Talk about the following:

Talk about timeseries, autocorrelation, ARIMA

On the section **Applying Machine Learning Model** (Cell 60) :

Talk about the different regression models used and how prediction were not accurate since they have not regarded the time series nature of the data

Search for this code if you couldn’t find cell numbers:

1. def regression(X\_train, X\_test, y\_train, y\_test): Regressor = {
2. 'Random Forest Regressor': RandomForestRegressor(n\_estimators = 200),
3. 'Gradient Boosting Regressor': GradientBoostingRegressor(n\_estimators = 500),
4. 'ExtraTrees Regressor': ExtraTreesRegressor(n\_estimators = 500, min\_samples\_split = 5),
5. 'Bayesian Ridge': BayesianRidge(),
6. 'Elastic Net CV': ElasticNetCV()
7. }

Explain the output of that code:

1. print('Bitcoin (BTC):')
2. regression(X\_train\_BTC, X\_test\_BTC, y\_train\_BTC, y\_test\_BTC)

Include the code for price prediction. Cell 64. Include only BTC and Ethereum

Now: We need to implement **ARIMA** to improve prediction:

Talk about stationarity and The [**Augmented Dickey-Fuller**](https://en.wikipedia.org/wiki/Augmented_Dickey%E2%80%93Fuller_test) test. You will find some info about it on the notebook. Please expand it a little

The below test shows that the BTC data we have is stationary:

## Explain whether stationary or not

1. from statsmodels.tsa.stattools
2. import adfuller X = bitcoin[0: -30]['daily\_avg'].values result = adfuller(X)
3. print('ADF Statistic: %f' % result[0])
4. print('p-value: %f' % result[1])
5. print('Critical Values:')
6. for key, value in result[4].items():
7. print('\t%s: %.3f' % (key, value))

Include the code of the ARIMA model right below. DO NOT explain the output. Still need to work on it

You will find bibliography on the draft doc. Please include your work in a separate document.

Please DO NOT use any colors for fonts or backgrounds.

Install the code format add on from MS Word.