FINAL PROJECT

Case Study: United Tractor

Alief Muhammad Amanu

Data Science Bootcamp
by Dibimbing. id

WHAT IS UNITED TRACTORS?

PT United Tractors Tbk or **United Tractors (UT/the Company)** is a company that was established on 13 October 1972 as the exclusive distributor of Komatsu Limited heavy equipment in Indonesia

Aside of being the largest distributor of heavy equipment in the country, the Company also plays an active role in the field of mining contracting and has recently ventured into coal and gold mining businesses

BIG COMPANY, DATA?

United Tractors went public and listed its shares in Indonesia Stock Exchange as UNTR. As one of the biggest heavy-equipment suppliers in Indonesia, United Tractors shares is one of the best company on the market. This makes its shares price fluctuate.

How to improve decision-making on buying the shares? The answer is **Forecasting.** We want to extract as much as possible information available inside our data to find a distinct pattern so we can create our decision more precise.

But **HOW**?

01 Doing some **Exploratory Data Analysis** with visualization using Lineplot, Correlogram, and other visualization method

---

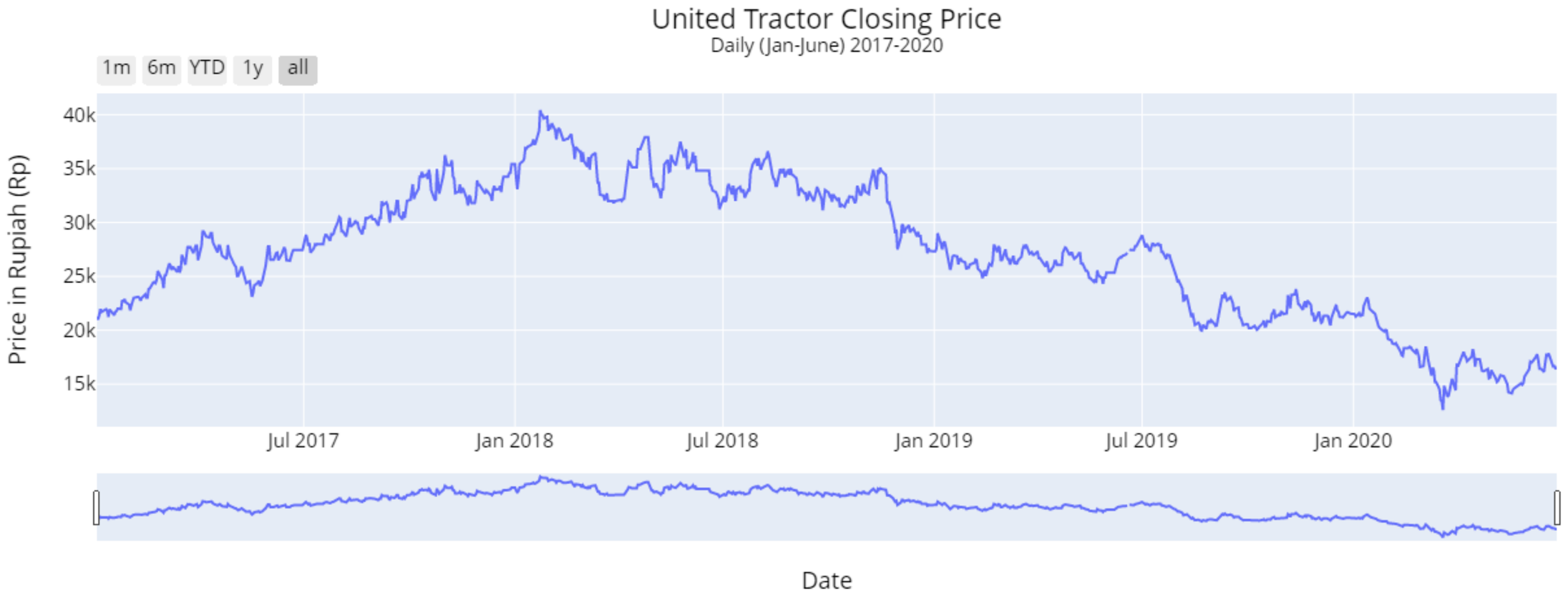
02 Create a **machine learning** model to help us create our business decision. For example, to predict shares price on features available in our data

---

03 By doing **Exploratory Data Analysis** and creating **machine learning** models, we can give a precise recommendation about what to do in the future.

UNITED TRACTORS

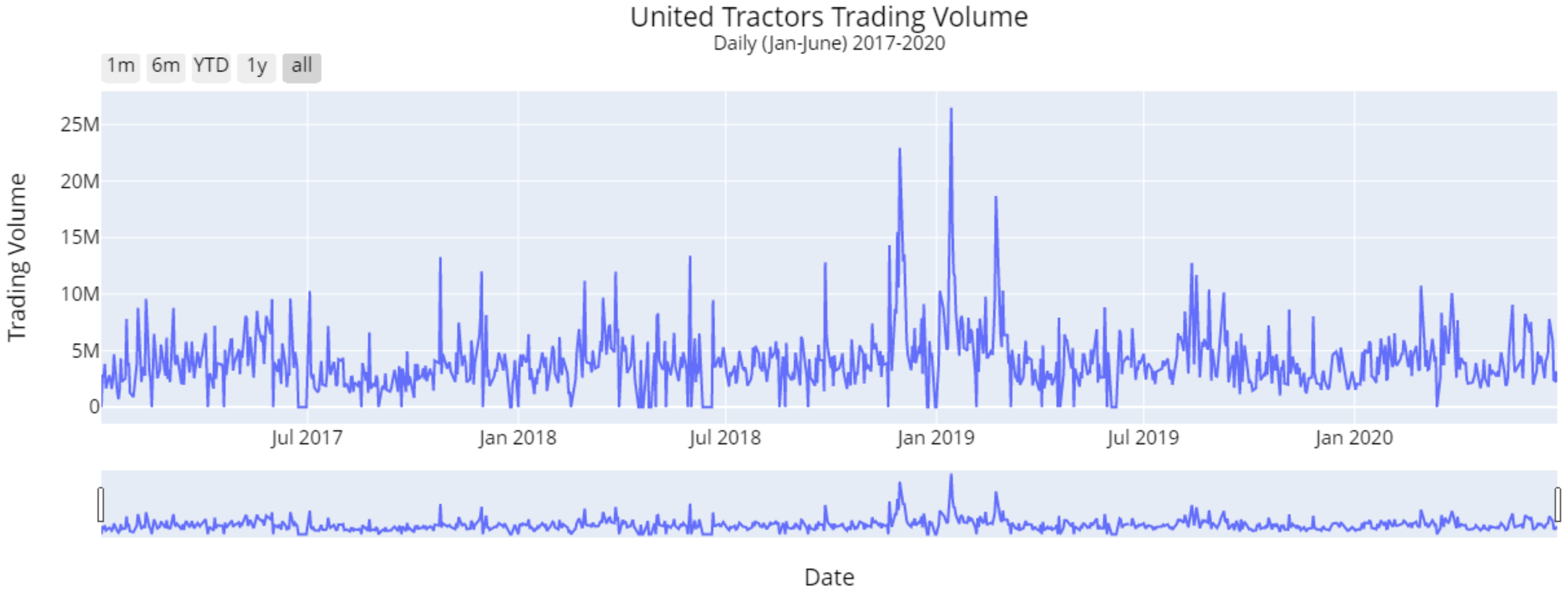
Exploratory Data Analysis



As shown by the plot above, the data has **uptrend** and **downtrend**. Highest price point is **Rp40.425** on 23 Jan 2018 and lowest is Rp12.600 on 19 Mar 2020. As for uptrend, closing price has positive changes of 88,25% and for downtrend is -59,03%

UNITED TRACTORS

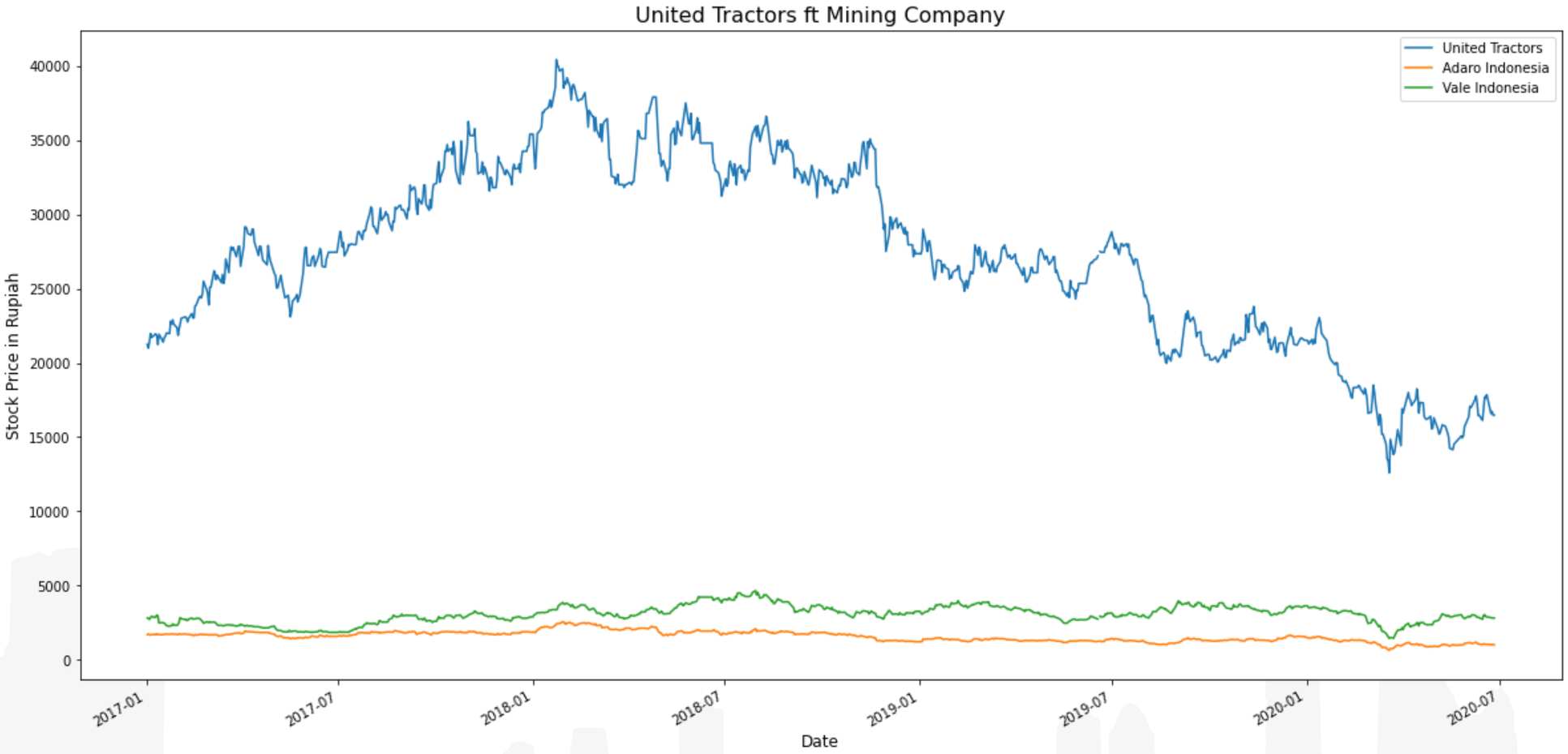
Exploratory Data Analysis



The highest trading volume is **26.4821 Million** on 14 Jan 2019. For some reason, trading volume has the lowest point at every middle of the year while the highest seems always at early year. This might require further action

UNITED TRACTORS

Exploratory Data Analysis



As for the relationship between United Tractors with Vale Indonesia and Adaro Indonesia, I believe that there is **no strong relationship** between those 3 companies for stock prices. Because it has no unique pattern or distinct pattern

UNITED TRACTORS

Data Modeling

Model used in this analysis

ARIMA Model

Prophet

Split the data into **Data Train** and **Data Test**

Proportion is 80%:20%

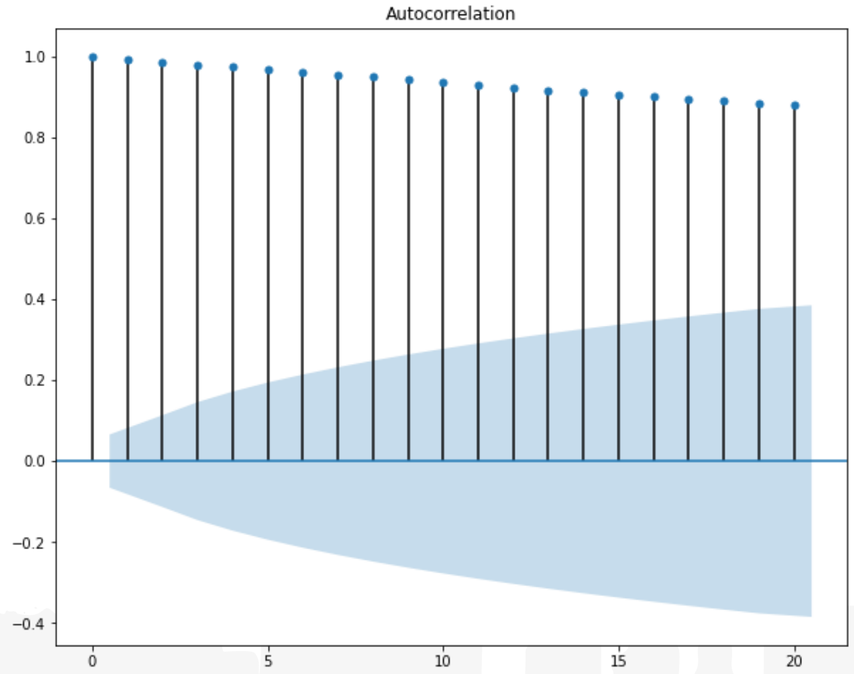
Data size	: 891
Data train	: 712
Data test/validation	: 179

UNITED TRACTORS

ARIMA Model

Stationarity Check

We can check stationarity thru ACF plot and Dickey-Fuller Test. We want to have a cuts-off pattern at ACF plot and p-value lower than significance level of 5% (0.05)



ACF Plot

From ACF plot above, we can saw a slow-decay or dying down exponentially pattern. That indicates that our data is not yet stationary. But, we need to test our data with Dickey-Fuller to make sure about stationarity condition

Results of Dickey-Fuller Test:

ADF Statistic	-0.912058
p-value	0.783976

From Dickey-Fuller test result, we can clearly say that our data is not stationary yet. Because p-value of the test is still greater than significance level of 5% or 0.05. We might want to fix this with differencing process

UNITED TRACTORS

ARIMA Model

Convert Series to Stationary

Differencing: is the transformation of the series to a new time series where the values are the differences between consecutive values. This procedure may be applied consecutively more than once, giving rise to the "first differences", "second differences", etc

The first differences Math image of a time series Math image are described by the following expression

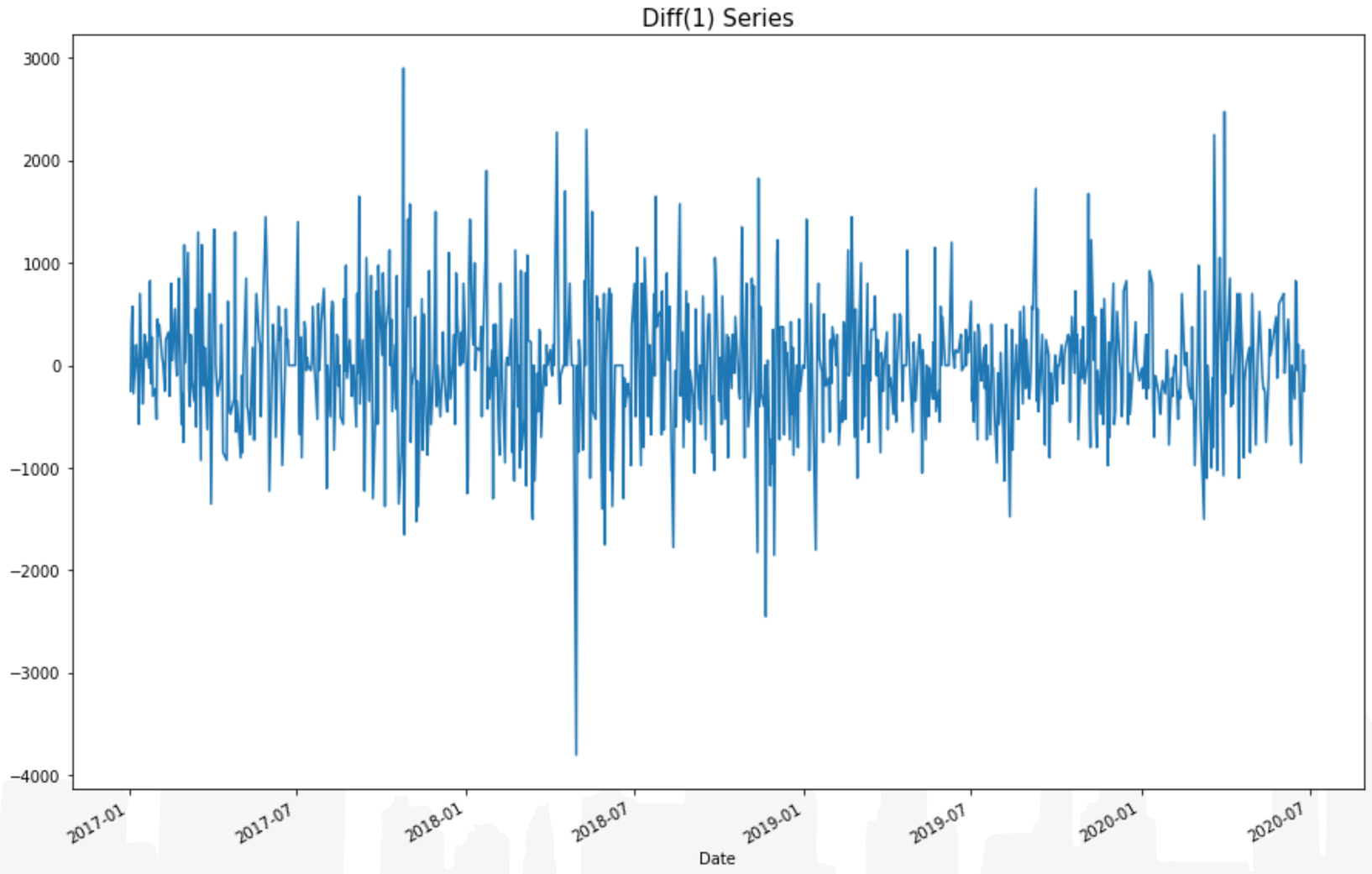
$$d(t) = x(t)-x(t-1)$$

Date	Close
2017-01-03	-250.0
2017-01-04	400.0
2017-01-05	575.0
2017-01-06	-275.0
2017-01-09	200.0

First 5 rows after data differenced

UNITED TRACTORS

ARIMA Model



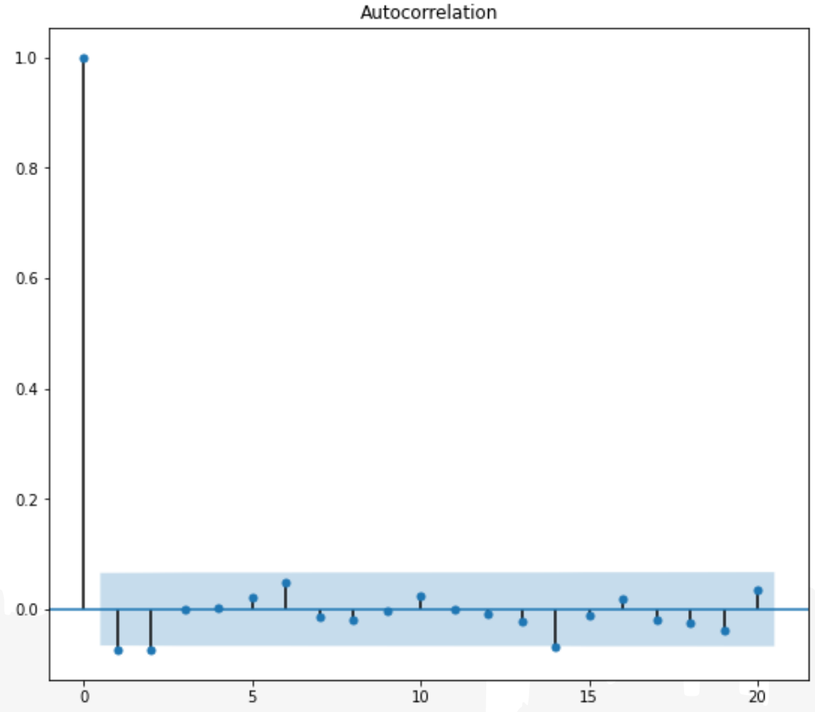
Data plot after first difference. The data seems like already satisfy stationarity assumption on mean

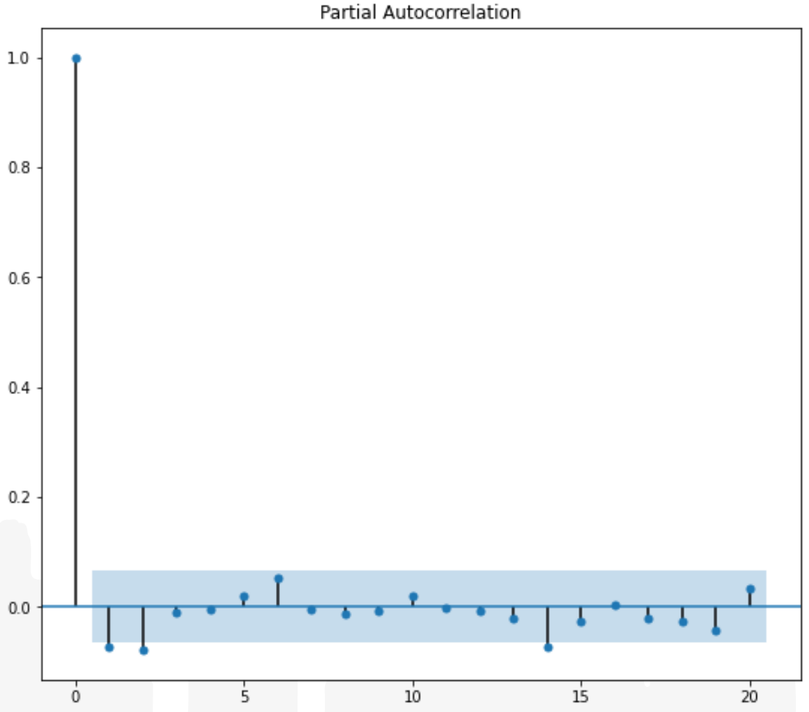
UNITED TRACTORS

ARIMA Model

Stationarity Check

We can check stationarity thru ACF plot and Dickey-Fuller Test. We want to have a cuts-off pattern at ACF plot and p-value lower than significance level of 5% (0.05)





ACF and PACF plot shows cuts-off pattern. As for sig lag, both plots significant at lag 0, 1, 2. Thus we have candidates of:

ARIMA(0,1,0), ARIMA(1,1,1), ARIMA(0,1,2), ARIMA(2,1,0), etc

UNITED TRACTORS

ARIMA Model

Stationarity Check

We can check stationarity thru ACF plot and Dickey-Fuller Test. We want to have a cuts-off pattern at ACF plot and p-value lower than significance level of 5% (0.05)

Results of Dickey-Fuller Test:

ADF Statistics : -23.5750
p-value : 0.0000

Dickey-Fuller results indicates our data is stationare with p-value is lower than significance level of 5% or 0.05

UNITED TRACTORS

ARIMA Model

ARIMA Model Results

Dep. Variable:	D.Close	No. Observations:	890
Model:	ARIMA(2, 1, 2)	Log Likelihood	-7003.143
Method:	css-mle	S.D. of innovations	632.542
Date:	Fri, 15 Oct 2021	AIC	14018.285
Time:	23:39:39	BIC	14047.033
Sample:	1	HQIC	14029.273

	coef	std err	z	P>|z|	[0.025	0.975]
const	-5.3751	18.749	-0.287	0.774	-42.123	31.372
ar.L1.D.Close	0.7963	0.291	2.734	0.006	0.225	1.367
ar.L2.D.Close	-0.5282	0.242	-2.184	0.029	-1.002	-0.054
ma.L1.D.Close	-0.8763	0.297	-2.950	0.003	-1.458	-0.294
ma.L2.D.Close	0.5234	0.241	2.170	0.030	0.051	0.996

Roots

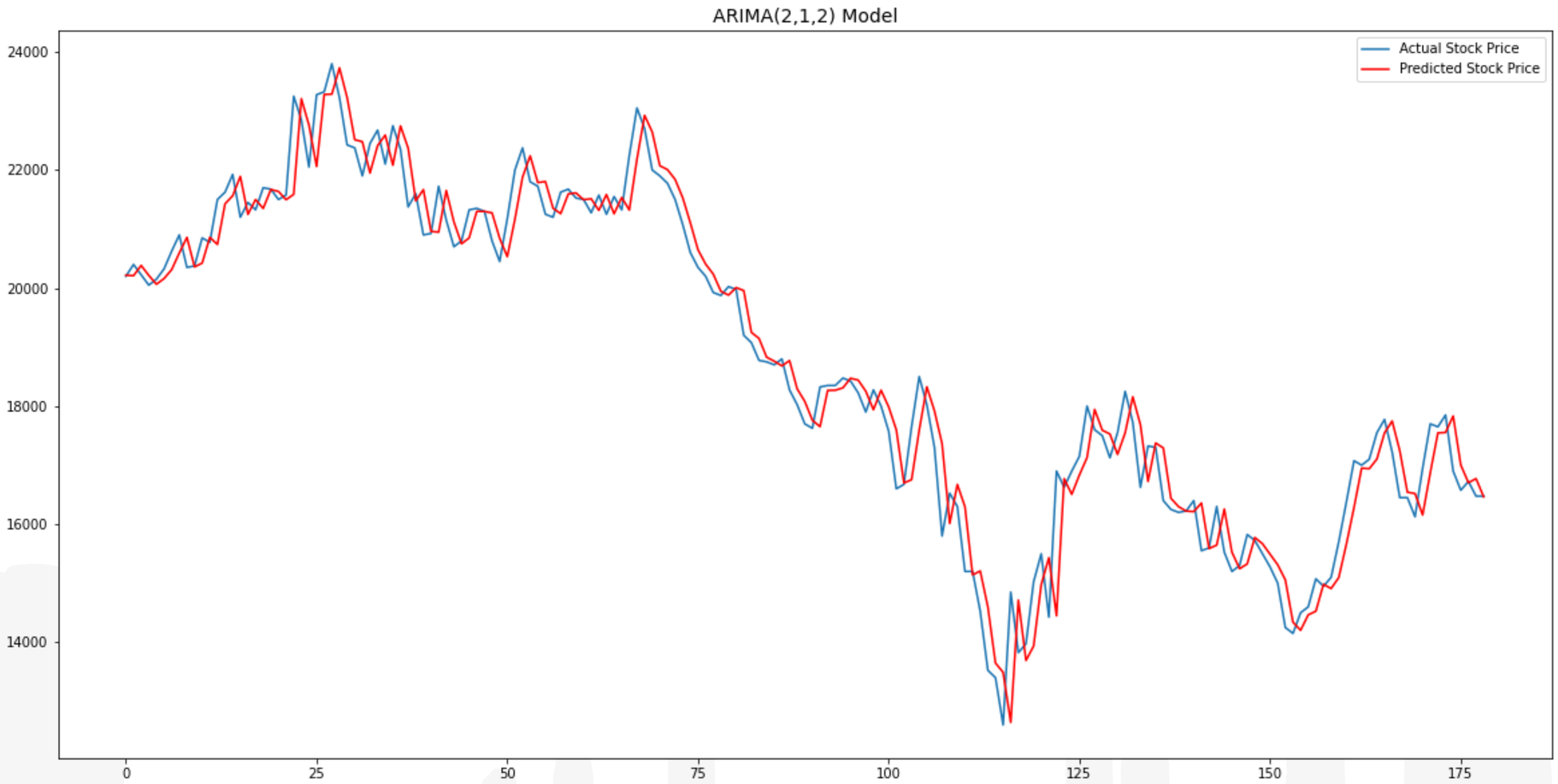
	Real	Imaginary	Modulus	Frequency
AR.1	0.7538	-1.1511j	1.3760	-0.1577
AR.2	0.7538	+1.1511j	1.3760	0.1577
MA.1	0.8371	-1.0999j	1.3822	-0.1465
MA.2	0.8371	+1.0999j	1.3822	0.1465

ARIMA(2,1,2) has AIC of 14018.205.
As for coefficient, we have:

AR(1): 0.7963
AR(2): -0.5282
MA(1): -0.8763
MA(2): 0.5234

UNITED TRACTORS

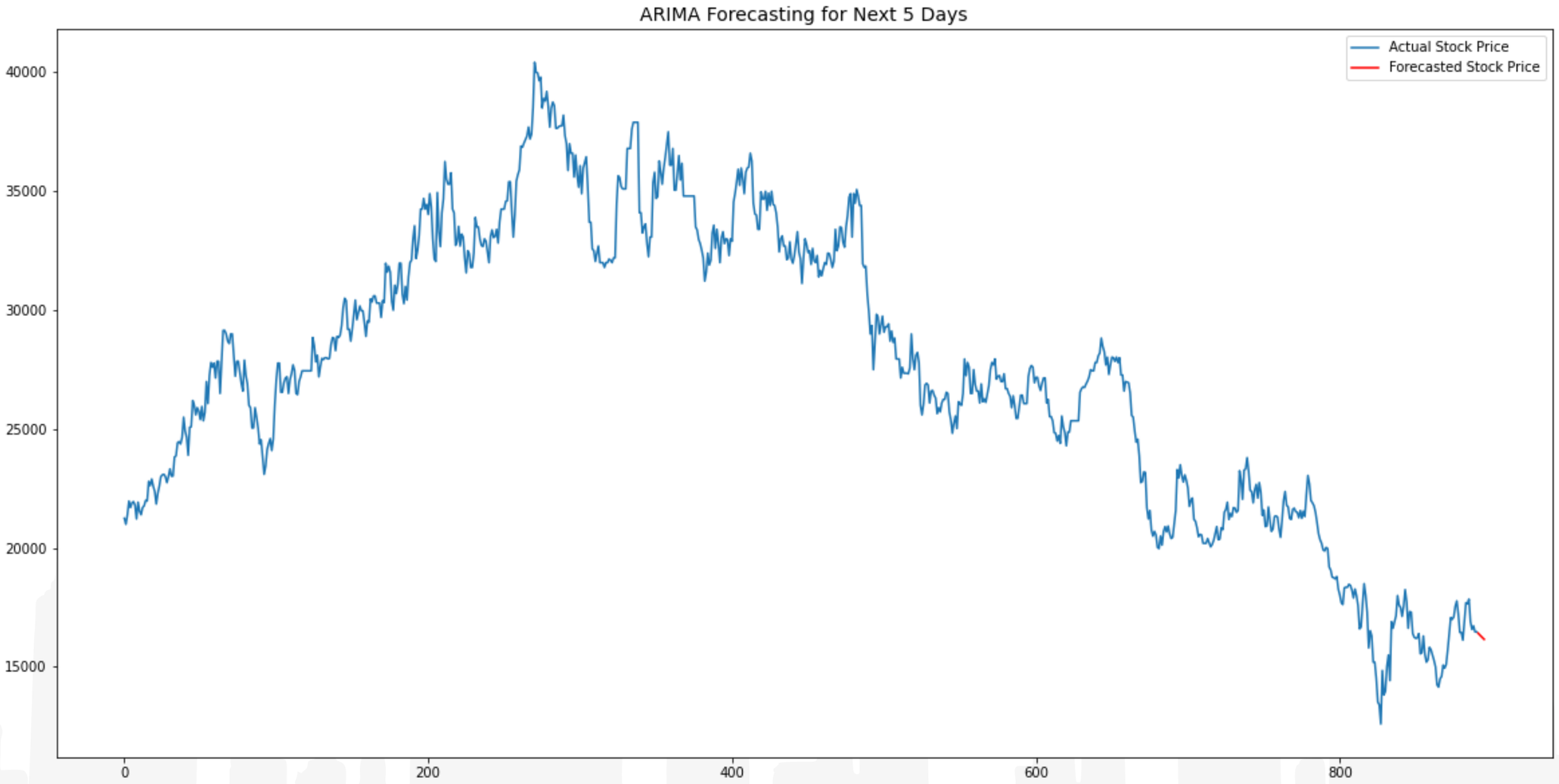
ARIMA Model



From plot above, we can see that actual value and predicted value is quite similar which means that ARIMA(2,1,2) is worked pretty good for this data. The more similar the pattern become, the better the model

UNITED TRACTORS

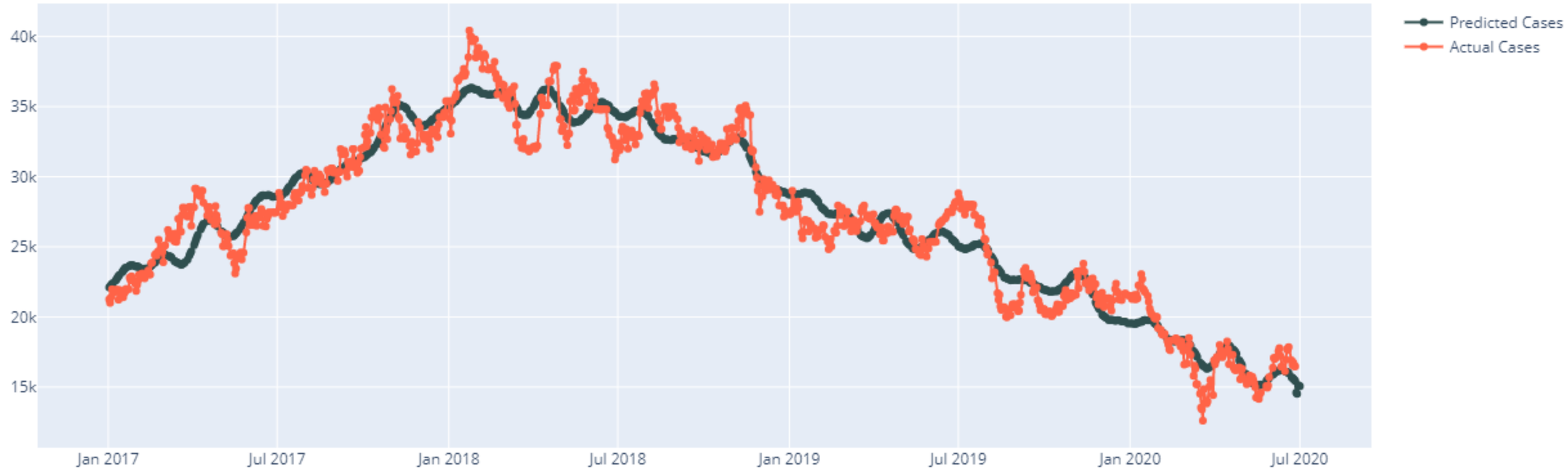
ARIMA Model



Forcasted value for next 5 days

UNITED TRACTORS

Prophet Model



Forcasted value for next 5 days

Insights and Recommendations

UNITED TRACTORS

DATA ANALYST/ DATA SCIENTIST

01

Both models work really well for our data, it is shown by the plot of actual value vs prediction value. For some reason, prediction performed by **ARIMA** model is not that good. Predicted value is quite similar, but we can enhance the model performance by tuning the model.

02

From stationarity process on ARIMA model, our data is already stationare on mean but it seems like the data is not stationare on variance. To achieve that goal, we can transform the data using log, ln, or another transform method. Transform the data may increase model performance.

Insights and Recommendations

UNITED TRACTORS

INVESTOR

01 Based on visualization in Exploratory Data Analysis, we might want to buy the shares at the end of the year or somewhere in the 7th-8th month. Because at that time point, the shares usually has lower price.

02 To enhance our decision making for buy or sell the shares, we might want to know about the upcoming price. So, we have to forecast the price. In order to achieve that, we can use forecasting method such as ARIMA model or Prophet. Based on data modeling in this analysis, ARIMA model works better. We can use this model to forecast the upcoming price.

Insights and Recommendations

UNITED TRACTORS

INVESTOR

03

As for action, we might want to buy the shares based on forecasting results. Because the share price is cheaper than earlier periods. But then, we also want to check another aspects such as company annual report or some news about the industry.

UNITED TRACTORS

Thank You!