Predicting Stock Prices Using Forecasting Feature Analysis

Kallam Ramesh Babu

kallamrameshbabu@gmail.com

Table of Contents

1. Introduction

Everyone knows how the stock is volatile and nonlinear. The stock price and its trend change due to the various external factors such as global politics and economy and a company's financial performance. But there are a number of methods and systems to predict the future stock prices by using lots of data. Microsoft excel makes it very easy for us to build the stock prediction. In this report I used the various Data Analytics methods and techniques like "Time Series Forecasting" and "Monte Carlo Simulation". Apart from these traditional approaches, I compare the stock monthly performance of any combination of them as well by using regression methods.

The objectives of the analysis using the historical stock prices are as follows,

1. To find the future values of the stock price using various variables from the dataset, validate the data tuning the hyper parameters, select the best model and evaluate the model performance.

2. Identify the price range of a stock in a particular month.

3. Predict stock performance over other stocks to invest and gain the good returns in the future.

2. Datasets

2.1 Dataset Description

For this case study, I am using Tesla, Ford, GM, Volkswagen historical stock price data from 31-01-2019 to 30-09-2021. The dataset contains 7 columns and 690 records.

1. Column 1 – Date

This column represents the date of stock price. This will help us identify the closing price of the stock on any particular day.

2. Column 2 – High

This column represents the highest price at which a stock traded during the course of a particular trading day.

3. Column 3 – Low

This column represents the lowest price at which a stock traded during the course of a particular trading day.

4. Column 4 – Open

This column represents the opening price of a stock i.e., the first transaction of a business day. It can help us to analyse the volatility of the stock.

5. Column 5 – Close

This column represents the closing price of a stock i.e., the last transaction of a business day.

6. Column 6 – Volume

This column represents the number of shares that have been traded in the security exchange for that particular day.

7. Column 7 – Adjacent Close

This column represents the adjacent closing price of the stock i.e., a stock's closing price to reflect that stock's value after accounting for any corporate actions.

2.2 Data Cleaning

For regression method, I deleted the open price, close price, volumes, High and Low and stored the remaining content in a separate excel sheet to find out the one stock price monthly performance with the other three stocks.

3. Data Analysis

3.1 Introduction to Time Series Forecasting

Most of the time in our lives we work and deal with time series data almost every day. In our daily life, we every so often make decisions about certain things based on our past observations and experiences.

Let's take an example, if the stock price of a specific company has been increasing consistently over the last 7 days, we can assume that the price will rise tomorrow too. Or if it has been raining every day for the past week, we can guess that it would rain today as well, and hence it's a good idea to carry a rain coat or umbrella.

The above examples tell us that recent past could give us a good idea about the future events. This is the main idea behind time series forecasting.

In a time series, each individual point is dependent on the previous value. Thus, we can use past values and estimate the values in the future. The "time" component is crucial here.

3.2 A Quick Look at the Different Time Series Components

To understand the exponential smoothing models and how they forecast future values, we must be aware of the different time series components. The time series has the following three components:

1. **Trend Component**
   The trend labels the general tendency of the data which could be increasing or decreasing or stable. For instance, at the time of covid-19 outbreak, we observed a decreasing trend for stock prices.

   The trend often represents the long-term movement of the series.

2. **Seasonal Component**

The next significant component is the seasonal component of the time series. For example, there could be a higher sale in clothing items and sweets around New Year's or Christmas every year. Likewise, there could be an increase in bus or train bookings around the holiday season(s). And this pattern could be observed throughout the year.

To some extent it's difficult to find the seasonality in the first series. But the seasonality of the second series is evident. We can have a particular pattern repeating every year, every week or every month which shows that we have a yearly, weekly or monthly seasonality for the second series.

3. **Residual Component**

Let's say we recognize the trend and seasonal component from a time series and remove these two. What remains after removing these two is the residual component. It does not have any pattern or trend. As the name suggests, the residual component is irregular.

Now that we got to know about the different components of a time series, let's find out how exponential smoothing algorithms use these to make predictions.

Exponential smoothing algorithms are popularly used for forecasting univariate time series. There are three types of exponential smoothing algorithms:

- Simple exponential smoothing
- Double exponential smoothing
- Triple exponential smoothing

3.2.1 Simple Exponential Smoothing

We found out that the data points in a time series will always depend on each other. Hence, we can use historical data to make forecasts for the future value.

The simple exponential smoothing model ponders the historical values and assigns weights to these values. We need to remember that weights are always higher for recent observations.

Look at the mathematical equations:

$$X_t = \alpha Y_t + \alpha(1-\alpha) Y_{t-1} + \alpha(1-\alpha)^2 Y_{t-2} + \alpha(1-\alpha)^3 Y_{t-3} + \ldots$$

Where,

- ➔ Y_t represents the historical values
- ➔ X_t is the forecast
- ➔ alpha α is the smoothing parameter

The alpha (α) value lies between 0 - 1. Alpha is a hyperparameter and we can choose the alpha value based upon the historical records we have. The table below will help us understand how changing the alpha value affects the forecasts:

Simple Exp Smoothing	Alpha	Simple Exp Smoothing	Alpha	Simple Exp Smoothing	Alpha
7.2842803	**0.4**	7.2842803	**0.2**	7.2842803	**0.9**
7.2842803		7.2842803		7.2842803	
7.24002142		7.26215086		7.184697819	
7.324113483		7.299771004		7.423696203	
7.452021954		7.368593735		7.621865813	

A greater number of historical values are considered for the forecast if we choose a lower alpha value. If we choose the alpha at a higher value, such as 0.8 or 0.9, very few observations are taken into consideration. I have chosen **alpha (α) value as 0.2** as I have more historical data.

Now, if we use the same equation for the second forecast, it will be:

$$Xt = \alpha Yt+1 + (1-\alpha)\ [Xt-1]$$

Likewise, we can write this equation for the other forecasts to predict the values. We have to understand that each new term has an additional (1-alpha).

Look at the below image how we implemented the formula in the excel sheet.

=($J$2*G10+(1-$J$2)*I10)

D	E	F	G	H	I	J
Open	**Close**	**Volume**	**Adj Close**	**Actual Close**	**Simple Exp Smoothing**	**Alpha**
7.53000021	7.900000095	47494400	7.2842803		7.2842803	**0.2**
7.96999979	7.78000021	39172400	7.173633099		7.2842803	
7.909999847	8.079999924	43039800	7.450251579		7.26215086	
8.109999657	8.289999962	40729400	7.643884659		7.299771004	
8.420000076	8.369999886	45644000	7.71764946		7.368593735	
8.449999809	8.720000267	48404900	8.040370941		7.43840488	
8.710000038	8.670000076	39490400	7.994268417		7.558798092	
8.770000458	8.819999695	41559900	8.132575989		7.645892157	
8.800000191	8.989999771	44833800	8.289326668		7.743228923	
9.020000458	8.840000153	65311700	8.151017189		7.852448472	
8.680000305	8.289999962	73869900	7.643884659		7.912162216	

Below Image shows the forecast result. The Grey line is forecast and remaining blue and orange lines are trained and validation values.

If we observe that we have a flat line at the end that's because simple exponential smoothing considers only historical data and trends are not included to forecast the values. This can be resolved using double exponential smoothing.

3.2.2 Double Exponential Smoothing

The double exponential smoothing (DES) algorithm works as same like as simple exponential smoothing. It is also using historical values for making the predictions and assign weights in an exponentially increasing manner (higher weight to recent observations as same like Simple Exponential Smoothing). Moreover, double exponential smoothing also considers the **"trend"** of the series.

Forecast (DES) = Level + Trend

'Level' is calculated by the weighted average of the historical data, this is the same method that we followed for simple exponential smoothing.

Mathematical equation,

$$\mathbf{L_{t+1} = \alpha L_t + (1-\alpha)\ [L'_t]}$$

In DES we have another component that is the trend. The "Trend" is calculated as below:

$$T_{t+1} = \beta(L_{t+1} - L_t) + (1-\beta)\ T_t$$

The beta here is a smoothing parameter for the trend component. The simple exponential smoothing alpha (α) is the smoothing parameter. The trend at any given time is calculated to be the difference between the level terms (Lt). In order to consider the weighted sum of past trend values, we use (1-β) Tt where Tt is the trend calculated for the previous time step.

So that the final forecast will be Xt = Lt + Tt. Let's look at the images below on how we use these equations in Excel!

=($J$2*G8+(1-$J$2)*(L7+M7))

G	H	I	J	K	L	M	N
Adj Close	**Actual Close**	**Simple Exp Smoothing**	**Alpha**	**Double Exp Smoothing**	**Level**	**Trend**	**Beta**
7.2842803		7.2842803	**0.2**	8.4842803	7.2842803	1.2	**0.0**
7.173633099		7.2842803		9.409044388	8.22215086	1.186893528	
7.450251579		7.26215086		10.18459143	9.017285826	1.1673056	
7.643884659		7.299771004		10.8183486	9.676450073	1.141898532	
7.71764946		7.368593735		11.30910032	10.19820878	1.110891541	
8.040370941		7.43840488		11.73355869	10.65535444	1.078204247	
7.994268417		7.558798092		12.02651198	10.98570063	1.040811344	
8.132575989		7.645892157		12.24959677	11.24772478	1.001871984	
8.289326668		7.743228923		12.41981203	11.45754275	0.962269283	
8.151017189		7.852448472		12.4856344	11.56605306	0.919581335	
7.643884659		7.912162216		12.38844829	11.51728445	0.871163838	
7.708428383		7.858506704		12.27680794	11.45244431	0.824363639	
7.911282539		7.82849104		12.18441125	11.40370286	0.780708385	
7.837516785		7.84504934		12.0522718	11.31503236	0.73723944	
7.689987659		7.843542829		11.87343157	11.17981497	0.693616599	

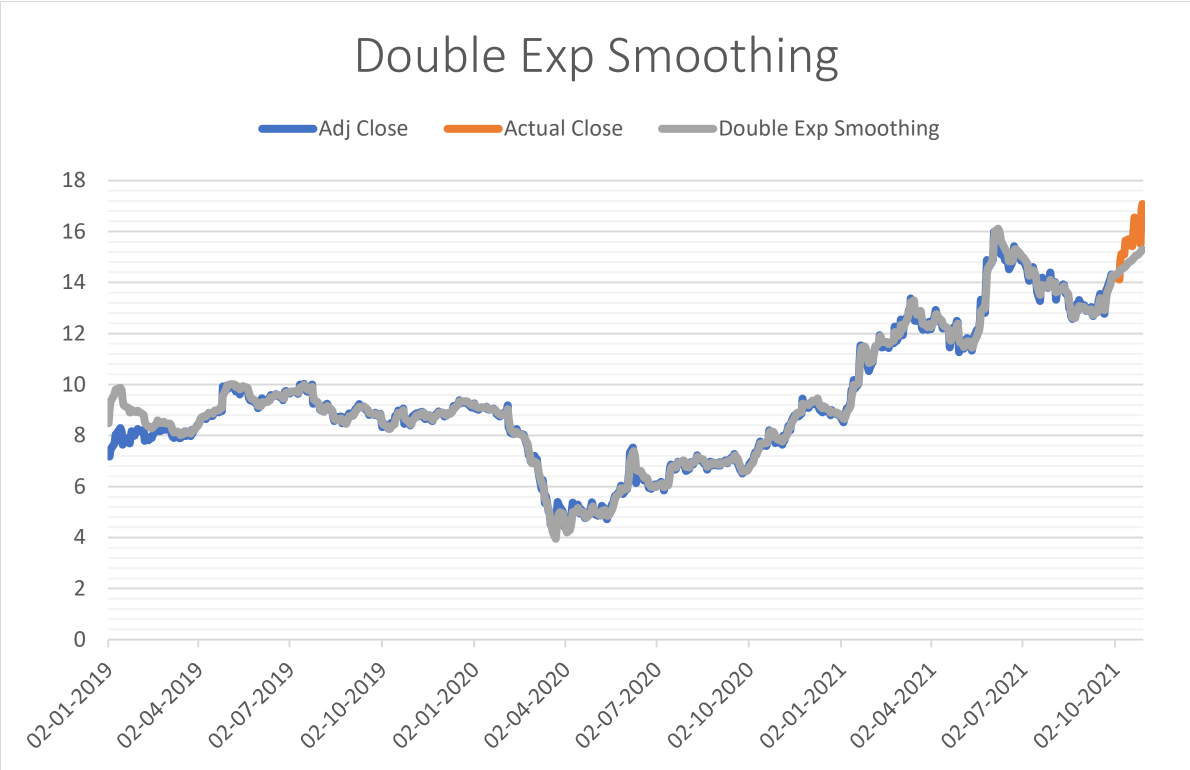
=($N$2*(L8-L7)+(1-$N$2)*M7)

G	H	I	J	K	L	M	N
Adj Close	**Actual Close**	**Simple Exp Smoothing**	**Alpha**	**Double Exp Smoothing**	**Level**	**Trend**	**Beta**
7.2842803		7.2842803	**0.2**	8.4842803	7.2842803	1.2	**0.05**
7.173633099		7.2842803		9.409044388	8.22215086	1.186893528	
7.450251579		7.26215086		10.18459143	9.017285826	1.1673056	
7.643884659		7.299771004		10.8183486	9.676450073	1.141898532	
7.71764946		7.368593735		11.30910032	10.19820878	1.110891541	
8.040370941		7.43840488		11.73355869	10.65535444	1.078204247	
7.994268417		7.558798092		12.02651198	10.98570063	1.040811344	
8.132575989		7.645892157		12.24959677	11.24772478	1.001871984	
8.289326668		7.743228923		12.41981203	11.45754275	0.962269283	
8.151017189		7.852448472		12.4856344	11.56605306	0.919581335	
7.643884659		7.912162216		12.38844829	11.51728445	0.871163838	
7.708428383		7.858506704		12.27680794	11.45244431	0.824363639	
7.911282539		7.82849104		12.18441125	11.40370286	0.780708385	
7.837516785		7.84504934		12.0522718	11.31503236	0.73723944	
7.689987659		7.843542829		11.87343157	11.17981497	0.693616599	
7.929722786		7.812831795		11.73886932	11.08468981	0.654179511	
8.169458389		7.836209993		11.64347254	11.02498713	0.618485401	
7.985046864		7.902859672		11.49368855	10.9117874	0.581901145	
8.077253342		7.919297111		11.3581383	10.81040151	0.547736793	

=L8+M8

I	J	K	L	M	N
Simple Exp Smoothing	**Alpha**	**Double Exp Smoothing**	**Level**	**Trend**	**Beta**
7.2842803	**0.2**	8.4842803	7.2842803	1.2	**0.05**
7.2842803		9.409044388	8.22215086	1.186893528	
7.26215086		10.18459143	9.017285826	1.1673056	
7.299771004		10.8183486	9.676450073	1.141898532	
7.368593735		11.30910032	10.19820878	1.110891541	
7.43840488		11.73355869	10.65535444	1.078204247	
7.558798092		12.02651198	10.98570063	1.040811344	
7.645892157		12.24959677	11.24772478	1.001871984	
7.743228923		12.41981203	11.45754275	0.962269283	
7.852448472		12.4856344	11.56605306	0.919581335	
7.912162216		12.38844829	11.51728445	0.871163838	
7.858506704		12.27680794	11.45244431	0.824363639	
7.82849104		12.18441125	11.40370286	0.780708385	
7.84504934		12.0522718	11.31503236	0.73723944	
7.843542829		11.87343157	11.17981497	0.693616599	

If we observe here, we have an increasing line. This is because DES considers the trend and historical data to make the forecasts. This can be adjusted by using Triple Exponential Smoothing or Holt Winters.



3.2.3 Triple Exponential Smoothing

Triple Exponential Smoothing is also known as **"Holt Winter's Algorithm".** Takes the seasonal component of the time series into consideration along with the Level and Trend.

Using TES, we can predict the final forecast in a multiplicative form.

I.e., Xt+1 = (Level + Trend) x Seasonality

Equation of levels can be found with the seasonality adjusted observations (Yt/st-m), along with the previous level and trend values.

$$Lt = \alpha(Yt/St\text{-}m) + (1\text{-}\alpha)\ [Lt\text{-}1 + Tt\text{-}1]$$

The equation of trends is similar to the double exponential smoothing model. Here's the equation:

$$Tt+1 = \beta\ (Lt+1\text{-}\ Lt) + (1\text{-}\beta)\ Tt$$

Finally, we have a seasonal component. This is the seasonal value at the particular time step t and the seasonal value at the t-m step. Here is how we can calculate the final value:

$$S_t = \gamma\,(Y_t / L_t) + (1-\gamma)\, S_{t-m}$$

Let's have a look at how we implement these formulas in excel.

=($J$2*(G9/O2)+(1-$J$2)*(R8+S8))

G	H	I	J	K	L	M	N	O	P	Q	R	S
Adj Close	Actual Close	Simple Exp Smoothing	Alpha	Double Exp Smoothing	Level	Trend	Beta	Seasonal	Holt Winters	Gamma	Level - Lt	Trend - Tt
31.89335632		31.89335632	0.5	33.09335632	31.89	1.2	0.06	0.988166		0.9		
30.57553101		31.89335632		32.95890891	31.83	1.124		0.947335				
31.59945488		31.23444366		33.36286351	32.28	1.084		0.97906				
32.57597733		31.41694927		34.02949545	32.97	1.06		1.009316				
33.00261307		31.9964633		34.54532282	33.52	1.029		1.022535				
33.35340118		32.49953818		34.94287292	33.95	0.994		1.033403				
32.92676163		32.92646968		34.86784485	33.93	0.933		1.020185				
35.24956131		32.92661566		36.00318215	35.06	0.944		1.880717	35.55684233		17.83585	1.070151
35.70463181		34.08808848		36.78942953	35.85	0.936		1.323644	39.61856021		28.29777	1.633657
35.60034943		34.89636014		37.09473964	36.19	0.9		1.098989	38.43512184		33.1466	1.826567
35.71411514		35.24835479		37.26285881	36.4	0.858		1.023594	37.89111023		35.17882	1.838907
36.27347946		35.48123496		37.59691918	36.77	0.829		1.003044	38.15429357		36.2459	1.792597
36.60531616		35.87735721		37.90011962	37.1	0.799		0.997242	38.33837336		36.7303	1.714105
36.1692009		36.24133669		37.78173465	37.03	0.747		0.98073	37.83006853		36.949	1.624381
35.71411514		36.20526879		37.4329707	36.75	0.685		1.214856	36.22503766		28.7815	1.036868

=$N$2*(R9-R8)+(1-$N$2)*S8

	H	I	J	K	L	M	N	O	P	Q	R	S
se	Actual Close	Simple Exp Smoothing	Alpha	Double Exp Smoothing	Level	Trend	Beta	Seasonal	Holt Winters	Gamma	Level - Lt	Trend - Tt
35632		31.89335632	0.5	33.09335632	31.89	1.2	0.06	0.988166		0.9		
53101		31.89335632		32.95890891	31.83	1.124		0.947335				
45488		31.23444366		33.36286351	32.28	1.084		0.97906				
97733		31.41694927		34.02949545	32.97	1.06		1.009316				
51307		31.9964633		34.54532282	33.52	1.029		1.022535				
40118		32.49953818		34.94287292	33.95	0.994		1.033403				
76163		32.92646968		34.86784485	33.93	0.933		1.020185				
56131		32.92661566		36.00318215	35.06	0.944		1.880717	35.55684233		17.83585	1.070151
53181		34.08808848		36.78942953	35.85	0.936		1.323644	39.61856021		28.29777	1.633657
34943		34.89636014		37.09473964	36.19	0.9		1.098989	38.43512184		33.1466	1.826567
11514		35.24835479		37.26285881	36.4	0.858		1.023594	37.89111023		35.17882	1.838907
47946		35.48123496		37.59691918	36.77	0.829		1.003044	38.15429357		36.2459	1.792597
31616		35.87735721		37.90011962	37.1	0.799		0.997242	38.33837336		36.7303	1.714105
92009		36.24133669		37.78173465	37.03	0.747		0.98073	37.83006853		36.949	1.624381
11514		36.20526879		37.4329707	36.75	0.685		1.214856	36.22503766		28.7815	1.036868
57279		35.95969197		37.45323861	36.81	0.647		1.260952	37.24572673		28.5755	0.962296
75473		36.06918238		37.66632902	37.04	0.623		1.174907	38.19871275		31.43592	1.076184
10043		36.35146855		37.65145022	37.06	0.587		1.080782	38.08336413		34.06737	1.169499
57996		36.40728449		37.61338447	37.06	0.551		1.025003	37.92778794		35.79938	1.20325
39047		36.43993223		37.87166704	37.34	0.535		1.001957	38.36576599		37.08276	1.208058
02237		36.75016135		37.94129493	37.43	0.508		0.976235	38.26530239		38.00586	1.19096
47949		36.87209186		37.82709297	37.35	0.473		1.050385	37.4498196		34.73044	0.922977

asting | Monte Carlo Simulation

In TES, along with the Level and Trend we have to find the seasonal value using the below formula.

Alignment | Number | Styles | Cells

=$Q$2*(G9/R9)+(1-$Q$2)*O8

H	I	J	K	L	M	N	O	P	Q	R
Actual Close	**Simple Exp Smoothing**	**Alpha**	**Double Exp Smoothing**	**Level**	**Trend**	**Beta**	**Seasonal**	**Holt Winters**	**Gamma**	**Level - Lt**
	31.89335632	**0.5**	33.09335632	31.89	1.2	**0.06**	0.988166		**0.9**	
	31.89335632		32.95890891	31.83	1.124		0.947335			
	31.23444366		33.36286351	32.28	1.084		0.97906			
	31.41694927		34.02949545	32.97	1.06		1.009316			
	31.9964633		34.54532282	33.52	1.029		1.022535			
	32.49953818		34.94287292	33.95	0.994		1.033403			
	32.92646968		34.86784485	33.93	0.933		1.020185			
	32.92661566		36.00318215	35.06	0.944		1.880717	35.55684233		17.83585
	34.08808848		36.78942953	35.85	0.936		1.323644	39.61856021		28.29777
	34.89636014		37.09473964	36.19	0.9		1.098989	38.43512184		33.1466
	35.24835479		37.26285881	36.4	0.858		1.023594	37.89111023		35.17882
	35.48123496		37.59691918	36.77	0.829		1.003044	38.15429357		36.2459
	35.87735721		37.90011962	37.1	0.799		0.997242	38.33837336		36.7303
	36.24133669		37.78173465	37.03	0.747		0.98073	37.83006853		36.949
	36.20526879		37.4329707	36.75	0.685		1.214856	36.22503766		28.7815
	35.95969197		37.45323861	36.81	0.647		1.260952	37.24572673		28.5755
	36.06918238		37.66632902	37.04	0.623		1.174907	38.19871275		31.43592
	36.35146855		37.65145022	37.06	0.587		1.080782	38.08336413		34.06737
	36.40728449		37.61338447	37.06	0.551		1.025003	37.92778794		35.79938
	36.43993223		37.87166704	37.34	0.535		1.001957	38.36576599		37.08276
	36.75016135		37.94129493	37.43	0.508		0.976235	38.26530239		38.00586

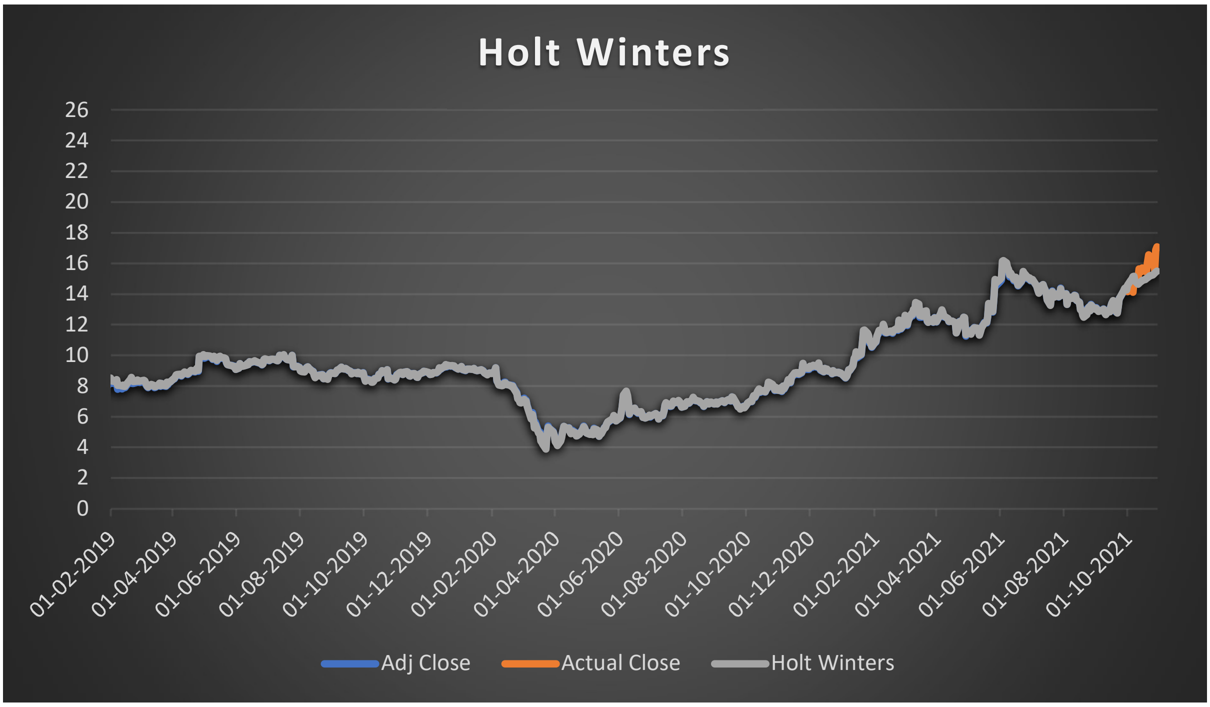
ont | Alignment | Number | Styles | Cells | Ed

f_x =(R9+S9)*O9

G	H	I	J	K	L	M	N	O	P	Q	R	S
Close	**Actual Close**	**Simple Exp Smoothing**	**Alpha**	**Double Exp Smoothing**	**Level**	**Trend**	**Beta**	**Seasonal**	**Holt Winters**	**Gamma**	**Level - Lt**	**Trend -**
39335632		31.89335632	**0.5**	33.09335632	31.89	1.2	**0.06**	0.988166		**0.9**		
57553101		31.89335632		32.95890891	31.83	1.124		0.947335				
59945488		31.23444366		33.36286351	32.28	1.084		0.97906				
57597733		31.41694927		34.02949545	32.97	1.06		1.009316				
00261307		31.9964633		34.54532282	33.52	1.029		1.022535				
35340118		32.49953818		34.94287292	33.95	0.994		1.033403				
92676163		32.92646968		34.86784485	33.93	0.933		1.020185				
24956131		32.92661566		36.00318215	35.06	0.944		1.880717	35.55684233		17.83585	1.0701
70463181		34.08808848		36.78942953	35.85	0.936		1.323644	39.61856021		28.29777	1.6336
50034943		34.89636014		37.09473964	36.19	0.9		1.098989	38.43512184		33.1466	1.8265
71411514		35.24835479		37.26285881	36.4	0.858		1.023594	37.89111023		35.17882	1.8389
27347946		35.48123496		37.59691918	36.77	0.829		1.003044	38.15429357		36.2459	1.7925
50531616		35.87735721		37.90011962	37.1	0.799		0.997242	38.33837336		36.7303	1.7141
.1692009		36.24133669		37.78173465	37.03	0.747		0.98073	37.83006853		36.949	1.6243
71411514		36.20526879		37.4329707	36.75	0.685		1.214856	36.22503766		28.7815	1.0368
17867279		35.95969197		37.45323861	36.81	0.647		1.260952	37.24572673		28.5755	0.9622
53375473		36.06918238		37.66632902	37.04	0.623		1.174907	38.19871275		31.43592	1.0761
46310043		36.35146855		37.65145022	37.06	0.587		1.080782	38.08336413		34.06737	1.1694
47257996		36.40728449		37.61338447	37.06	0.551		1.025003	37.92778794		35.79938	1.203
06039047		36.43993223		37.87166704	37.34	0.535		1.001957	38.36576599		37.08276	1.2080
99402237		36.75016135		37.94129493	37.43	0.508		0.976235	38.26530239		38.00586	1.190
76647949		36.87209186		37.82709297	37.35	0.473		1.050385	37.4498196		34.73044	0.9229

recasting | Monte Carlo Simulation

Plot that shows the Actual values and forecasts. From the plot, we can see that the forecasted values fit the tested data more than the other two algorithms.



3.3 Monte Carlo Simulation

In the previous model we conducted stock price prediction through time series forecasting, so for the second method I am interested in using a Monte Carlo Simulation to analyse the risk and decision making. Before making any decision, everyone will do risk analysis. Since we constantly faced a lot of ambiguity. Monte Carlo Simulation takes random possible outcomes and probabilities that can occur for any action. It calculates the possible outcomes over and over with a set of random numbers.

In this method I am studying on the possible outcome of the stock prices.

To do the Monte Carlo Simulation, we need to find out the Historical Volatility. To calculate the (μ) we simply average the closing price of the present and yesterday's values with log function LN.

➔ =Ln (P (t) / P (t-1)

Look at the below image how we entered in a excel,

=LN(B3/B2)

B	C	D	E	F
Adj Close	Actual Close			Historical Volatility (µ)
7.2842803				
7.173633099				-0.015306408
7.450251579				0.037835566
7.643884659				0.025658136
7.71764946				0.009603906
8.040370941				0.040965375
7.994268417				-0.005750382
8.132575989				0.017152886
8.289326668				0.01909102
8.151017189				-0.016826016
7.643884659				-0.064236791
7.708428383				0.008408388
7.911282539				0.025975585
7.837516785				-0.009367863
7.689987659				-0.019002869
7.929722786				0.030698899
8.169458389				0.029784537
7.985046864				-0.022831963
8.077253342				0.011481231
8.171065331				0.011547414

I found the historical volatility and need to find out the "Daily Variance", "Annualized Variance" and Standard Deviation. To compute all these, I need to find out the SQRT(µ) and average (µ) will give us the required Daily Variance.

For Annualized Variance we need to figure out the trading days that a stock market typically has, according to the past data a stock market can have the 250-254 trading days, so I have taken 252 trading days for our calculation.

By doing the SQRT of Annualized Variance we can get the Standard Deviation value.

Now that we have the trend of past daily results and the standard deviation (Volatility), we can now use the above values and can find the simulated stock prices by applying the below formula.

= L2 * (1 + $F$2 * (43) + $H$2 *SQRT (1/43) *NORMSINV (RAND ()))

=L657*(1+$F$652*(1/43)+$H$655*SQRT(1/43)*NORMSINV(RAND()))

C	D	E	F	G	H	I	J	K	L
Actual Close			Historical Volatility (μ)	Historical Volatility (μ)^2					
13.75				Mean Prices	13.73977128			6	12.55532282
13.81999969				Median Prices	12.63700927			7	11.96647427
13.93000031				Percentile				8	12.21781409
13.89999962				5%	6.327356175			9	12.72598409
13.59000015				95%	25.22730726			10	13.44553781
13.46000004				25%	9.362229632			11	13.11557974
12.98999977				75%	16.96303302			12	11.40029796
13								13	11.3209467
12.67000008								14	10.95891823
12.56999969								15	12.3143493
12.72999954								16	11.58294769
13.07999992								17	10.92040949
13.17000008								18	11.60546927
12.89999962								19	11.4935029
13.31000042								20	12.15674129
13.05000019								21	11.40378054
13.02999973								22	10.89214923
13.10999966								23	10.89683945
13.01000023								24	11.64441844
12.89000034								25	10.15502652
12.94999981								26	11.22590765
13.02999973								27	10.13231286

As I am going to predict the next 43 days of the stock prices, therefore dt = 1/43. My starting point will be the last closing price.

As I simulated the possible stock prices from the above equation, now I want to find out the range of the stock price that can be a closing price. For that I'm going to collect the data for the 43rd day outcomes and I will start calculating the statistics for what the price will look like in 43 days now since the outcomes.

I'm referencing the ending price of the 43rd day and to simulate 1000 possible outcomes. I used the data table feature to find the 1000 outcomes. Now, I can use this data and find the statistics.

Statistics to find out

Mean Price

Median Price

Percentile

5%

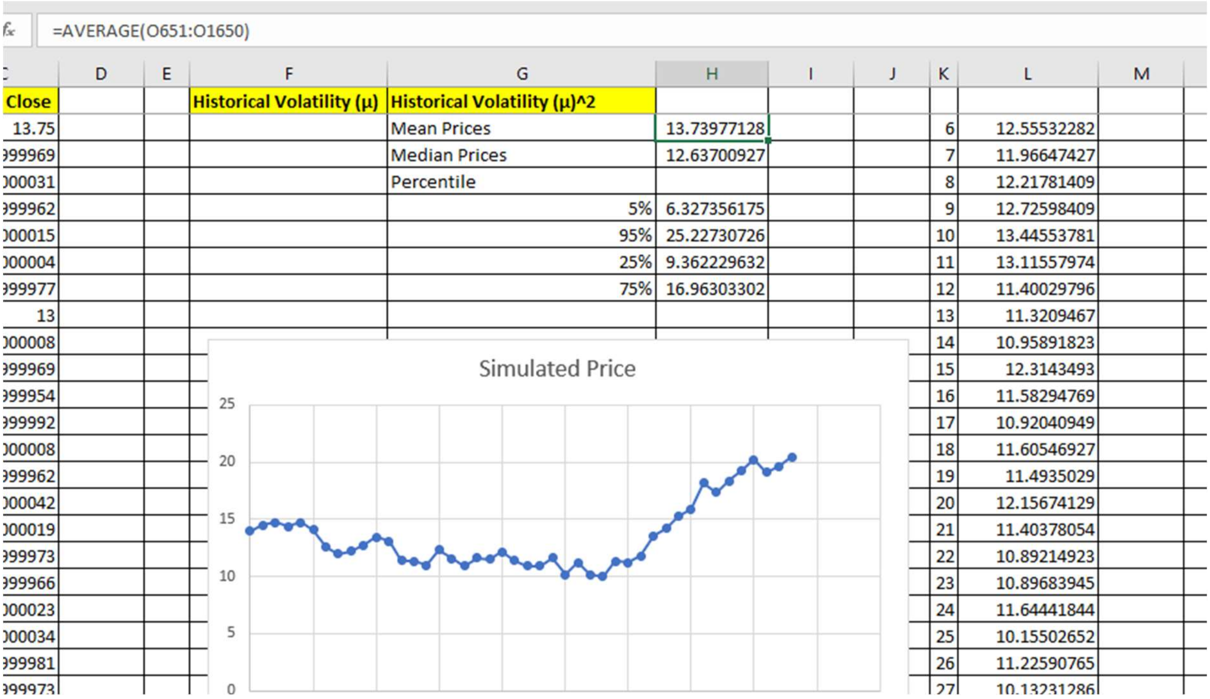
25%

75%

95%

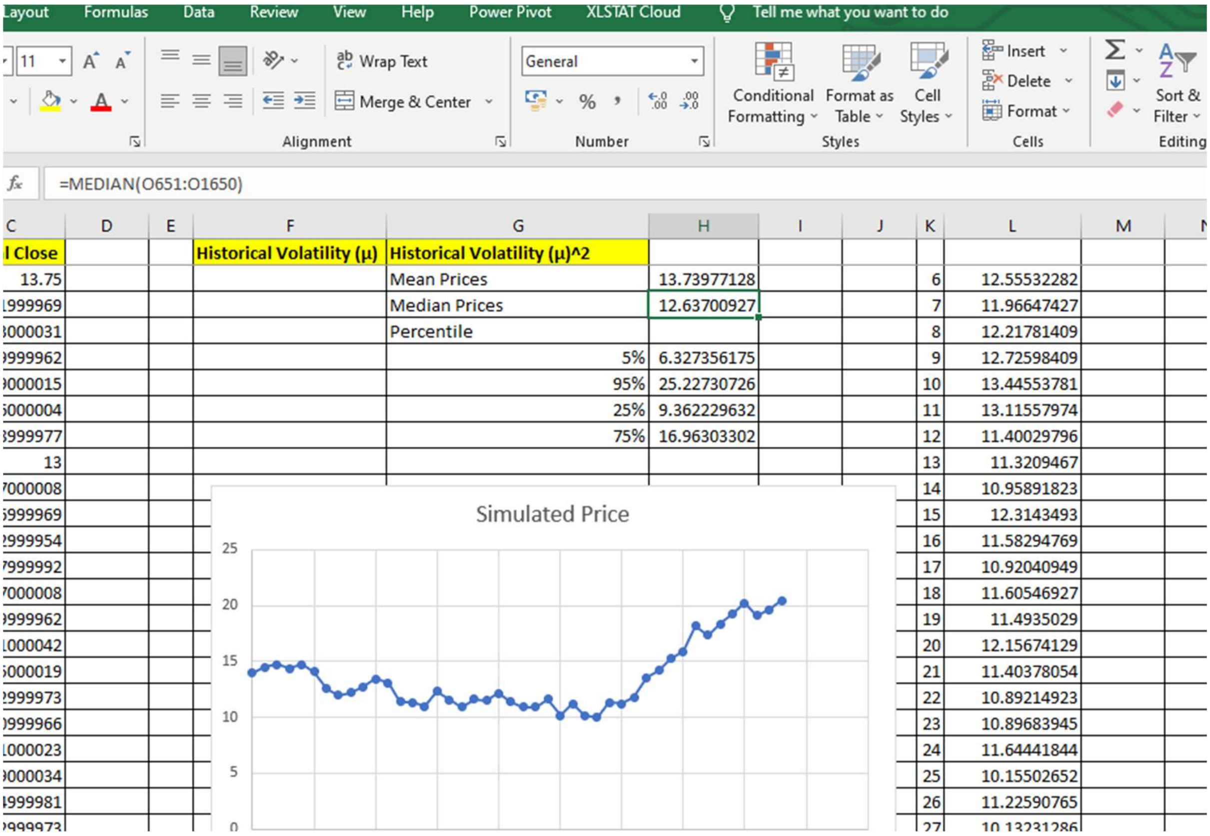
Mean Price can be found by doing an average of 1000 possible outcomes.

=Average (H1:H1000)



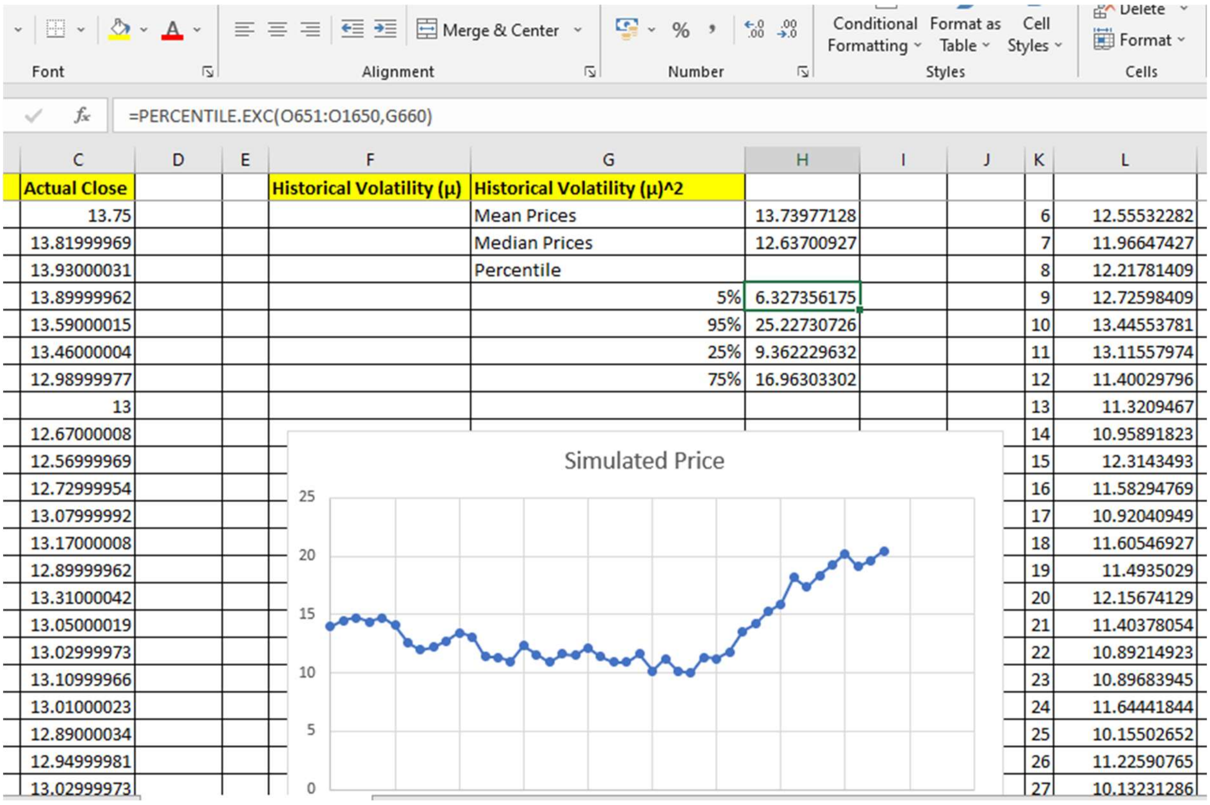
Median Price can be calculated by

=Median (H1:H1000)



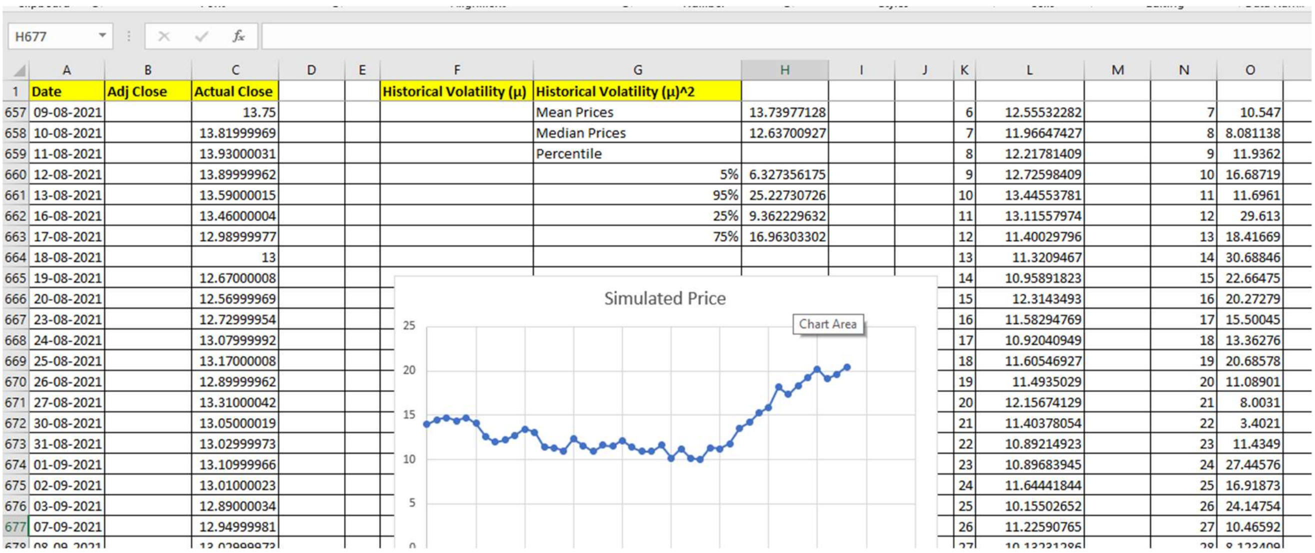
Percentile can be calculated by

=Percentile.EXC (H1:H1000,5%)



=PERCENTILE.EXC(O651:O1650,G660)

C	D	E	F	G	H	I	J	K	L
Actual Close			Historical Volatility (µ)	Historical Volatility (µ)^2					
13.75				Mean Prices	13.73977128			6	12.55532282
13.81999969				Median Prices	12.63700927			7	11.96647427
13.93000031				Percentile				8	12.21781409
13.89999962				5%	6.327356175			9	12.72598409
13.59000015				95%	25.22730726			10	13.44553781
13.46000004				25%	9.362229632			11	13.11557974
12.98999977				75%	16.96303302			12	11.40029796
13								13	11.3209467
12.67000008								14	10.95891823
12.56999969								15	12.3143493
12.72999954								16	11.58294769
13.07999992								17	10.92040949
13.17000008								18	11.60546927
12.89999962								19	11.4935029
13.31000042								20	12.15674129
13.05000019								21	11.40378054
13.02999973								22	10.89214923
13.10999966								23	10.89683945
13.01000023								24	11.64441844
12.89000034								25	10.15502652
12.94999981								26	11.22590765
13.02999973								27	10.13231286



H677

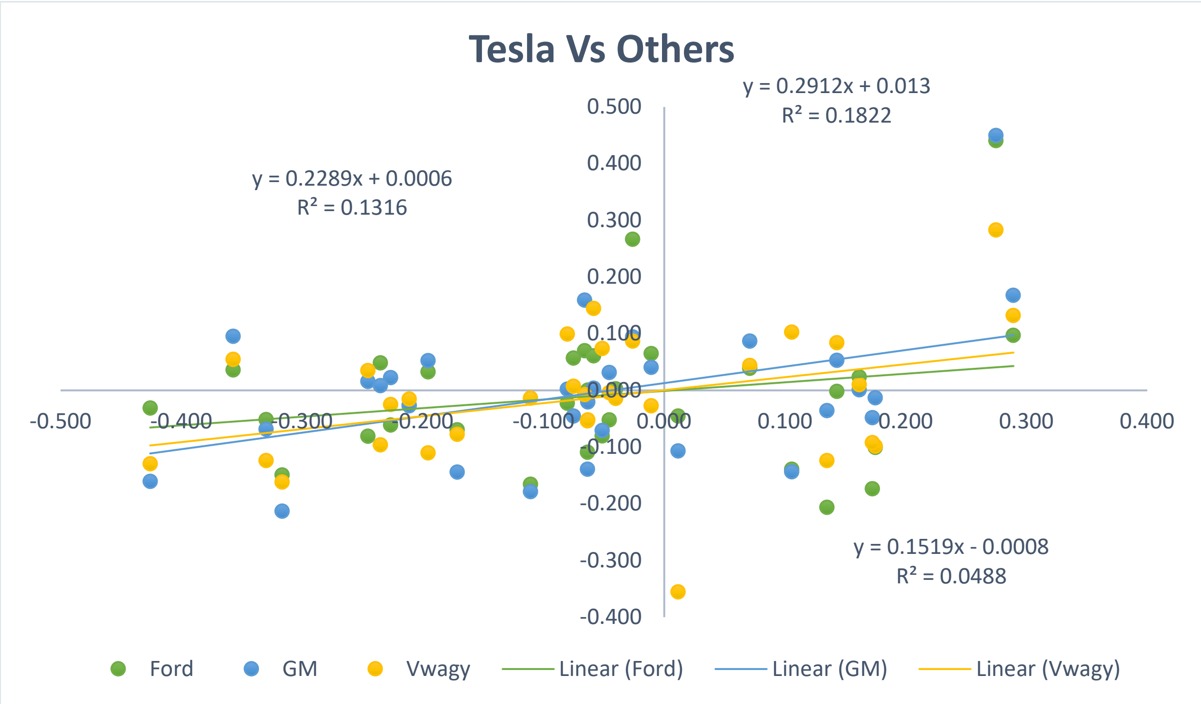
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	Date	Adj Close	Actual Close			Historical Volatility (µ)	Historical Volatility (µ)^2								
657	09-08-2021		13.75				Mean Prices	13.73977128			6	12.55532282		7	10.547
658	10-08-2021		13.81999969				Median Prices	12.63700927			7	11.96647427		8	8.081138
659	11-08-2021		13.93000031				Percentile				8	12.21781409		9	11.9362
660	12-08-2021		13.89999962				5%	6.327356175			9	12.72598409		10	16.68719
661	13-08-2021		13.59000015				95%	25.22730726			10	13.44553781		11	11.6961
662	16-08-2021		13.46000004				25%	9.362229632			11	13.11557974		12	29.613
663	17-08-2021		12.98999977				75%	16.96303302			12	11.40029796		13	18.41669
664	18-08-2021		13								13	11.3209467		14	30.68846
665	19-08-2021		12.67000008								14	10.95891823		15	22.66475
666	20-08-2021		12.56999969								15	12.3143493		16	20.27279
667	23-08-2021		12.72999954								16	11.58294769		17	15.50045
668	24-08-2021		13.07999992								17	10.92040949		18	13.36276
669	25-08-2021		13.17000008								18	11.60546927		19	20.68578
670	26-08-2021		12.89999962								19	11.4935029		20	11.08901
671	27-08-2021		13.31000042								20	12.15674129		21	8.0031
672	30-08-2021		13.05000019								21	11.40378054		22	3.4021
673	31-08-2021		13.02999973								22	10.89214923		23	11.4349
674	01-09-2021		13.10999966								23	10.89683945		24	27.44576
675	02-09-2021		13.01000023								24	11.64441844		25	16.91873
676	03-09-2021		12.89000034								25	10.15502652		26	24.14754
677	07-09-2021		12.94999981								26	11.22590765		27	10.46592

From the plot, we can say that in the next 43 days 95% we’ll see the price can be higher that 6.32 and at the same time 95% the price can be less than the 25.22.

4. Linear Regression

In this model, I calculated the monthly returns of each stock, to calculate the performance of the stock price of Tesla. From the plot, we can see that the tesla stock price performance with the other competitor stocks. Later, I picked the Ford stock price to calculate its performance with the other competitor stocks and so on.

I compute the monthly returns of each stock. I used the scatter plot to plot the "Tesla Stock" monthly returns against the other 3 stocks and repeated the process for the other 3 stocks.



In the above plot, we can see that the trendline, this trendline or regression line gives us the information on how tesla varies with other stocks. Regression line is sloping upward so it tells us that as Tesla returns increase along with the other stocks. If we observe carefully, there is huge variation in returns and that tells us the stock price varies for several other reasons and that can be found out by the equation presented there.

[y=0.2912x + 0.013

R^2 = 0.1822]

The equation here Y represents the Tesla returns and x represents the other stock returns. For instance, 1% increase in GM returns, which is a 0.3042 increase in Tesla's price.

5. Conclusion

The historical data from the year 31-10-2019 to 30-09-2021 were taken into account for analysis. The time forecasting series methodology is trained on the level, trend and seasonality of the stock price and predicts the future stock prices on the test datasets. In the next model I used the Monte Carlo Simulation to simulate that uncertainty the affects the stock price and then I calculated a representative value given those possible values of the underlying inputs.

In the last model linear regression, I used the monthly returns of one company's stock with the other 3 competitor's stock price to predict the price trend and correlation between them.

In conclusion, I only used historical data to predict the price, and we can't create a loop for further values since we do not use the predicted values as our data or features.

5.1 Remarks for future analysis

For future analysis, if we have the data related to the financial news like company quarter results, on diverse political and economic factors, change of leadership, investor sentiment, and many other factors could help us in formulating a better model. Along with the data, advanced data models like vector mechanism, naive Bayers regression and deep learning. We can improve the accuracy of stock price predictions by collecting the huge amount of time series data and analyze that with the related to news articles, using deep learning models.

6. References

1. https://github.com/Austin-Li-1123/Tesla_forcasting_feature_analysis

2. https://finance.yahoo.com/lookup