Trader Behavior Insights Report

Final report on project

On

Trader Behavior Insight

Submitted by:

Shubhank Awasthi

Mail

Github

Objective

Explore the relationship between trader performance and market sentiment, uncover hidden patterns, and deliver insights for smarter trading strategies.

Data Overview

- ➔ Two datasets were taken up for exploration of the relationship between trader performance and market sentiment.
- ➔ The two datasets were-
  - Historical Trader Data (Link): The dataset contains 211,224 trades spanning from May 2023 - May 2025.
  - Sentiment Data (Link): The dataset contains 2644 rows of sentiment data from February 2018 - May 2025 with sentiment being classified as values ranging from Extreme Fear to Extreme Greed.
- ➔ **Key Columns**: Closed PnL (profit/loss), Size USD (trade size), Execution Price, value (sentiment score), classification (sentiment category: Extreme Fear, Fear, Neutral, Greed, Extreme Greed).
- ➔ **Preprocessing**: The datasets were joined on date after aligning the date values from both the tables and thereafter dealing with missing sentiment values by forward-filling, and filtering for same date overlap (May 2023–May 2025).

➔ **Feature Engineering:**

- log_size: The logarithm (adding small value for zeros) of size USD forms a right-skewed distribution, indicating that most trades are very small (histogram).
- pnl_per_usd: The profit/loss per USD traded value (added protection from division by zero) has a wide range for sentiment.
- abs_pnl and ClosedPnL_winsor: Both absolute and winsorized absolute PnL show a range of extreme values and were winsorized down to 1% tails to minimize the effect of outliers.

Methodology

➔ **EDA**: Analysed distributions, correlations, and aggregations by sentiment.

➔ **Visualizations**: Made use of histograms, scatterplots, countplots, barplots, heatmaps, and cross-correlation plots.

➔ **Statistical Tests**: Conducted Welch's t-test on Greed vs. Fear PnL, and chi-square test on profitability across sentiments.

➔ **Cross-Correlation**: Investigated the lagged impact of sentiment on daily total_pnl.

➔ **Tools**: Python (pandas, numpy, seaborn, matplotlib, scipy).

Key Findings

➔ **Profitability and Sentiment :**

- PnL per USD based on Sentiment: Extreme Greed- Mean 0.040 (4% Profit), high sd (1.929), Extremely Low (-384), therefore denotes High-Risk and High-Rewards trades. Fear- Mean 0.015 (1.5% Profit), SD is low (0.072) is stable and less profit.

  The median is 0.00 across every sentiment, pointing out that most of the trades in the dataset are break-even; however, profitability is derived from outliers.
- Scatterplot (Closed PnL vs Sentiment) (Fig-3)- No clear linear relationship, there is high variance here in terms of Closed PnL across value- Extreme Greed has some high gains (disregarding outliers), but it is rare (~$80,000).
- Profitable Trades vs. Not Profitable (Fig-6)- Looking at the barplot of diff (True - False)- We see we have more non-profitable trades in terms of Fear (-9,799) and Greed (-11,581), Extreme Greed shows slightly positive (+2,803) therefore, Sentiment likely impacted indicators' probability of profitability

➔ **Correlation and lag effects**:

- Daily sentiment is positively associated with profitability and volume, albeit weakly.
- According to cross-correlation analysis, lag effects could occur, meaning that changes in market sentiment could occur two to three days before trading PnL results.

➔ **Statistical Tests:**

- Welch's t-test (Greed vs. Fear): P-value 0.0868 (close to < 0.05) means there isn't enough evidence to suggest a significant mean difference in Closed PnL.
- Chi-square Test: χ^2 value 1.96e-176 (< 0.05) determines there is a significant difference in rates of profitable trades across sentiments and provides evidence that confirms the breakout in the heatmap (Fig-7) with higher non-profitable trade counts in Fear (35,818) versus Greed (30,945).

➔ **Strategic Consequences:**

- High-risk opportunities: Extreme Greed calls for risk management of extreme losses (-384 pnl_per_usd) as well as high mean PnL ($67.89) and profitability (46.49%).
- Times of Stability: Fear provides consistent trade sizes ($735.96 median) and is less volatile (std 0.072), suggesting a solid approach for conservative-based approaches.
- Timing Cognizance: The -24 day lag (Fig-8) suggests that better entry/exit timing can be realized by factoring in changes in sentiment directionality.
- Loss Mitigation: Tighter stop-losses or filtering sentiment-based trading will be required in Fear and Greed due to large numbers of non-profitable trades (-9,799, -11,581).

Conclusion

Extreme Greed offers high returns (mean PnL $67.89, 46.49% profitability) with risks (min pnl_per_usd -384), while Fear offers stability (mean pnl_per_usd 0.015) with more non-profitable trades (-9,799), according to an analysis of 211,224 trades made between May 2023 and May 2025. Timing and sentiment effects are highlighted by a significant profitability variation (p-value 1.96e-176) and a -24-day lagged correlation (-0.181). Returns can be maximised by employing tactics like stabilising in Fear, leveraging lagged sentiment, and utilising risk controls to target Extreme Greed.