Process and Programming for Analytics

An Exploratory dashboard to monitor crypto markets for decision-making.

Graphical user interface

Description automatically generated with low confidence

# Abstract:

The cryptocurrency market is a highly unregulated, technically complex and financially very risky market. Most cryptocurrencies are based on blockchain technology enabling it nearly impossible to counterfeit, with complete automation and anonymity. The project analyses different cryptocurrencies over various years and visualizes essential time series price fluctuations on an interactive dashboard for better tracking of performance indicators and other metrics.

It is very crucial to understand the risk factor and trends of the cryptocurrency market for an investor/trader. The dashboard enables a user from a subset of cryptocurrencies to check price features, candlestick charts, line charts, bar charts etc.

## Data Collection:

Datasets could be collected by using Yahoo Finance/ mplfinance library, Crypto API or BeautifulSoup’s web scraping. To make things less tricky, I got help from GitHub where data was scraped into multiple files. From a GitHub repository, 23 CSV files of data from different cryptocurrencies are used. The combined and cleaned dataset is of 37082 rows x 9 columns with columns containing Name, Symbol, Date, High, Low, Open, Close, Volume and Marketcap.

Graphical user interface, text, application

Description automatically generated

Multiple CSV files are concatenated using the ‘concat’ function from the “Pandas” library within a loop parsing across all the CSV files in the Data folder. The SNo. Column is removed using the “drop” function and with the use of the “DateTime” library, the date data has been used for indexing and sorting the df data frame (i.e the combined data of all 23 cryptocurrencies considered). The final data looks like this:

Graphical user interface, table

Description automatically generated

It is a time series data starting from 2013 in some cases, but with a one-day interval of each value in the dataset. All the price metrics are in USD ($) for ease of comparison.

## Data Analysis:

To generate graphs over the timeline, ‘dt.datetime’ is used to set a starting date and the closing date. A function “get\_input” imports the cryptocurrency symbol, and the starting date and the closing date are taken as input to make the data analysis for any user more efficient. All the crypto symbols are also mapped to the real name of the cryptocurrency to be used for better data description and calling using “get\_crypto\_name”. Also “get\_data” function is used to import the corresponding CSV file for exploratory data charts of the specific cryptocurrencies. The function also creates a corresponding dataframe “dt\_cryptoname” for plotting various useful graphs.

“Matplotlib.pyplot” library is used to generate a bar graph that sorts the data in descending order and groups data by each cryptocurrency(here symbol) and market cap to show the biggest cryptocurrencies by their market value.

Graphical user interface

Description automatically generated with low confidence

Output is:

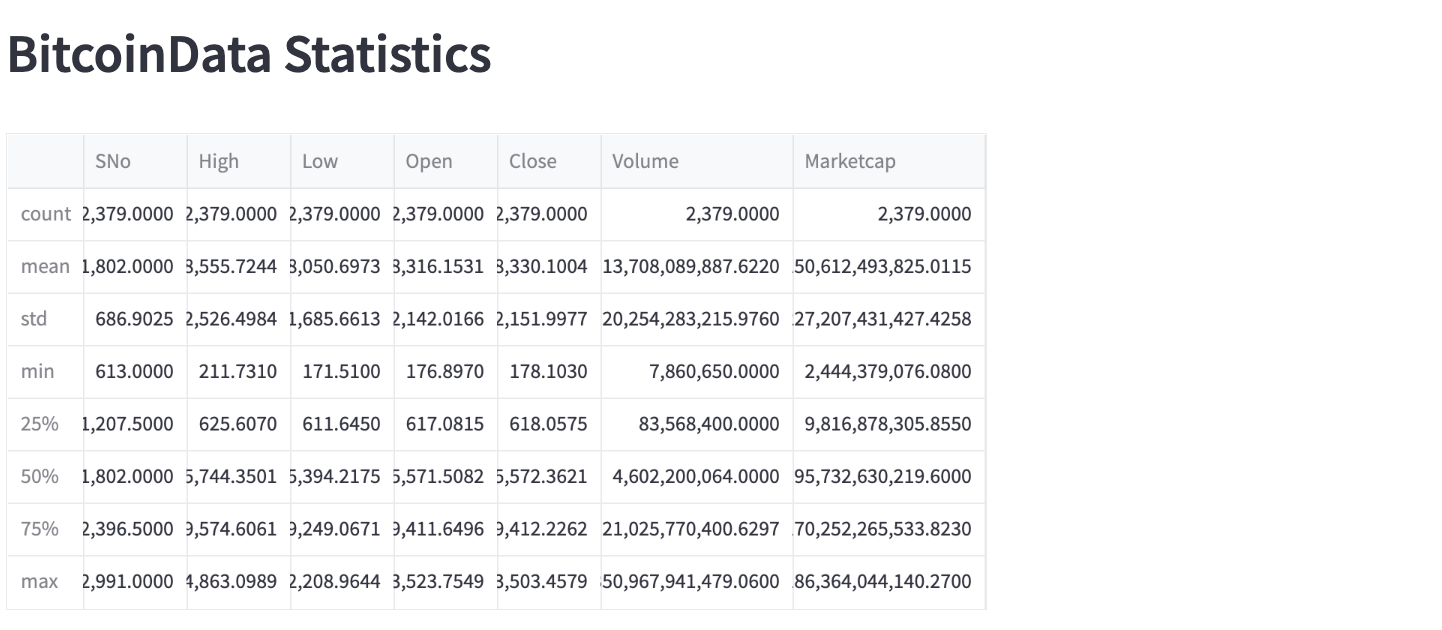
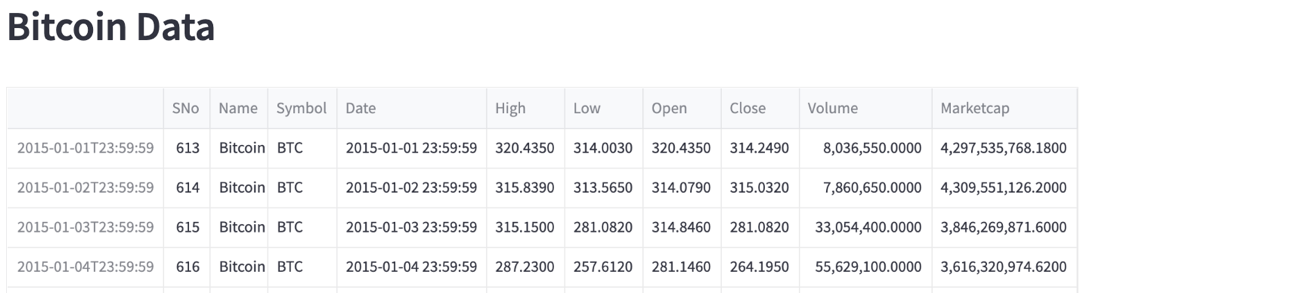
Graphical user interface, application

Description automatically generated

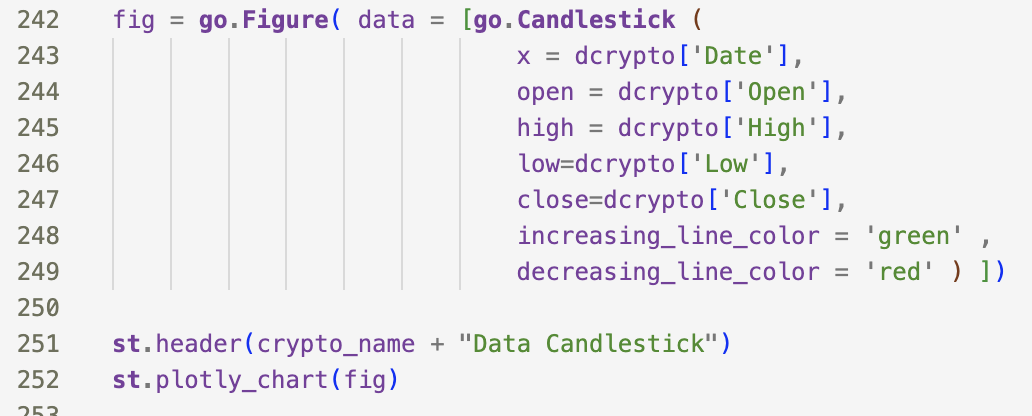
For people to understand the intensity of volatility and the latest data, both data and data statistics are also described. Line charts for the daily closing price, bar charts of the volume being traded on the day and the market cap of the selected cryptocurrency are made using inbuilt functions of the “Streamlit” library.

Graphical user interface, text, application, chat or text message

Description automatically generated

Output is: 

   “Plotly.graph\_objects” library is used to create a candlestick-style graph (used by a significant number of traders to gain insights into the trends and opportunities.)



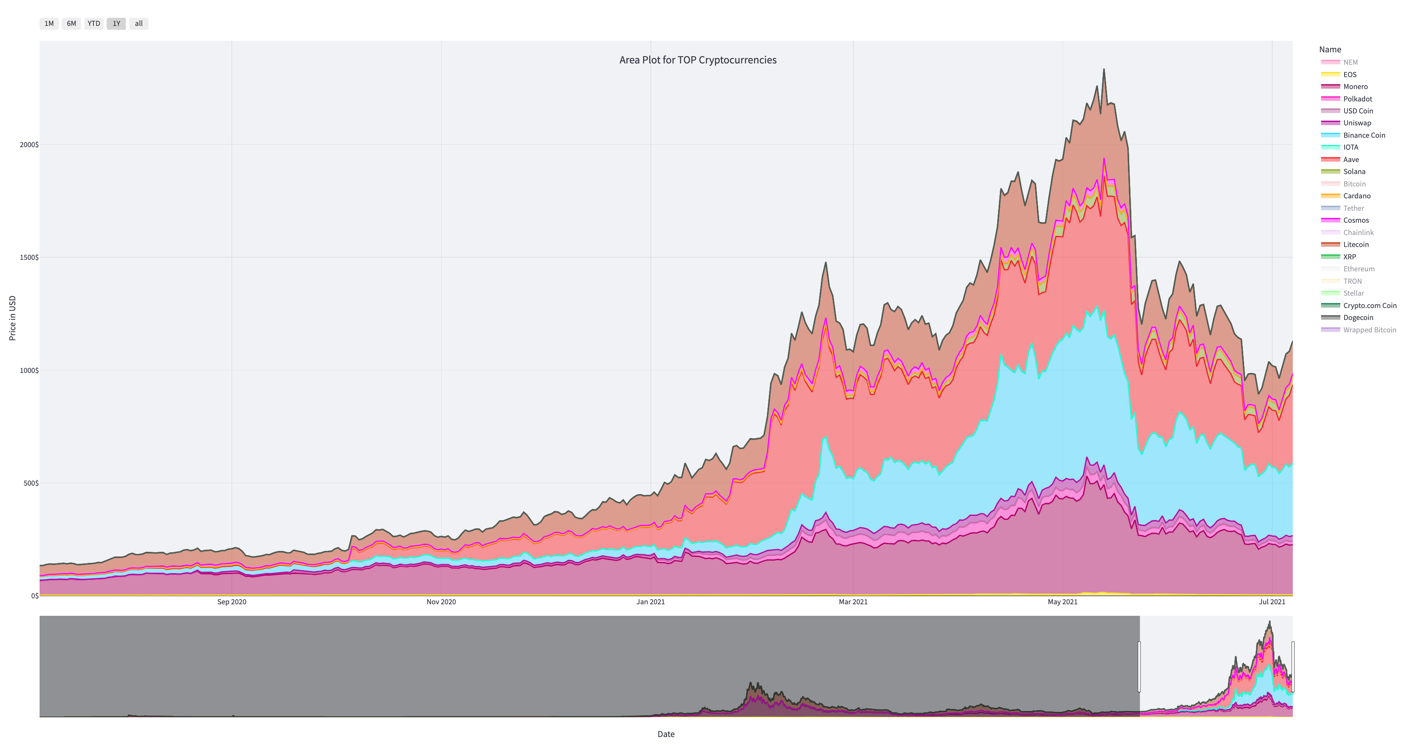
Output is:



“Plotly.express” library functions are used on our dataframe df(with merged data) to build two interactive area plot charts that describe the market cap and highest price of crypto over time. The duration of the data can be changed by the user using the buttons defined or the slider below. Also, a user can select any of the cryptocurrencies within the graph itself to form a better understanding of the cryptocurrency.



Output is:



Note:

1. All the graphs are described in an interactive dashboard using the “Streamlit” library.
2. Since seaborn graphs take some time to be computed each time the code is run. These graphs are described at the end so that other functionalities of the dashboard don’t suffer much.

Four line-plot graphs are created using “seaborn” library functions, which describe the closing prices of 4 top cryptocurrencies, the graph shows the comparison in the values of the tokens and how much have they increased as compared to others in the same period. Each graph omits the top one and compares the next top four.

Graphical user interface, text, application

Description automatically generated

Output is:

Graphical user interface, application

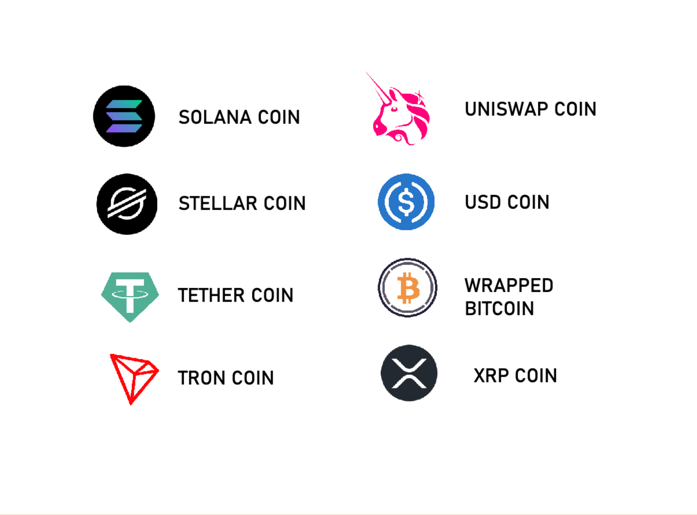
Description automatically generated

Histogram

Description automatically generated

## Conclusions:

The detail in the dataset includes – Date (The date of observations), Open (The opening price on the given day), High (The highest price), Low (The lowest price), Close (The closing price), Volume (the volume of transactions), Market Cap (The market capitalization in USD). The following cryptocurrencies were used for the data analysis.



The graphs and data show insights like the final close price, highest close price, lowest close price, average daily volume, comparison of close price within selected time, average volume based on date range, and price comparison table of various cryptocurrencies.

Based on the chart describing the top 5 cryptocurrencies, Bitcoin has the highest market cap value and Ethereum is halfway behind, the rest are very less in comparison. USDT is 3 times lesser than Ethereum. Also, in the candlestick chart of USDT, it is visible that it is very volatile.

The Value of Bitcoin reached an all-time high of around 60,000 USD, which is almost 15 times more than 2nd highest. Among the top cryptocurrencies, the growth of BTC, ETH, and USDT over the last five years was beneficial for investors.

## Recommendations:

This analysis is just for a data analytics project, this exploratory analysis of data is in no way a piece of financial advice. The cryptocurrency market is very volatile, unregulated and thus, very risky. It is recommended to invest in cryptocurrency if you are aware of the risks or the asset class and the factors that may affect its value.

With regards to the analysis done, more statistical, machine learning predictive models could also have been used. The data set used is a demo data set, the analysis can be made better with real time data imported using Yahoo Finance/ mplfinance library, Crypto API or BeautifulSoup’s web scraping. The analysis can be further improvesd by integrating sentimental analysis of data from twitter or digital media in general. An earlier research paper published by the author below discusses the relation of social factors affecting the cryptocurrency market. The value prediction models using machine learning and other complex methods can be integrated better to understand the crypto market better when used in tandem with the tools employed by traders in stock market.

## Reference:

. 2022. [PDF] Understanding the Social Factors Affecting the Cryptocurrency Market | Semantic Scholar. [ONLINE] Available at: <https://www.semanticscholar.org/paper/Understanding-the-Social-Factors-Affecting-the-Aggarwal-Patel/b0f2741f846fe08040a77c6bd59f0f9e8d6eff34>.

GitHub. 2022. Cryptocurrency-BTC-Data-Analysis/Data at main · pik1989/Cryptocurrency-BTC-Data-Analysis · GitHub. [ONLINE] Available at: https://github.com/pik1989/Cryptocurrency-BTC-Data-Analysis/tree/main/Data.