

**Project Name: Exploratory Data Analysis of Bitcoin Price in Python**

**By:** Rantu Adhikary

**Objective**

The objective of this project is to perform an exploratory data analysis on the historical prices of Bitcoin. The project provides a general overview of the relationships among the variables, explores their changes over time and gain insights about the data through coding and visualization techniques available in python libraries. It also cleans and prepares the data so that the data can be used for complex statistical analysis, machine learning or other data science projects.

**About the Data:**

The data comes in a .csv file which can be downloaded [here](https://github.com/Rantu7/Exploratory-Data-Analysis-of-Bitcoin/blob/main/coin_Bitcoin.csv). The dataset contains 2991 entries with the following columns:

SNo: Serial number

Name: Cryptocurrency name (Bitcoin)

Symbol: Cryptocurrency symbol (BTC)

Date: Date and time of the record

High: Highest price of the day

Low: Lowest price of the day

Open: Opening price of the day

Close: Closing price of the day

Volume: Trading volume

Marketcap: Market capitalization

**Data Analysis Steps:**

1. **Importing Python libraries** :

For data manipulation and mathematical calculation: Numpy & Pandas

For Visualization: Seaborn and Matplotlib.pyplot

1. **Importing Data**: The data was downloaded from the internet in .csv format. Then it was directly loaded in Python from the computer.
2. **Cleaning the data**: Though there are numerous ways to clean dirty data, the following actions were taken to clean the downloaded data:
   1. Modifying column values to improve readability
   2. Setting index column
   3. Changing data types for calculation and analysis
   4. Identifying and handling null values
   5. Identifying and handling duplicates
   6. Removing unnecessary columns
3. **Exploratory Data Analysis**:

The attached python file briefly explains each step of the EDA along with the code.

**Observations and Insights**:

* The market capitalization of Bitcoin, which is the total value of all Bitcoins in circulation, ranged from around **$778 million** to over **$1.1 trillion** during the given period. This highlights the growing importance and adoption of Bitcoin as a digital asset.
* The dataset shows significant price fluctuations in Bitcoin over the given time period. The price ranged from a low of around **$65.526** to a high of around **$64,863** indicating the high volatility of the cryptocurrency market.
* The trading volume of Bitcoin also varied significantly, with the highest daily volume reaching over **$126 billion** on May 19, 2021. This suggests that Bitcoin is a highly liquid asset, with a large number of market participants actively trading it.
* Bitcoin reached its peak closing price of **63503.458$** which is a **926.99%** increase from its lowest value of **68.431$**
* From the Heatmap graph, it is seen that Market capitalization of bitcoin has equal and a high degree correlation of **0.81** with ‘High’, ‘Low’, ‘Close’ and ‘Open’ price columns. This relation is also observed in the upward slopes of the ‘Pair plot’ in the file.
  + About correlation:
  + Correlation in statistics refers to the degree of relationship between two variables. It measures how changes in one variable are associated with changes in another variable. Three types of correlation are usually observed:
  + Positive Correlation (Correlation Coefficient = +1): A positive correlation indicates that as one variable increases, the other variable also tends to increase. In this scenario, the correlation coefficient takes a value of +1, indicating a perfect positive linear relationship between the variables.
  + Zero Correlation (Correlation Coefficient = 0): A correlation coefficient of 0 suggests that there is no linear relationship between the variables. Changes in one variable are not associated with changes in the other variable.
  + Negative Correlation (Correlation Coefficient = -1): A negative correlation indicates that as one variable increases, the other variable tends to decrease. In this case, the correlation coefficient takes a value of -1, indicating a perfect negative linear relationship between the variables.
* Though the prices, trading volume and market capitalization have been almost non-existing and stagnant from 2013 to around 2017, it shows dramatic rise and changes of each of these parameters since then till the end of the data.
* From 2016 to 2017, there was a rise of the highest and lowest closing price of bitcoin. However, highest closing price faced a decline till 2019. Since 2019, the highest and lowest closing price kept increasing almost proportionally until the highest price reached its peak in 2021.
* Unlike the closing price, trading volume has seen nothing but growth since 2016. In 2021, the highest volume of bitcoin sold was **350billion**, surpassing previous records.
* In the last 12 months of the data (August, 2020- July, 2021), the average close, open and high prices reached their highest in November 2020. Then they drastically kept dropping till April, 2021. However, these prices started seeing an upward trend from then till July, 2021.
* Though the close price of bitcoin declined between 2017 and 2019, the trading volume have always shown an upward trend since 2016. This means, buyers who invested in bitcoin and sold them within this time period faced major financial loss.

Thank you for reading!