

Trader Behavior Analysis Project

1. Introduction

This project analyzes historical trader data with the goal of uncovering patterns in trader behavior, coin performance, profitability, and risk. The study proceeds through a detailed pipeline of data exploration, feature engineering, visualization, and statistical analysis.

2. Exploratory Data Analysis (EDA)

- Dataset loaded and inspected for structure, statistics, and missing values.
- Statistical analysis included mean, median, variance, and outlier detection (notably in 'Size Tokens' and 'Execution Price').
- Correlation heatmaps examined relationships between numerical variables.
- Categorical frequencies visualized for 'Account', 'Coin', 'Side', and 'Direction'.
- Dropped irrelevant columns for clarity.

3. Feature Engineering

- Converted 'Timestamp IST' to datetime for time-based analysis.
- Categorized 'Closed PnL' into: Profit, Loss, Break-Even.
- Created new features:
 - 'Total Volume' from 'Size Tokens' and 'Size USD'
 - 'Risk Indicator' as absolute value of 'Closed PnL'

4. Data Visualization

- Heatmaps showed trade counts by account and coin.
- Count plots explored combinations of 'Side' and 'Direction'.
- Scatter/line plots visualized trade 'Direction' over time.
- Box plots illustrated 'Start Position' and 'Closed PnL' for top coins.
- Scatter plots explored:
 - 'Start Position' vs. 'Closed PnL'

Trader Behavior Analysis Project

- 'Size Tokens' vs. 'Size USD'
- 'Total Volume' vs. 'Closed PnL'

5. Profitability and Risk Analysis

- Identified accounts with most losing trades.
- Total profit/loss analyzed per account and side.
- Listed top profitable and top losing accounts.
- Distribution of 'Closed PnL' visualized.

6. Summary of Findings

- Profitability: Wide 'Closed PnL' range; some consistently profitable or loss-making accounts.
- Risk: 'Risk Indicator' showed trades with high gains/losses; frequent losing accounts flagged.
- Volume: Combined token and USD size into 'Total Volume', which correlated with 'Closed PnL'.
- Outliers: Found in 'Closed PnL' and trade sizes; few large-impact trades observed.

7. Recommendations & Future Work

- Use PCA or clustering to discover deeper insights.
- Build predictive models to forecast profit/loss or risk.