

Project Report

GitHub URL

https://github.com/ivansusac/UCDPA_ivansusac

ABSTRACT

In this project I am using “All Cryptocurrencies price 2015-2021” dataset from Kaggle.com. I will clean and analyze data so I can extract insights and attempt to predict future prices with help of backtesting (general method for seeing how well a strategy or model would have done ex-post).

INTRODUCTION

I choose this project as I have personal interest in Cryptocurrencies and have been using trading BOT's in the past. It might be fun to learn how BOT's work and attempt to create my own predictions.

DATASET

<https://www.kaggle.com/ysthehurricane/all-cryptocurrencies-price-20152021>

The dataset includes 11 types of cryptocurrencies prices in USD currency for the last 6 years (2015-2021).

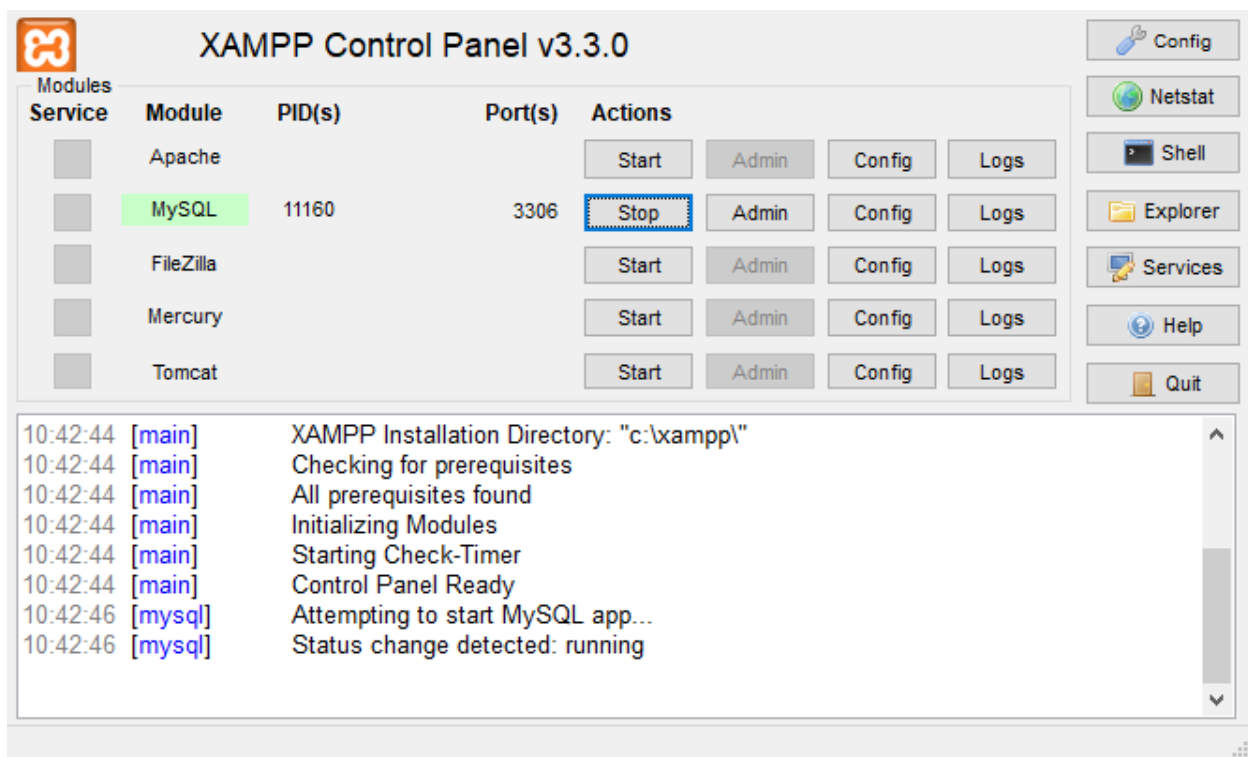
I will be analyzing only BTC, ETH, DOGE and XRP.

IMPLEMENTATION PROCESS

Importing data

Data was imported from csv file and local SQL database.

For SQL database I found that XAMPP is simple to install and easy to use:



Analyzing data

Firstly, I wrote a function that uses Regex and Iterators to check if all the dates are in correct format.

Checking datatypes

SQL datasets (BTC and DOGE) are imported as “strings”. It was fixed by converting all columns but first one to numeric value and first column is then converted to datetime.

Handling null values

I checked for missing values and replace them with a “backfill” method.

Handling duplicate data

There was no duplicated in the dataset. If we encounter duplicates I would drop them with `df.drop_duplicates()`.

Visualizing the data

Four plots are created for each cryptocurrency based on different values: Close price, Volume and Daily return. “Adjusted Close price” was not used at it is only relevant for stocks.

Merging Dataframes

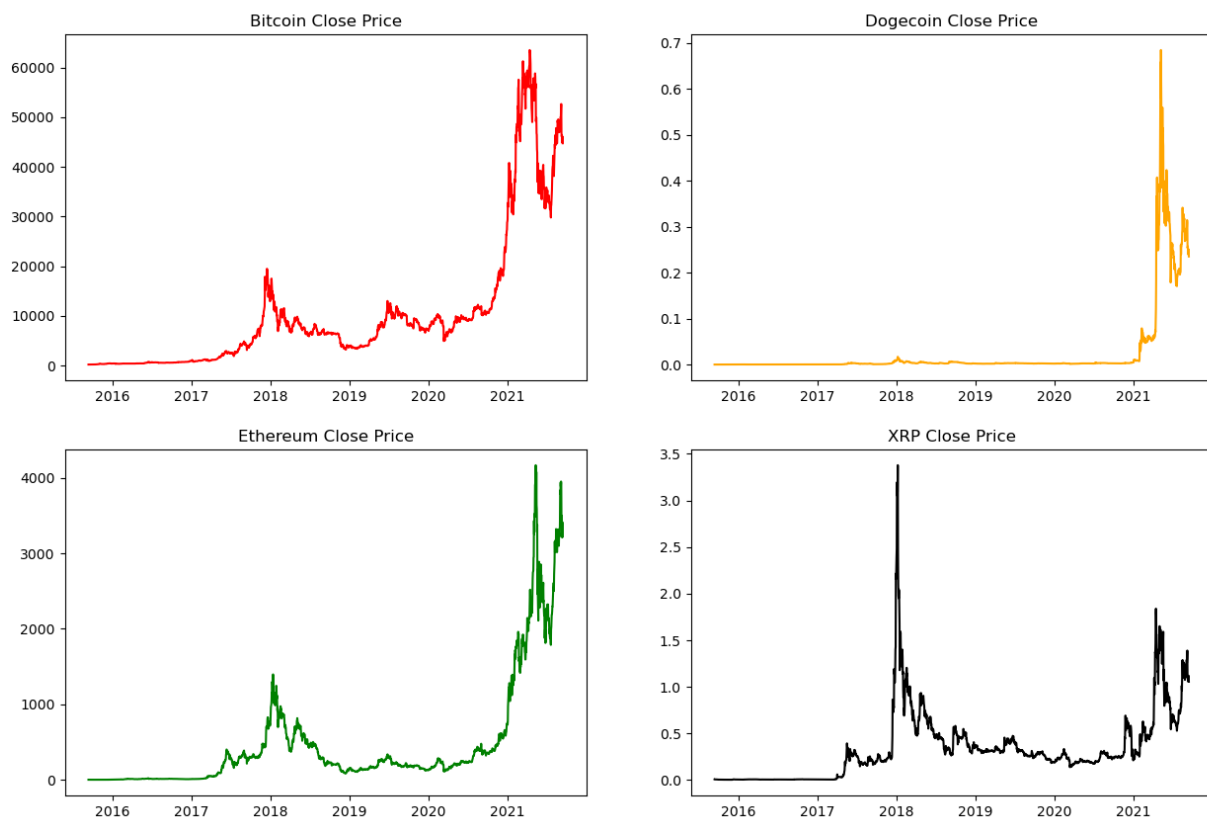
We merged dataframes keeping only “Date” and “Close” columns so we can check the correlation between currencies.

Machine learning

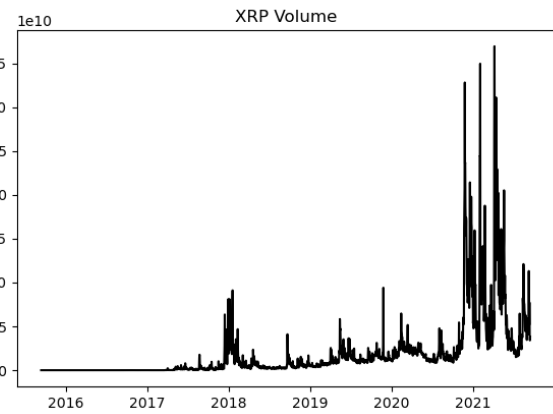
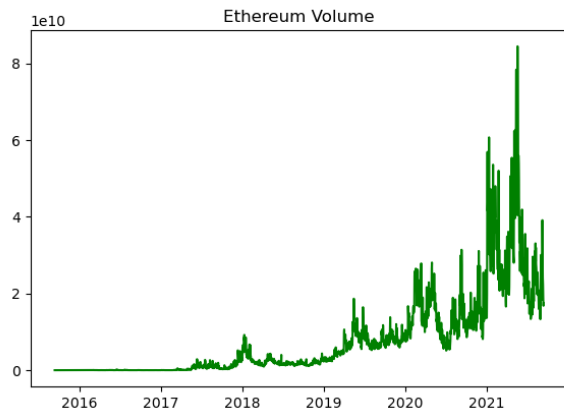
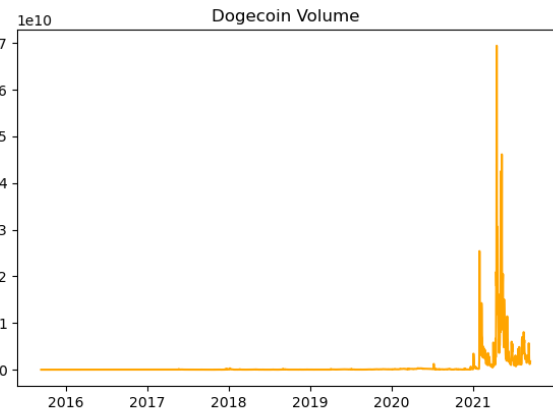
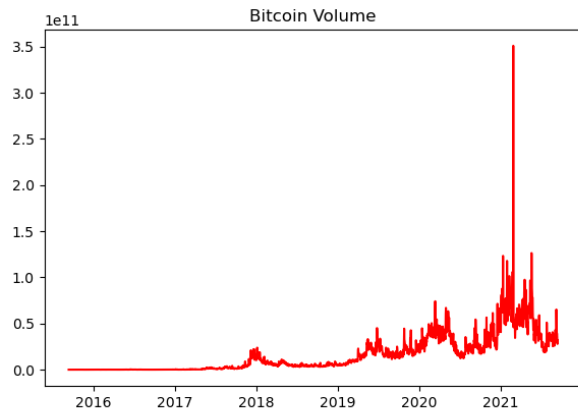
I did not manage to finish my model on time. I will need to focus more on this subject in the future.

RESULTS

Graph below is showing “**Close Price**” across all four currencies.

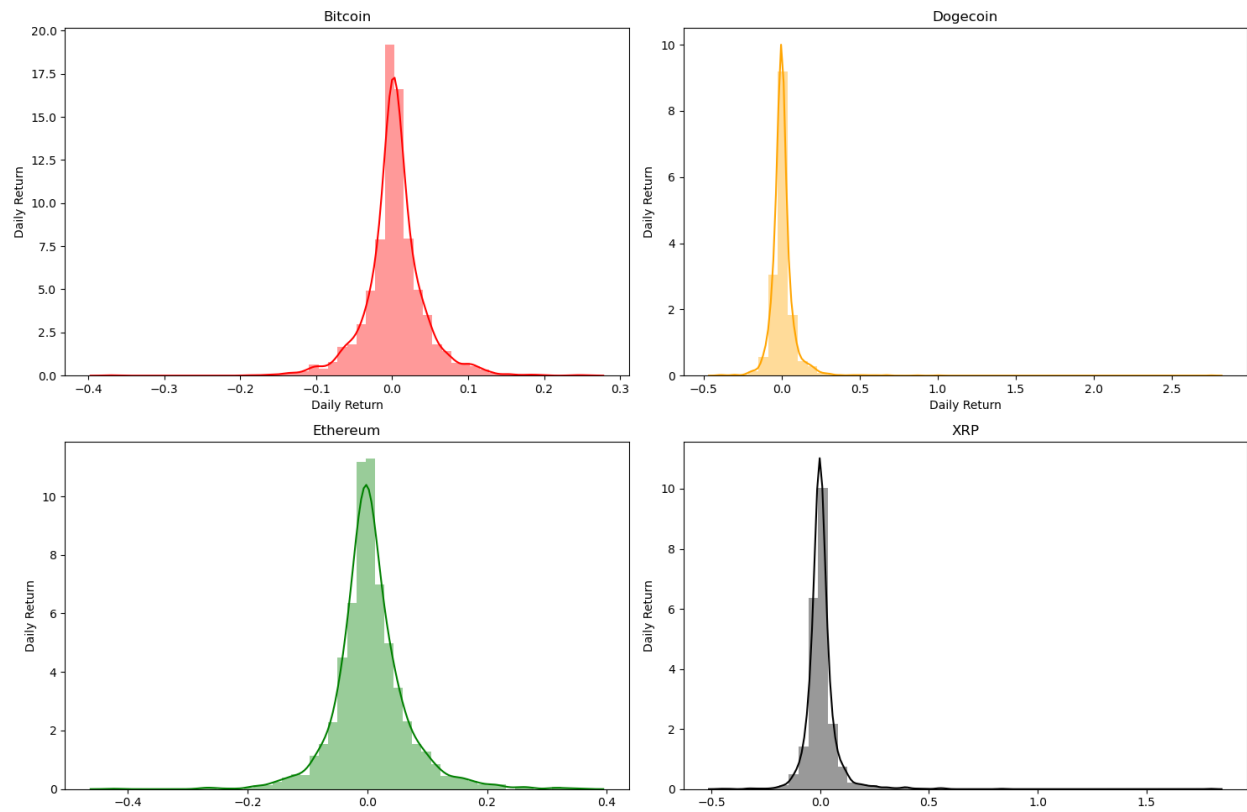


Graph below is showing “**Volume**” traded across all four currencies.



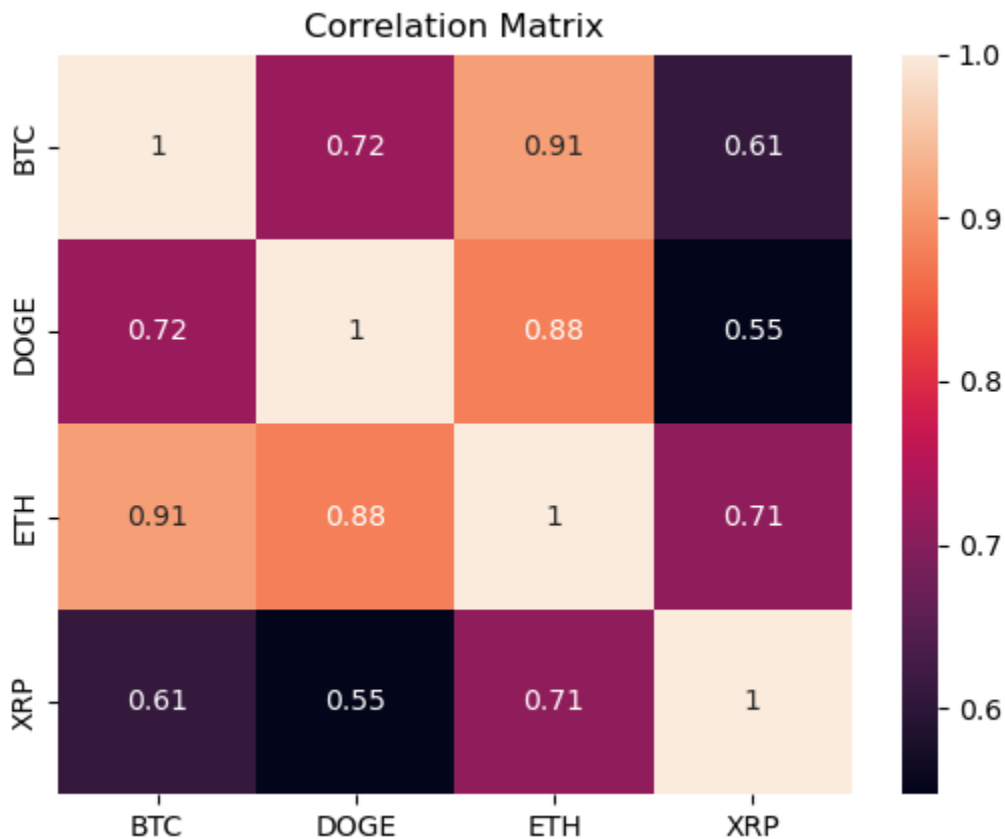
Graph below is showing “**Daily Return**” across all four currencies.

It shows how much in the % prices change daily.



INSIGHTS

1. BTC is trading at highest price and volume.
2. XRP had a peak value in 2018 while all others had a peak in 2021.
3. Traded "Volume" was highest in 2021 across all four currencies. This might indicate increasing interest from general population.
4. While all currencies are somewhat correlated, the highest correlation is between BTC and ETH.



5. DOGE and XRP are least correlated.