Analysis of Stock Prices using Time Series Model

Shaswata Banik

Sem-VI

Roll No: 3-14-21-0440

Reg No: A01-1112-0853-21

Introduction



Data Description

Daily data on stock prices on trading days (namely date, opening, high, low, closing, adjusted closing prices & volume) of -

- Amazon (Jan,1999 Dec,2022)
- Microsoft (Jan, 2000 Nov, 2023)
- Netflix (May,2002 July,2020)

Sources:

- https://www.kaggle.com/datasets/prajwaldongre/microsoft-stock-price2000-2023
- https://www.kaggle.com/datasets/aayushmishra1512/netflix-stockdata
- https://www.kaggle.com/datasets /sriharshaeedala/amazon-stockprice-from-1999-to-2022

4	Α	В	С	D	Е	F	G
1	Date	Open	High	Low	Close	Adj Close	Volume
2	23-05-2002	1.156429	1.242857	1.145714	1.196429	1.196429	104790000
3	24-05-2002	1.214286	1.225	1.197143	1.21	1.21	11104800
4	28-05-2002	1.213571	1.232143	1.157143	1.157143	1.157143	6609400
5	29-05-2002	1.164286	1.164286	1.085714	1.103571	1.103571	6757800
6	30-05-2002	1.107857	1.107857	1.071429	1.071429	1.071429	10154200
7	31-05-2002	1.078571	1.078571	1.071429	1.076429	1.076429	8464400
8	03-06-2002	1.08	1.149286	1.076429	1.128571	1.128571	3151400
9	04-06-2002	1.135714	1.14	1.110714	1.117857	1.117857	3105200
10	05-06-2002	1.110714	1.159286	1.107143	1.147143	1.147143	1531600
11	06-06-2002	1.15	1.232143	1.148571	1.182143	1.182143	2305800
12	07-06-2002	1.177857	1.177857	1.103571	1.118571	1.118571	1369200
13	10-06-2002	1.135	1.175	1.134286	1.156429	1.156429	484400
14	11-06-2002	1.156429	1.188571	1.128571	1.153571	1.153571	1003800
15	12-06-2002	1 152571	1 1821/13	1 029226	1.092857	1 092857	1799000
NFLX (+)							

Methodology

We'll take log returns of the stock prices and divide them in 8:2 ratio for training & testing purposes

ARIMA(p,d,q) fitting

Check Stationarity through ADF test and go on differencing until satisfactory results come and choose 'd' appropriately

We'll fit appropriate ARIMA model to the training dataset and plot the fitted & true values together.

We'll obtain error created from the forecasted values.

Choose 'p' and 'q' from PACF & ACF plots respectively

We'll obtain forecasted and upper & lower 95% CLs for testing dataset and plot them together. We'll explain how the obtained error values are behaving through several measures like MAD, MPE, MAPE.

ARCH(m) fitting

We'll plot the stationarity datapoints to check if volatility clustering is present in the dataset or not.

We'll obtain 'm' in ARCH from the PACF plot.

We'll obtain 'q'
parameter for moving
avgs. in ARMA, which is
taken as mean
equation here, from
the ACF plot.

We'll use ARCH Lagrange's Multiplier Test to test for ARCH effect.

If ARCH effect is present, we'll fit it to the training data and produce two SDs superimposed graph to look at how well volatility is in them

We'll find out the forecasted volatility series and plot them to see how they're varying over the time point in the testing data.

through Augmented-Dickey Fuller Test and go on differencing the data if unsatisfactory

result comes out.

Check for stationarity

Results

We'll look at the results obtained through analysis on log returns of opening prices of Amazon.

Stationarity Checking (through ADF test)

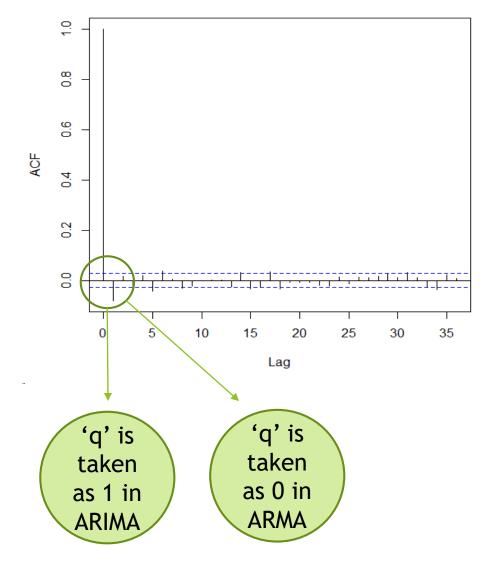
• p-value based on the true values: 0.2751

• p-value based on the values after 1st order differencing: 0.01

'd' is taken to be 1

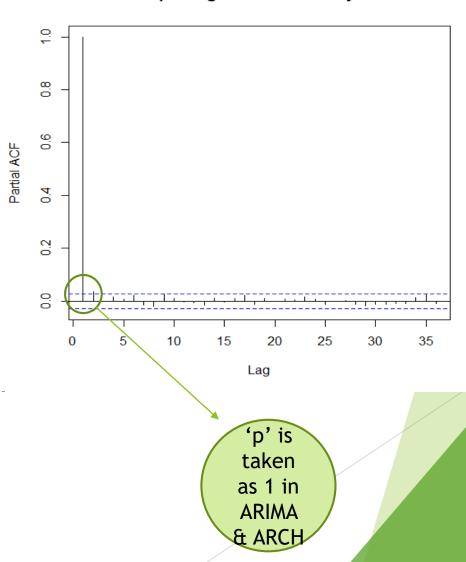
ACF Plot

ACF plotting on stationary series



PACF Plot

PACF plotting on non-stationary series

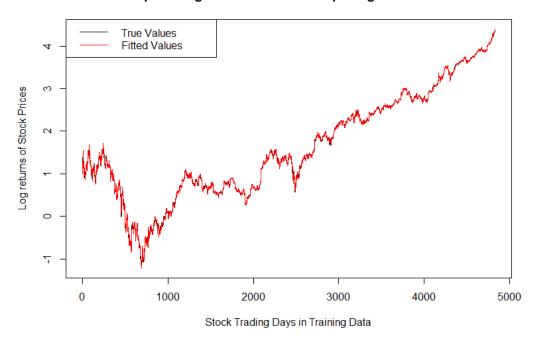


Fitting ARIMA(1,1,1) model

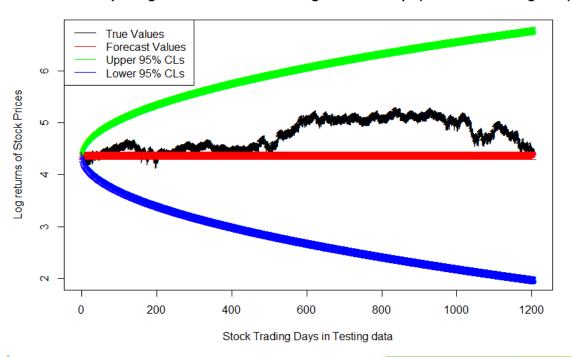
The fitted model is given by $(\hat{Y}_t - Y_{t-1}) = -0.6922(Y_{t-1} - Y_{t-2}) + e_t + 0.6304(e_{t-1})$ where Y_t is the Amazon's. opening stock prices' log returns at t.

Plotting forecasted & 95% CLs' values along with true testing data's values

Graph of Log returns of Amazon Opening Stock Prices



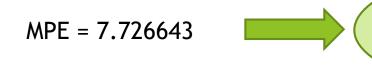
Amazon Opening & forecasted Prices' Log returns' Graph(based on Testing data)



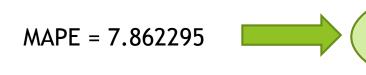
Measures on error based on the forecasted values



The forecasted values is on an average deviated from the real testing values in the absolute magnitude of 0.389801 units.



The average percentage deviation of the forecasted values from the real testing values is of 7.726643.

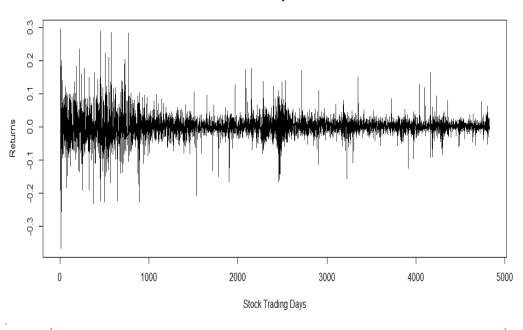


The average percentage absolute deviation of the forecasted values from the real testing values is of 7.862295.

ARCH fitting

Checking for volatility clustering



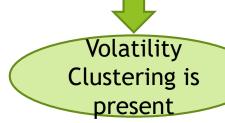


Fitting ARMA(1,0) as the mean equation

The model is given by, $r_t = \mu_t + a_t$ where,

 μ_t is the mean equation at time point t & μ_t = 0.0006239 + 1.0000467 $r_{t\text{-}1}$

 $\ensuremath{r_t}$ is the log return of the stock prices at time point t



Fitting linear regression eqn. to sq. shock values

We'll fit sq. shock values on its prev. m(=1) lag values

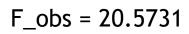


The equation is given by, $a_t^2 = 0.001351 - 0.008867a_{t-1}^2$; t=2,...,4831

Performing ARCH LM test for ARCH effect

$$SSR_0 = 0.1213478$$
 where $\bar{w} = 0.001351208$

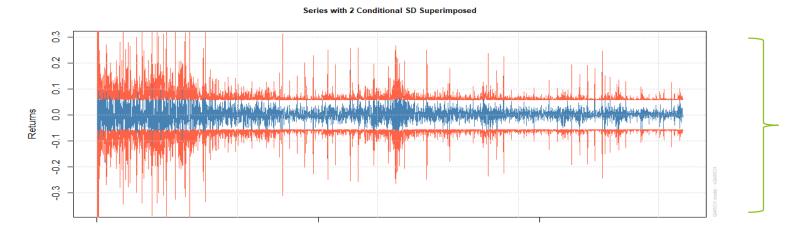
$$SSR_1 = 0.1208329$$



F_obs > $\chi^2_{m;\alpha} = 3.841459$ when $\alpha = 0.05$

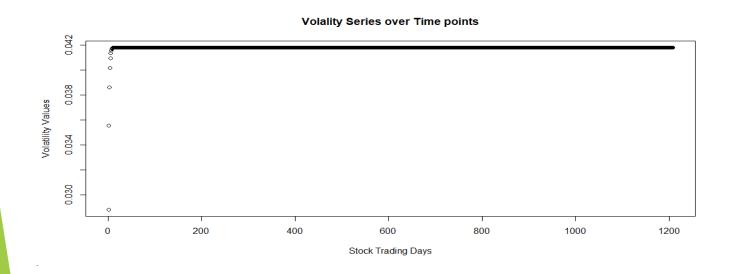
Null hyp. Is rejected at α l.o.s. So, ARCH effect is present

Plotting the series with two conditional SDs



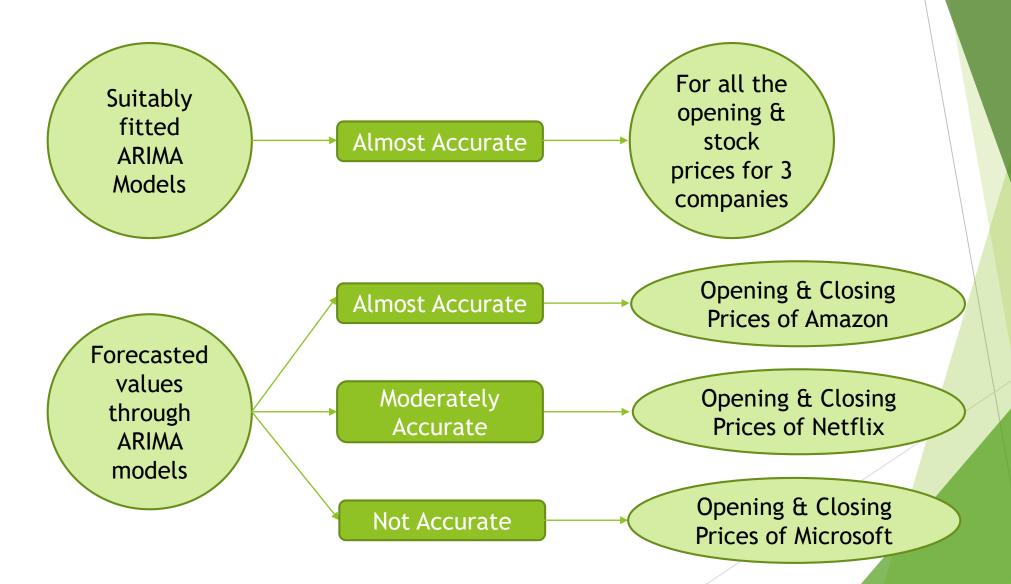
Most of volatility component is inside the interval formed by the SDs

Plotting forecasted volatility series over time points

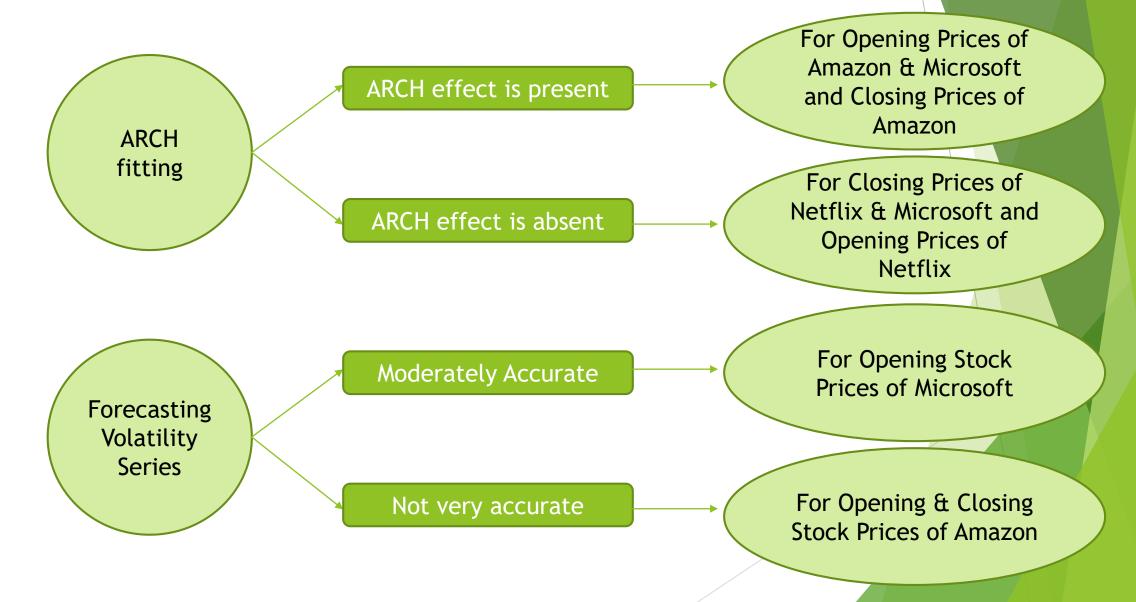


Volatility of stock prices increases for 13 days, then stay at same level.

Conclusions



Conclusions Contd.



Acknowledgement

Date: 04/05/2024

I, hereby, affirm that I have acknowledged any materials that I've used in my dissertation from outside sources. So, to the best of my knowledge, any parts of my dissertation paper doesn't use any unacknowledged materials.

Shaswala Banik

Shaswata Banik Student (3rd Year) Dept. of Statistics

St. Xavier's College (Autonomous), Kolkata

References

- https://www.businessnewsdaily.com/2871-how-most-millionaires-got-rich.html
- https://www.visualcapitalist.com/cp/stock-market-and-gdp-shares/
- https://help.indiainfoline.com/portal/en/kb/articles/what-are-open-high-low-prev-close-prices
- https://www.idrisstsafack.com/post/garch-models-with-r-programming-a-practical-example-with-tesla-stock
- https://saturncloud.io/blog/what-are-logarithmic-returns-and-how-to-calculate-them-in-pandas-dataframe/#:~:text=Logarithmic%20returns%20are%20important%20in,over%20a%20period%20of%20time.
- https://en.wikipedia.org/wiki/Volatility_clustering#:~:text=In%20finance%2C%20volatility%20clustering%20refers,be%20followed%20by%20small%20changes.%22
- https://www.youtube.com/watch?v=Li95a2biFCU
- https://www.youtube.com/watch?v=x81kTJ1afA8&list=PLzAfHlPtM1I4bPhNE5FKOi 0S7erZqm5pg
- ► Tsay, R. S. (2005) Analysis of Financial Time Series, 2nd ed., Wiley, New York.(https://cpb-us-w2.wpmucdn.com/blog.nus.edu.sg/dist/0/6796/files/2017/03/analysis-of-financial-time-series-copy-2ffgm3v.pdf)

THANK YOU!!