Exploratory Data Analysis (EDA) Report

Dataset Overview:

- Dataset used: coin_gecko_2022-03-16.csv and coin_gecko_2022-03-17.csv
- Key columns: price, 24h change, market cap, volume

EDA Performed:

- Missing value handling using dropna()
- Summary statistics using describe()
- Data types inspection and column listings
- Histograms plotted for all numeric columns using matplotlib

Observations:

- Price and volume data showed right-skewed distributions
- Strong variance observed in 24h change and market cap
- Large scale values in volume required log transformation for modeling

Conclusions:

EDA helped identify key features for model training (price, 24h change, market cap) and guided the decision to apply feature scaling and log transformation on volume.

Source Code Summary

Notebook: sachin_Crypto_Liquidity_Project.ipynb

Code Modules:

1. Data Upload and Loading

- Used files.upload() in Colab
- Loaded with pandas.read_csv()

2. Data Preprocessing

- Null value handling with dropna()
- Feature selection: price, 24h, market cap
- Normalization using StandardScaler

3. Modeling

- Linear Regression
- Random Forest with log-transformed target

4. Evaluation

- Metrics: RMSE, MAE, R^2 Score
- Comparison between linear, random forest, and log models

5. Exporting

- Trained model and scaler saved using pickle

Code is well-commented and written for clarity, following clean ML pipeline structure.