CRYPTOCURRENCY PRICE PREDICTION

Presented by: Mime LIU

Agenda

Key Talking Points

- Why do i do this project
- Tech stack used and idea flows
- key points/background knowledge
- Model
- Findings and Reflections
- Challenges and key take-aways

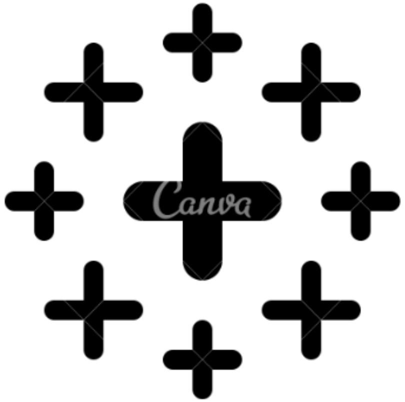
Huge market

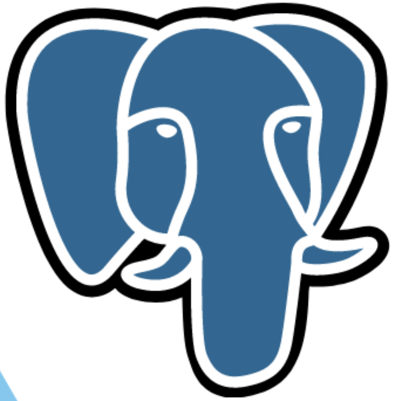
Pertaning to Machine Learning

Personal interest

Readily available big data

Tech Stack





Idea flow / Framework

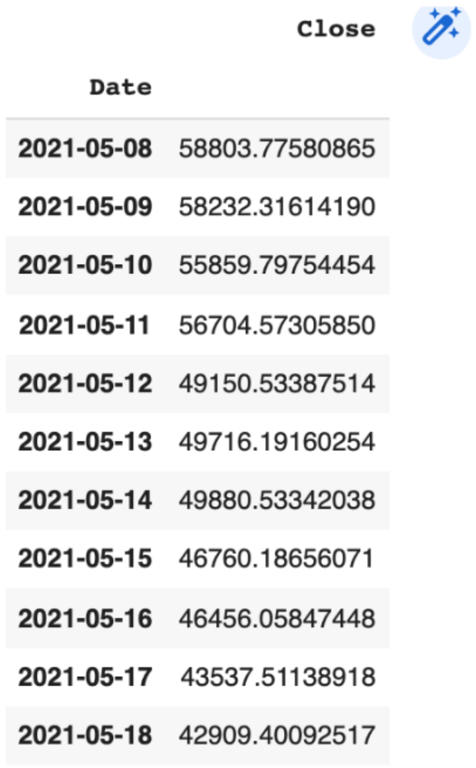
01 HTML- user interface

02 General data Analysis and Visualisation

03 Machine learning

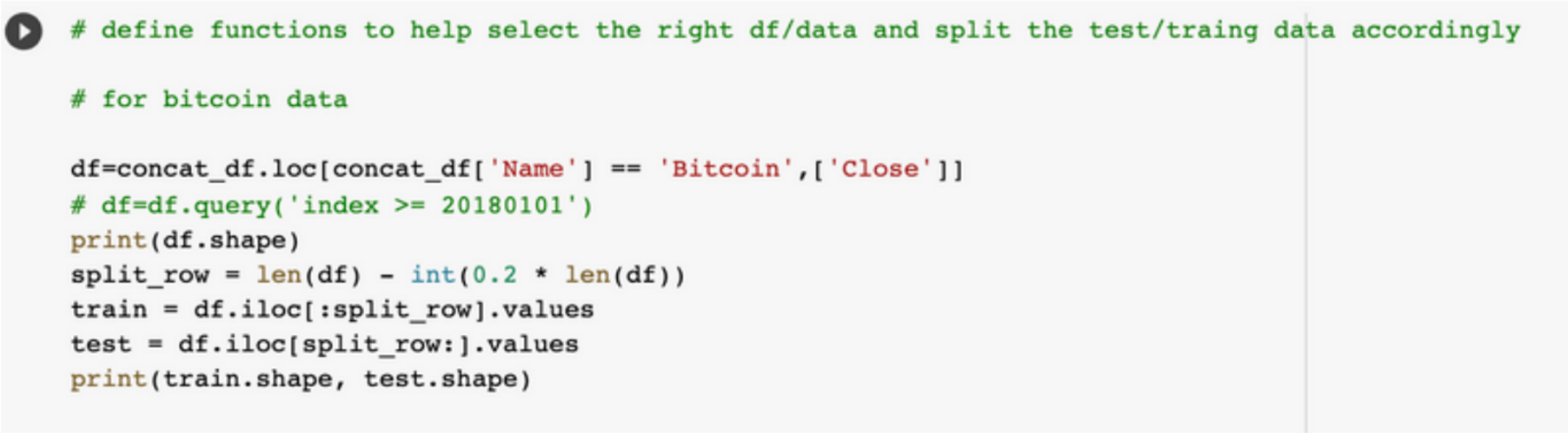
Key points

LSTM/RNN



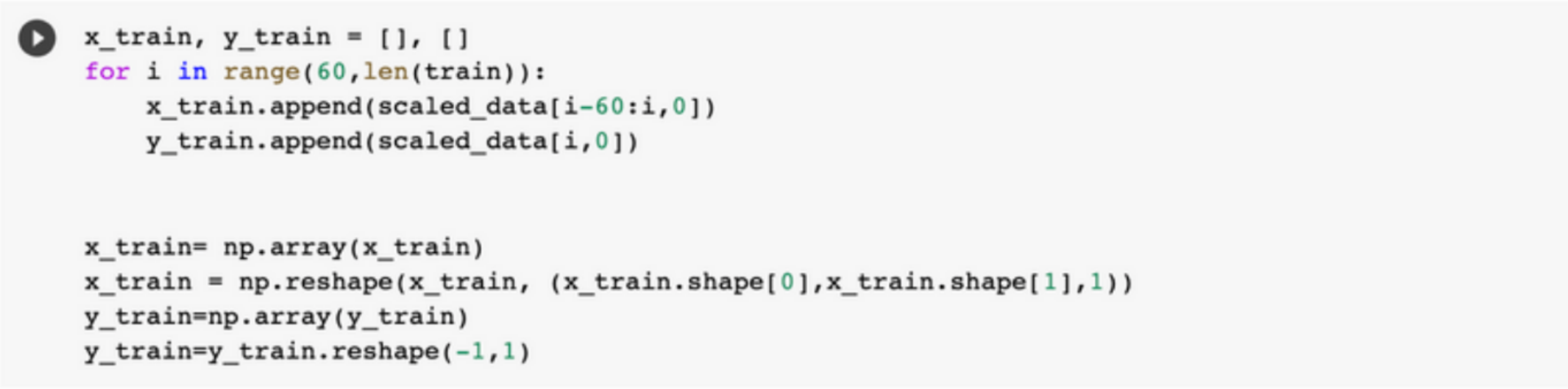
	Close
Date	
2021-05-08	58803.77580865
2021-05-09	58232.31614190
2021-05-10	55859.79754454
2021-05-11	56704.57305850
2021-05-12	49150.53387514
2021-05-13	49716.19160254
2021-05-14	49880.53342038
2021-05-15	46760.18656071
2021-05-16	46456.05847448
2021-05-17	43537.51138918
2021-05-18	42909.40092517

11 data
only 1 variable
date is either index
or not able to be used in study

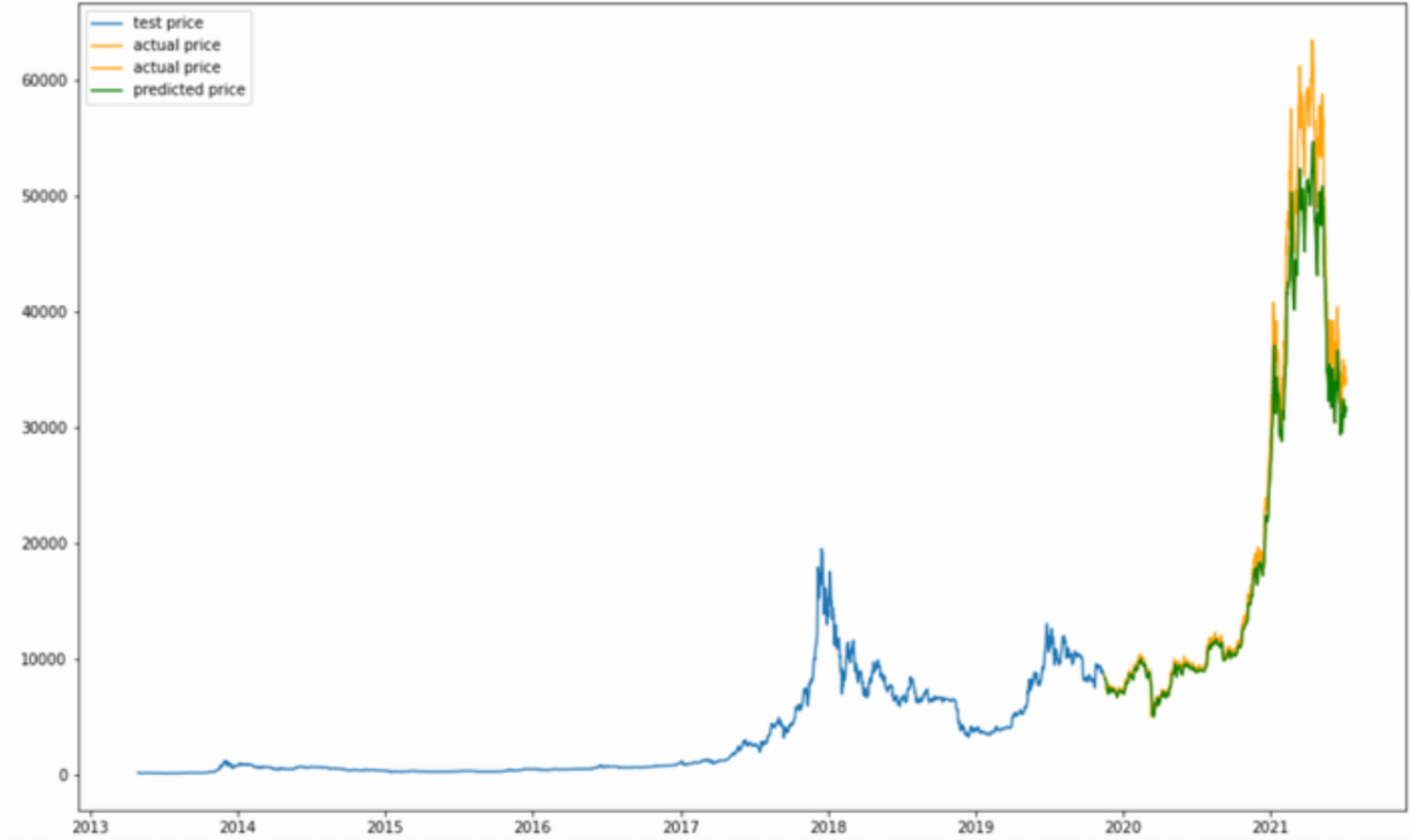


```
# define functions to help select the right df/data and split the test/traing data accordingly

# for bitcoin data

df=concat_df.loc[concat_df['Name'] == 'Bitcoin',['Close']]
# df=df.query('index >= 20180101')
print(df.shape)
split_row = len(df) - int(0.2 * len(df))
train = df.iloc[:split_row].values
test = df.iloc[split_row:].values
print(train.shape, test.shape)
```

Have to set a time_step/lookback

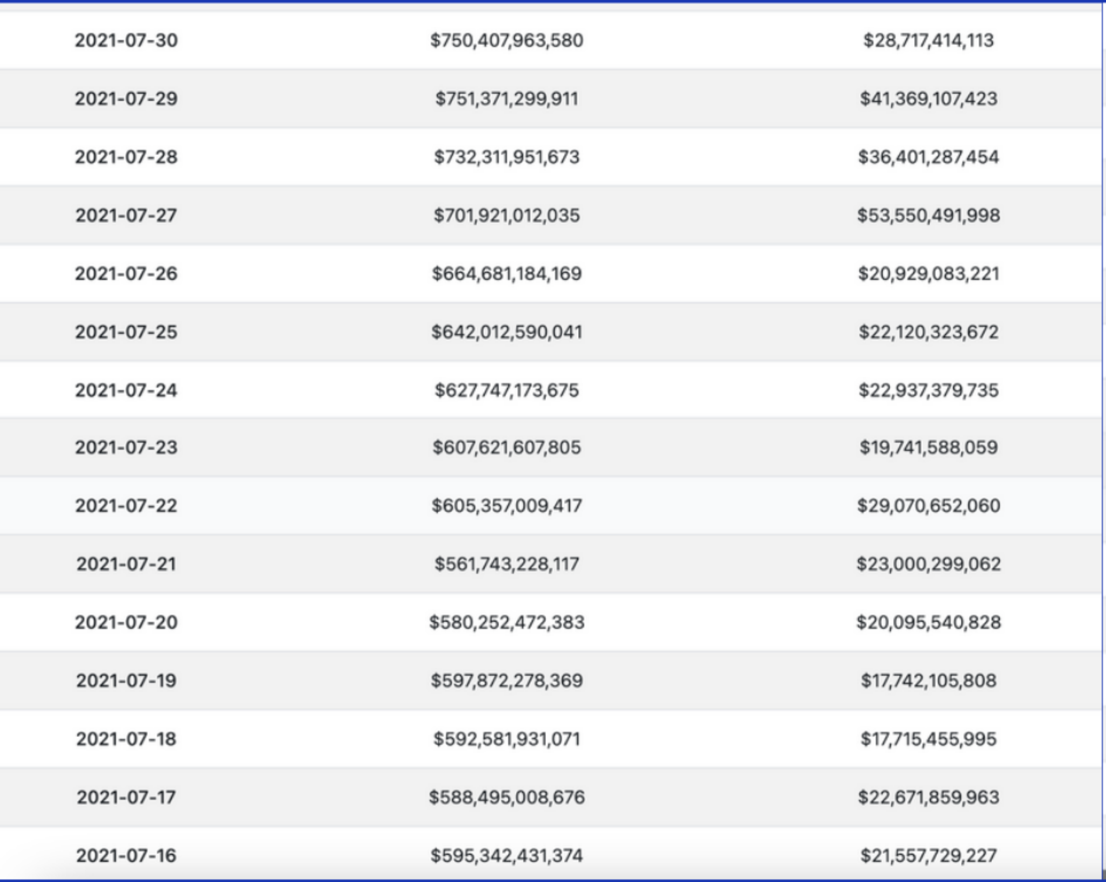


```
x_train, y_train = [], []
for i in range(60,len(train)):
    x_train.append(scaled_data[i-60:i,0])
    y_train.append(scaled_data[i,0])


x_train= np.array(x_train)
x_train = np.reshape(x_train, (x_train.shape[0],x_train.shape[1],1))
y_train=np.array(y_train)
y_train=y_train.reshape(-1,1)
```



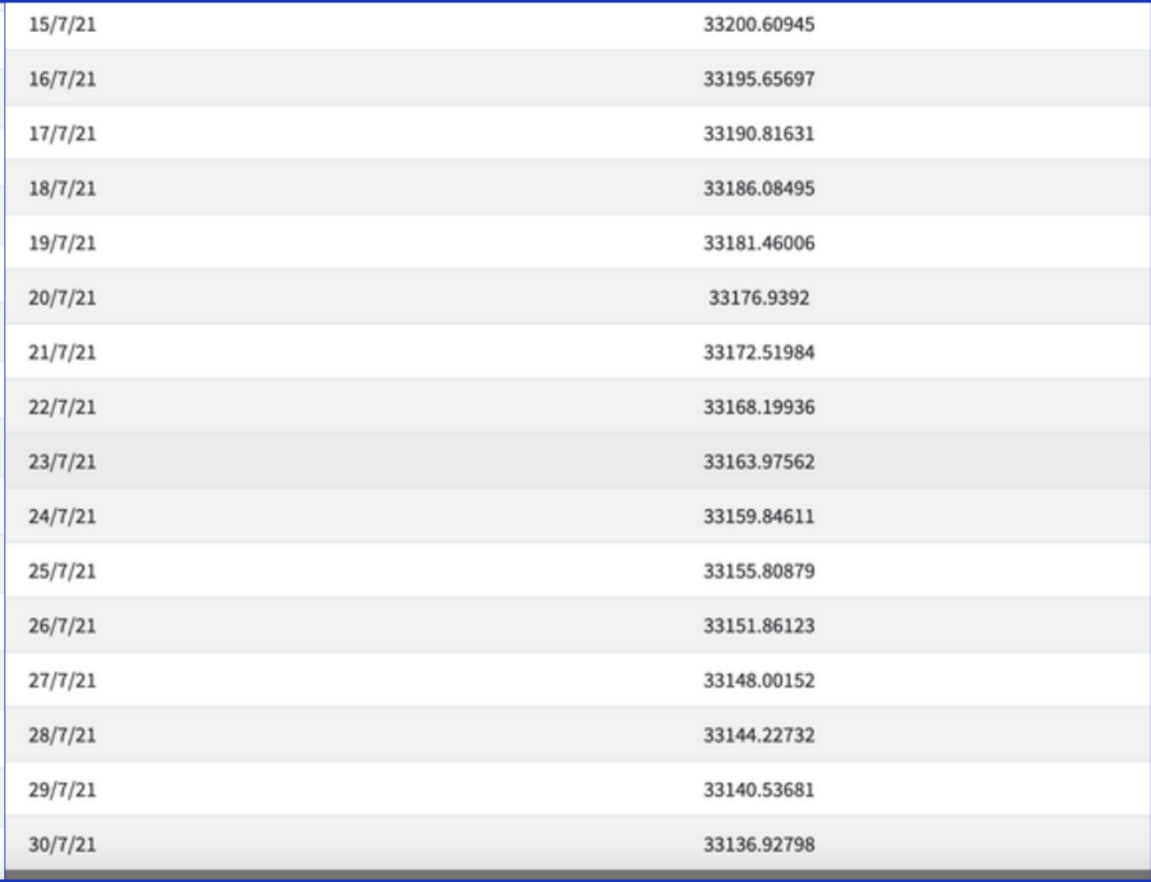
Machine Learning Training Outcome

Actual data

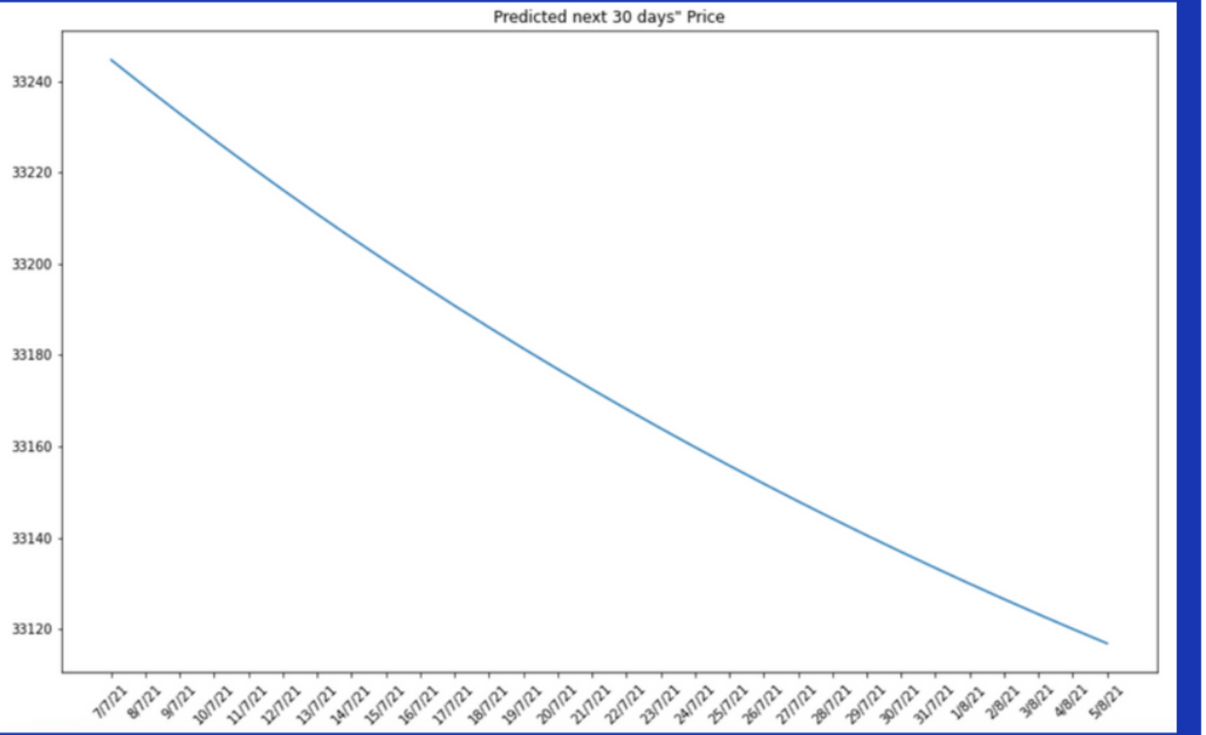


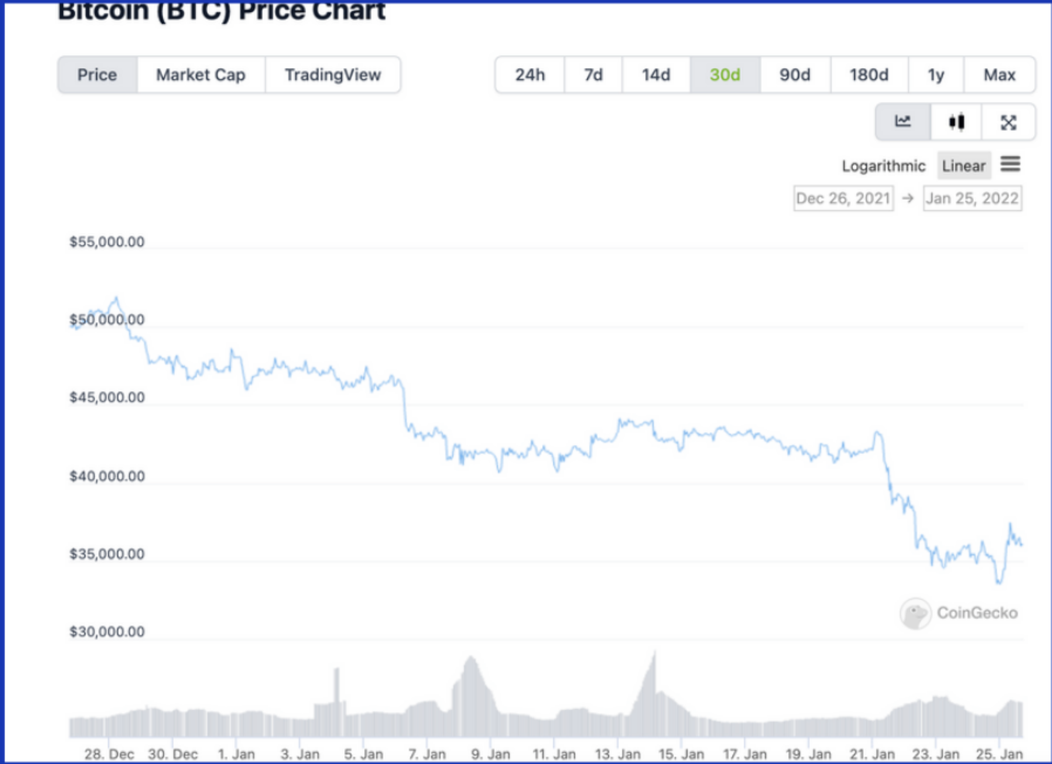
2021-07-30	$750,407,963,580	$28,717,414,113
2021-07-29	$751,371,299,911	$41,369,107,423
2021-07-28	$732,311,951,673	$36,401,287,454
2021-07-27	$701,921,012,035	$53,550,491,998
2021-07-26	$664,681,184,169	$20,929,083,221
2021-07-25	$642,012,590,041	$22,120,323,672
2021-07-24	$627,747,173,675	$22,937,379,735
2021-07-23	$607,621,607,805	$19,741,588,059
2021-07-22	$605,357,009,417	$29,070,652,060
2021-07-21	$561,743,228,117	$23,000,299,062
2021-07-20	$580,252,472,383	$20,095,540,828
2021-07-19	$597,872,278,369	$17,742,105,808
2021-07-18	$592,581,931,071	$17,715,455,995
2021-07-17	$588,495,008,676	$22,671,859,963
2021-07-16	$595,342,431,374	$21,557,729,227

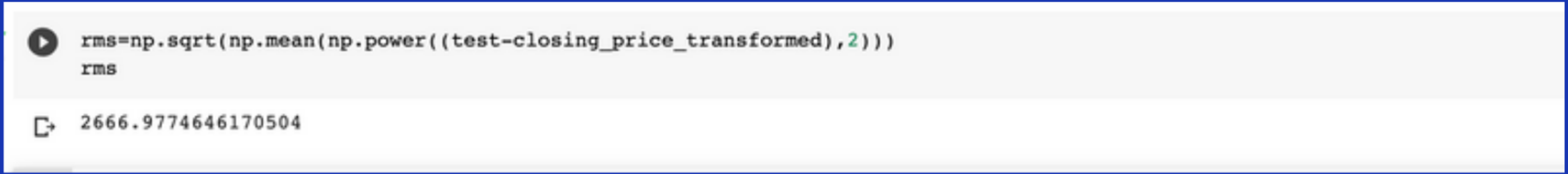
My prediction



15/7/21	33200.60945
16/7/21	33195.65697
17/7/21	33190.81631
18/7/21	33186.08495
19/7/21	33181.46006
20/7/21	33176.9392
21/7/21	33172.51984
22/7/21	33168.19936
23/7/21	33163.97562
24/7/21	33159.84611
25/7/21	33155.80879
26/7/21	33151.86123
27/7/21	33148.00152
28/7/21	33144.22732
29/7/21	33140.53681
30/7/21	33136.92798


Predicted next 30 days" Price
33240
33220
33200
33180
33160
33140
33120
7/7/21
8/7/21
9/7/21
10/7/21
11/7/21
12/7/21
13/7/21
14/7/21
15/7/21
16/7/21
17/7/21
18/7/21
19/7/21
20/7/21
21/7/21
22/7/21
23/7/21
24/7/21
25/7/21
26/7/21
27/7/21
28/7/21
29/7/21
30/7/21
31/7/21
1/8/21
2/8/21
3/8/21
4/8/21
5/8/21

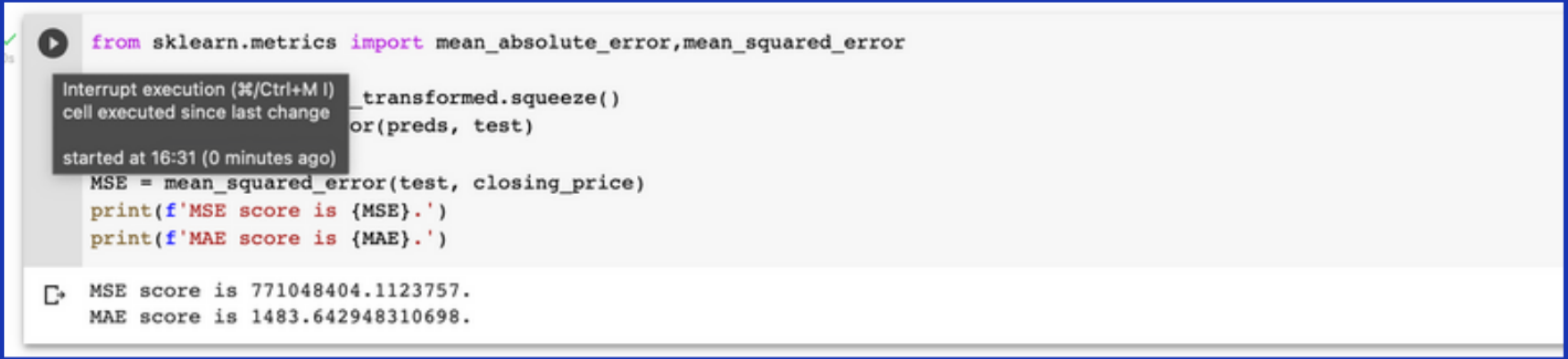

Bitcoin (BTC) Price Chart
Price
Market Cap
TradingView
24h
7d
14d
30d
90d
180d
1y
Max
Logarithmic
Linear
Dec 26, 2021
→
Jan 25, 2022
$55,000.00
$50,000.00
$45,000.00
$40,000.00
$35,000.00
$30,000.00
CoinGecko
28. Dec
30. Dec
1. Jan
3. Jan
5. Jan
7. Jan
9. Jan
11. Jan
13. Jan
15. Jan
17. Jan
19. Jan
21. Jan
23. Jan
25. Jan



```
rms=np.sqrt(np.mean(np.power((test-closing_price_transformed),2)))
rms
```

```
2666.9774646170504
```

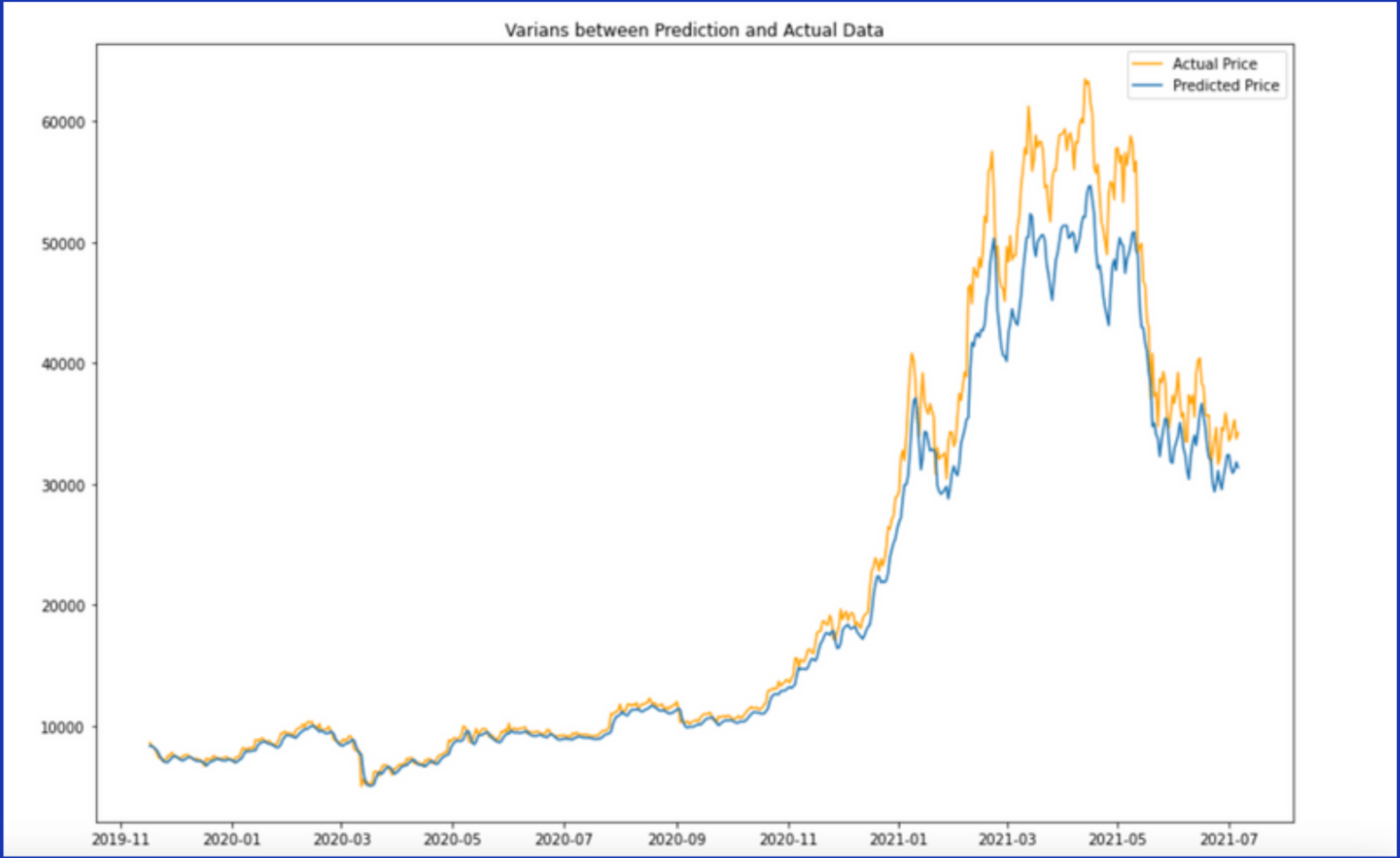
Errors



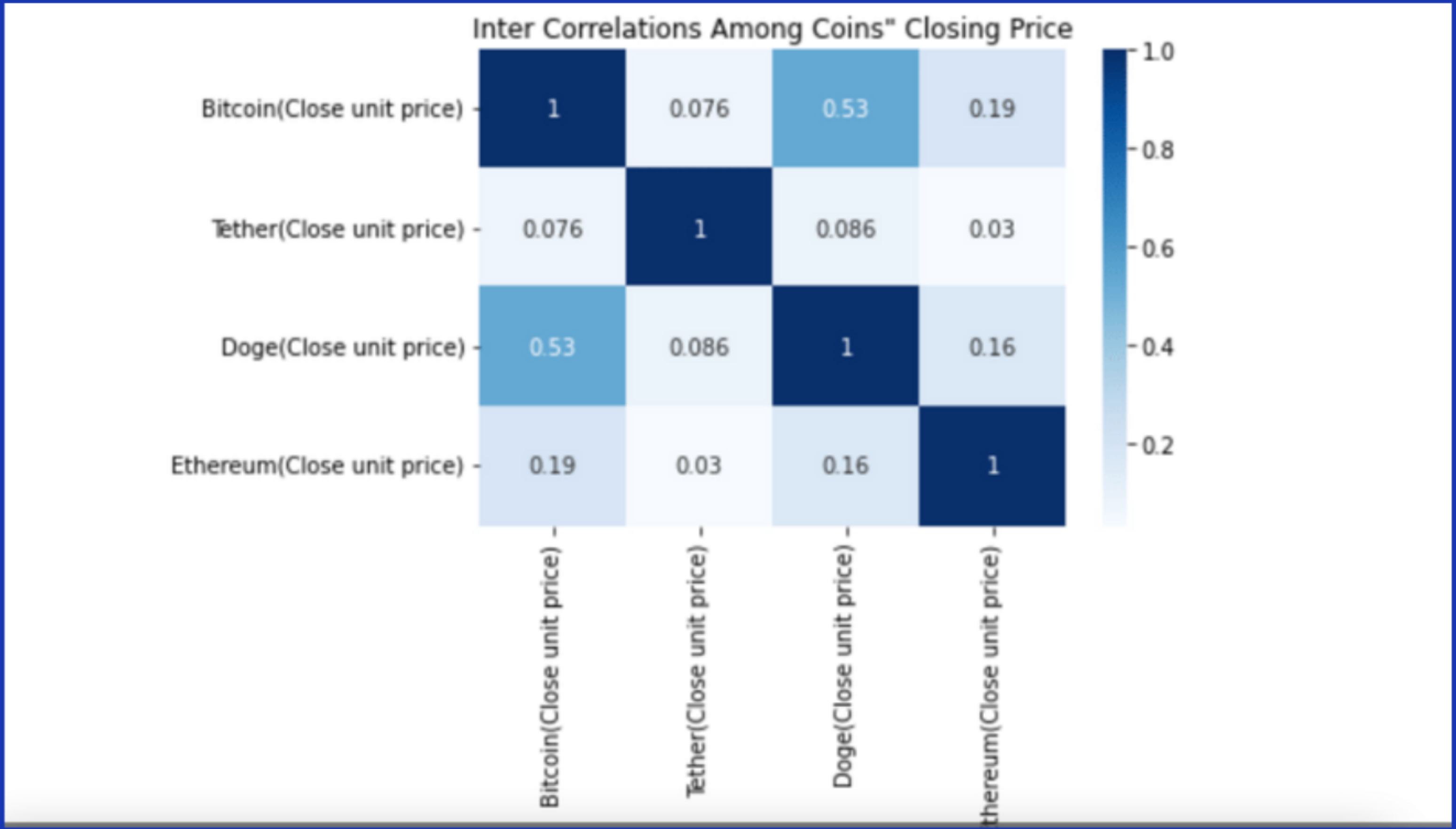
Interrupt execution (⌘/Ctrl+M I)
cell executed since last change
started at 16:31 (0 minutes ago)

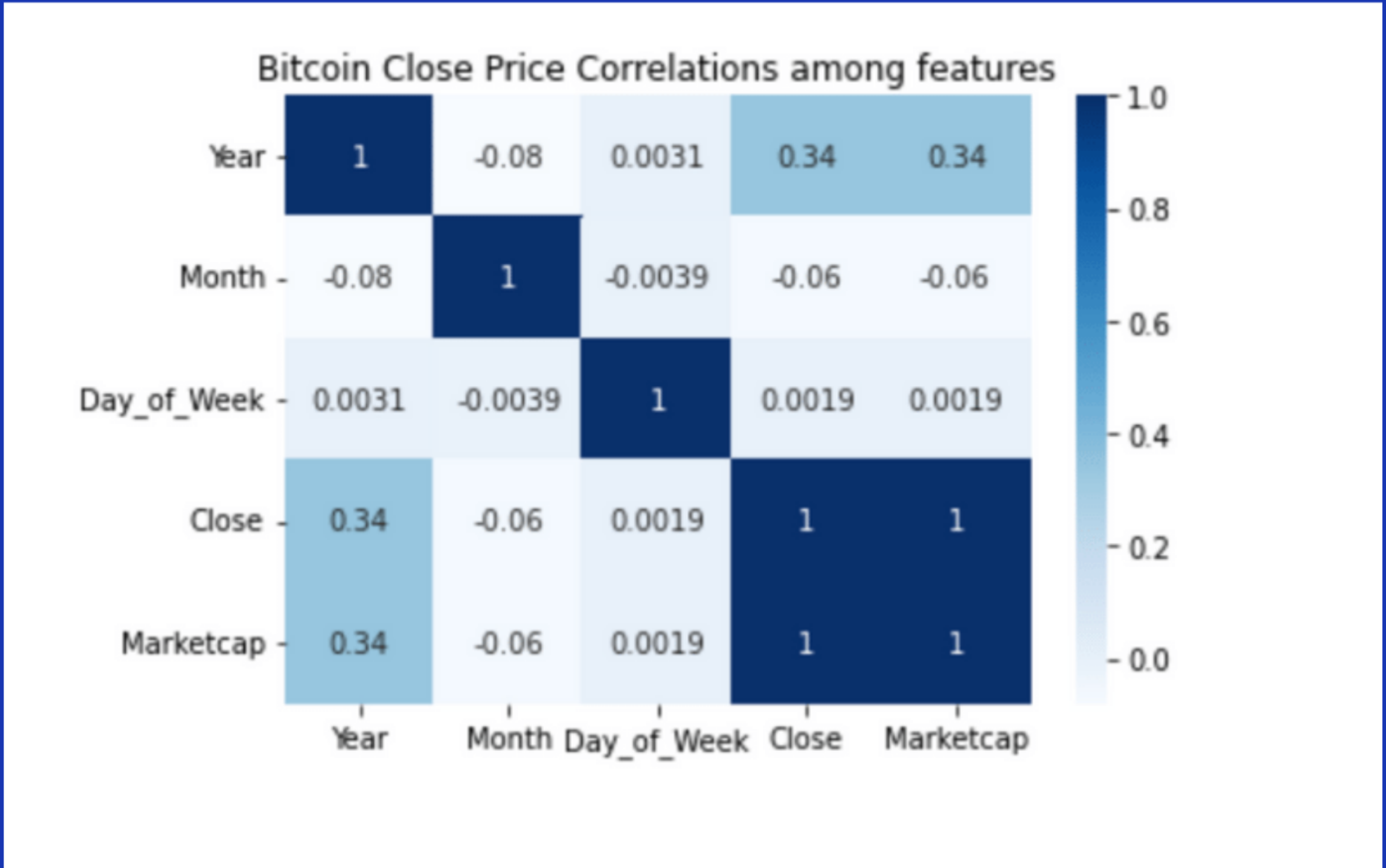
```
from sklearn.metrics import mean_absolute_error,mean_squared_error

                   _transformed.squeeze()
                  or(preds, test)

MSE = mean_squared_error(test, closing_price)
print(f'MSE score is {MSE}.')
print(f'MAE score is {MAE}.')
```

```
MSE score is 771048404.1123757.
MAE score is 1483.642948310698.
```


Varians between Prediction and Actual Data
Actual Price
Predicted Price
60000
50000
40000
30000
20000
10000
2019-11
2020-01
2020-03
2020-05
2020-07
2020-09
2020-11
2021-01
2021-03
2021-05
2021-07

Other findings





Bitcoin Close Price Correlations among features

	Year	Month	Day_of_Week	Close	Marketcap
Year	1	-0.08	0.0031	0.34	0.34
Month	-0.08	1	-0.0039	-0.06	-0.06
Day_of_Week	0.0031	-0.0039	1	0.0019	0.0019
Close	0.34	-0.06	0.0019	1	1
Marketcap	0.34	-0.06	0.0019	1	1

Reflections

Acceptable test result however prediction was not ideal, indicating many other overlooked factors?

Issues of overtfitting

Need a bit more trialling and erroring to find out within what date range the prediction would be more accurate etc.

Challenges

Time series Machine learning was a brand new thing

Unable to use Tensorflow from the local machine, faced with unfamiliarity with Colab, and time pressure

Stuck at the front end to back end part

Thank you!