

Exploratory Data Analysis (EDA) Report

Goal: To predict the 24-hour trading volume (liquidity) of cryptocurrencies using market-based features like price, market cap, and price movements

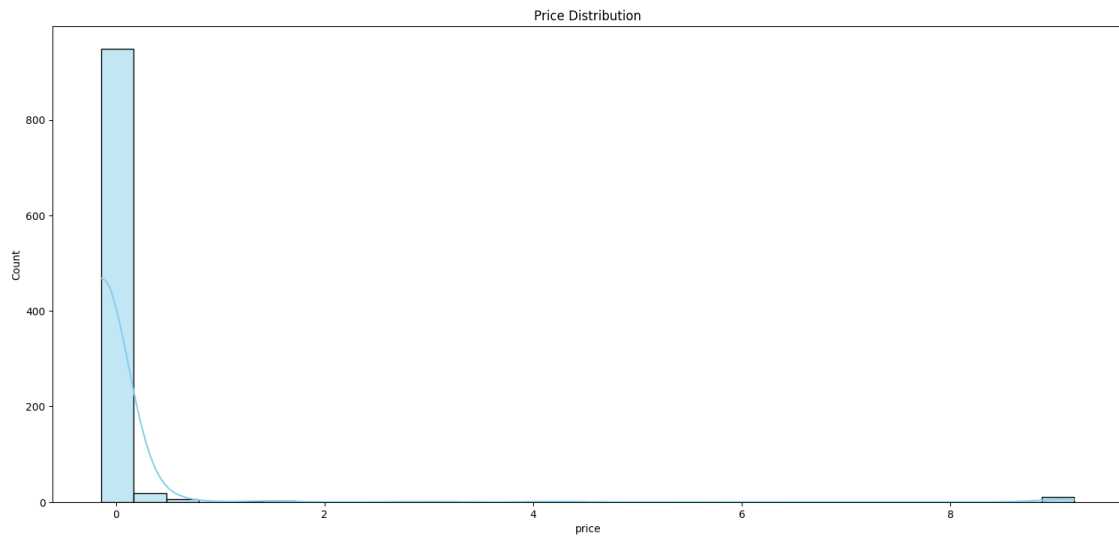
Summary of Dataset Statistics

The dataset contains cryptocurrency data (mainly from CoinGecko) across two days with 992 valid records after cleaning. It includes the following key numerical features:

Feature	Min	Max	Mean	Std. Dev
Price (scaled)	-0.14	9.19	~0	1.00
24h Volume (scaled)	-0.01	21.00	~0	1.00
Market Cap (scaled)	0.00	20.00	~0	1.00
Price Change Score	-9.79	12.05	~0	0.78
Volume-to-Market Cap Ratio	-94.28	63.79	0.88	4.19

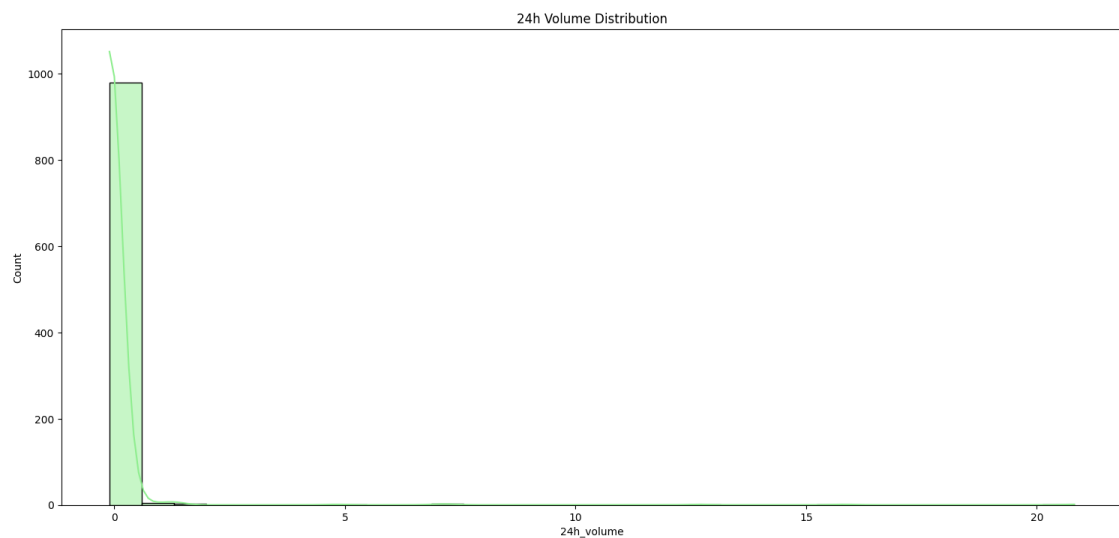
Basic Visualizations

1. Price Distribution



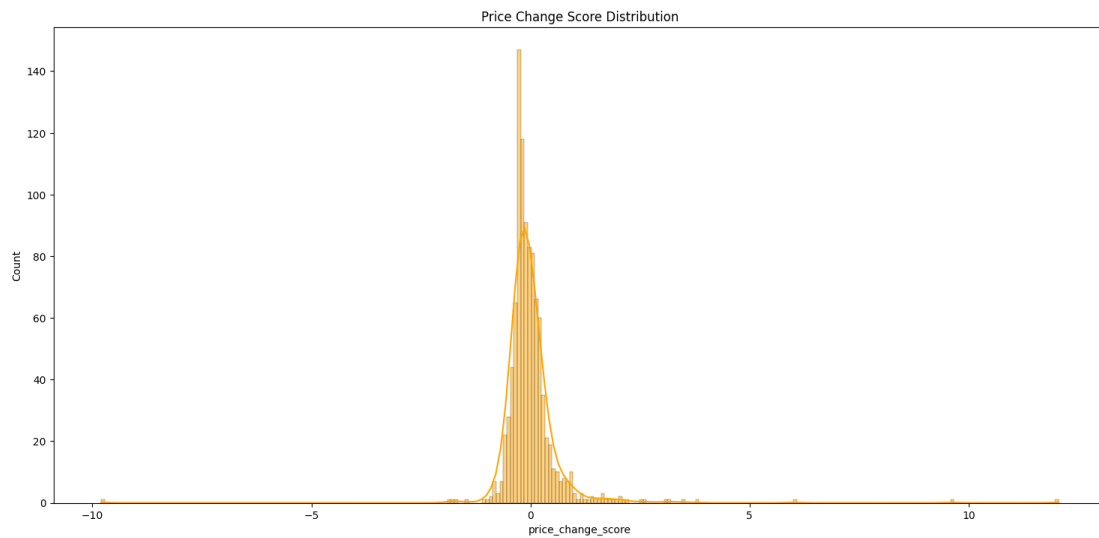
- Most prices are tightly clustered around 0 after scaling.
- A few cryptocurrencies (like Bitcoin) have much higher prices, creating right-skewed distribution.
- This skew suggests large price disparities between major coins and smaller altcoins.

2 .24h Volume Distribution



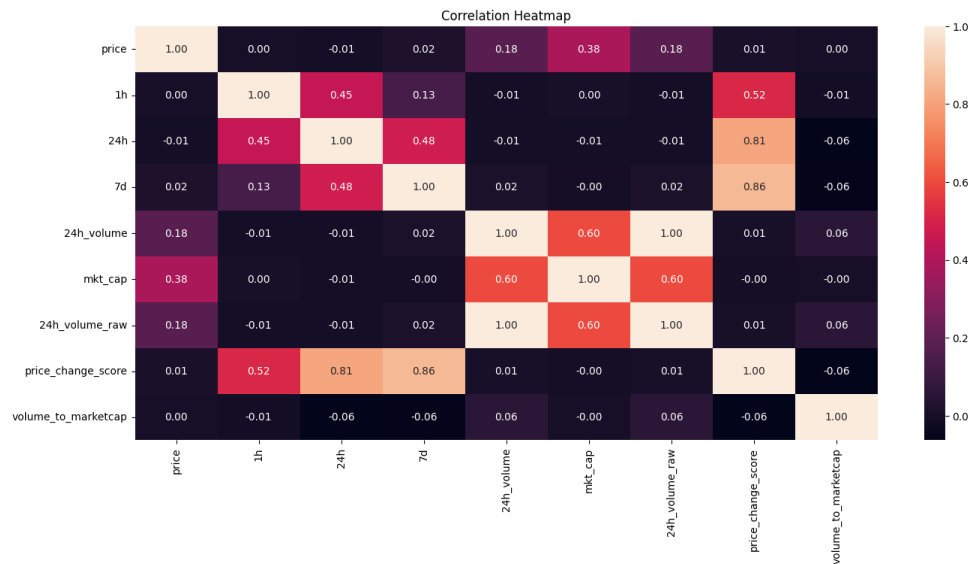
- Strong right-skew: most coins have low 24h trading volume, while a few have exceptionally high volume.
- Liquidity distribution is heavily imbalanced across coins.

3.Price Change Score Distribution



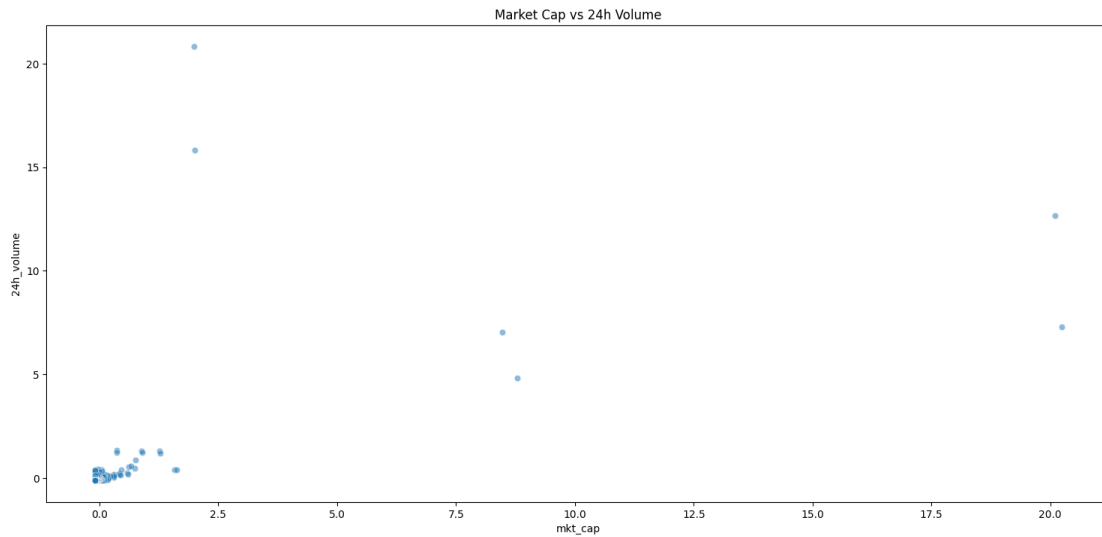
- Bell-shaped curve centered around zero — shows that most cryptocurrencies had stable short-term price movement.
- A few outliers with extreme volatility exist (both positive and negative).

4. Correlation Heatmap



- price_change_score is **highly correlated** with its components: 24h (0.81), 7d (0.86), 1h (0.52).
- 24h_volume is moderately correlated with mkt_cap (0.60), but only weakly with price.
- volume_to_marketcap has low correlation with other features, making it a unique predictor.
- No signs of perfect multicollinearity

5. Market Cap vs 24h Volume



- Shows a weak but visible upward trend — as market cap increases, 24h volume tends to rise too.
- Several high market cap coins show wide variability in liquidity.

Key takeaways

- Data is right-skewed in price and volume, meaning most coins are small-cap with low liquidity.
- Engineered features like price_change_score and volume_to_marketcap bring clarity and help the model.
- Correlation analysis supports model reliability — features contribute uniquely without redundancy.