📊 EDA Report: Cryptocurrency Liquidity Prediction for Market Stability

🔍 1. Introduction

This EDA explores historical cryptocurrency market data with a focus on understanding **liquidity patterns**, **price volatility**, and **volume trends**. The goal is to extract meaningful features that contribute to liquidity prediction for enhanced market stability.

📁 2. Dataset Overview

We use a historical dataset containing the following columns:

- `timestamp`: Datetime of the record
- `price`: Cryptocurrency price (e.g., BTC/USDT)
- `volume`: Trading volume
- `order_book_spread`: Bid-ask spread (used as liquidity proxy)
- `market_cap`: Total market capitalization
- `num_trades`: Number of trades executed

```
python
Copy code
import pandas as pd

df = pd.read_csv('data/crypto_data.csv')
df.info()
df.describe()
```

🧼 3. Data Cleaning & Preprocessing

- **Missing Values**: Checked and removed rows with `NaN`.
- **Data Types**: Converted `timestamp` to datetime.
- **Indexing**: Set `timestamp` as the index for time-series analysis.

```
python
Copy code
df['timestamp'] = pd.to_datetime(df['timestamp'])
df.set_index('timestamp', inplace=True)
df = df.dropna()
```

📈 4. Time Series Plots

📉 Price over Time

python
Copy code
```
import matplotlib.pyplot as plt

df['price'].plot(figsize=(12, 5), title='Price Over Time')
plt.ylabel("Price")
plt.show()
```
📊 Volume over Time

python
Copy code
```
df['volume'].plot(figsize=(12, 5), title='Volume Over Time',
color='orange')
plt.ylabel("Volume")
plt.show()
```
🔁 Order Book Spread (Liquidity Proxy)

python
Copy code
```
df['order_book_spread'].plot(figsize=(12, 5), title='Bid-Ask
Spread (Liquidity)')
plt.ylabel("Spread")
plt.show()
```
📉 5. Volatility Analysis

Calculated using rolling standard deviation of returns:

python
Copy code
```
df['returns'] = df['price'].pct_change()
df['volatility'] = df['returns'].rolling(window=5).std()
df['volatility'].plot(title='Rolling Volatility (5-period)')
plt.show()
```
🔍 6. Correlation Matrix

To identify strong relationships between features:

python

Copy code

```
import seaborn as sns

corr = df.corr()
plt.figure(figsize=(8,6))
sns.heatmap(corr, annot=True, cmap='coolwarm')
plt.title("Correlation Matrix")
plt.show()
```

💡 *Observation:* Strong correlation often exists between price, volume, and market cap. Spread shows an inverse correlation with volume in many cases, suggesting lower liquidity during low trading activity.

⚙️ 7. Feature Ideas from EDA

From the analysis, we propose using:

- **Rolling Mean/STD of Volume**: liquidity signal
- **Volatility**: market uncertainty indicator
- **Bid-Ask Spread**: direct liquidity proxy
- **Returns**: price movement patterns

✅ 8. Conclusion

This EDA confirms the presence of meaningful indicators for modeling cryptocurrency liquidity. The derived features will feed into the ML pipeline to predict high or low liquidity periods — a key component for maintaining market stability.