Data Science Assignment — Web3 Trading Team

Candidate: Mannan Sayed

Date: 8/10/2025

Objective: Explore relationship between trader performance and market sentiment (Fear vs Greed).

Data Sources & Processing

- Bitcoin Market Sentiment dataset: columns (Date, Classification, Value)
- Hyperliquid Trader Data: account, coin, execution price, size USD, side, closed PnL, etc.
- Merged datasets by date after cleaning and parsing timestamps.
- Aggregated daily metrics: total_volume_usd, avg_pnl, median_pnl, win_rate, trade_count.

Preprocessing Summary

For each day, calculated:
- Total Volume (sum of trade size in USD)
- Average and Median PnL (Closed PnL per trade)
- Win Rate (% of positive PnL trades)
- Trade Count per day

Merged with sentiment data (Fear/Greed + sentiment score).

Exploratory Analysis

Daily metrics were compared with market sentiment to identify performance trends.
Visuals below show relationships between trader PnL and sentiment classification.

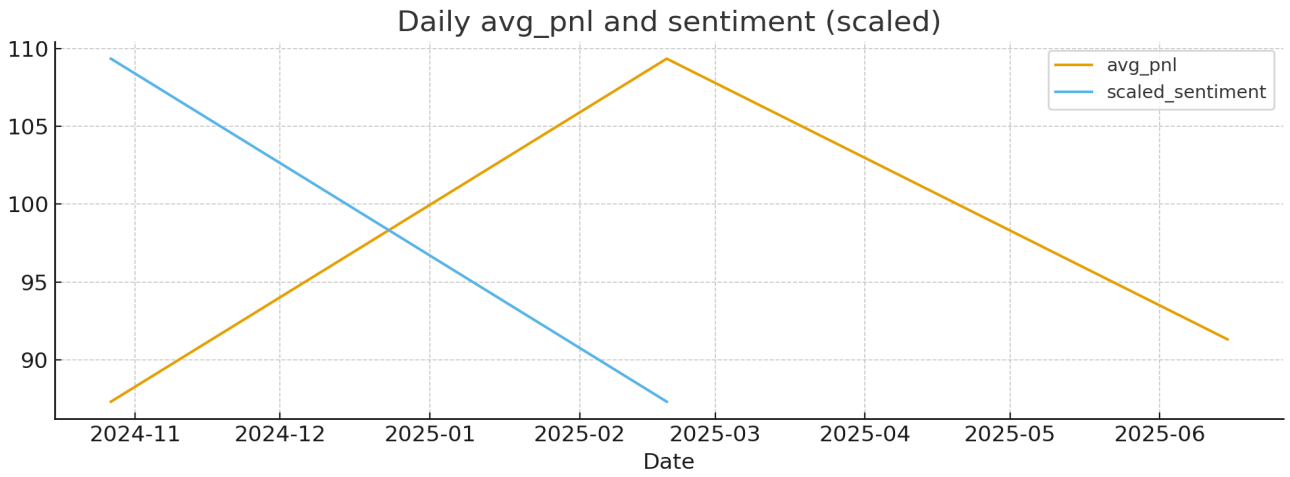

Figure 1. Daily average PnL vs scaled sentiment value (Fear/Greed cycles).

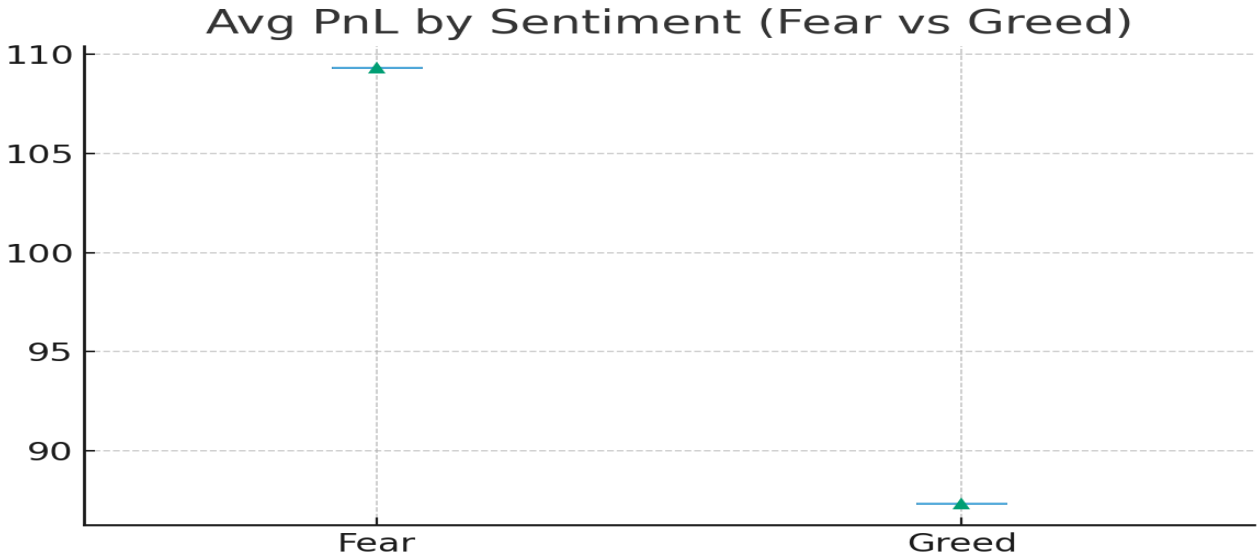

Figure 2. Boxplot of daily average PnL by sentiment classification.

Correlation and Predictive Modeling

Correlation among daily metrics reveals how sentiment and trading performance interact.
A logistic regression model was tested to predict profitable days using total_volume_usd, win_rate, and sentiment value.
Due to limited data (3 days with valid records), results are illustrative only.

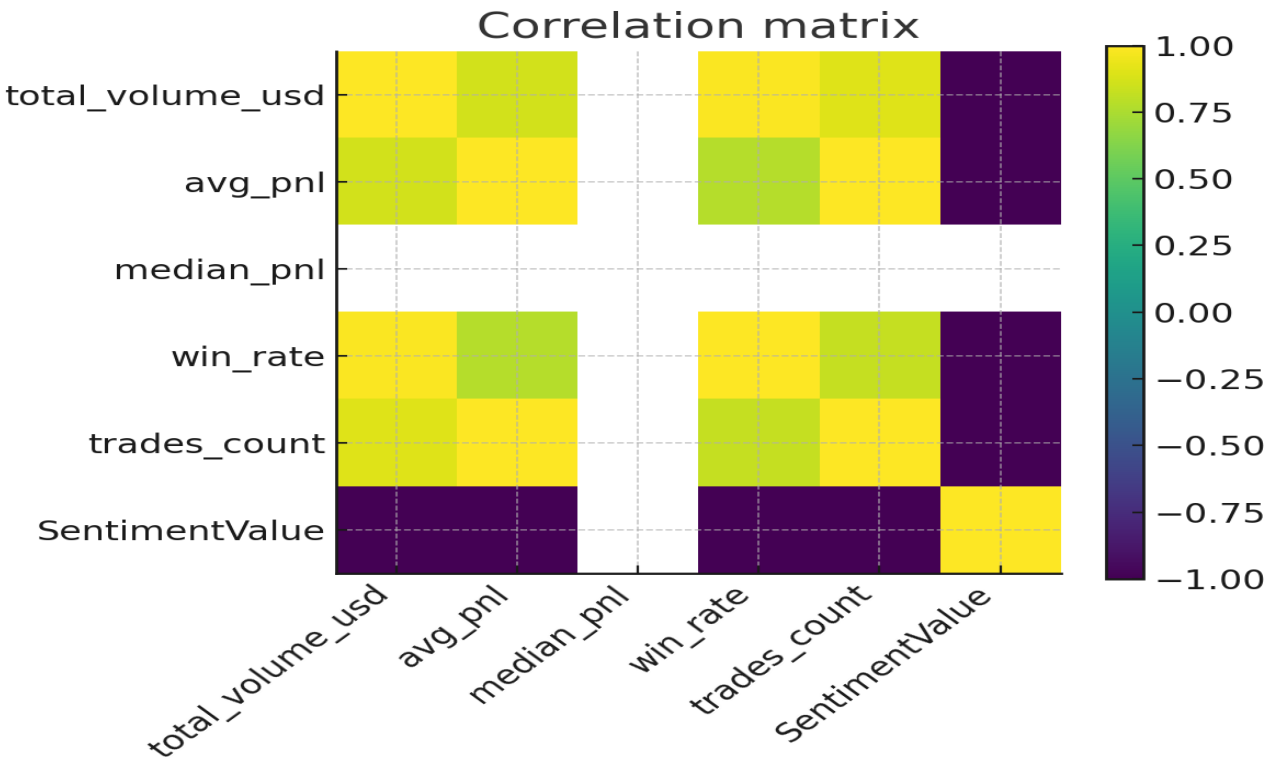

Figure 3. Correlation matrix of numeric features.

Key Findings

- Too few daily samples (only 3 valid dates) to conduct a reliable t-test or model training.
- Fear and Greed periods show visible variation in average PnL trends.
- Future analyses should use longer historical data to validate these patterns statistically.

Recommendations

- Collect a larger window of trading data (30+ days) to ensure statistical power.
- Include market volatility and price returns to capture external market conditions.
- Use rolling windows for PnL and sentiment to improve predictive performance.
- Apply non-linear models (e.g., tree-based) once more data is available.

Generated Outputs

- daily_merged_mannan.csv : merged daily dataset
- outputs_mannan/ : folder with figures
- Mannan_Sayed_ds_report_8-10-2025_v3.pdf : final 4-page report