USD/INR Exchange Rate Prediction

Project Goal

- FORECASTING AND UNDERSTANDING DIRECTIONS AND RISKS
- STUDY THE BEHAVIOR OF EXCHANGE RATES OF USD AND INR AND FORECAST THE EXCHANGE RATES FOR THE NEAR FUTURE

Index

- Introduction
- Data
- Model
- Conclusion

Introduction

- The time series data we analyse is the USD/INR Exchange rate from Jan 01, 2010 to Dec 31, 2019

- The data source is: www.investing.com

- In this project, we expect to fit a forecasting model for the daily USD/INR exchange rate.

Data

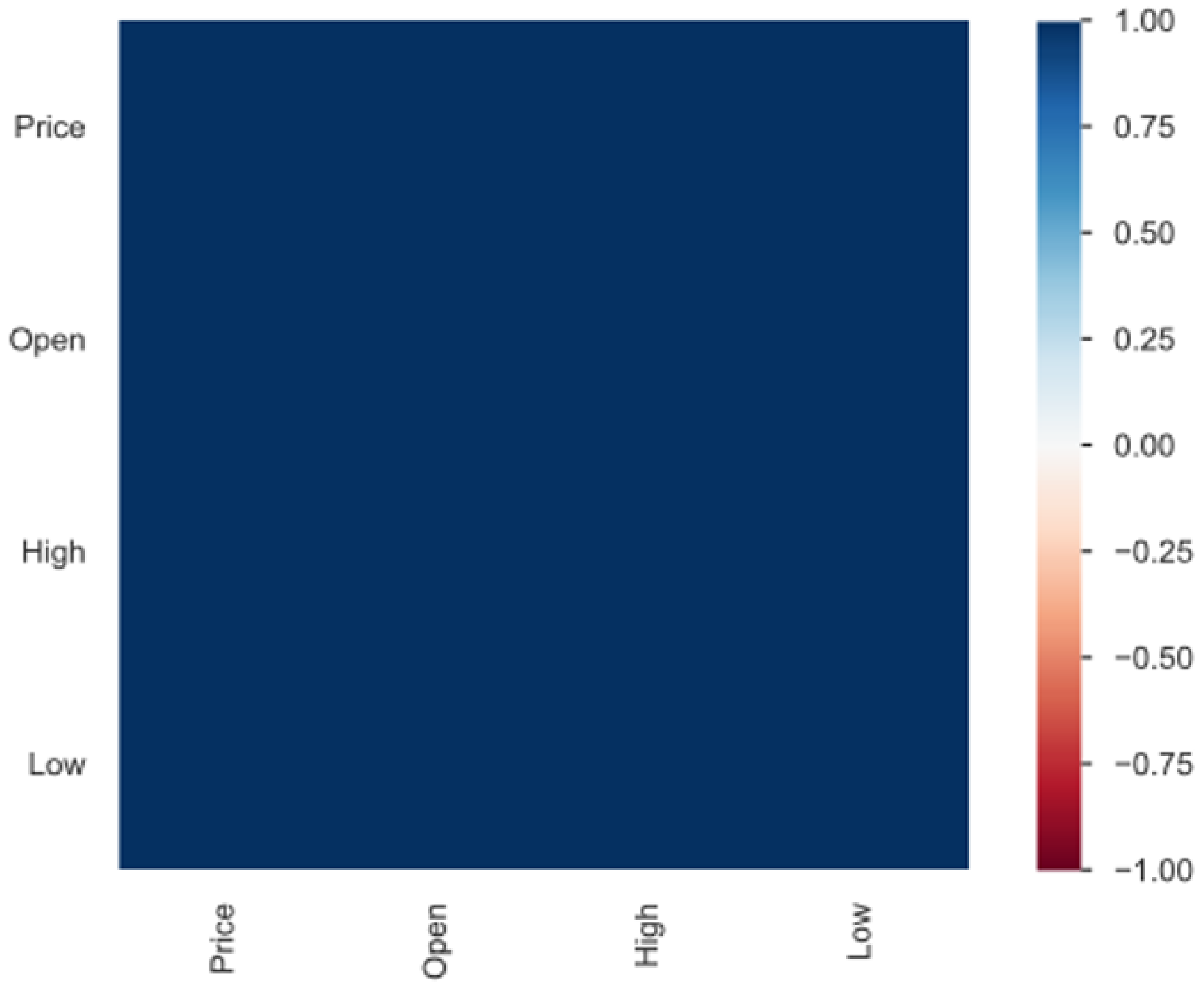
- **Variables of the dataset**
  - Date
  - Price
  - Open
  - High
  - Low
- **Number of Observations: 2608**

	Date	Price	Open	High	Low	Change %
0	Dec 31, 2019	71.35	71.295	71.385	71.225	0.06%
1	Dec 30, 2019	71.31	71.340	71.427	71.290	-0.18%
2	Dec 27, 2019	71.44	71.315	71.505	71.175	0.21%
3	Dec 26, 2019	71.29	71.270	71.348	71.225	0.01%
4	Dec 25, 2019	71.28	71.280	71.280	71.280	0.01%

Data

- High correlation between all variables
  - Univariate analysis

- No missing values



Models

- Linear Regression

- Time Series Forecasting
  - ARMA
  - ARIMA

Linear Regression

- Built a simple linear Regression model to predict exchange rate with lagged exchange rate
- Added new variable 'Lag_1' which is has the exchange rates of previous day
- Split the dataset into train (Jan 2010 - Dec 2017) and test (Jan 2018 - Dec 2019) data
- Input Variable -> Lagged Price
- Output Variable -> Price

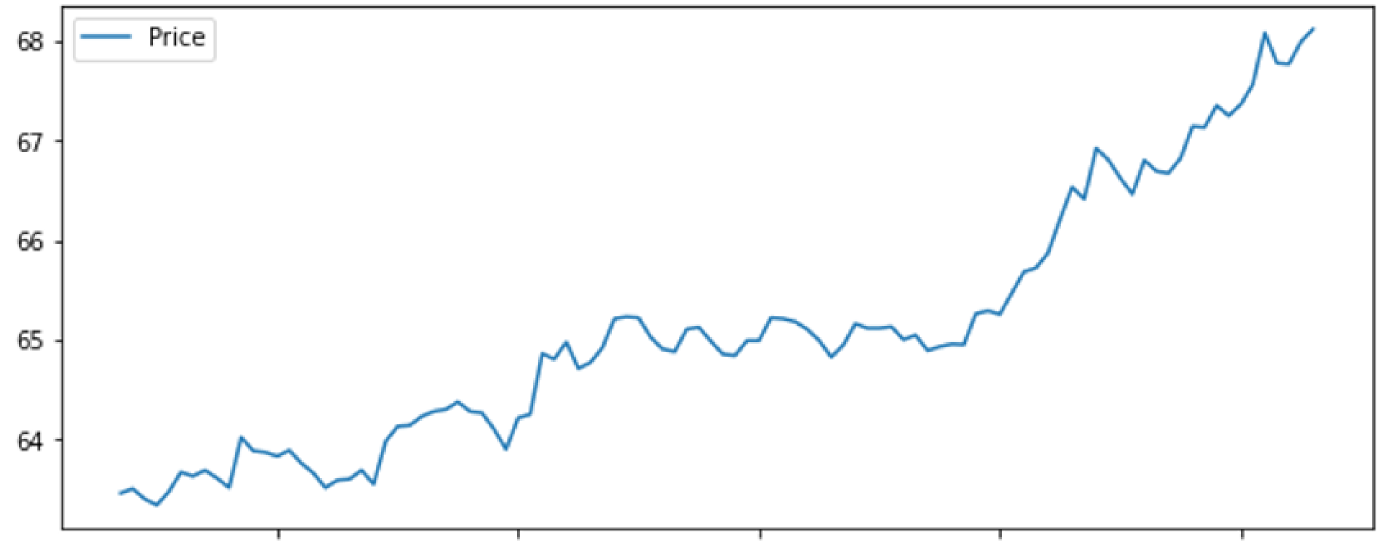
	Date	Price	Open	High	Low	Change %	Lag_1
2	Jan 05, 2010	46.205	46.305	46.305	46.045	-0.19	46.295
3	Jan 06, 2010	45.695	46.165	46.205	45.695	-1.10	46.205
4	Jan 07, 2010	45.650	45.610	45.890	45.570	-0.10	45.695
5	Jan 08, 2010	45.470	45.680	45.900	45.470	-0.39	45.650
6	Jan 11, 2010	45.260	45.510	45.510	45.230	-0.46	45.470

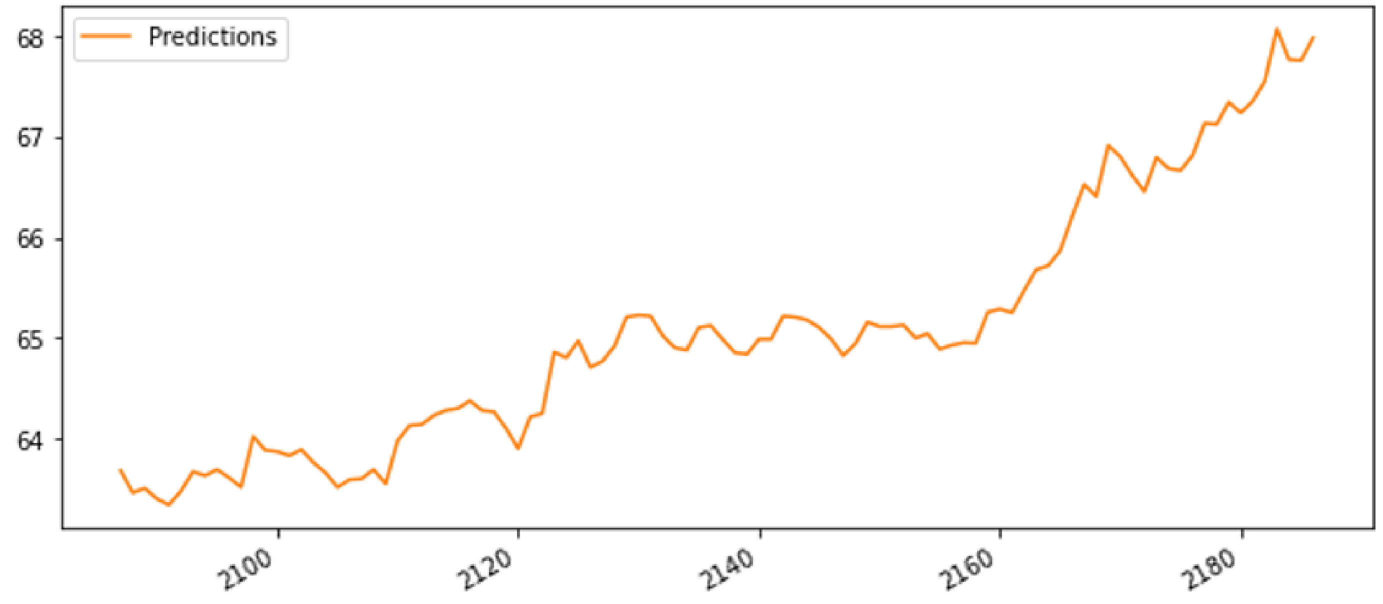
<u>Final Model</u>

Price = 0.062 + 0.999 Lagged Price

Prediction using linear regression

- Model was build on training data and predictions were made on the test data
- MSE for test data = 0.073
- Plot of actual test data and predicted values





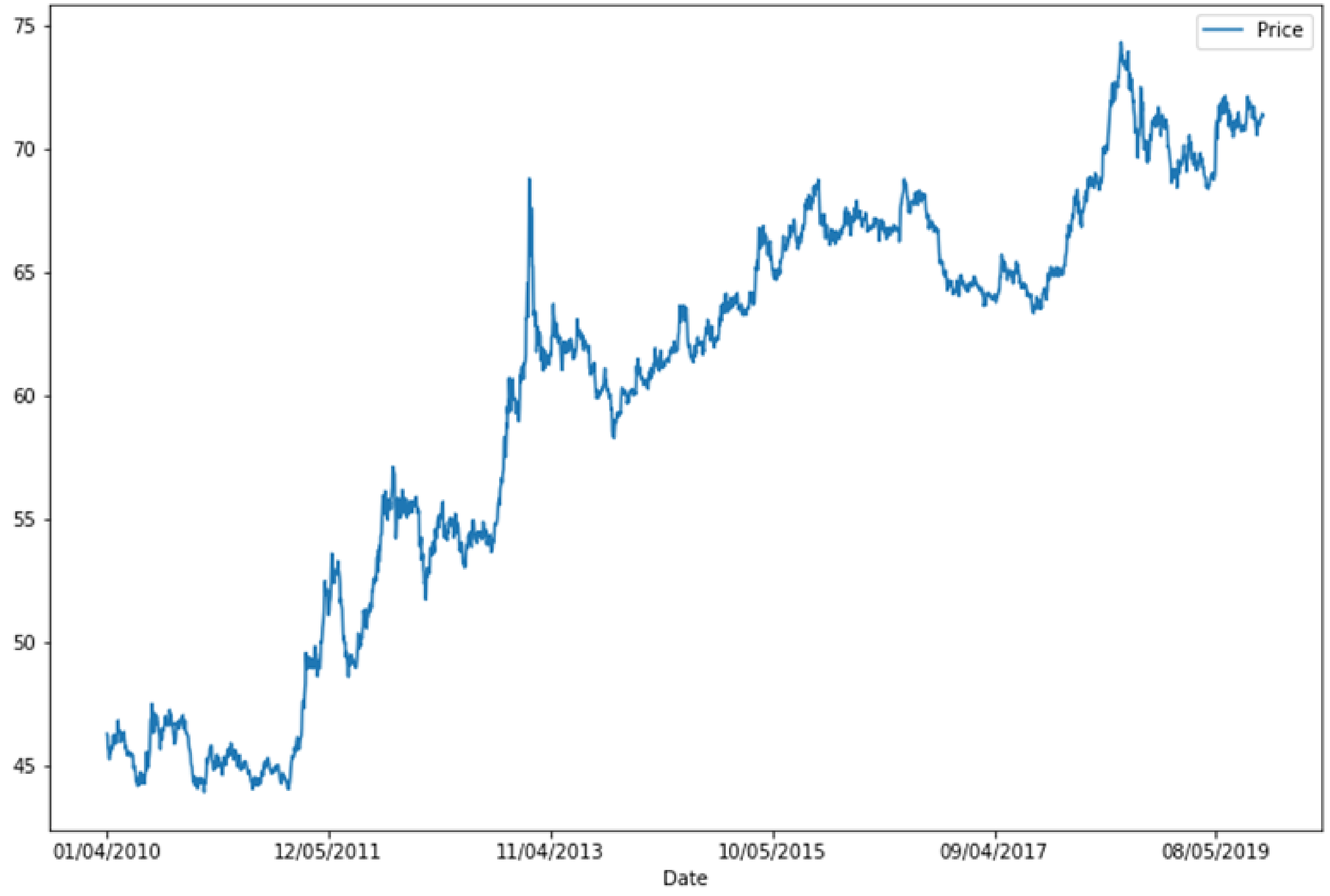
	Price	Predictions
2087	63.460	63.682779
2088	63.505	63.462984
2089	63.400	63.507942
2090	63.340	63.403040
2091	63.475	63.343096

Time Series Analysis

Time Series Analysis

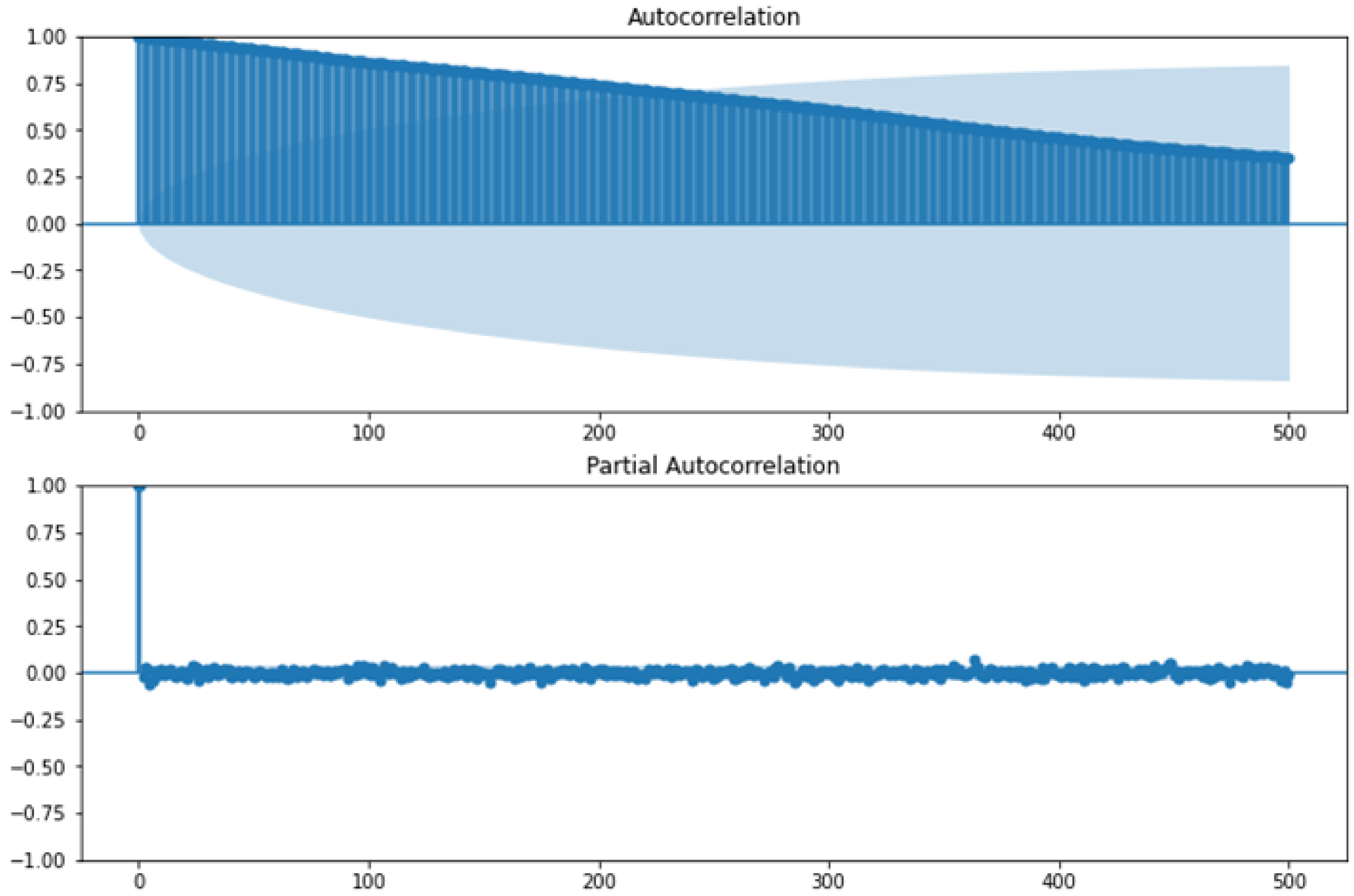
TREND AND SEASONALITY INSPECTION

- **Upward Trend**
- **Non-Stationary**



ACF and PACF

- Indicates the relationship of current observations with the previous observations
- Plot of ACF and PACF for 500 lags
- Helps in understanding the type of model that can be built
- ACF graph shows correlation with other lags and a decay
- PACF shows a spike at lag 1



ARMA Model

- Auto Regressive Moving Average Model
- AR parameter p = 1
- MA parameter q = 0
- Built an ARMA model with p=1 and q=0 on the train data
- Obtain the forecast from ARMA model
- Compare with the test data
- **Final model**
    - Y(t) = 55.997 + 0.999Y(t-1)
    - Y(t) corresponds to Price

```
## ARMA model (p=1,q=0)

model_1 = sm.tsa.ARMA(train['Price'], (1,0)).fit()
print(model_1.params)
```

```
const           55.997039
ar.L1.Price      0.999483
dtype: float64
```

ARMA Model Results

Dep. Variable:	Price	**No. Observations:**	2085
Model:	ARMA(1, 0)	**Log Likelihood**	-261.991
Method:	css-mle	**S.D. of innovations**	0.274
Date:	Wed, 08 Dec 2021	**AIC**	529.983
Time:	16:23:23	**BIC**	546.911
Sample:	0	**HQIC**	536.185

	coef	std err	z	P>|z|	[0.025	0.975]
const	55.9970	6.890	8.127	0.000	42.492	69.502
ar.L1.Price	0.9995	0.001	1926.275	0.000	0.998	1.000

Roots

	Real	Imaginary	Modulus	Frequency
AR.1	1.0005	+0.0000j	1.0005	0.0000

ARMA Model

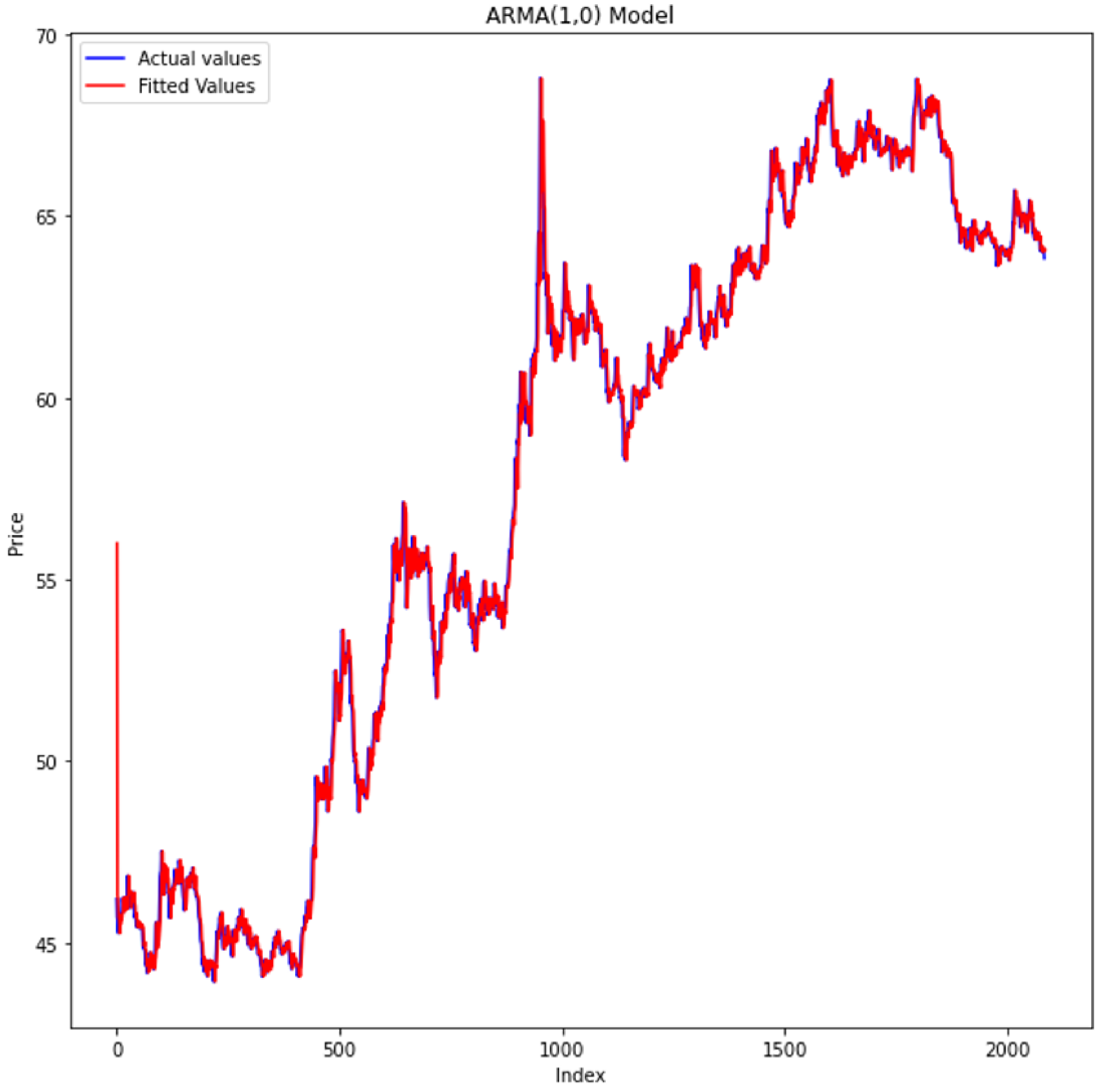
ARMA Model Results

Dep. Variable:	Price	**No. Observations:**	2085
Model:	ARMA(1, 0)	**Log Likelihood**	-261.991
Method:	css-mle	**S.D. of innovations**	0.274
Date:	Wed, 08 Dec 2021	**AIC**	529.983
Time:	16:23:23	**BIC**	546.911
Sample:	0	**HQIC**	536.185

	coef	std err	z	P>|z|	[0.025	0.975]
const	55.9970	6.890	8.127	0.000	42.492	69.502
ar.L1.Price	0.9995	0.001	1926.275	0.000	0.998	1.000

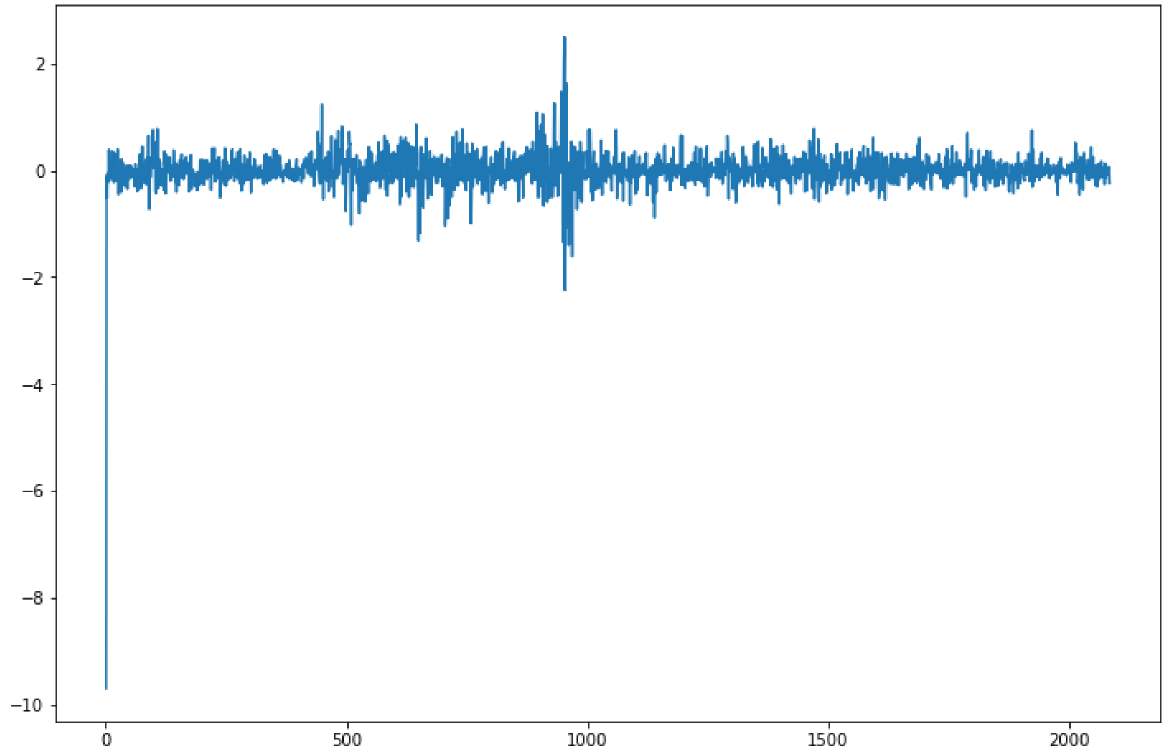
Roots

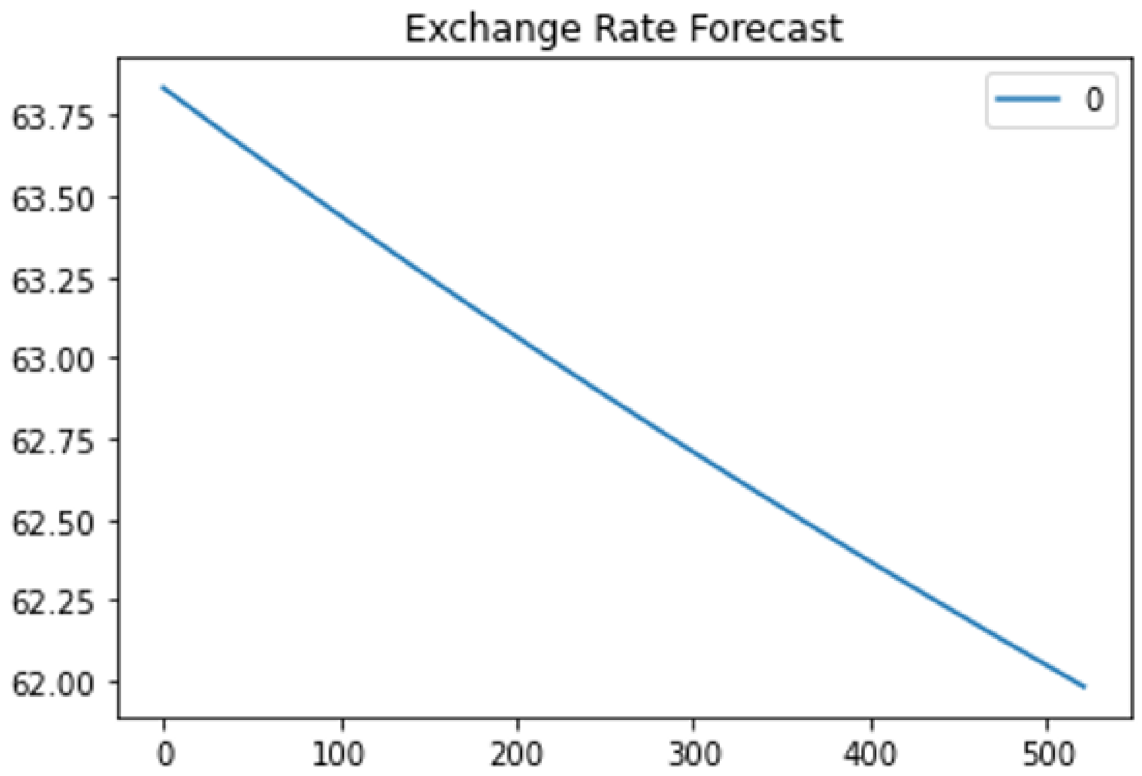
	Real	Imaginary	Modulus	Frequency
AR.1	1.0005	+0.0000j	1.0005	0.0000



ARMA Model

- Residual plot
- Plot of forecasted values for next 521 days
- MSE for train data = 0.12
- MSE for test data is very high
- p-value of AR parameter <0.05
- AIC = 530





ARIMA Model

- Auto Regressive Integrated Moving Average model
- AR parameter p = 1
- MA parameter q = 1
- Differencing parameter d = 1
- Built an ARIMA model with p=1, d=1 and q=1 on the train data
- Obtained the forecast from ARIMA model
- Compared with the test data

```
## ARIMA MODEL

from statsmodels.tsa.arima_model import ARIMA

arima_1= ARIMA(train['Price'],order=(1,1,1)).fit()
```

ARIMA Model Results

Dep. Variable:	D.Price	**No. Observations:**	2084
Model:	ARIMA(1, 1, 1)	**Log Likelihood**	-253.618
Method:	css-mle	**S.D. of innovations**	0.273
Date:	Wed, 08 Dec 2021	**AIC**	515.236
Time:	16:55:42	**BIC**	537.804
Sample:	1	**HQIC**	523.505

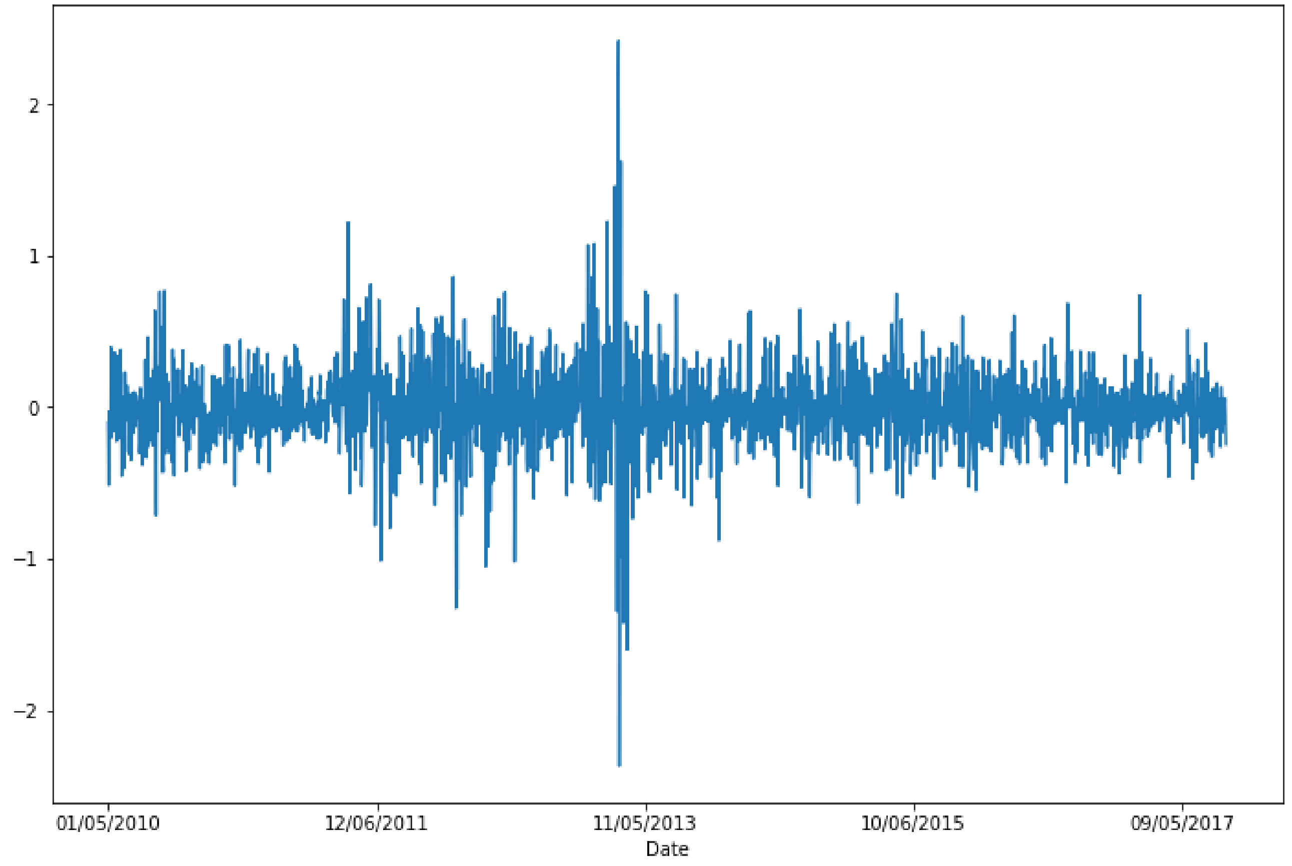
	coef	std err	z	P>|z|	[0.025	0.975]
const	0.0084	0.006	1.352	0.176	-0.004	0.021
ar.L1.D.Price	-0.4186	0.154	-2.720	0.007	-0.720	-0.117
ma.L1.D.Price	0.4755	0.148	3.208	0.001	0.185	0.766

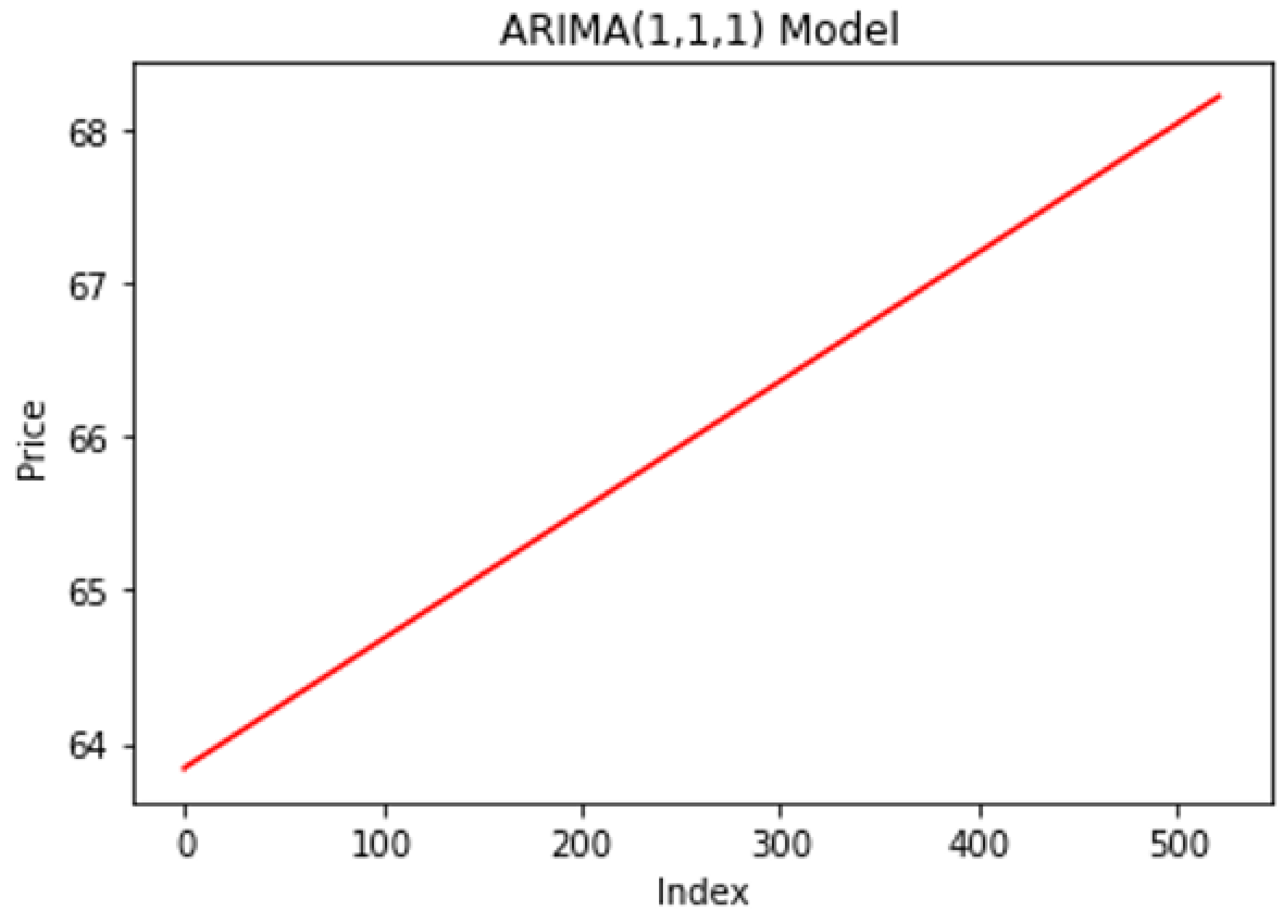
Roots

	Real	Imaginary	Modulus	Frequency
AR.1	-2.3890	+0.0000j	2.3890	0.5000
MA.1	-2.1032	+0.0000j	2.1032	0.5000

ARIMA Model

- Residual plot
- Forecast made for next 521 days
- MSE for test data = 15.342
- p-value of AR component and MA component <0.05
- AIC = 515.25





ARIMA Model

- AR parameter p = 2
- MA parameter q = 2
- Differencing parameter d = 1
- Built another ARIMA model with different set of parameters

```
## ARIMA (2,1,2)

arima_2= ARIMA(train,order=(2,1,2)).fit()
arima_2.summary()
arima_2_pred = arima_2.forecast(steps=522)[0]
```

ARIMA Model Results

Dep. Variable:	D.Price	**No. Observations:**	2084
Model:	ARIMA(2, 1, 2)	**Log Likelihood**	-237.294
Method:	css-mle	**S.D. of innovations**	0.271
Date:	Tue, 07 Dec 2021	**AIC**	486.589
Time:	23:56:43	**BIC**	520.441
Sample:	01-05-2010	**HQIC**	498.993
	- 12-29-2017		

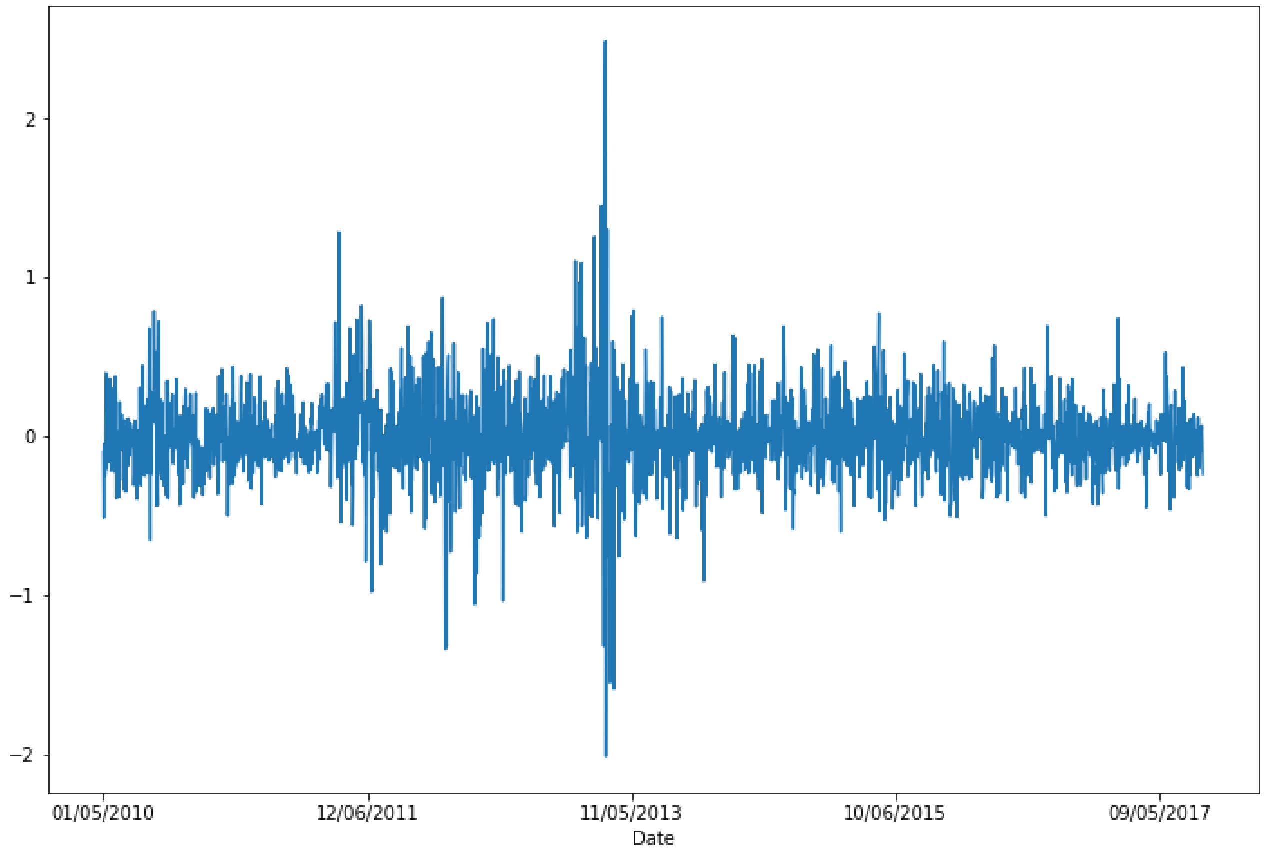
	coef	std err	z	P>|z|	[0.025	0.975]
const	0.0085	0.006	1.532	0.126	-0.002	0.019
ar.L1.D.Price	0.2180	0.158	1.379	0.168	-0.092	0.528
ar.L2.D.Price	-0.5031	0.093	-5.433	0.000	-0.685	-0.322
ma.L1.D.Price	-0.1883	0.170	-1.109	0.267	-0.521	0.144
ma.L2.D.Price	0.3822	0.098	3.911	0.000	0.191	0.574

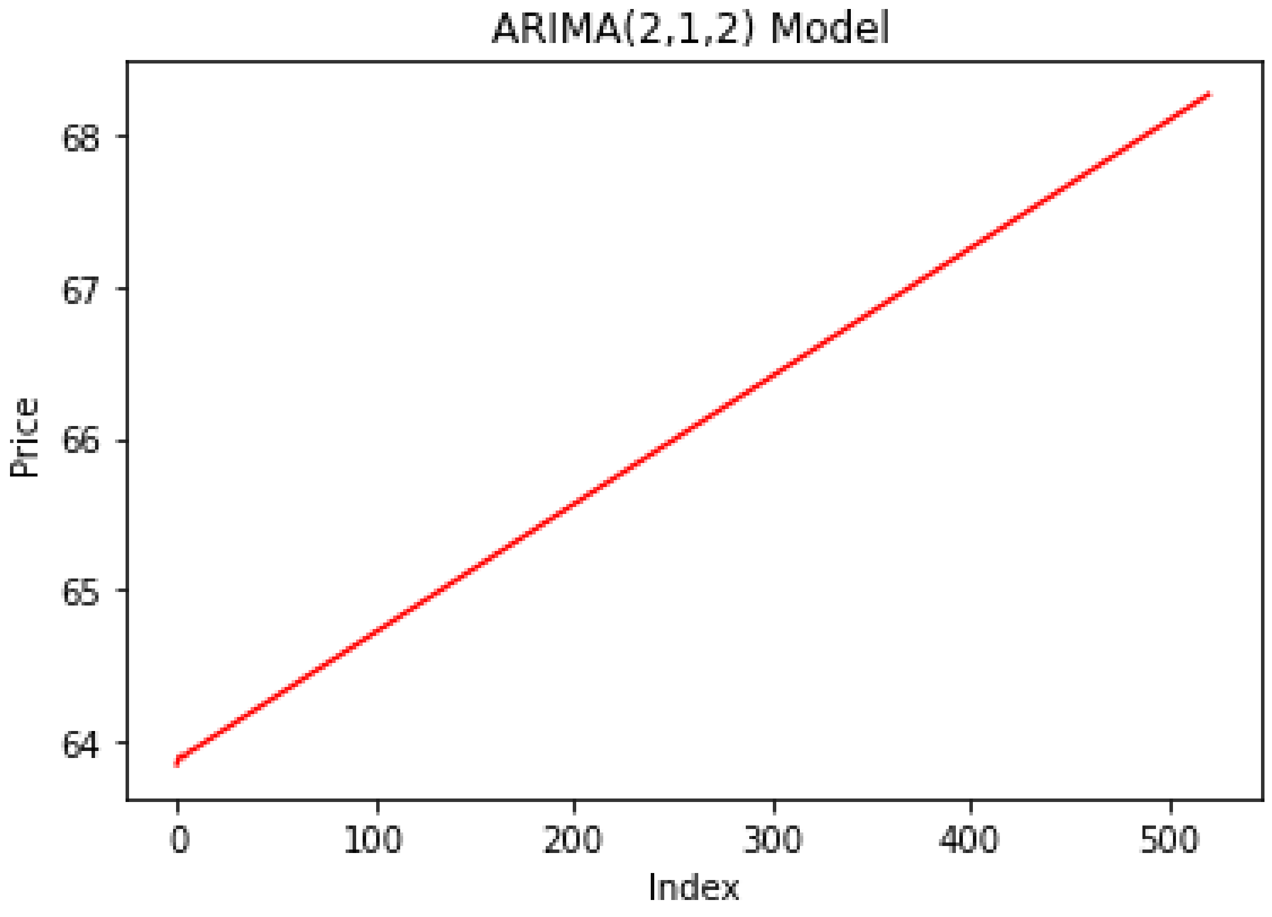
Roots

	Real	Imaginary	Modulus	Frequency
AR.1	0.2166	-1.3930j	1.4098	-0.2254
AR.2	0.2166	+1.3930j	1.4098	0.2254
MA.1	0.2463	-1.5986i	1.6175	-0.2257

ARIMA Model

- Forecast made for the length of test data
- MSE for test data = 15.05369
- p-value of AR and MA component at lag 1 is >0.05
- AIC = 486.589





Conclusion

- Linear Regression model gave the least mean squared error value for test data and was able to capture the trend of the time series data well.
- ARMA model parameters were significant but the mean squared error for test data was very high.
- ARIMA(1,1,1) model's all the AR and MA parameters were significant and AIC value was also not high
- ARIMA(2,1,2) model's AR and MA parameters at lag 1 were insignificant but AIC value and MSE was lower than ARIMA(1,1,1)

Thank you!

Any questions?