

Exploratory Data Analysis (EDA) Report

Project Title: Cryptocurrency Liquidity Prediction for Market Stability

1. Introduction This report details the exploratory data analysis (EDA) conducted on the cryptocurrency market dataset sourced from CoinGecko, dated March 16, 2022. The primary objective of this EDA is to uncover underlying patterns, identify key trends, and examine correlations between different market features. The insights derived will inform the subsequent feature engineering and model development phases of the project.

2. Dataset Overview The dataset comprises market data for various cryptocurrencies. Key features include:

- **coin:** The name of the cryptocurrency.
- **price:** The trading price in USD.
- **1h, 24h, 7d:** Percentage price change over the last 1, 24, and 7 days.
- **24h_volume:** The trading volume over the last 24 hours.
- **mkt_cap:** The total market capitalization of the cryptocurrency.

3. Key Findings & Visualizations

3.1. Top Cryptocurrencies by Market Capitalization To identify market leaders, the top 5 cryptocurrencies were filtered based on their market capitalization. As of March 2022, the leading assets were Bitcoin, Ethereum, Tether, BNB, and USD Coin. The visualization below clearly illustrates Bitcoin's dominance in the market.

3.2. Feature Correlation Analysis A heatmap was generated to analyze the linear relationships between the numerical features in the dataset.

- **Key Insight:** A moderate positive correlation of **0.64** was observed between a cryptocurrency's price and its 24h_volume. This suggests that higher-priced assets tend to have higher trading volumes, which is a logical market dynamic. Other correlations were weak, indicating that a non-linear model, such as the chosen Random Forest, is appropriate for this problem.

3.3. Price vs. 24-Hour Volume A scatter plot was used to further investigate the relationship between price and 24h_volume.

- **Observation:** The plot confirms that assets with the highest prices (e.g., Bitcoin) also command the highest trading volumes, reinforcing the finding from the correlation heatmap.

4. Conclusion The EDA provided critical insights into the market structure and feature relationships. The dominance of a few key cryptocurrencies and the correlation between price and volume are important factors for the predictive model. The analysis confirms that the selected features are relevant for predicting market dynamics, including liquidity.