Time Series Analysis Project

Statistics Major

Submitted to DR Eman Mahmoud and DR Yousra Hassan

20th of May 2023

By: **Nour Mohamed Nashat Eletreby 5200917**

Haneen Ahmed Abdelaziz Elgabry 5200199

Foreign Direct Investment in The United States of America

Content Table

Introduction

First, we would like to tackle the topic by answering multiple of questions why is it important to analyze the foreign direct investment in the USA? Why is it important to model it? Why is it important to forecast this time series?

Foreign direct investment “FDI” is critical and extremely important as it diversifies investors’ portfolios, promotes stable long term lending, provides financing to developing countries and provides technology to developing countries. When it comes to the USA; it is extremely important to recognize this topic as the USA tops the world when it comes to foreign direct investment (top destination). So, here comes the importance of time series analysis to reach decisions from the data.

Time series analysis can be used here to reach a good model that can be used to forecast business metrics such as sales, turnover, stock market price, etc. Moreover, it allows us to reach conclusions about timely patterns in data and analyze trends. Let: • T represents the trend component in time period t, • Ct represents the cyclical component in time period t, • St represents the seasonal component in time period t, and • It represents the random component in time period t. It is important to forecast this time series as it is a powerful method for predicting future trends and values; hence, predicting values in the FDI will be useful for the US economy no doubts.

Literature Review and Background

Talking about foreign direct investment (FDI) in the United States, which means when a company from another country invests money in the US to create jobs and grow their business. It is proved that that the US is the most popular country in the world for this type of investment, with $4.3 trillion invested in the US as of 2019. This type of investment is really important because it helps create jobs, supports new ideas and technology, and helps the economy grow. Moreover, it was also proved that FDI created 8.1 million jobs in the US in 2019, which is a lot more than a few years ago. We have also searched for which sectors is getting more benefits from it; it was found that manufacturing is the most popular, followed by finance and insurance, and then professional, scientific, and technical services. The USA is known as a good place for this kind of investments for its big markets, strong legal system and skilled workforce.

Now we would like to tackle this topic from its pros and cons:

The pros of the FDI:

Job opportunities, advancement of technology and knowledge, increases economic growth, improves infrastructure, and enhances skills and education as well. FDI can help boost the host country's economy.

The cons of the FDI:

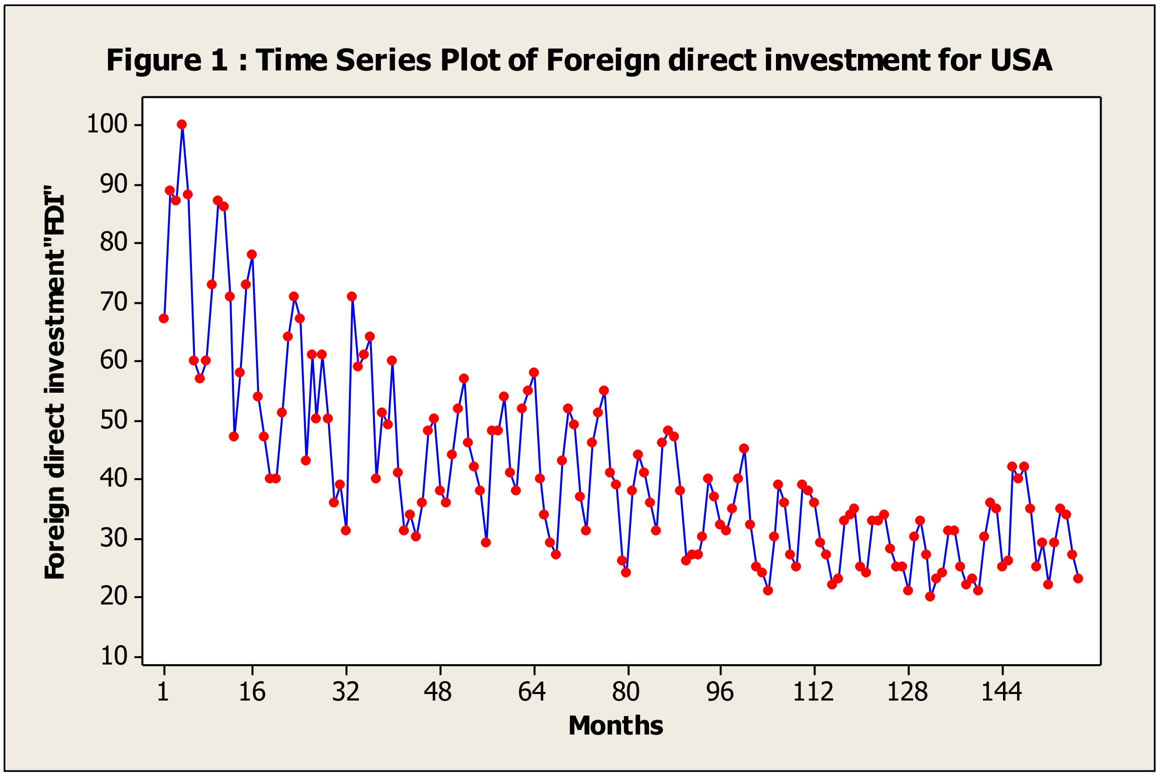
Dependence on foreign investment, loss of domestic control, competition with domestic firms, and loss of profits. Host countries should be aware of these risks and implement policies that promote responsible investment and protect local interests.

Moreover, the FDI has 4 types which are horizontal, vertical, conglomerate and platform. But, horizontal FDI is the most used and most well-known type.

In summary, time series analysis can be a powerful tool for countries seeking to optimize their foreign direct investment strategy. By providing insights into FDI trends, risks, and performance, countries can make more informed decisions and improve their economic competitiveness.

Data Analysis

Our data is about the foreign direct investment in the USA; we got the data from google trends-the link can be found in the references page-, the variable measured is the foreign direct investment and its scale unit is months (monthly time series) and the sample size is 157 observations.



Comment: This is the time series plot for our data; obviously we can see here a negative trend in our data; where the foreign direct investment in the USA decreases over months, also we can see seasonality in this graph. We can assume that our model is multiplicative as we can see S=percentage of T*check the appendix for more information*. Also, we cannot assume that it is additive as we have no proof that S and T are independent. There exist very few outliers.

After applying the decomposition method, we reached:

`Yt = 62.66 - 0.266147*t`

Comment: Here is our fitted regression equation, this equation is made by deseasonalised time series not with the original variable.

B0= 62.88 , the FDI=62.66 on average when holding time constant and β1=-0.266147 this means when time increases by 1 unit the FDI decreases by 0.266147 on average

Seasonal Indices

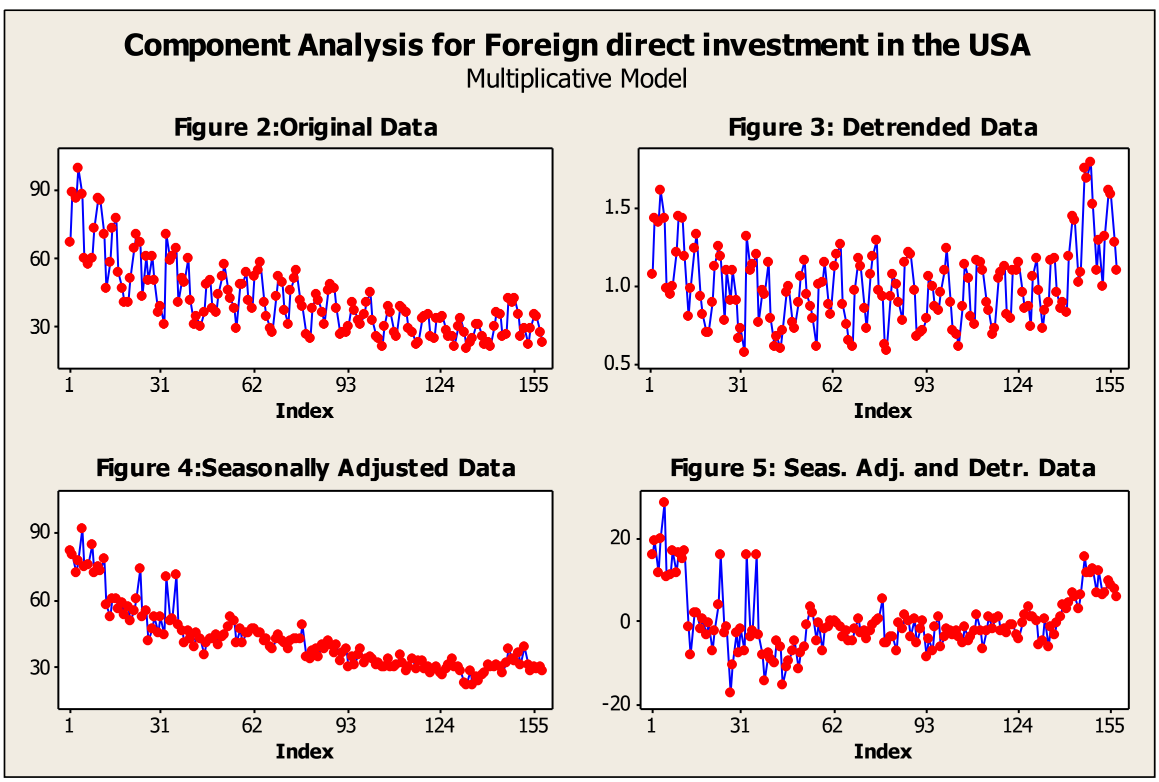
Period	Index
1	0.82087
2	1.12484
3	1.21611
4	1.30158
5	0.96762
6	0.80553
7	0.75643
8	0.71109
9	1.02066
10	1.17472
11	1.19189
12	0.90867

Comment: As it is multiplicative model: for season 1 (0.82087*100)= 82.087 % then season 1 is reducing FDI by 17.913% of the trend , for season 2 (112.484%) then season 2 is increasing FDI by 12.484% , for season 3 (121.611%) then season 3 is increasing FDI by 21.611% of the trend, for season 4 (130.158%) then season 4 is increasing FDI by 30.158% of the trend , for season 5 (96.762%) then season 5 is decreasing FDI by 3.238% of the trend, for season 6 (80.553%) then season 6 is decreasing FDI by 19.447% of the trend, for season 7 (75.643%) then season 7 is decreasing FDI by 24.457 % of the trend , for season 8 (71.109%) then season 8 is decreasing FDI by 28.891% of the trend, for season 9 (102.066%) then season 9 is increasing the trend by 2.066% of the trend, for season 10 (117.472%) then season 10 is increasing FDI by 17.472% of the trend, for season 11 (119.189%) then season 11 is increasing FDI by 19.189% of the trend, for season 12 (90.867%) then season 12 is decreasing FDI by 9.133% of the trend.

And here we can see the future forecast for the FDI in the upcoming 12 months (1 year)

Forecasts

Period	Forecast
158	23.1774
159	24.7343
160	26.1264
161	19.1653
162	15.7405
163	14.5797
164	13.5165
165	19.1293
166	21.7040
167	21.7041
168	16.3048
169	14.5109



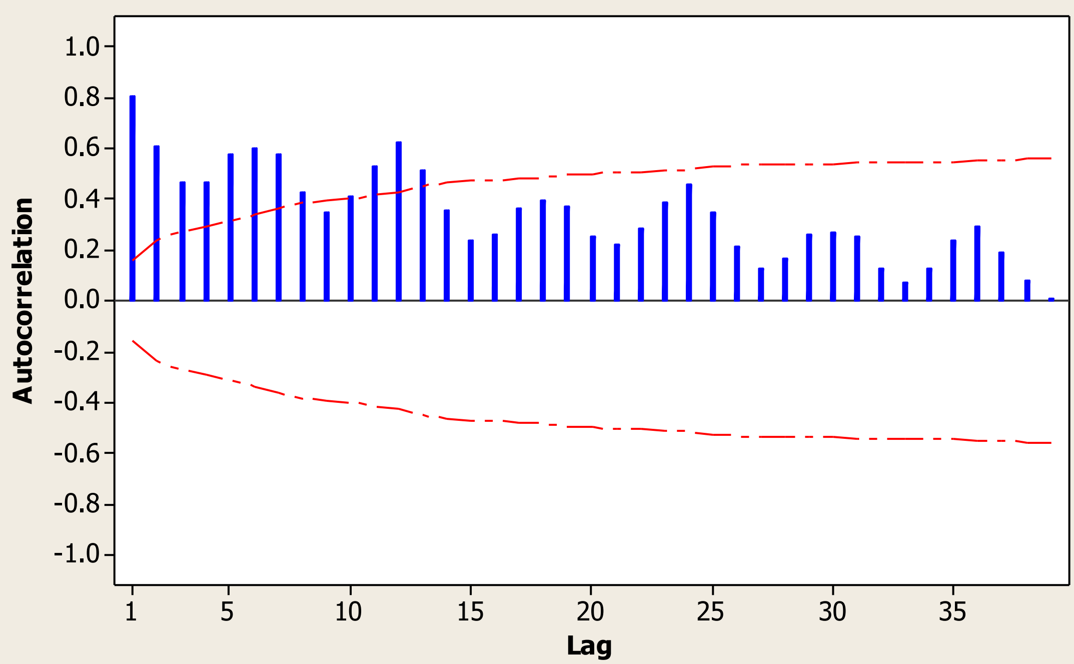
Comment: In figure 2 we can see the original data with negative trend and seasonal effect, then in figure 3 we reached the detrended data; here we can see the data without the trend effect and mainly constant with no changes in the trend, but still with seasonal effect, then we reached figure 4 the deseasonalised data (seasonally adjusted) here we got rid of the seasonal effect but the negative trend is still there. Finally, we reached figure 5 the deseasonalised and detrended data here we got rid of the trend effect as well as the seasonal effect.

Now in order to apply the modern approach we should check whether or series is stationary or not:

We can check with 2 methods, the first method from the time series plot -check figure 1- so we can assume from the plot that our series is not stationary as there is a trend (so the series is not constant in mean) also we can see fluctuations (so the series is not constant in variance)

The second method is the ACF method:

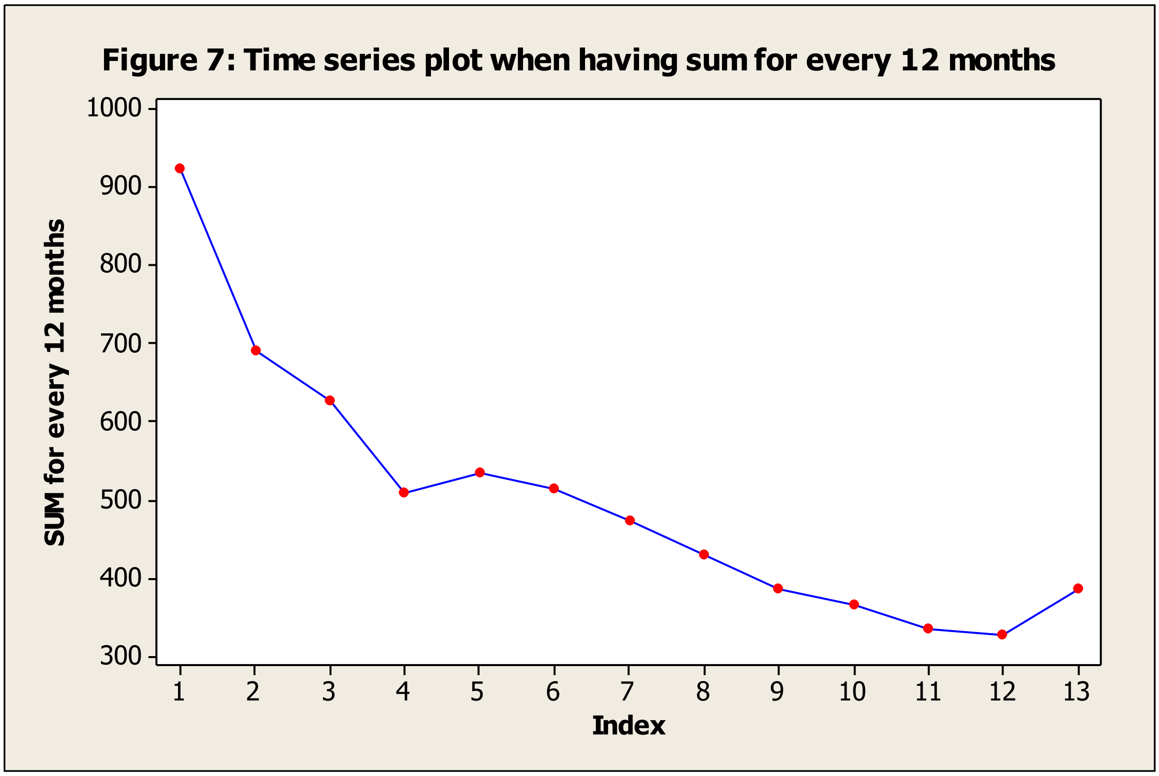
Figure 6: ACF for data



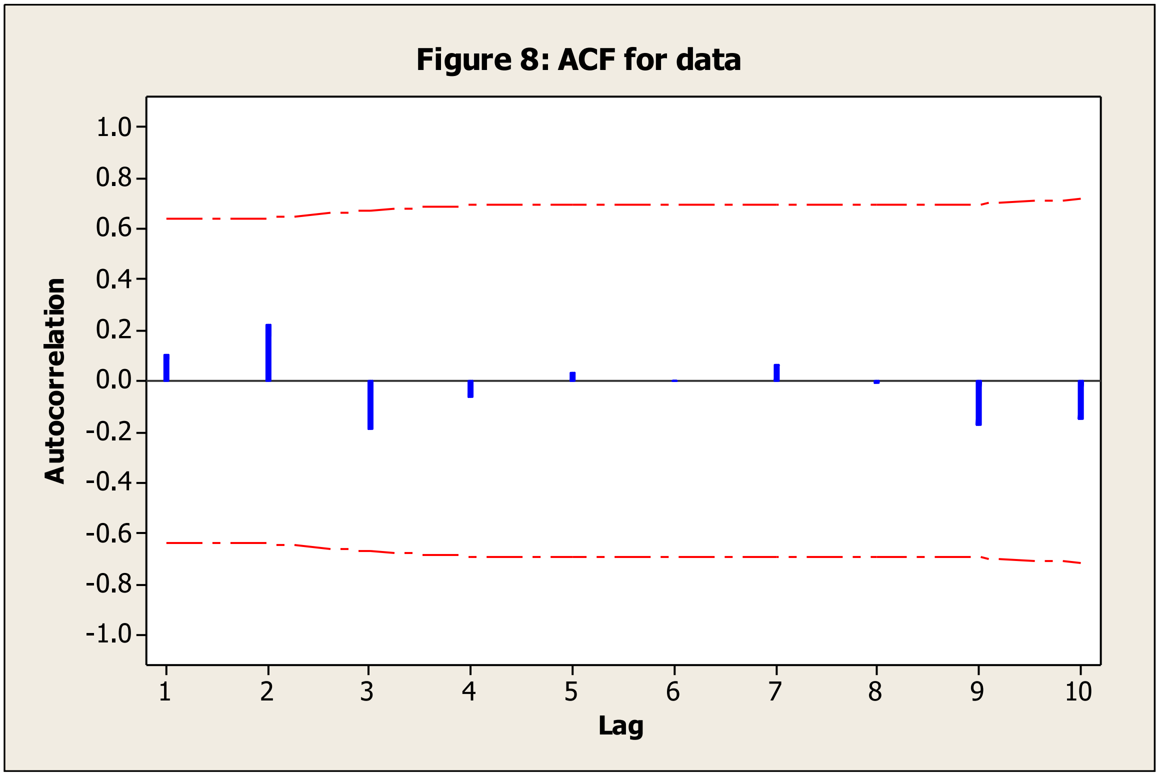
From figure 6 we can see from the ACF that is not stationary as the graph did not die quickly (did not decrease)

So, now we should transform it to stationary one

We will take summation for every 12 months (every year) to have a constant variance



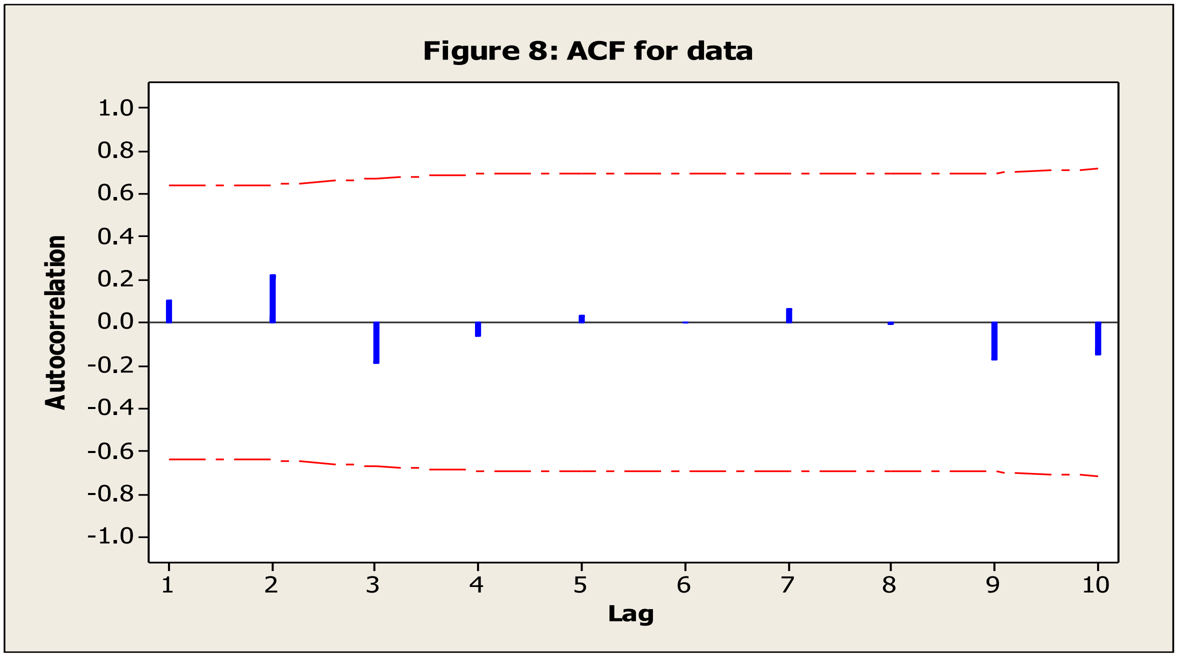
It is clear here in figure 7 we have trend (non-constant mean) but we have constant variance so we need to get 1st difference in order to detrend the series and reach the stationary model

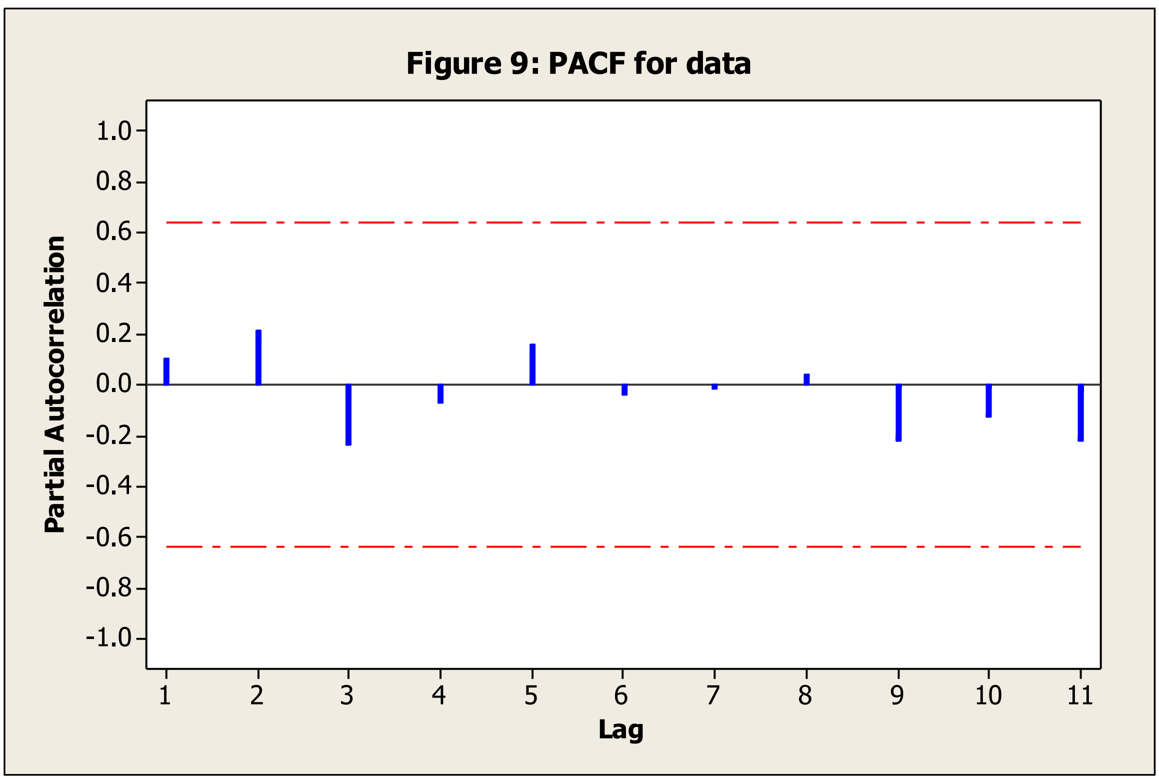


Now after we got summation for the variable then we had 1st order difference ; now our model is stationary as it became constant in variance and mean as well and we can see this in figure 8 that the model is now stationary from the ACF

We will start to apply the modern approach "Box and Jenkins Analysis" as our model is now stationary

Identify the initial model by identifying order of ARIMA this is by ACF and PACF





We can see the ACF and PACF cuts off after 1 lag and also we can assume our model is stationary.

Now we will start with trying several models to reach the best one:

As our number of observations decreased after making summation for our data, therefore the p value in Ljung box could not be conducted. Hence, we had to make our decision to know the best model based on the ACF and PACF for the residuals and the p value of the parameters to be less than 0.05

After testing several models on minitab we had only 2 models which are AR(1) , MA(2) The two models had their ACF and PACF residuals are insignificant and the p value for the parameters is less than 0.05

Now we will compare between them using MAD , MAPE and MSE

MAD for AR(1)= 49.4634

MSE for AR(1)= 4339.93

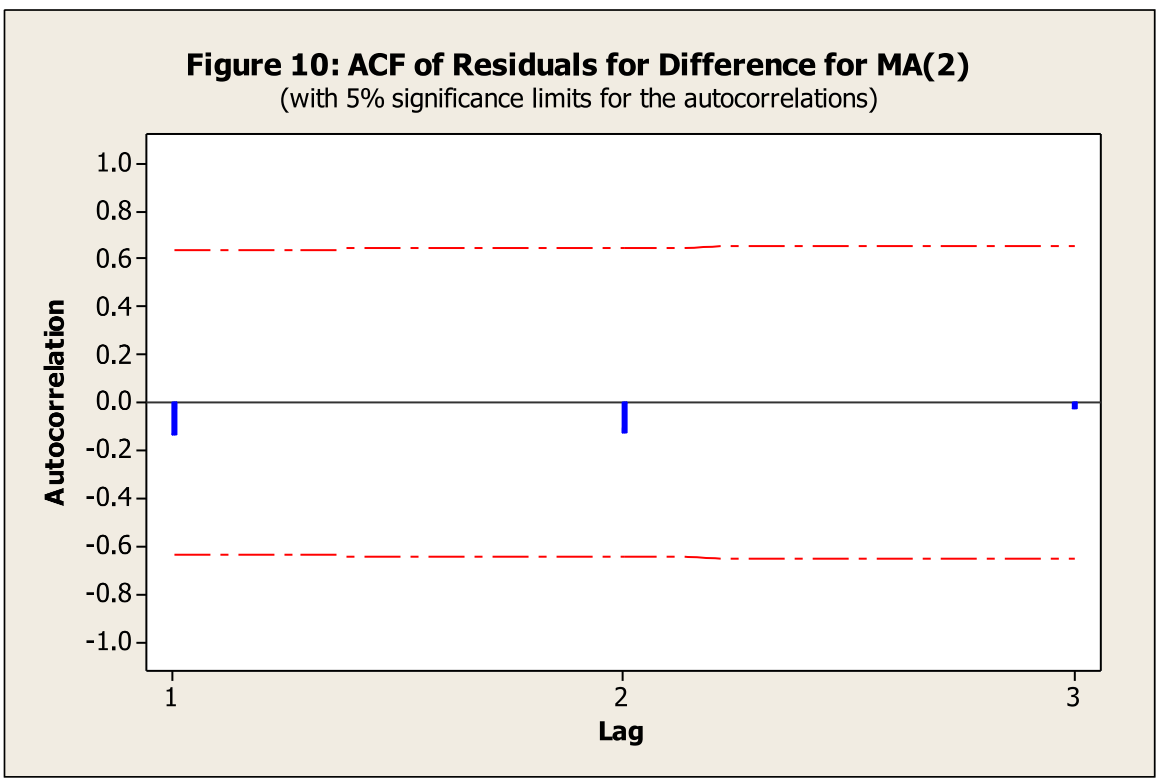
MAPE for AR(1)= 0.315416

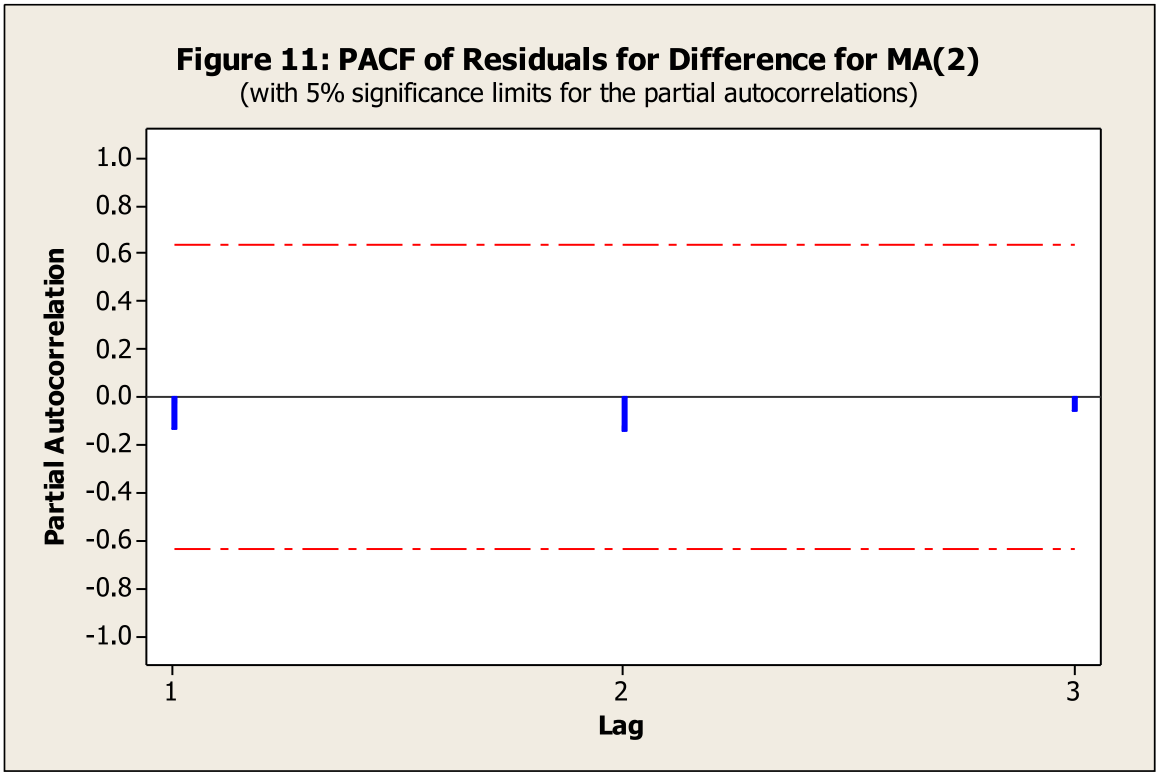
MAD for MA(2)= 30.1326

MSE for MA(2)= 2574.62

MAPE for MA(2)= 0.318321

As we can see that MA(2) is the appropriate model is it had the least MAD and MSE





Comment: From the ACF and PACF of residuals of MA(2) we can say that they are insignificant.

```
Final Estimates of Parameters


Type        Coef   SE Coef       T       P
MA   1   -0.4962    0.2196   -2.26   0.047
MA   2   -0.9646    0.2243   -4.30   0.002
```

And we can see here the p value for the parameters is less than 0.05

Hence, MA(2) is the appropriate model after we compared it with AR(1) by accuracy measures.

Now we will check the invertibility of our model

1) -0.4962+(-0.4962)= -1.4608 which is less than 1
2) -0.9646-(-0.4962)=-0.4684 which is less than 1
3) -1<θ2<1 which is -1<-0.9646<1

Therefore, we can assume that our model MA(2) is invertible

And here are the forecasts:

```
Forecasts from period 13

                      95% Limits
Period  Forecast     Lower    Upper  Actual
    14    33.997   -74.969  142.963
    15    65.896   -55.744  187.537
    16     0.000  -160.761  160.761
    17     0.000  -160.761  160.761
    18     0.000  -160.761  160.761
    19     0.000  -160.761  160.761
    20     0.000  -160.761  160.761
    21     0.000  -160.761  160.761
    22     0.000  -160.761  160.761
    23     0.000  -160.761  160.761
    24     0.000  -160.761  160.761
    25     0.000  -160.761  160.761
    26     0.000  -160.761  160.761
    27     0.000  -160.761  160.761
    28     0.000  -160.761  160.761
    29     0.000  -160.761  160.761
    30     0.000  -160.761  160.761
    31     0.000  -160.761  160.761
    32     0.000  -160.761  160.761
    33     0.000  -160.761  160.761
    34     0.000  -160.761  160.761
    35     0.000  -160.761  160.761
    36     0.000  -160.761  160.761
    37     0.000  -160.761  160.761
    38     0.000  -160.761  160.761
    39     0.000  -160.761  160.761
    40     0.000  -160.761  160.761
    41     0.000  -160.761  160.761
    42     0.000  -160.761  160.761
    43     0.000  -160.761  160.761
    44     0.000  -160.761  160.761
    45     0.000  -160.761  160.761
    46     0.000  -160.761  160.761
    47     0.000  -160.761  160.761
    48     0.000  -160.761  160.761
```

49	0.000	-160.761	160.761
50	0.000	-160.761	160.761
51	0.000	-160.761	160.761
52	0.000	-160.761	160.761
53	0.000	-160.761	160.761
54	0.000	-160.761	160.761
55	0.000	-160.761	160.761
56	0.000	-160.761	160.761
57	0.000	-160.761	160.761
58	0.000	-160.761	160.761
59	0.000	-160.761	160.761
60	0.000	-160.761	160.761
61	0.000	-160.761	160.761
62	0.000	-160.761	160.761
63	0.000	-160.761	160.761
64	0.000	-160.761	160.761
65	0.000	-160.761	160.761
66	0.000	-160.761	160.761
67	0.000	-160.761	160.761
68	0.000	-160.761	160.761
69	0.000	-160.761	160.761
70	0.000	-160.761	160.761
71	0.000	-160.761	160.761
72	0.000	-160.761	160.761
73	0.000	-160.761	160.761
74	0.000	-160.761	160.761
75	0.000	-160.761	160.761
76	0.000	-160.761	160.761
77	0.000	-160.761	160.761
78	0.000	-160.761	160.761
79	0.000	-160.761	160.761
80	0.000	-160.761	160.761
81	0.000	-160.761	160.761
82	0.000	-160.761	160.761
83	0.000	-160.761	160.761
84	0.000	-160.761	160.761
85	0.000	-160.761	160.761
86	0.000	-160.761	160.761
87	0.000	-160.761	160.761
88	0.000	-160.761	160.761
89	0.000	-160.761	160.761
90	0.000	-160.761	160.761

91	0.000	-160.761	160.761
92	0.000	-160.761	160.761
93	0.000	-160.761	160.761
94	0.000	-160.761	160.761
95	0.000	-160.761	160.761
96	0.000	-160.761	160.761
97	0.000	-160.761	160.761
98	0.000	-160.761	160.761
99	0.000	-160.761	160.761
100	0.000	-160.761	160.761
101	0.000	-160.761	160.761
102	0.000	-160.761	160.761
103	0.000	-160.761	160.761
104	0.000	-160.761	160.761
105	0.000	-160.761	160.761
106	0.000	-160.761	160.761
107	0.000	-160.761	160.761
108	0.000	-160.761	160.761

And the forecasts by the decomposition method for the upcoming year can be found in page 7 in the decomposition section.

Conclusion

As a conclusion we can see a negative trend in our data where FDI in the USA decreases over months and this can be extremely useful for the country while analyzing the data in order to see the overall trend. Moreover, after applying the modern approach we tried several models AR(1),AR(2),MA(1),MA(2),ARMA(1,1),ARMA(1,2),ARMA(2,1),ARMA(2,2); we reached out that MA(2) is the appropriate model for our data set and we found that our model is invertible. Also, we can see in the decomposition method we had forecasts for the upcoming year (next 12 months), on the other hand while measuring the forecast in MA(2) we only had 2 periods forecasts. We can assume that these forecasts are super beneficial for the country to take decisions and forecast the future. Finally, we can see that time series has a great role in the analysis of the phenomena; for example, our dataset on the FDI we can see this analysis beneficial for the countries to understand the situation and forecast future.

In conclusion, the time series analysis of FDI provided valuable insights into the trends, seasonality, cyclical patterns, and forecasting of FDI. The analysis highlighted the significance of FDI fluctuations over time and identified key factors influencing FDI flows. These findings contribute to a better understanding of FDI dynamics and can assist policymakers, investors, and stakeholders in making informed decisions related to FDI planning and investment strategies.

Appendix

Here are the accuracy measures when we assumed that the model is additive

```
Accuracy Measures

MAPE  14.7004
MAD    5.8784
MSD   60.2192
```

And here are the accuracy measures when we assumed that the model is multiplicative

```
Accuracy Measures

MAPE  13.1110
MAD    5.4602
MSD   55.3376
```

We can assume that the model is multiplicative as all accuracy measures are least when multiplicative.

-Our dataset: From google trends : Foreign Direct Investment in the USA

Link :

https://trends.google.com/trends/explore?date=2010-01-01%202023-01-01&geo=US&q=%2Fm%2F02n7nb&hl=en

References:

-Countries with highest foreign direct investment (FDI) position in the United States in 2021

https://www.statista.com/statistics/456713/leading-fdi-countries-usa/

-FDI Strong: The U.S. Tops the World in Foreign Direct Investment

https://www.gray.com/insights/fdi-strong-the-u-s-tops-the-world-in-foreign-direct-investment/

-U.S. FOREIGN DIRECT INVESTMENT (FDI) STOCK

https://www.trade.gov/data-visualization/us-foreign-direct-investment-fdi-stock

-Our dataset: From google trends : Foreign Direct Investment in the USA

Link :

https://trends.google.com/trends/explore?date=2010-01-01%202023-01-01&geo=US&q=%2Fm%2F02n7nb&hl=en