The Faculty of Engineering in Foreign Languages

**ARIMA model analysis and forecast for the  
shares of MICROSOFT**

Student: Tanase Nicolae-Costin

Bucharest 2019

Contents

[Introduction 3](#_Toc105502506)

[Data description 3](#_Toc105502507)

[Candlestick Analysis 4](#_Toc105502508)

[Stationarity analysis 7](#_Toc105502509)

# Introduction

Microsoft Corporation is an American multinational technology company with headquarters in Redmond, Washington. It develops, manufactures, licenses, supports and sells computer software, consumer electronics, personal computers, and services. Its best known software products are the Microsoft Windows line of operating systems, the Microsoft Office suite, and the Internet Explorer and Edge web browsers. Its flagship hardware products are the Xbox video game consoles and the Microsoft Surface tablet lineup. As of 2016, it is the world's largest software maker by revenue, and one of the world's most valuable companies. The word "Microsoft" is a portmanteau of "microcomputer" and "software". Microsoft was founded by Bill Gates and Paul Allen on April 4, 1975, to develop and sell BASIC interpreters for the Altair 8800. It rose to dominate the personal computer operating system market with MS-DOS in the mid-1980s, followed by Microsoft Windows. The company's 1986 initial public offering (IPO), and subsequent rise in its share price, created three billionaires and an estimated 12,000 millionaires among Microsoft employees. Since the 1990s, Microsoft has increasingly diversified from the operating system market and has made a number of corporate acquisitions. In May 2011, Microsoft acquired Skype Technologies for $8.5.

# Data description

In the analysis, the shares of Apple Inc. will be studied. The start point will be 01.01.2019 and the endpoint will be 29.05.2022. The source of the data is Yahoo Finance and the information  
contained for each share is:  
- Date  
- High, the highest price for the share for the day  
- Low, the lowest price for the share for the day  
- Open, the opening price for the share for the day  
- Close, the closing price for the share for the day  
- Volume, the number of shares transacted for the day  
  
The data isn’t continuous, having skip dates when the stock exchange is closed, for example on the weekends.

# Candlestick Analysis

For starters, the candlestick graph will be analyzed for the stated period. A candlestick chart  
(also called Japanese candlestick chart) is a style of financial chart used to describe price  
movements of a security, derivative, or currency. Each "candlestick" typically shows one day,  
thus a one-month chart may show the 20 trading days as 20 "candlesticks".

Chart, line chart

Description automatically generated

Figure 1 Candlestick chart of Microsoft shares

As seen in the chart above, the shares of Microsoft have fluctuated throughout the period while maintaining a strong trend. The highest value was reached on 2021-11-22 , and the lowest was reached in 2017-07-06. Microsoft also became a trillion dollar company in April 2019 and gaining another trillion only 2 years later. The chart also shows that the fluctuations happen even throughout the day, as you’d expect from a such important stock.

Chart, line chart

Description automatically generated

Figure 2 Candlestick chart with the linear regression of the high values

The regression line of the High series shows the upward trend of the share value for the period and it also shows that stationarity exists.

Text

Description automatically generatedThe High series can be described as having:  
- A mean of 172.04  
- A standard deviation of 81.35  
- A minimum value of 64.42  
- A maximum value of 348.13  
- The first quartile of 102.46  
- The median or second quartile of 147.09  
- The third quartile of 237.48

Figure 3 Descriptive statistics for high series

Next, we will be looking at the visual representation of the distribution of values:

Chart, line chart

Description automatically generatedChart, histogram

Description automatically generated‘’

Figure 4 Distribution of high series

It can be seen from the previous chart that the series doesn’t follow a Gaussian distribution. This  
means that in order to be analyzed the series will most likely have to be differentiated.

# Stationarity analysis

Firstly, the autocorrelation coefficient must be studied using a autocorrelation function:

Text

Description automatically generated

Figure 5 Autocorrelation function

Chart, line chart

Description automatically generated

Figure 6 High values series

A picture containing text, building

Description automatically generated

Figure 7 high values acf chart

For a more detailed view of the autocorrelation plot, the notebook shows more clearly that the series is not stationary and must be differentiated.

Chart

Description automatically generated

Figure 8 High values differenced

Graphical user interface

Description automatically generated with low confidence

Figure 9 Acf chart for diffrenced values

From the previous graphs it is clear that the series needs no more differencing in order to become stationary.

A picture containing chart

Description automatically generated

Figure 10 PACF for high differenced values

After being differentiated once the series becomes stationary. There are no significant values  
present in the ACF and PACF charts. Therefore, an ARIMA model can now be built for the series. Now, the choice of the three parameters, P, D and Q remains. D is clearly 1, since it only took 1 differencing to achieve stationarity. I have chosen P and Q to be 1 since in both the ACF and PACF charts for 100 lags there are very few outliers and are all close to the limit. Also, the ARIMA summary gives me a poorer result when increasing P and Q.

# Arima model and forecasting

After the model is built, this is the summary:

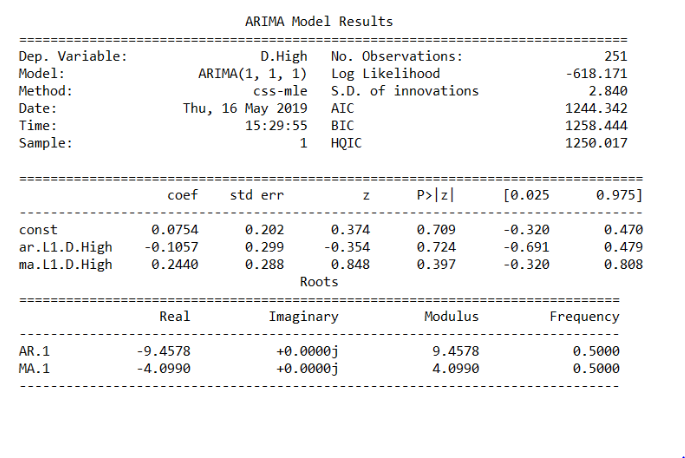


Figure 11 Arima model results

For the final part of this paper, a forecast will be made for the built model. The original series  
will be split into two, a training series and a testing series. 80% of the original series will be used  
for training and 20% for testing the predictions.

Chart, line chart

Description automatically generated

Figure 12 Forecasting for every value

Chart, line chart

Description automatically generated

Figure 13 Forecasting with a 80/20 train/test split

From the chart above we can conclude that the prediction the model offered is satisfactory.

# Conclusions

In conclusion, this paper serves as a guide to create an ARIMA models for any stock. The paper analyzed the Microsoft stock from the beginning on 2019 up until the present. The model obtained was satisfactory given the approach taken.