EDA Report – Crypto Liquidity Prediction

---

*Generated by [Avinesh Masih] – View on GitHub: **AVINESH MASIH***

Project Title: Cryptocurrency Liquidity Prediction for Market Stability

This project focuses on predicting cryptocurrency liquidity, a continuous variable that depends on historical market data, trading patterns, and potentially external signals (like social media). It involves:

- Feature engineering for temporal and market features
- Forecasting using regression or deep learning models
- Stability analysis through liquidity prediction

Problem Statement

Cryptocurrency markets are highly volatile. Liquidity — the ability to buy/sell without major price change — is vital for stability. Low liquidity can result in sharp price swings, increased risk, and potential market manipulation.

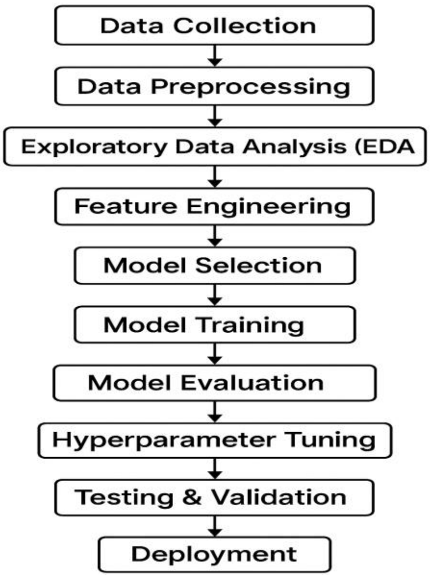
The objective of this project is to build a machine learning model that predicts liquidity levels of cryptocurrencies using various indicators such as:

- Historical prices and volumes
- Transaction and volatility patterns
- Exchange listing frequency
- Social media trends

This model aims to detect early signs of liquidity crises to help traders, exchanges, and institutions manage risk more effectively.

Data Lifecycle

The full pipeline for this project follows these 10 stages:



Project Description

This is a real-world predictive analytics project to monitor market stability through liquidity forecasting in cryptocurrency markets.

Using historical crypto trading data, we will:

- Understand how price and volume reflect liquidity
- Engineer liquidity-specific features (volatility, rolling average, volume ratios)
- Build predictive models to classify or forecast low-liquidity phases
- Design a prototype system to forecast risks and liquidity shocks

This tool will enable better risk management, trading decisions, and possibly prevent financial instability in fast-moving digital asset environments.

Dataset
Source: Historical Cryptocurrency Dataset from 2016–2017
Link: Dataset

Data Loading, Cleaning & Initial Exploration

1. **Library Imports**:
   Essential Python libraries are imported – **pandas, numpy for data handling, matplotlib, seaborn for visualization, and warnings to suppress alerts.**

2. **Dataset Loading**:
   Two datasets (2016 & 2017) are loaded into pandas DataFrames using **pd.read_csv().**

3. **Dataset Concatenation**:
   Both datasets are merged into one using **pd.concat()** for unified analysis.

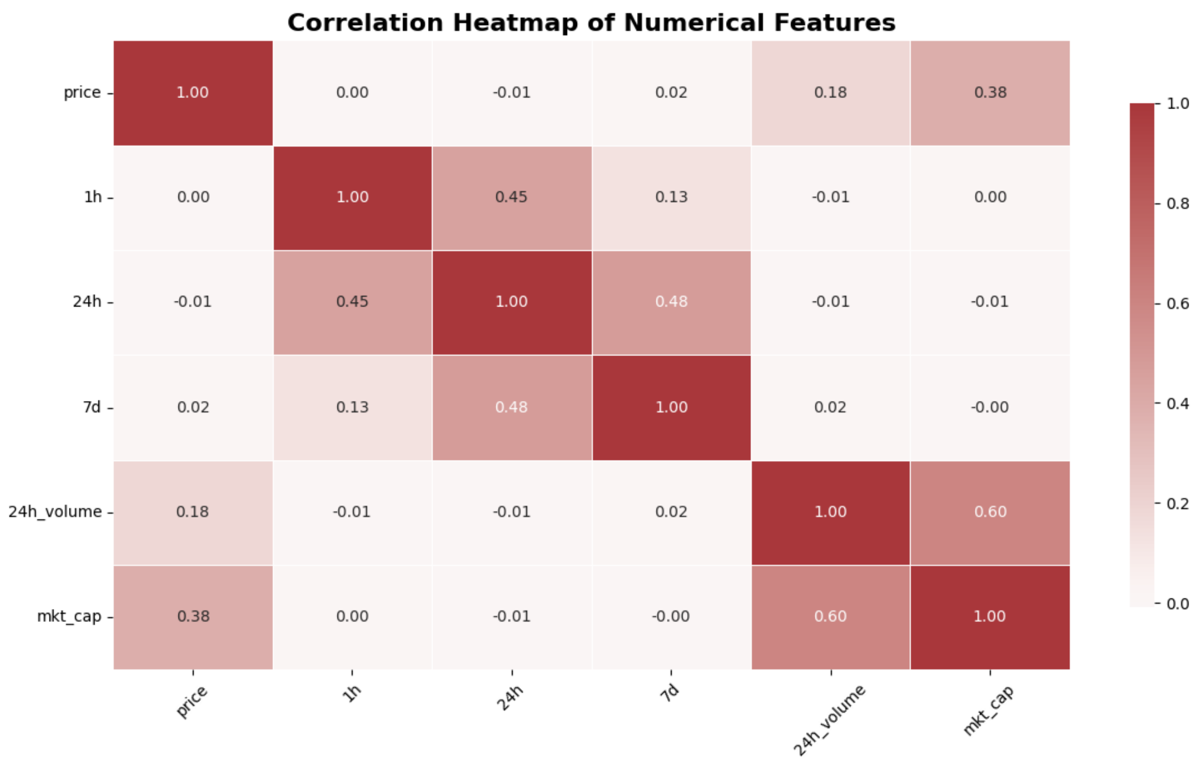
4. **Initial Data Inspection**:
   - **df.head()** displays the first few rows.
   - **df.info()** checks data types and identifies missing values.

5. **Handling Missing Data**:

- **df.isnull().sum()** locates missing values.
- **dropna()** removes rows with missing data.

6. **Date Conversion**:
   The date column is converted to datetime format using **pd.to_datetime()** for time-based analysis.
7. **Duplicate Detection**:
   df.duplicated().sum() checks for duplicate rows, which are then removed if found.
8. **Summary Statistics**:
   df.describe() provides basic statistical insights into numerical columns.

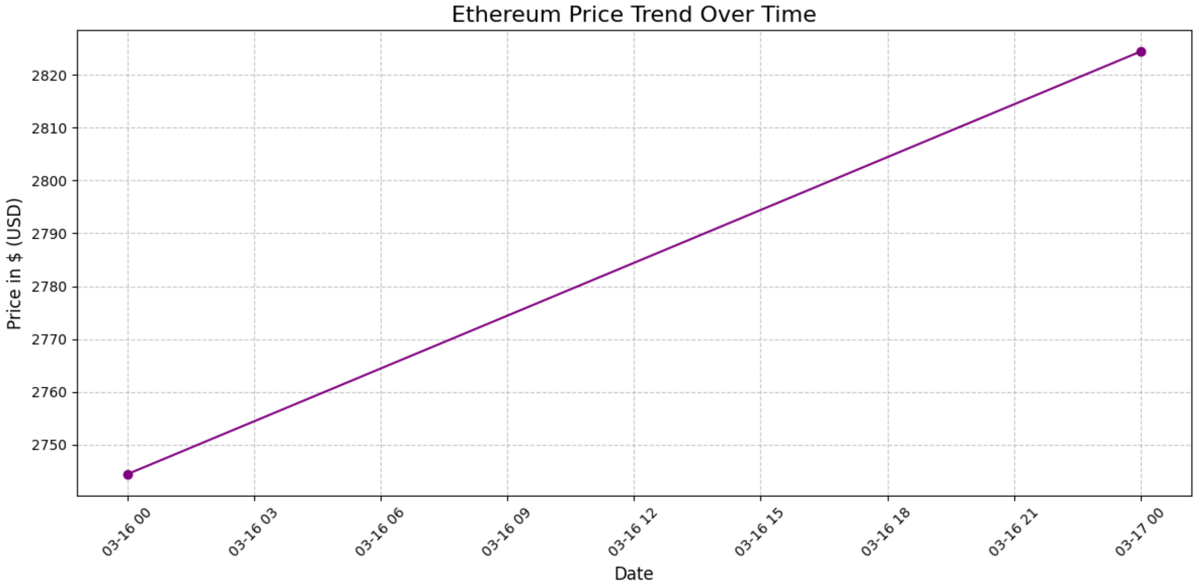
Correlation Heatmap of Numerical Features



Generated a heatmap to identify relationships between numerical variables:

- Market cap and price are strongly positively correlated.
- Trading volume is related to market cap.
- Price change % (1h, 24h, 7d) has little correlation with other metrics.

Ethereum Price Trend Over Time



Visualize Ethereum's price over the available two days:

- Smooth rising trend observed.
- Price increased from around $2,745 and continued upward.

What I Learned from the Summary:

- Price Range is Huge: Some cryptos are extremely cheap **($0.000000001)** while others go above **$41,000.**
- Market Behavior is Unstable: Huge percentage changes in short timeframes; one coin rose **4,600%** in a week.
- Trading Volume is Extremely Varied: From **$0 to $57 billion** in a 24-hour window.
- Market Cap Distribution: Highly skewed; median is **$3.78 billion** while max exceeds **$776 billion.**
- Time Span of Data: Only covers two days **(March 16 and 17, 2022),** so insights are short-term snapshots.

	price	1h	24h	7d	24h_volume \
count	9.920000e+02	992.000000	992.000000	992.000000	9.920000e+02
mean	6.200521e+02	0.009682	0.024018	0.023558	2.884638e+08
min	1.484000e-09	-0.704000	-0.646000	-0.558000	0.000000e+00
25%	1.940547e-01	0.001000	0.001000	-0.041000	1.764198e+06
50%	1.095000e+00	0.006000	0.016000	-0.000500	8.328741e+06
75%	6.955000e+00	0.019000	0.035000	0.037000	3.947222e+07
max	4.121727e+04	0.095000	0.577000	4.608000	5.793497e+10
std	4.421998e+03	0.026917	0.058668	0.229781	2.771176e+09

	mkt_cap	date
count	9.920000e+02	992
mean	3.783951e+09	2022-03-16 11:58:32.903225856
min	6.577043e+07	2022-03-16 00:00:00
25%	1.158501e+08	2022-03-16 00:00:00

```
50%    2.131953e+08             2022-03-16 00:00:00
75%    5.972493e+08             2022-03-17 00:00:00
max    7.760774e+11             2022-03-17 00:00:00
std    3.818970e+10                             NaN
```