Cryptocurrency Volatility Prediction

---

1. Introduction

Exploratory Data Analysis (EDA) is a critical step in understanding the structure, patterns, and behavior of the dataset before building a machine learning model.
This report presents an analysis of historical cryptocurrency market data to identify trends, relationships, and volatility behavior.

---

2. Dataset Overview

2.1 Dataset Description

The dataset contains **daily historical market data** for more than 50 cryptocurrencies. Each record represents one day of trading information for a cryptocurrency.

2.2 Features Included

- Date
- Cryptocurrency Symbol
- Open Price
- High Price
- Low Price
- Close Price
- Trading Volume
- Market Capitalization

---

3. Data Quality Check

3.1 Missing Values

- Missing values were identified in rolling calculations and initial records
- Rows with missing values were removed to maintain data consistency

3.2 Data Types

- Date column converted to datetime format
- All price-related columns stored as numerical values

---

4. Descriptive Statistics

Key statistics were computed to understand data distribution:

- Mean, median, and standard deviation of prices
- Range of trading volume
- Market capitalization spread across cryptocurrencies

These statistics indicate **high variability**, which is expected in cryptocurrency markets.

---

5. Price Trend Analysis

5.1 Closing Price Trends

Line plots were used to analyze the closing price trends of major cryptocurrencies.

Observations:

- Prices show sharp fluctuations over time
- Sudden spikes and drops indicate high market volatility
- Long-term trends vary across different cryptocurrencies

---

6. Volatility Analysis

6.1 Volatility Distribution

A histogram of rolling volatility was plotted.

Observations:

- Volatility is right-skewed
- Most days show moderate volatility
- Extreme volatility events occur less frequently

---

7. Volume and Market Capitalization Analysis

7.1 Trading Volume

- Trading volume varies significantly across cryptocurrencies
- Higher volume often corresponds to increased volatility

7.2 Market Capitalization

- Large-cap cryptocurrencies show relatively stable volatility
- Small-cap cryptocurrencies experience higher volatility fluctuations

---

8. Correlation Analysis

A correlation heatmap was generated to examine relationships between numerical features.

Key Observations:

- Strong correlation between open, high, low, and close prices
- Moderate correlation between volatility and price movement
- Liquidity-related features show meaningful influence on volatility

---

9. Feature Engineering Insights

During EDA, the following insights supported feature engineering:

- Rolling statistics effectively capture volatility behavior
- Technical indicators add valuable information beyond raw prices
- Liquidity ratio provides insight into market activity

---

10. Outlier Analysis

- Extreme price movements were observed during certain periods
- These outliers represent real market events and were retained
- Random Forest model is robust to such outliers

---

11. Key EDA Findings

- Cryptocurrency markets exhibit high volatility
- Price movement strongly influences volatility
- Liquidity and trading volume play important roles
- Feature engineering is essential for accurate volatility prediction

---

12. Conclusion

The EDA provided critical insights into the behavior of cryptocurrency market data. Understanding price trends, volatility distribution, and feature correlations helped guide feature engineering and model selection for the volatility prediction task.