

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | | | | | | | |
|  | | PROJECT SYNOPSIS | | | | |  | |
|  |  | | | | | | |  |
|  | | | |  |  | | | |
|  | | | | Real-Time EDA of Cryptocurrency |  | | | |
|  | | | | 28\02\2022—Abhishek Rana & Tanish Vermani—Mr. RAHUL RASTOGI |  | | | |
|  | | |  | | |  | | |

* Project Title

**Real-Time EDA of Cryptocurrency**

* Project Domain
* Data Analysis
* Data Science
* Regression Study
* Data Wrangling
* Project Description

* ***Exploratory Data Analysis (EDA)*** is an approach for analyzing data using a variety of techniques. The purpose of Exploratory Data Analysis is to summarize visualizations and statistics to better understand data, its quality and structure. The goal here is to try and answer all questions with data.

EDA is considered to be an iterative process since assumptions are first made on the exploratory visualizations, then build some models. We then make visualizations of the model results and tune our models. The various techniques used in EDA for analyzing data are:

1. Maximize insight into a data set

2. Understand structure of data

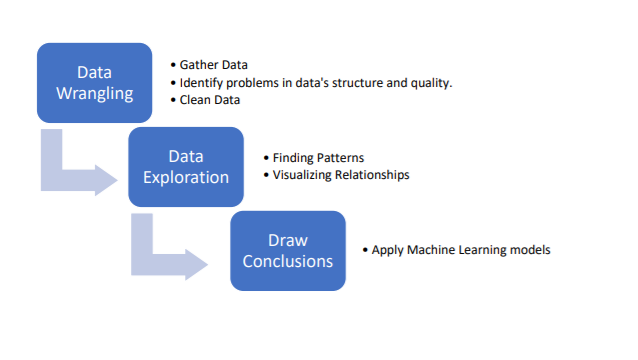
3. Extract important variables

4. Detect outliers and anomalies

5. Develop machine learning models

6. Determine optimal factor settings

* *Data Preparation , Cleaning And Visualization*
* Loading dataset into a data frame using pandas.
* Exploring the number of rows & columns , ranges of value etc.
* Handling missing , incorrect and invalid data.
* Exploring relationships between columns using scatter plots , bar charts , histograms etc.



* A ***Time Series Analysis* is nothing but a sequence of various data points that occurred in a successive order for a given period of time.**

Time Series Analysis is the way of studying the characteristics of the response variable with respect to time, as the independent variable. To estimate the target variable in the name of predicting or forecasting, use the time variable as the point of reference.

* ***Cryptocurrency*** is a digital payment system that doesn't rely on banks to verify transactions. [It’s a peer-to-peer system that can enable anyone anywhere to send and receive payments](https://www.fool.com/investing/2018/03/11/what-is-cryptocurrency.aspx). Instead of being physical money carried around and exchanged in the real world, cryptocurrency payments exist purely as digital entries to an online database describing specific transactions. When you transfer cryptocurrency funds, the transactions are recorded in a public ledger. Cryptocurrency is stored in digital wallets.

Cryptocurrencies run on a distributed public ledger called blockchain, a record of all transactions updated and held by currency holders. Units of cryptocurrency are created through a process called mining, which involves using computer power to solve complicated mathematical problems that generate coins. Users can also buy the currencies from brokers, then store and spend them using cryptographic wallets.

* *Data Science In Cryptocurrency*

A straightforward utilization of big data and data science in the crypto space is to perform [cryptocurrency](https://www.analyticsinsight.net/top-10-cryptocurrency-exchanges-to-trade-at-lower-fees-in-2022/) analytics. Big data infrastructure can handle the massive volume of [cryptocurrency](https://www.analyticsinsight.net/cryptocurrency-scams-5-ways-to-spot-avoid-and-protect/) data generated from transactions. Data science techniques can generate useful investment insights and predict future outcomes. By taking transaction data for analysis, it is possible to identify the price fluctuation of any given crypto (doing Bitcoin future predictions, for example), enabling investors to improve profitability and prevent substantial losses. In addition, crypto forecasting can also be trained using social-based data. Information like user activities and participation, combined with transaction data, current market price and computational powers, better prediction on the market volatility over time can be generated.

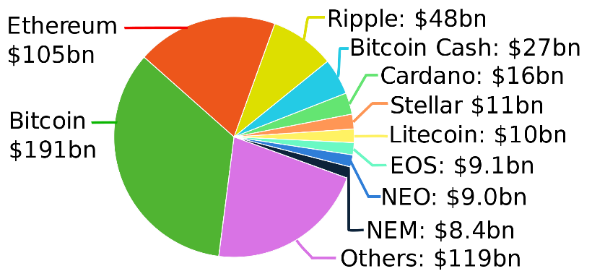
* Literature Study

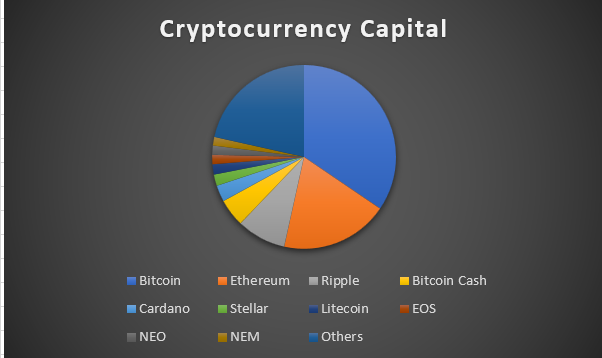
1. A pseudonymous computer scientist called Satoshi Nakamoto proposed bitcoin for the first time in 2008. The main purpose of this digital currency has been to use it as means of exchange, independent of any central authority. By encrypting every single bitcoin with advanced mathematical principles, secure electronic transactions are ensured by cryptocurrency miners. Although it is a legit usage, any digital cryptocurrency might usually be called as ‘Bitcoin’ among the traders since Bitcoin is the first cryptocurrency with the highest amount of market capital 194 billion dollars in the market in February 2018.
2. Many cryptocurrencies have emerged since it was invented in 2008. Names of some the major market “players” (Tokens) are,

[ “Ethereum”,” Ripple”,” Bitcoin”,”Litecoin”].

1. Some of the top Cryptocurrency in 2022, (w.r.t market capital)

[(Revised Data: 2021)](https://en.wikipedia.org/wiki/List_of_cryptocurrencies)





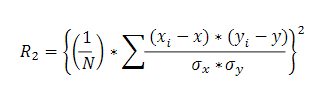
1. **Description**:

To study the changes in the cryptocurrency, we adopted the EDA approach to plot the real-time changes in the of (approximately) all major cryptocurrency (*criteria* -> at least $10Bn of invested capital).

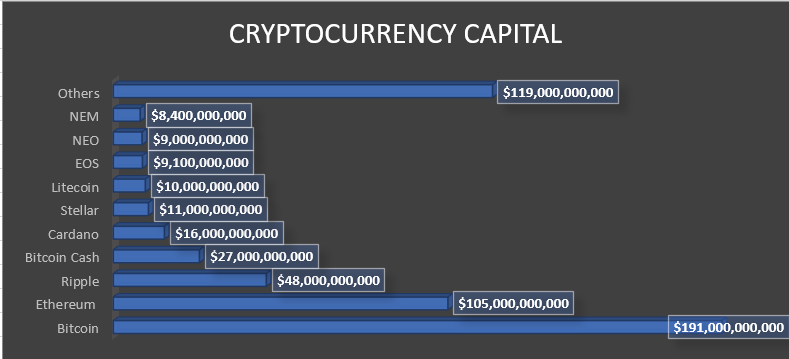
1. When using regression on cryptocurrency, we have to take note of few crucial aspect of the correctness of the regression model results.

Namely:

* Data Preparedness
* Data Understanding
* Data Modeling



* Evaluation
* Conclusion



1. **End Note --**The cryptocurrency market is a rapidly expanding canvas of trade and investment that has garnered the attention of traders, investors, entrepreneurs on a worldwide scale that is unprecedented in this century. By providing comparative studies and findings from the price data of cryptocurrency markets, it will further help document the behavior and habits of such a lucratively challenging and rapidly expanding market. In conclusion, this research deals with regression-based analysis algorithms in order to predict the price direction of Bitcoin.
2. The Microsoft Azure Machine Learning Studio were used in conducting the experiments with the Bitcoin dataset. Future work may explore other datasets with more features that could contribute in predicting a more reliable and accurate bitcoin price rate.
3. This Web app will provide the following enhancing features to the consumers.

* Readily access of the app
* Easy to use & Understand UI/UX
* Comparing two or more cryptocurrency would be possible
* Regression model[in beta stage] but available to process real time data to predict at least 89% accuracy.
* Software Specification
* Vs Code
* Python
* Stream-lit
* NumPy, Pandas
* Google Collab
* Hardware Specification
* Windows or mac(64-bit)
* Web browser (Web 2.0)
* Internet Connection
* Feasibility Study
* Implementation can be done using python and web-app stream-lit.
* Data set can be retrieved from various sources like Kaggle, Google API, etc.
* This case study can be practically implement using front end languages and API to create a dynamic and interactive web page.
* Team information

|  |  |  |
| --- | --- | --- |
| UID | Name | Signature |
| 20BCS3815 | Tanish Vermani |  |
| 20BCS3818 | Abhishek Rana |  |

* References:

1) Cryptocurrency History - [FULLTEXT01.pdf (diva-portal.org)](https://www.diva-portal.org/smash/get/diva2:1295584/FULLTEXT01.pdf)

2) Cryptocurrency Data - <https://en.wikipedia.org/wiki/List_of_cryptocurrencies>

3) Regression Analysis Study - <https://www.researchgate.net/publication/334452977_Regression_based_Analysis_for_Bitcoin_Price_Prediction>