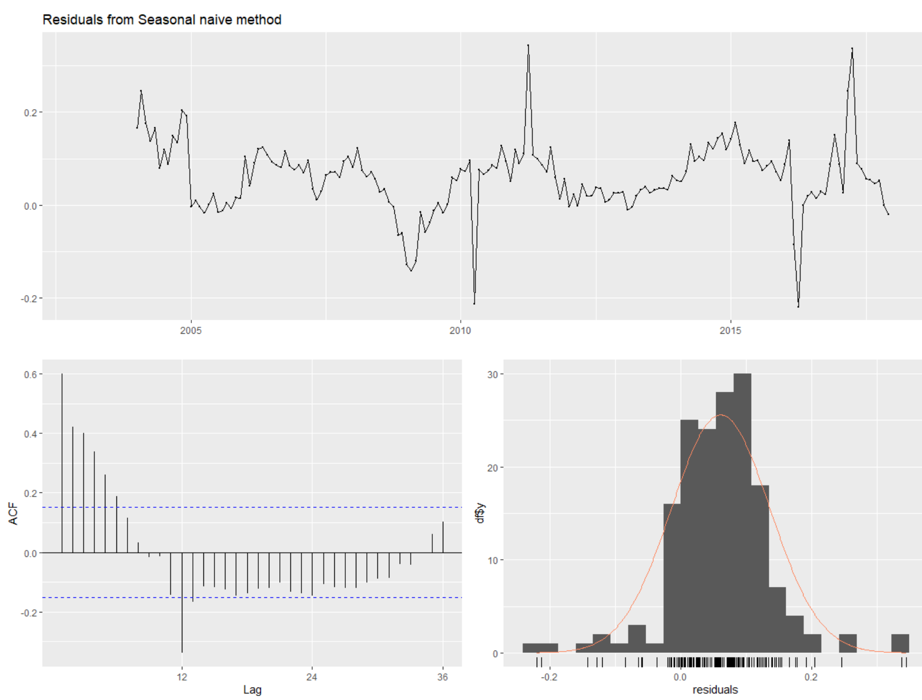


Figure 7

```
        Ljung-Box test

data:  Residuals from Seasonal naive method
Q* = 225.73, df = 24, p-value < 2.2e-16

Model df: 0.   Total lags used: 24
```

Figure 8

Figure 7 shows the Residual from the seasonal naive method and figure 8 is the screen shot of Ljung-Box test. From the figures we can see that the p-value is less than 0.05, so we can reject the null hypothesis of white noise which means there are residuals that has not been captured by the model.

Figure 9 gives the accuracy with the measures RMSE, MAE, MAPE and MASE.

```
> accuracy(air_m1,test)[,c(2,3,5,6)]
                  RMSE      MAE     MAPE      MASE
Training set 124232.13 95011.47 7.225186 1.0000000
Test set      55167.45 45156.54 2.518184 0.4752746
```

Figure 9

STL DECOMPOSITION

Use an STL decomposition to forecast the time series. Use the various underlying forecasting methods for the seasonally adjusted data (naive, rwdrift, ets, arima). Check the residual diagnostics and the forecast accuracy and select the best performing STL decomposition.

```
                   nr      RMSE       MAE     MAPE      MASE
STL naive lambda    1  59037.29  34250.34 2.739155 0.3604864
STL rwdrift lambda  2  58765.18  34077.38 2.736876 0.3586660
STL ets lambda      3  54557.10  34490.72 2.762746 0.3630164
STL arima lambda    4  52726.11  33295.08 2.701382 0.3504322
> a_test
                   nr      RMSE       MAE     MAPE      MASE
STL naive lambda    1  70269.29  56938.81 2.850634 0.5992836
STL rwdrift lambda  2 196051.25 167543.67 8.430373 1.7634046
STL ets lambda      3 165625.23 139695.92 6.975942 1.4703059
STL arima lambda    4 166868.02 141480.70 7.060734 1.4890907
> round(res_matrix, digits = 4)
                   nr      Q* df p-value
STL naive lambda    1 28.2128 24  0.2512
STL rwdrift lambda  2 28.2128 23  0.2078
STL ets lambda      3 15.4375 20  0.7509
STL arima lambda    4 10.9429 20  0.9477
```

Figure 10

Figure 10 shows the results of STL decomposition to forecast the time series using the forecasting methods, naive, rwdrift, ets and arima. Looking at the figure, one can say that STL with naive method is giving the good results. But when I checked the residual diagnostics, it had a residual at lag 1, so I choose STL with ETS method.

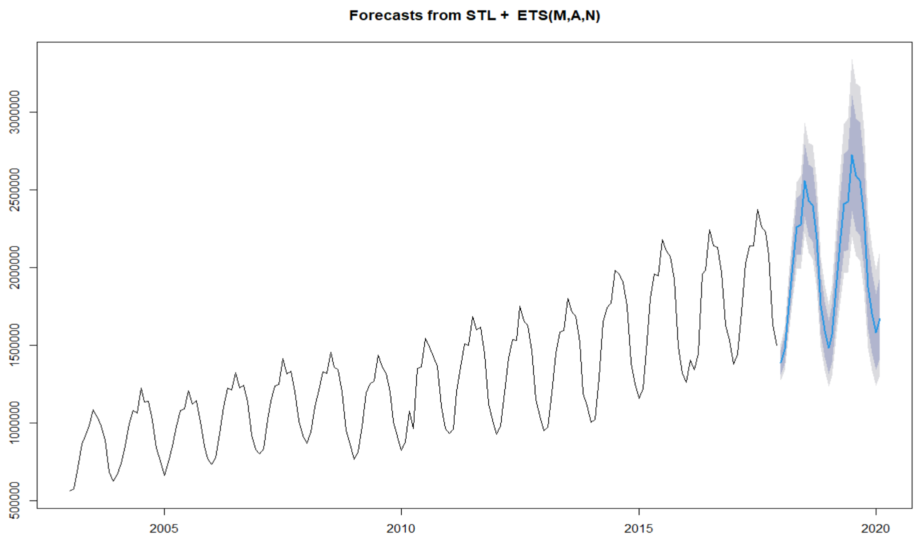


Figure 11

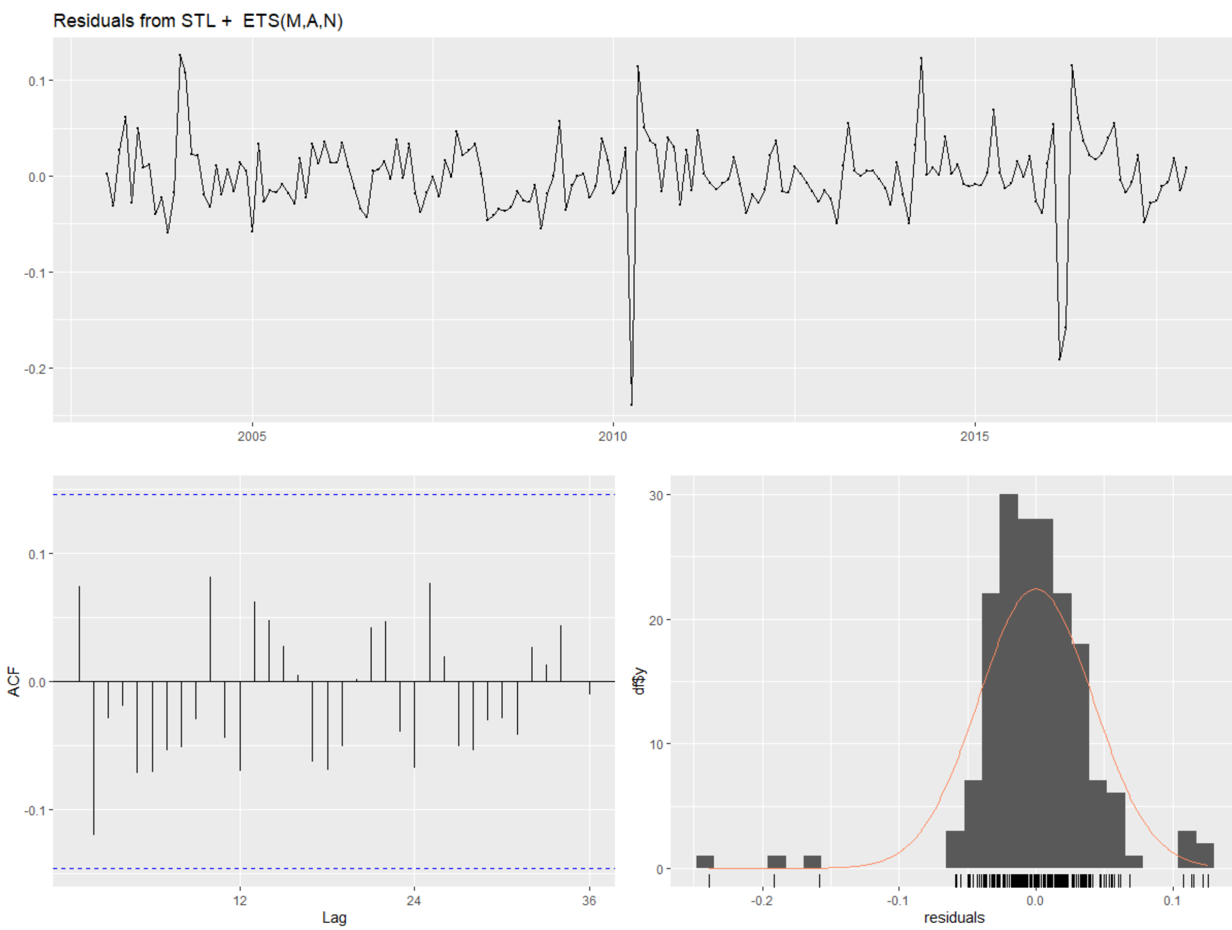


Figure 12

Figure 11 shows the forecasts using STL with ETS method. Figure 12 shows the residual plots for the model. From the figures we can see that the p-value is greater than 0.05 for the model, so we can accept the null hypothesis of white noise which means there are no residuals or model captured all data.

EXPONENTIAL SMOOTHING

For exponential smoothing, I have used 5 different methods and one auto ETS procedure. From the figure 13 we can see that ETS with Additive Error, Additional Trend and Additional Seasonality, and with damped true and lambda 0 and auto ETS procedure are giving the best results. For further investigation, I checked the residual diagnostics and Ljung-Box test.

Figure 14 and Figure 15 shown below are the residual plots and Ljung-Box test respectively, where we can see that the p-value is very close to 0.05 for the model and in the ACF plot we can see no residuals, so we can accept the null hypothesis of white noise which means there are no residuals or model captured all data.

```
> a_train
            nr     RMSE      MAE     MAPE      MASE
AAdA         1 65752.56 43419.15 3.636354 0.4569885
MAdA         2 65459.53 43085.22 3.583580 0.4534739
MAdM         3 55242.84 35047.47 2.778160 0.3688762
AAA lambda   4 59609.50 36333.46 2.909500 0.3824114
AAdA lambda  5 55655.79 34808.48 2.748344 0.3663608
Auto ETS     6 55655.79 34808.48 2.748344 0.3663608
> a_test
            nr      RMSE       MAE     MAPE      MASE
AAdA         1 168077.03 142881.88 6.914275 1.5038383
MAdA         2 168084.79 142833.78 6.910386 1.5033320
MAdM         3  69085.62  56001.95 2.812480 0.5894231
AAA lambda   4 195142.17 161405.00 7.946873 1.6987949
AAdA lambda  5  66708.89  53878.57 2.750050 0.5670744
Auto ETS     6  66708.89  53878.57 2.750050 0.5670744
> round(res_matrix, digits = 4)
            nr      Q* df p-value
AAdA         1 86.3902  7  0.0000
MAdA         2 98.9526  7  0.0000
MAdM         3 16.1686  7  0.0236
AAA lambda   4 18.1259  8  0.0203
AAdA lambda  5 14.2157  7  0.0475
Auto ETS     6 14.2157  7  0.0475
```

Figure 13

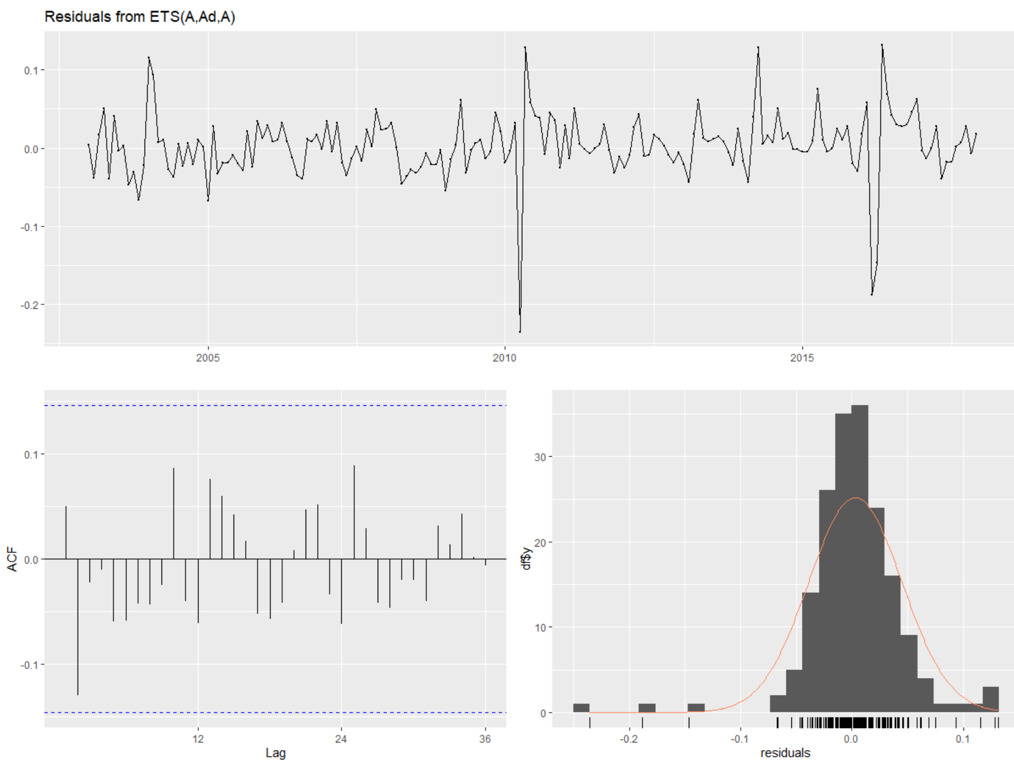


Figure 14

```
        Ljung-Box test

data:  Residuals from ETS(A,Ad,A)
Q* = 14.216, df = 7, p-value = 0.04747

Model df: 17.   Total lags used: 24
```

Figure 15

Since auto ETS procedure got the same accuracy, I am looking at the residuals of that model as well. Figures 16 and 17 show the residual plots and Ljung-Box test respectively. From the figures we can see that the auto ETS procedure select the same model as before as the best. So, we are the ETS model with Additive Error, Additional Trend and Additional Seasonality, and with damped true and lambda 0.

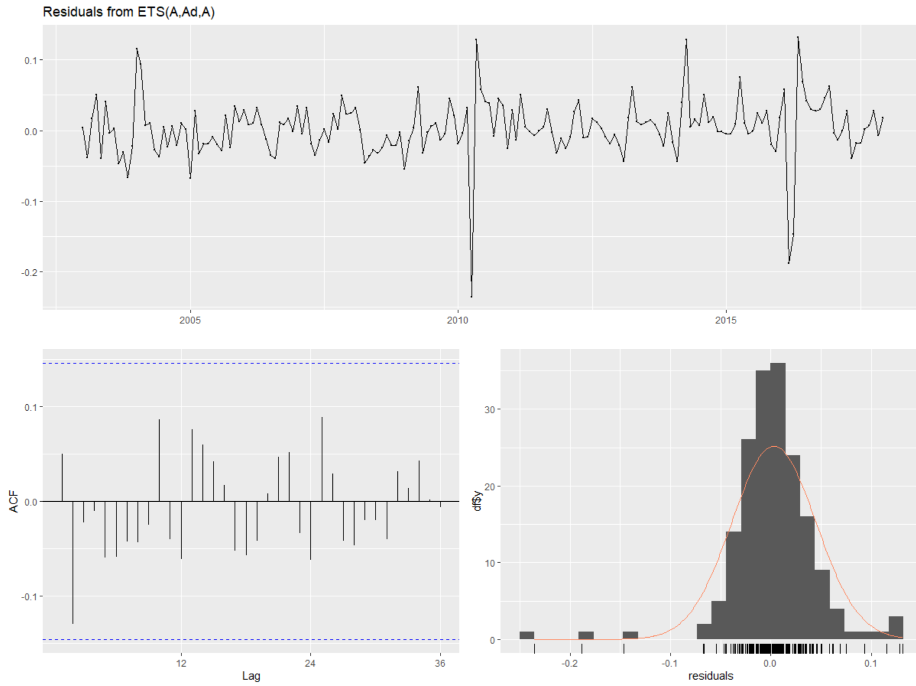


Figure 16

```
        Ljung-Box test

data:  Residuals from ETS(A,Ad,A)
Q* = 14.216, df = 7, p-value = 0.04747

Model df: 17.   Total lags used: 24
```

Figure 17

Figure 18 shows the forecast plot for ETS(AAdA).

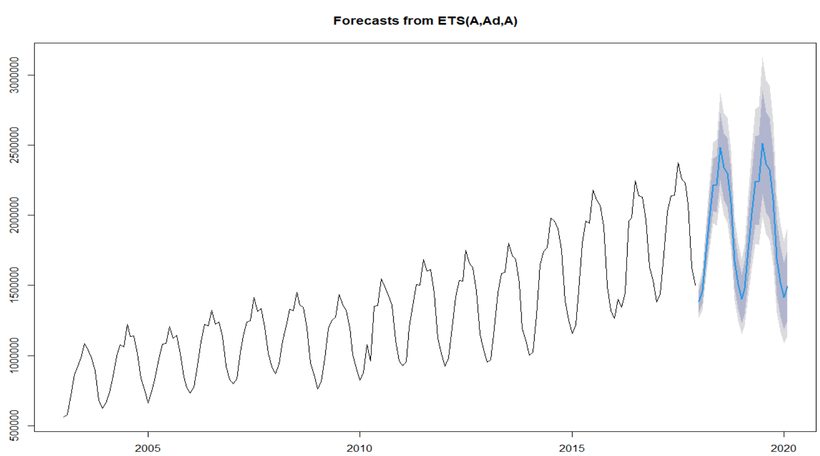


Figure 18

SEASONAL ARIMA METHOD

Later, Auto ARIMA method was used to find the forecasts. I chose 2 methods with different parameters for auto ARIMA method. They are as follows

1. stepwise = FALSE, approximation = FALSE, lambda = l, biasadj = TRUE
2. stepwise = FALSE, approximation = FALSE

The results of the both the models are shown in figure 19. When compared to accuracy results of both the models, the second model performed well. So, further investigation has conducted by looking at the residual diagnostics and Ljung-Box test for the second model.

```
> accuracy(air_m18, test)[,c(2,3,5,6)]
                  RMSE       MAE     MAPE      MASE
Training set  60317.75  38219.13 3.029713 0.4022581
Test set     164323.70 144693.79 7.389589 1.5229087
> accuracy(air_m17, test)[,c(2,3,5,6)]
                 RMSE       MAE     MAPE      MASE
Training set  58863.7  36708.87 2.874457 0.3863626
Test set     145139.1 127891.13 6.730419 1.3460599
```

Figure 19

The result of was ARIMA (1, 1, 1) (1, 1, 2) [12] with the following coefficients show in figure 20.

```
Coefficients:
         ar1      ma1    sar1     sma1    sma2
      0.5356  -0.8744  0.8507  -1.6106  0.7158
s.e.  0.1358   0.0882  0.3492   0.3834  0.2314

sigma^2 = 3.85e+09:  log likelihood = -2083.93
AIC=4179.86   AICc=4180.38   BIC=4198.56
```

Figure 20

Figure 21 shows the forecast from the model.

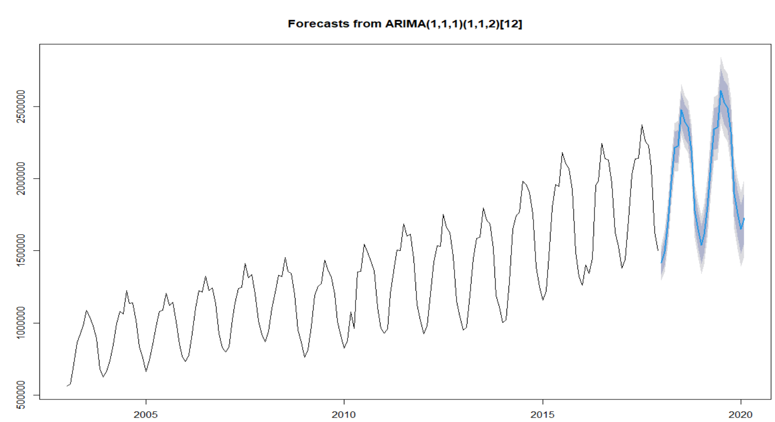


Figure 21

Figures 22 and 23 are Ljung-Box test results and Residual diagnostic plot. From the figures we can see that the p-value is greater than 0.05 for the model, so we can accept the null hypothesis of white noise which means there are no residuals or model captured all data.

```
        Ljung-Box test

data:  Residuals from ARIMA(1,1,1)(1,1,2)[12]
Q* = 18.044, df = 19, p-value = 0.5195

Model df: 5.   Total lags used: 24
```

Figure 22

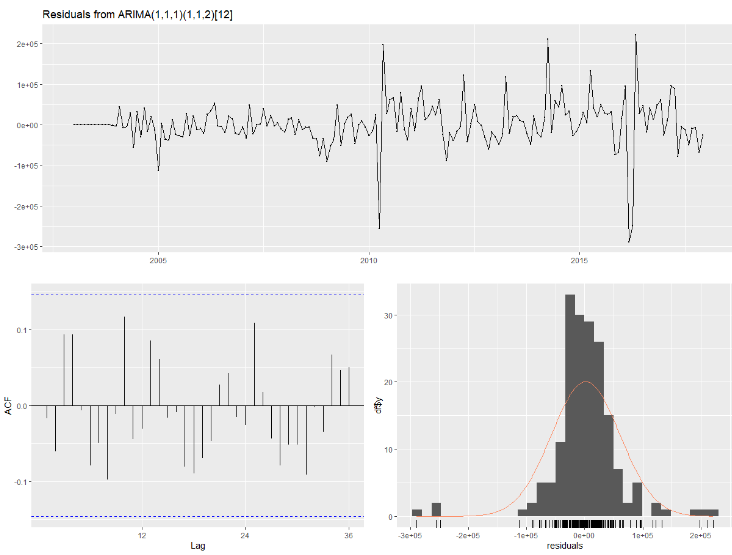
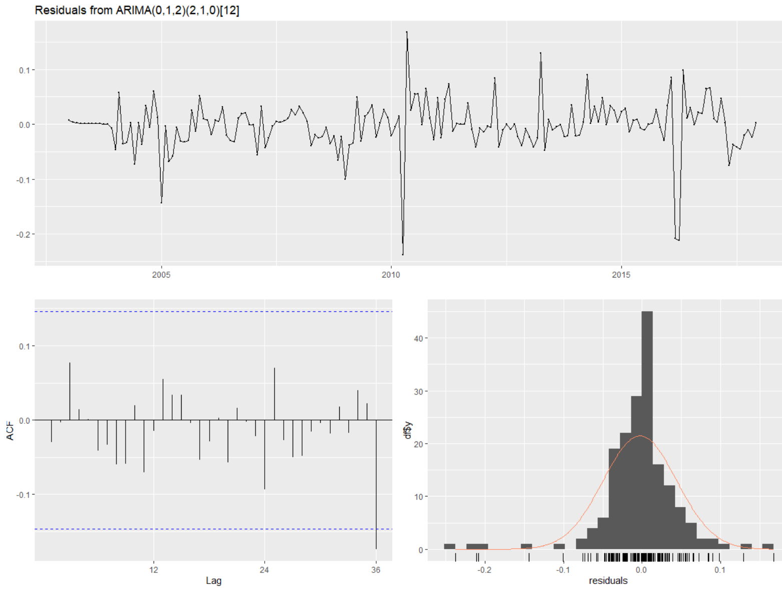


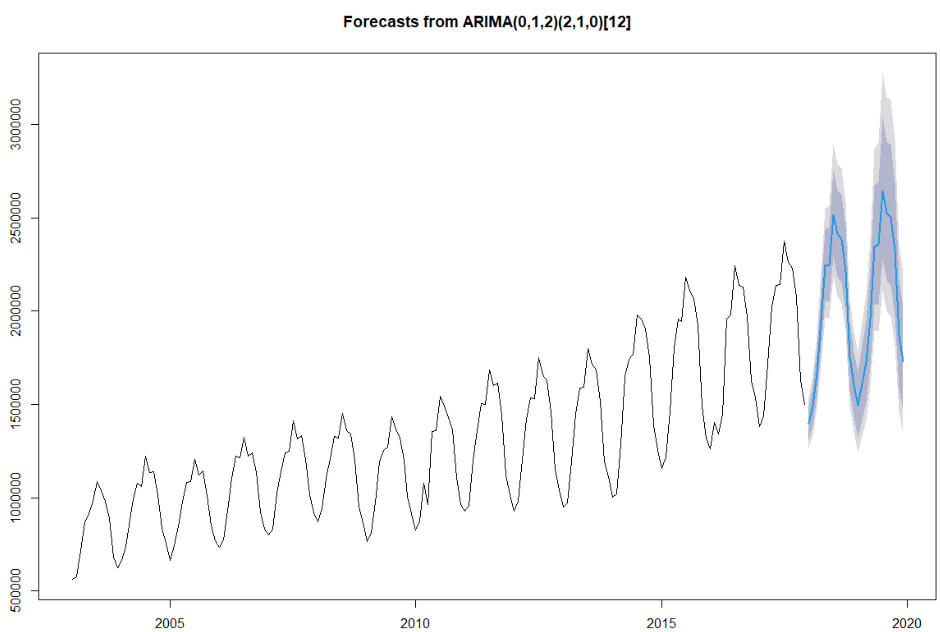
Figure 23

I also used Arima model with backward shift, by giving the values D = 1 and d = 1 with lambda 0 for the above Arima models. Below figure show the accuracy and Ljung-Box test and Residual Diagnostics for the model.

```
> accuracy(air_m20, test)[,c(2,3,5,6)]
                  RMSE       MAE     MAPE      MASE
Training set  62975.74  37881.59 2.963114 0.3987054
Test set     147749.69 129024.89 6.408476 1.3579928
> checkresiduals(air_m20)

        Ljung-Box test

data:  Residuals from ARIMA(0,1,2)(2,1,0)[12]
Q* = 8.7155, df = 20, p-value = 0.986

Model df: 4.   Total lags used: 24
```



From the figures we can see that the p-value is greater than 0.05 for the model, so we can accept the null hypothesis of white noise which means model captured all data. The performed like 2nd Arima model mentioned above. But by looking at the ACF plot we can see that there is a spike down the confidence interval which is not in case of 2nd model, so choosing the 2nd Arima model in this case. Below is the forecast plot for Arima with backward shift.



SELECTION OF BEST MODEL

Out of all the methods selected in previous steps, best model is selected out of them by looking at their MASE and checking their residual diagnostics. The best model selected in our case is ETS(AAdA) with lambda 0 or auto ETS procedure as the MASE of the models for test is 0.57 and p-value is 0.0475 which is closer to 0.05 which is acceptable and there were no residuals for the model which can be seen in figure

14. Figure 24 shows the summary of all the model selected. The highlighted ones are best ones in this case.

```
> a_train
                     nr      RMSE      MAE     MAPE      MASE
Seasonal Naive method  1 124232.13 95011.47 7.225186 1.0000000
STL ets lambda         2  54557.10 34490.72 2.762746 0.3630164
AAdA                   3  55655.79 34808.48 2.748344 0.3663608
Auto ETS               4  55655.79 34808.48 2.748344 0.3663608
ARIMA                  5  58863.70 36708.87 2.874457 0.3863626
> a_test
                     nr      RMSE       MAE     MAPE      MASE
Seasonal Naive method  1  55167.45  45156.54 2.518184 0.4752746
STL ets lambda         2 165625.23 139695.92 6.975942 1.4703059
AAdA lambda            3  66708.89  53878.57 2.750050 0.5670744
Auto ETS               4  66708.89  53878.57 2.750050 0.5670744
ARIMA                  5 145139.11 127891.13 6.730419 1.3460599
> round(res_matrix, digits = 4)
                     nr       Q* df p-value
Seasonal Naive method  1 225.7336 24  0.0000
STL ets lambda         2  15.4375 20  0.7509
AAdA lambda            3  14.2157  7  0.0475
Auto ETS               4  14.2157  7  0.0475
ARIMA                  5  18.0440 19  0.5195
```

Figure 24

INFERENCE FROM THE FORECAST

By looking at the figure 25 we can see that the forecasted values are in line with the previous seasonality and trend but because of COVID there is big difference in number of passengers from March 2020 when compared the actual numbers with the forecasted numbers which was unexpected.

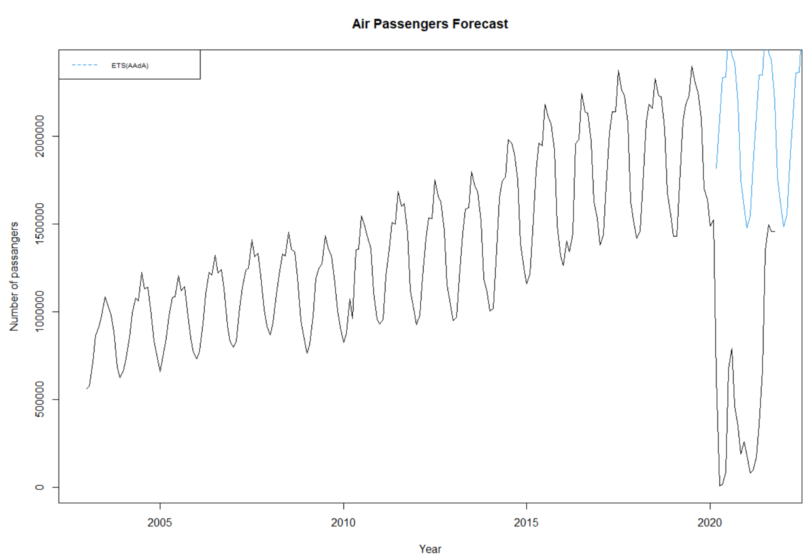


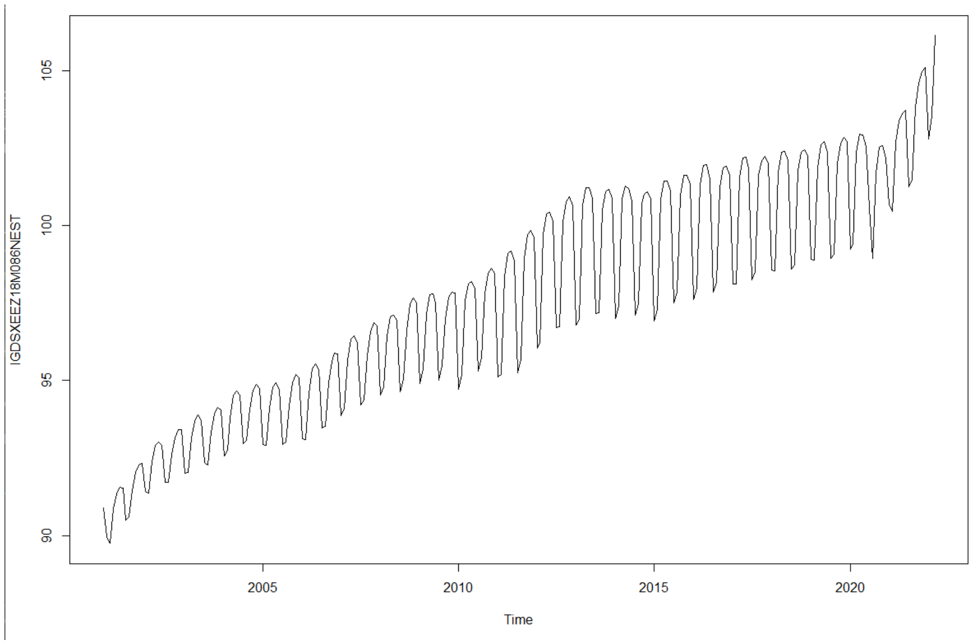
Figure 25

TASK 2

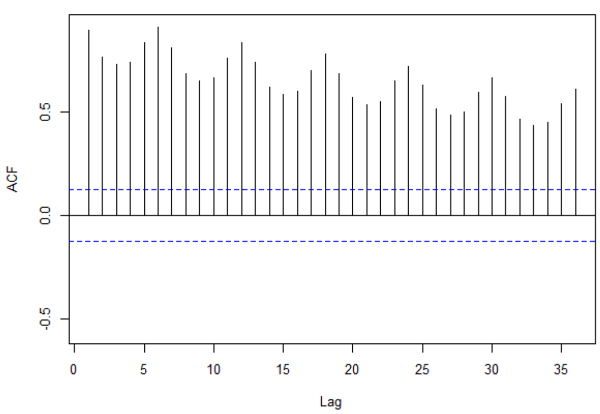
For task two, I have chosen the Harmonized Index of Consumer Prices: Non-Energy Industrial Goods for Euro Area dataset from https://fred.stlouisfed.org/series/IGDSXEEZ18M086NEST. The data set has the monthly frequency, and it starts from December 2000 and ends in March 2022.

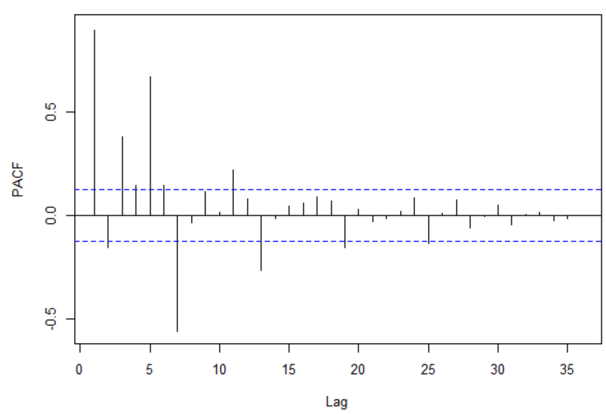
The dataset is split into 3 parts, from December 2000 to December 2017 as trainset, from January 2018 to December 2019 as test set, the rest of the data is used for further reference.

DATA EXPLORATION



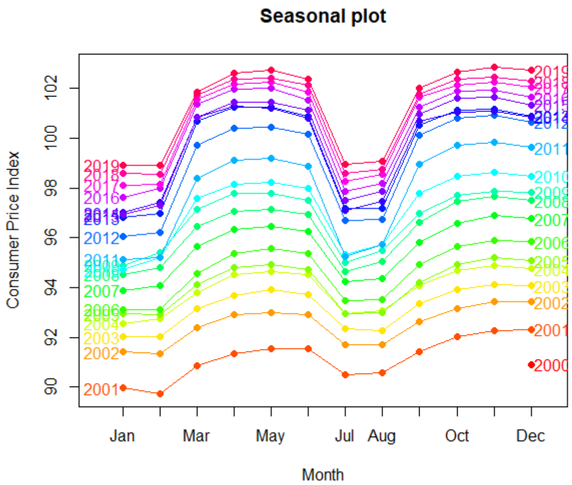
From the above figure we can see that there is an increasing trend in the chosen dataset. To check that in detail I plotted the ACF and PACF plots.

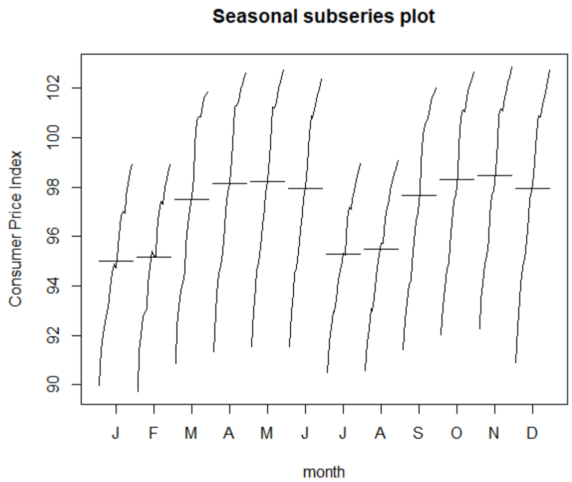




From the above we can see that there is a strong autocorrelation because we can see many high values in the ACF graph and even a strong component of seasonality. By looking at the PACF plot we can say that there is not much noise component.

I have also plotted the seasonal and season subseries plots which will shows the show the seasonality and trends. The below figure shows the seasonal plots and seasonal subseries plot, which strongly says that the data is seasonal, and the consumer price index is increasing over years.





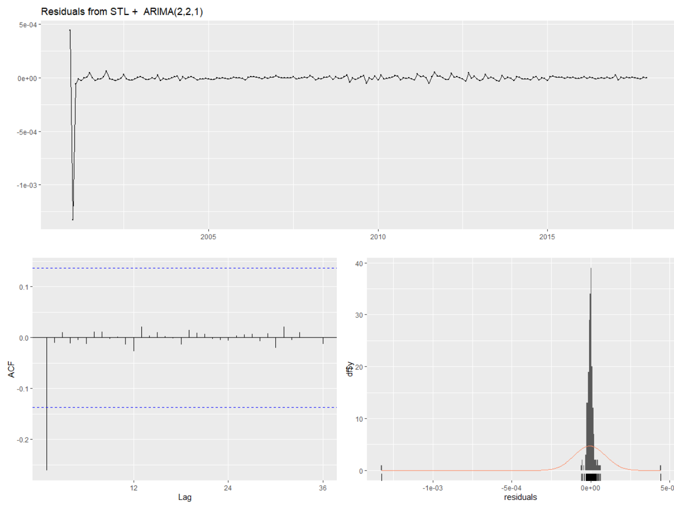
BOX-COX TEST

To check the optimal lambda value for the transformation, I have conducted the Box-Cox test which gave me the optimal value of -0.9999 which very close to -1, so I am choosing the lambda value as -1 for transformation which is inverse plus 1.

STL DECOMPOSITION

STL Decomposition is conducted with the naive, rwdrift, ETS and Arima methods with lambda, -1. The below figure shows the results of the STL decomposition with different methods. From the figure we can see that the STL decomposition with Arima method is doing better than any other methods. So, further the residual diagnostic and Ljuag-Box test is conducted to check the model.

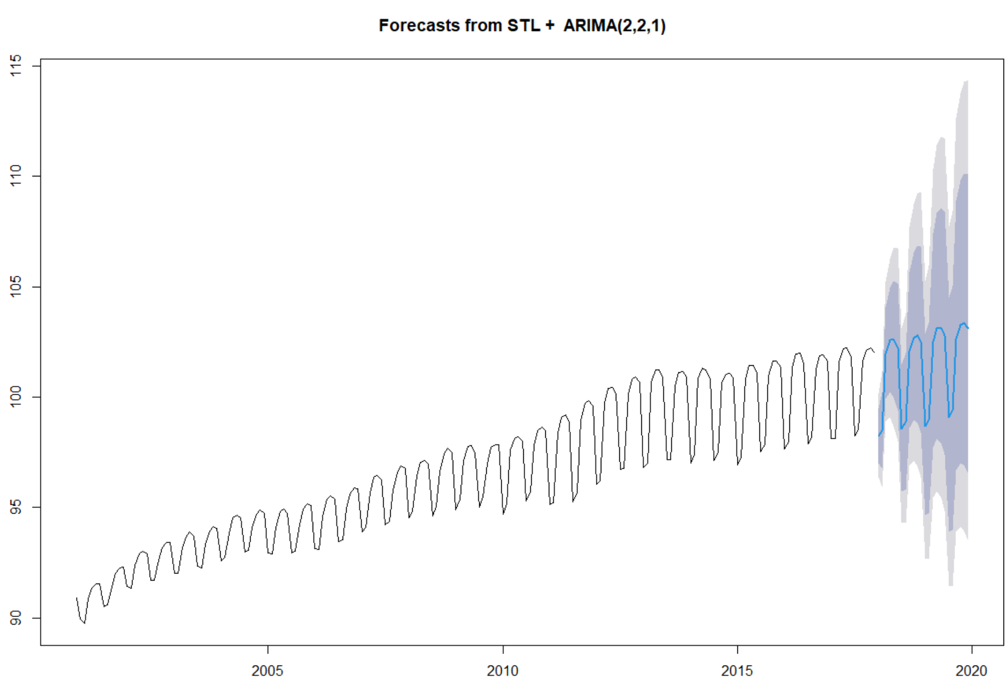
```
> a_train
                   nr      RMSE       MAE      MAPE      MASE
STL naive lambda    1 0.1654262 0.1136403 0.1182641 0.1854606
STL rwdrift lambda  2 0.1575242 0.1059376 0.1102487 0.1728899
STL ets lambda      3 0.1562669 0.1072796 0.1116404 0.1750800
STL arima lambda    4 0.8989584 0.1874707 0.2003338 0.3059517
> a_test
                   nr      RMSE       MAE      MAPE      MASE
STL naive lambda    1 0.3552371 0.2869632 0.2848329 0.4683232
STL rwdrift lambda  2 0.5582130 0.4856250 0.4772949 0.7925387
STL ets lambda      3 0.5508804 0.4792211 0.4709979 0.7820876
STL arima lambda    4 0.3517772 0.3047298 0.2995426 0.4973182
> round(res_matrix, digits = 4)
                   nr       Q* df p-value
STL naive lambda    1 197.2440 24  0.0000
STL rwdrift lambda  2 197.2440 23  0.0000
STL ets lambda      3 212.2387 20  0.0000
STL arima lambda    4  14.7982 21  0.8329
```




```
        Ljung-Box test

data:  Residuals from STL +  ARIMA(2,2,1)
Q* = 14.798, df = 21, p-value = 0.8329

Model df: 3.   Total lags used: 24
```

The above two figures show the residual plots and Ljung-Box test respectively. From the figure we can see that the p-value is greater than 0.05 for the model, so we can accept the null hypothesis of the white noise, that the model captured all the data. There is only one down spike in the ACF plot, but it is fine as it is in first lag, so we accept the model, STL decomposition with Arima method.

The below plot mention shows the forecast for the chosen model.



EXPONENTIAL SMOOTING METHOD

Since Holt-Winter's capture both the seasonality and trend, I went with Holt-Winter's method with damped and without damped and lambda 0.

The below figures the accuracy and Ljung-Box test for the Holt-Winter's method with lambda 1 respectively. Though the MASE here is better than the STL decomposition with Arima, the p-value is less than 0.05, which means we must reject the null hypothesis of white noise. So, I am rejecting this model.

```
                  RMSE       MAE      MAPE      MASE
Training set 0.2904222 0.2092963 0.2185066 0.3415711
Test set     0.2304878 0.1981920 0.1949566 0.3234488
```

```
        Ljung-Box test

data:  Residuals from Holt-Winters' additive method
Q* = 221.02, df = 8, p-value < 2.2e-16
```

Further, I have checked the accuracy and Ljung-Box test for the Damped Holt-Winter's method with lambda 1 respectively. Below figures show the same. Here also it is same that the MASE here is better than the STL decomposition with Arima, the p-value is less than 0.05, which means we must reject the null hypothesis of white noise. So, I am rejecting this model.

```
                  RMSE       MAE      MAPE      MASE
Training set 0.2328155 0.1755676 0.1822255 0.2865259
Test set     0.1745666 0.1323698 0.1317824 0.2160271
```

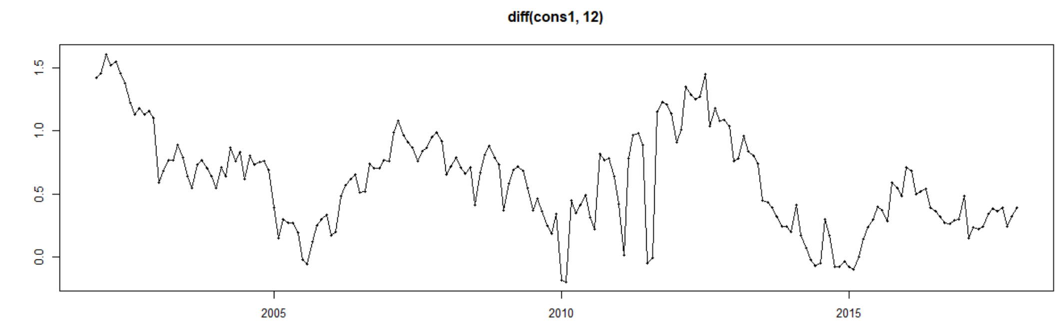
```
        Ljung-Box test

data:  Residuals from Damped Holt-Winters' additive method
Q* = 176.43, df = 7, p-value < 2.2e-16

Model df: 17.   Total lags used: 24
```

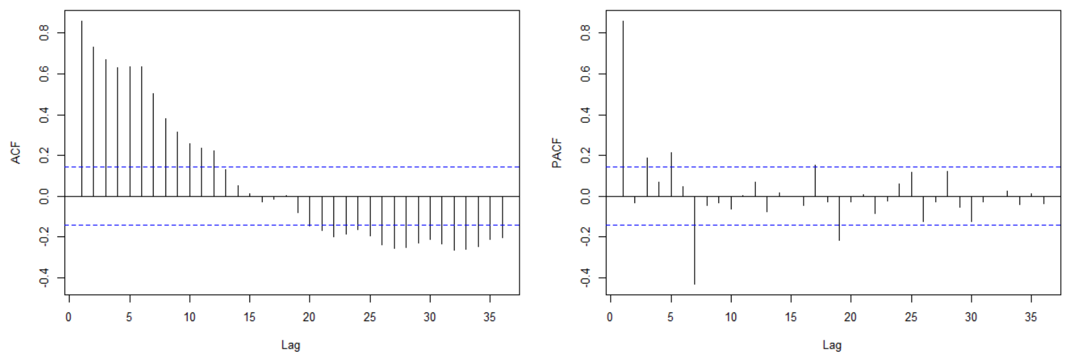
ARIMA MODELS

I have use set of parameters for Arima model, they are shown below

order = c (2,1,0), seasonal = c (2,1,1), include.drift = TRUE, lambda = l, biasadj = TRUE

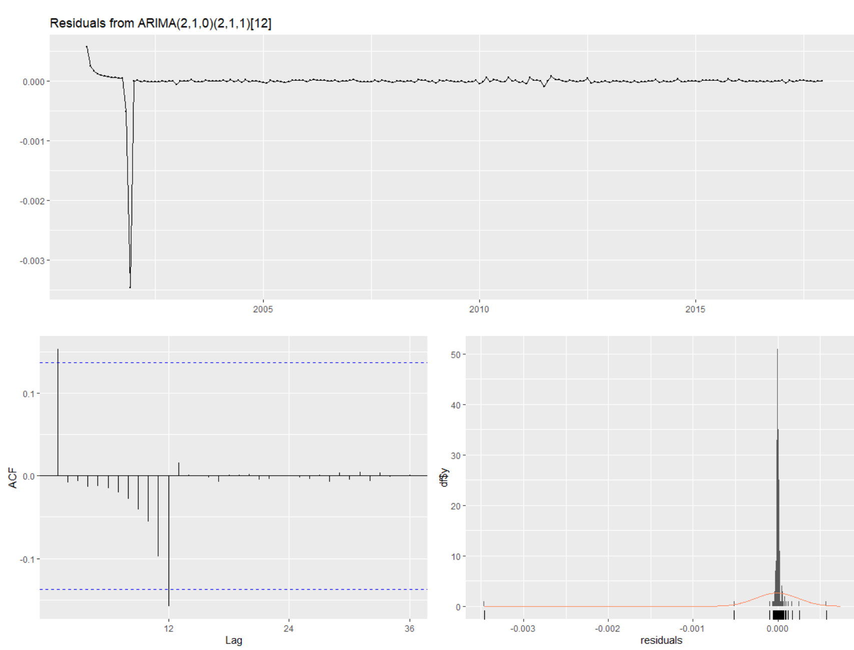
Below is the timeseries plot. By seeing that I ran the Arima (2,1,0) (2,1,1) [12]. I played with different values with p and q to get the best results. The mentioned combination gave me the best results. The difference in both order and seasonal were kept as 1 always.





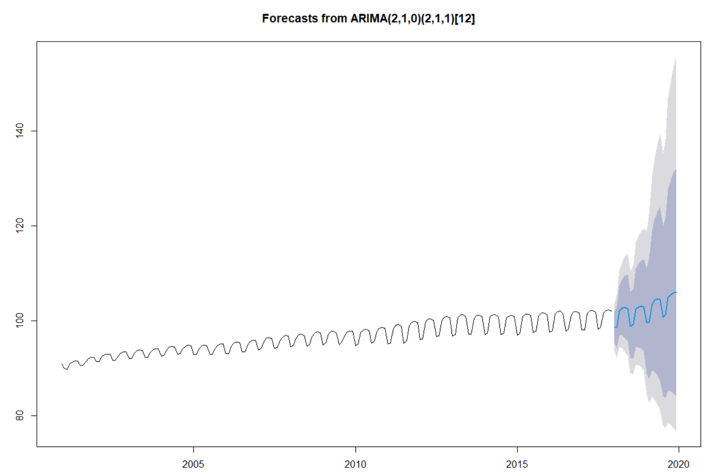
Below figures show the accuracy and residual diagnostics of the model. Looking at the figures we can see that the p-value is greater than 0.05 for the model, so we can accept the null hypothesis of the white noise, that the model captured all the data. There are only 2 spikes in the ACF plot, but it is acceptable, so we accept the model.

	RMSE	MAE	MAPE	MASE
Training set	3.083499	0.4277979	0.4576532	0.698165
Test set	1.577992	1.2246979	1.2044759	1.998704



```
        Ljung-Box test

data:  Residuals from ARIMA(2,1,0)(2,1,1)[12]
Q* = 13.965, df = 19, p-value = 0.7857

Model df: 5.   Total lags used: 24
```

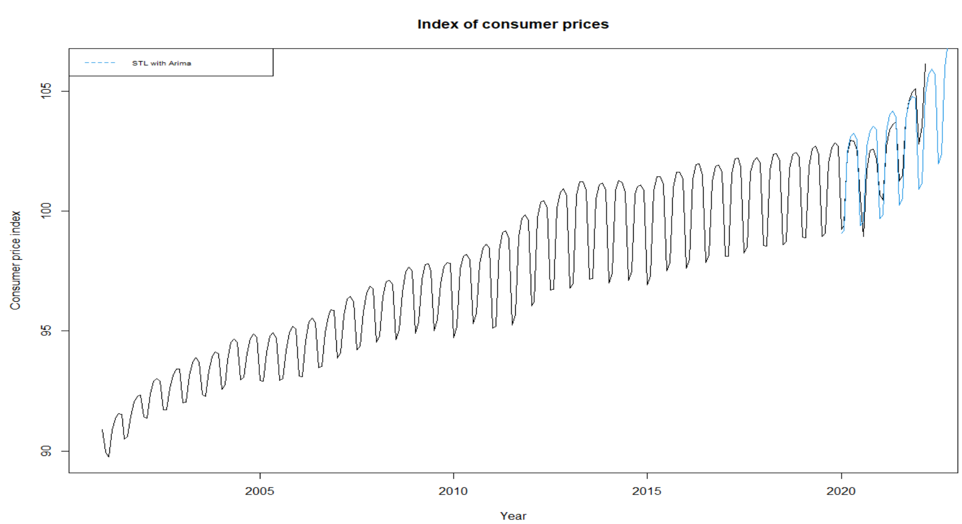
The below plot is the forecast using the model, which is not that.



FORECAST USING THE BEST MODEL

After comparing the results of different models, I conclude that the STL with Arima model was the best model. So, I did forecast for the time window January 2020 to December 2022 using the STL with Arima model. The model gave the accuracy of 0.5 in terms of MASE and p-value is greater than 0.05.

The below plot shows the forecasted values for the mentioned time window. The forecasted values have higher variations in seasonality when compared with actual values, but they somehow overlapped with the actual values which is good sign that the model forecasted well. The forecasted values are also having the increasing trend as the actual values.



REFERENCES

- https://www.statology.org/ljung-box-test/
- https://en.wikipedia.org/wiki/Ljung%E2%80%93Box_test
- https://otexts.com/fpp2/residuals.html
- https://towardsdatascience.com/identifying-ar-and-ma-terms-using-acf-and-pacf-plots-in-time-series-forecasting-ccb9fd073db8
- Professor's materials
- https://otexts.com/fpp2/