TURKISH AIRLINES PRICE BETWEEN 2014 - 2023

PROJECT REPORT SUBMITTED
IN FULFILMENT OF THE REQUIREMENTS FOR COURSE STAT 497 –
APPLIED TIME SERIES ANALYSIS
DEPARTMENT OF STATISTICS OF MIDDLE EAST TECHNICAL UNIVERSITY

BY

SERHAT HABİL ÇELİK – 2361202

January 2024

Abstract

This report explains Turkish Airlines Price and its forecast between December 2014 and December 2023. By doing this analysis, various methods such as ACF/PACF graphs, TBATS, and ETS forecasts were used in the report. Before doing these steps, data cleaning and outlier detection were done appropriately. Lastly, the best model was chosen with respect to RMSE and ACF1 values of the forecast accuracy.

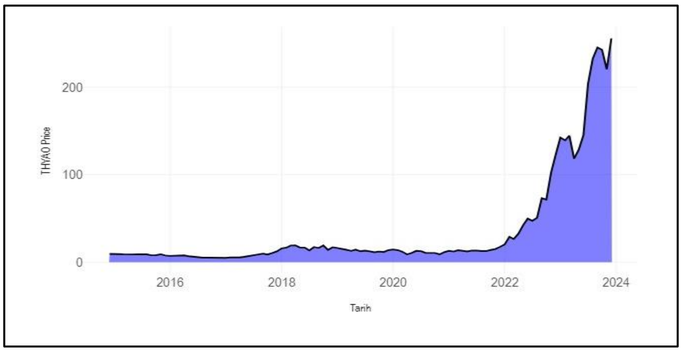
1. Introduction

Turkish Airlines is an aviation company. Its share first started to be sold in Istanbul Stock Exchange on 20th December 1990. It has proved to be a lasting and renowned company since its beginning. They have been in force in aviation sector by offering their flight seats to millions of people for long years. On the other hand, its price has been increasing day by day, except for the pandemic season. Necessary tests and forecasts for THYAO price have been done in R- Studio 4.3.2. Firstly, ACF and PACF graphs have been analyzed, and then unit root tests have been done regularly. After all these processes, the best model has been obtained in the analysis. In addition, you can learn what the prices of THYAO will be in detail in the future in this report.

2. Data Description and Pre-processing

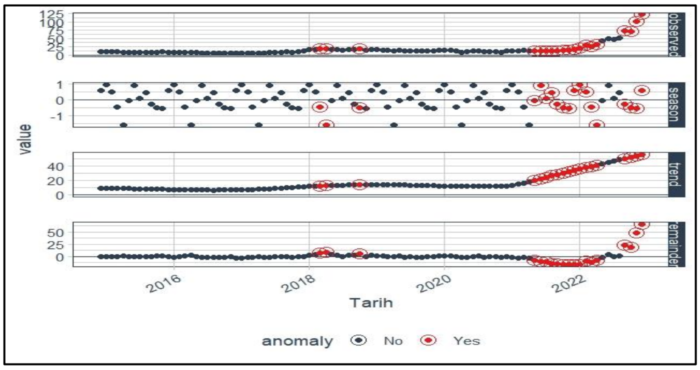
My data consist of 109 observations and 2 variables, one of which is date. The data is taken from https://tr.investing.com/equities/turk-hava-yollari-historical-data. I have taken the last nine years' price of THAYO starting from December 2014 to December 2023.

Graph 1: Time Series Plot of THYAO Price



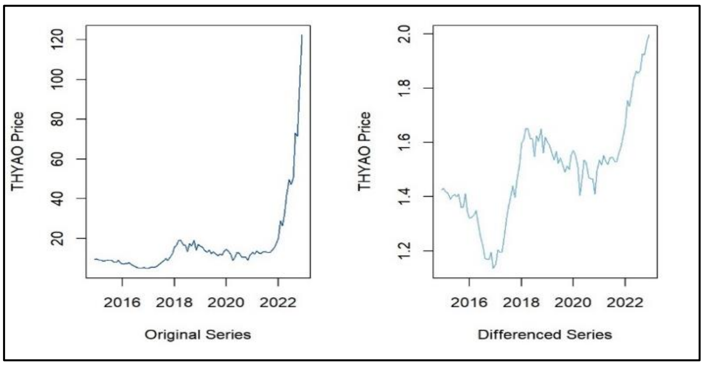
It appears to be obvious that there is an increasing trend in the graph. There is no seasonality in the data. The data are not stationary because of the increasing trend. If we examine the type of trend, it could be stochastic trend. This is because the slope of the trend is changing month by month.

Graph 2: Anomaly Detection Plot



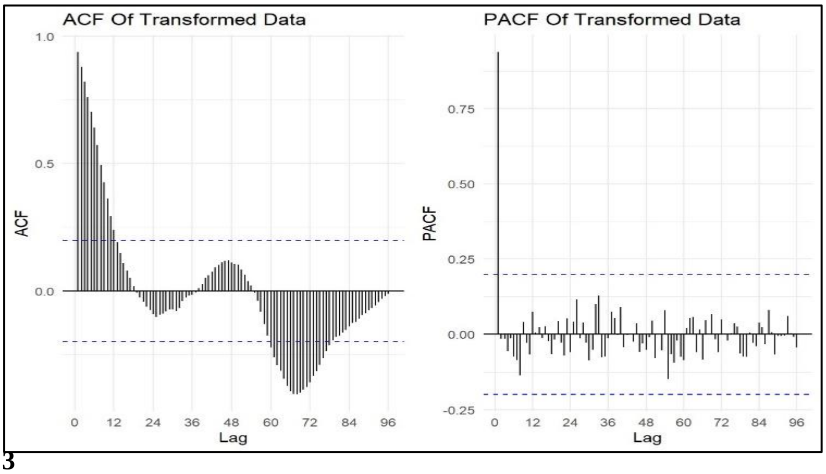
In the first step, the data are divided into train and test set. Test set consists of the last 12 observations of the data. I will work with train and test data for all analyses. As it can be seen, anomaly points should be inside the data. These values should not be removed from the data. So I will keep them in the data.

Graph 3: Original Series vs. Differenced Series



Box-Cox transformation is done to get better results in analysis. Lambda value is -0.44. When we do that, nothing changes. The data are still stationary. In the original series, there is an increasing trend; however, in the differenced series, there is a decreasing and increasing trend at the same time.

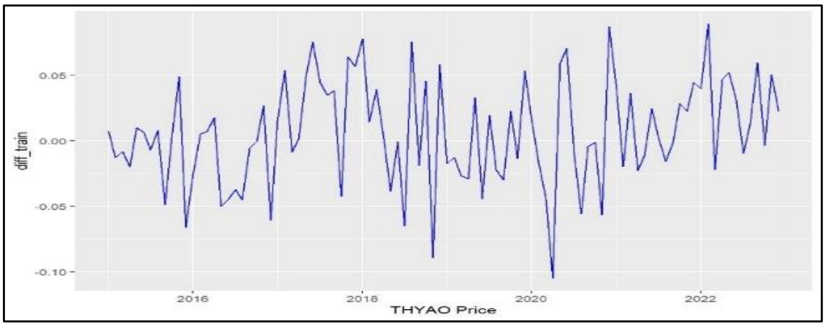
Graph 4: ACF and PACF of Transformed Data



As it is seen, there is a slow decay in ACF graph, meaning that the series is not stationary. PACF cuts off after the first lag. There is no need to check PACF.

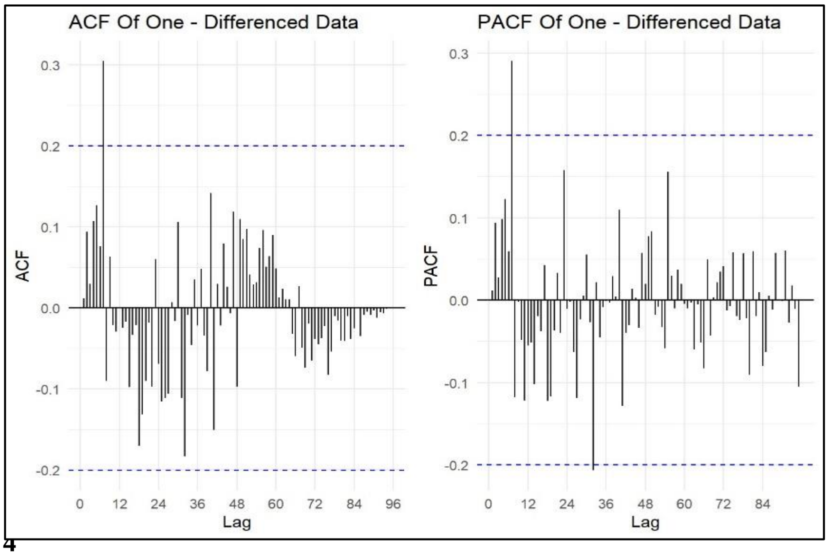
KPSS Level and Trend Test are applied to the series to be sure about stationary. Its p-value for level is 0.01, which is smaller than 0.05. Hence, the series is not stationary. Also, its p-value for trend is 0.0412, which is less than 0.05; this indicates that the series has a stochastic trend. We need to take difference to remove stochastic trend. In what follows, PP test is applied to the series. Its p-value is 0.8868, concluding that we do not have enough evidence to claim that we have a stationary system. Finally, ADF test with “ct” is applied so as to be able to say whether the series has a stochastic trend or not. Its p-value is 0.9136. It means that we have nonstationary system with a stochastic trend. After doing all these tests, one-differencing is needed to remove unit root, causing the series not to be stationary. One-differencing is enough. After differencing, p-value of KPSS Test is bigger than 0.05 and also p-value of ADF Test with “nc” is smaller than 0.05. We can conclude that the series has become stationary because of one-differencing.

Graph 5: Time Series Plot of One- Differenced Data



As you can see above, it seems like stationary. The values are around zero.

Graph 6: ACF and PACF Plot of One - Differenced Data



The series becomes stationary after taking one difference. Since there is no seasonality in the data, there is no need to check HEGY.TEST. One differencing is enough. I will suggest a model according to the results of ACP-PACF graphs.

We can suggest ARIMA (0,1,0), called random walk process. ARIMA (7,1,7) can also be suggested here. When I fit them accordingly, ARIMA (0,1,0) gives the best result compared to ARIMA (7,1,7). Random walk has the smallest AIC and BIC values. A random walk can be considered to be a time-series model in which there is an equality between the current observation and the previous one with a random step up or down.

The last method to identify model is auto.arima function built in R .

Table 1: Auto Arima Table

Series : train_box	
ARIMA(0,1,0)	
Sigma^2 = 0.001657	log likelihood = 171.11
AIC = -340.22	BIC = -337.65

3. Modelling and Diagnostic Checking

After finding the best model, residual analysis is done with different tests such as Shapiro Wilk or Breusch – Godfrey Test.

Firstly, we should check Normality Assumption with Shapiro Wilk and Jarque – Bera tests. Shapiro Wilk and Jarque – Bera Test give the same result. P-value of both is higher than 0.05, meaning that normality assumption is satisfied.

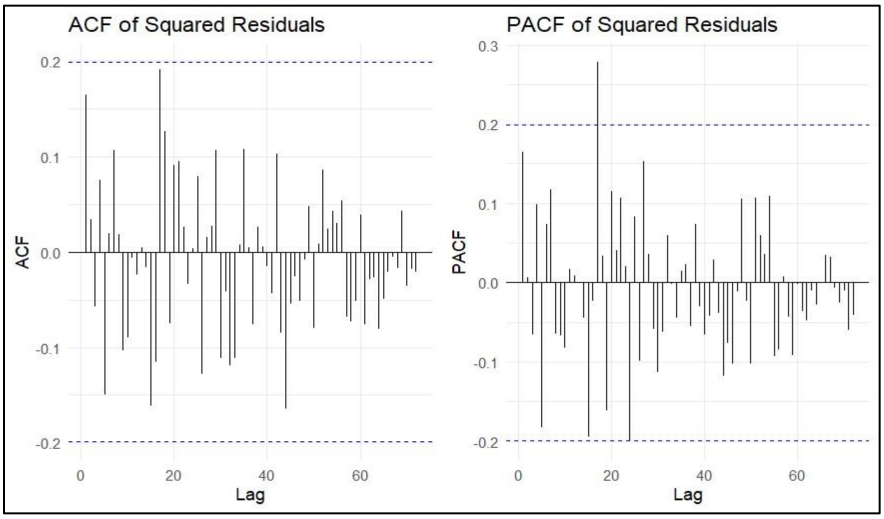
We will continue with Breusch-Godfrey to check whether the residuals of models are correlated or not. According to the result of this test, p-value is equal to 0.2589, which is higher than 0.05. So the residuals are uncorrelated because we fail to reject H_0.

Table 2: ARCH Engle's Test Table

ARCH LM – test; Null hypothesis : no ARCH effects		
data : r1		
Chi – squared : 9.6194	df = 12	p-value = 0.6493

Lastly, we will apply ARCH Engle's Test for detecting presence of ARCH effect. Since p - value = 0.6493 > 0.05, we fail to reject H_0. After all, we can say that there is no presence of ARCH effects. Consequently, we do not need to use ARCH and GARCH model.

Graph 7: ACF and PACF of Squared Residuals



Almost all spikes are in 95% of the White Noise Bands that is an indication of homoscedasticity. After fitting ARIMA model, we will decide the best exponential smoothing method for our data.

Table 3: ETS Table

ETS (M, A ,N)		
Smoothing Parameters :		
alpha = 0.7781		
beta = 0.1759		
Initial States :		
l = 9.6632		
b = -0.3211		
sigma : 0.1365		
AIC	AIC_C	BIC
550.3300	550.9893	563.2035

Exponential Smoothing Method has a multiplicative error and additive according to the result of Table3. After fitting ETS, normality of residuals is checked and it is seen that normality assumption is satisfied for it because p-value of Shapiro Wilk is equal to 0.4536.
After ETS, TBATS Model is conducted for the data. It is actually used with multiple seasons. Although we don’t have seasonality in my data, I have preferred to use it to forecast the data.

Table 4: TBATS Table

BATS (0 , {0,0} ,1 -)
Call : tbats (y=train)
Parameters:
Lambda : 0
Alpha : 0.801937
Beta : 0.152849
Designing Parameter : 1
sigma : 0.1270689
AIC : 545.1156

After fitting the TBATS, normality assumption for residuals ought to be checked. Jarque Bera Test says that normality assumption is satisfied because p-value is higher than 0.05. As a last model, NNETAR model is conducted.

Table 5: NNETAR Table

Series: train
Model: NNAR(1,1)
Call: nnetar(y = train, P=0)
Average of 20 networks, each of which is
a 1-1-1 network with 4 weights
options were - linear output units
sigma^2 estimated as 12.59

We take P equal to be zero because our series is not seasonal. That is why we write P=0. It is repeated twenty times on average to get a better result. When we check its residual for normality, we reject H_0 since p-value is lower than alpha (0.05). Hence, Normality Assumption is not satisfied for NNETAR.

We do not apply prophet model because our model does not have seasonality. We have preferred not to use this model.

We fit all the necessary models. We obtain train and forecast values from each of them with accuracy function. Now, we will compare their accuracy and decide which one is the best model.

Table 6: Train Accuracy of Models

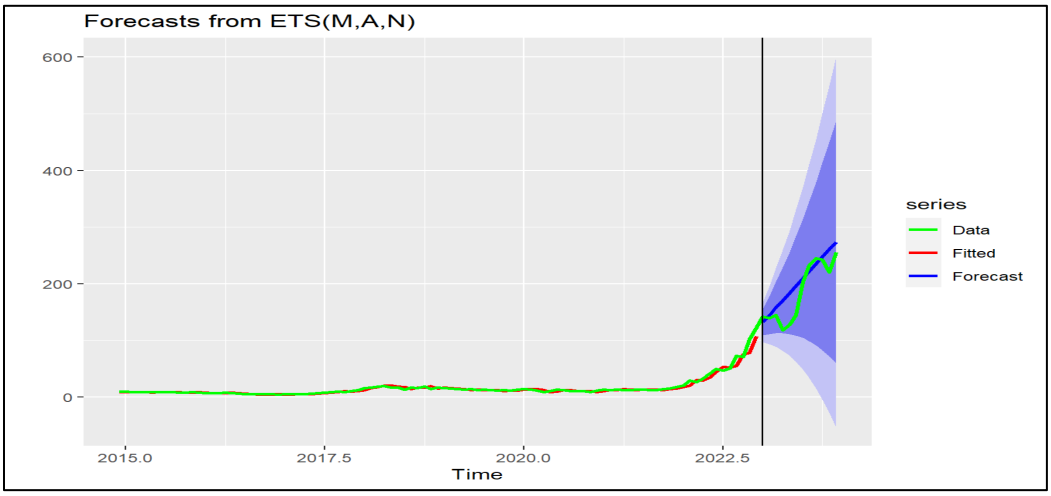
	ME	RMSE	MAE	MPE	MAPE	ACF1
ARIMA	0.01	0.04	0.03	0.31	2.18	0.01
ETS	0.77	4.03	1.98	1.55	10.01	0.08
NNETAR	0.35	3.65	1.94	-0.03	10.27	-0.47
TBATS	0.01	3.55	1.92	-1.39	10.33	-0.17

Table 7: Forecast Performance of Models

	ME	RMSE	MAE	MPE	MAPE	ACF1
ARIMA	182.75	189.58	182.75	98.83	98.83	6.47
ETS	-18.23	30.34	23.29	-12.64	15.25	0.50
NNETAR	26.30	54.56	45.55	7.82	22.73	0.77
TBATS	-231.19	299.18	231.99	-112.82	112.82	0.69

According to Table 6, the best forecasting model could be selected as ARIMA because of the lowest RMSE and ACF1 values when compared to other models. However, when we look at Table 7, ETS outperforms the other methods with the lowest RMSE, MAE, MAPE and ACF1. The second-best forecasting model is TBATS according to Table 7.

Graph 8: Forecasting Plot of ETS



4. Conclusion

In this study, Turkish Airlines Price between 2014 – 2023 was examined with different tests and plots. At first, the data were introduced in detail and pre-processing was done. After that, time series plot of the series was drawn and whether it was stationary or not was checked.

Then, anomalize detection and Box–Cox transformation was applied to the model to make data more clear. Later, ACF and PACF graphs were checked and the model of the series was selected. It is ARIMA (0,1,0). After all these things, residuals of ARIMA were checked with Shapiro–Wilk and Breusch–Godfrey Test. The results were plausible. The next step was to determine the best forecast model. ETS, TBATS and NNETAR were done for the data. PROPHET was not useful for the data because of non-seasonality. ARIMA and ETS were the best forecasting model compared to the others.

We have learnt the structure of THYAO price and its future situation. It could be beneficial to invest in this share in the future.

REFERENCES

Turkish Airlines (THYAO) geçmiş fiyatları - investing.com. Investing.com Türkiye. (n.d.). https://tr.investing.com/equities/turk-hava-yollari-historical-data