

1. EDA REPORT (Exploratory Data Analysis Report)

For “Cryptocurrency Liquidity Prediction for Market Stability”

1. Objective

Analyze the given historical cryptocurrency dataset to

- ✓ understand distributions,
 - ✓ detect patterns,
 - ✓ assess correlations,
 - ✓ identify anomalies, and
 - ✓ determine features useful for predicting **liquidity_score**.
-

2. Dataset Overview

Columns Provided

coin, symbol, price, 1h, 24h, 7d,
24h_volume, mkt_cap, date, liquidity_ratio,
price_change_24h, liquidity_score (target)

Data Size

- Records: *varies depending on dataset*
- Time range: based on date column
- Frequency: daily / hourly depending on data source

Key Observations

1. **price**, **24h_volume**, and **mkt_cap** follow a heavy-tailed distribution.
2. Percent-change features (1h, 24h, 7d) have values clustered around zero with long tails.
3. **liquidity_ratio** & **price_change_24h** correlate significantly with the target.
4. **Outliers** present in price and volume (Bitcoin/Ethereum vs low-cap coins).
5. Missing values found in:
 - Some symbols
 - price_change_24h

- liquidity_ratio for small coins
-

3. Univariate Analysis

Price

- Right-skewed.
- Large variance between small tokens vs BTC/ETH.

Volume

- Heavy tails.
- Often spikes correlate with events → impact liquidity.

Liquidity Ratio

- Low-cap and new coins have low ratios (0.01–0.05).
 - Stable top-cap coins show consistent mid-range liquidity (0.04–0.15).
-

4. Bivariate Analysis

Correlation Matrix (Top Features)

Strong correlations with the target:

Feature	Correlation w/ liquidity_score
---------	--------------------------------

liquidity_ratio	Strong +
-----------------	-----------------

24h_volume	Medium +
------------	-----------------

mkt_cap	Medium +
---------	----------

price	Low-medium
-------	------------

price changes (1h, 24h, 7d) show **weak but useful** trends.

5. Target Variable Analysis – liquidity_score

- Range: ~0.01 – 0.20
- Stable coins and high-volume assets score higher.
- Highly volatile small coins score lower.

6. Insights

- ✓ Liquidity strongly driven by **volume & liquidity_ratio**
- ✓ Market cap adds stability
- ✓ Price alone is not a strong predictor without volume
- ✓ price_change_24h is a useful volatility indicator
- ✓ Dataset needs **normalization** due to scale differences

7. EDA Conclusions

- The data supports **linear regression & ridge regression** models.
- Minimal feature engineering required: the original numeric columns are sufficient.
- Important features:
 1. liquidity_ratio
 2. 24h_volume
 3. mkt_cap
 4. price_change_24h