

Primetrade.ai

Title: Data Science Assessment — Shaikh Amaan

Date: 19 Wednesday 2025

1. Executive Summary

This assessment explores the relationship between trading behaviour and the Bitcoin Fear & Greed Index. After cleaning and preparing the data, daily aggregates were generated, merged with sentiment indicators, visualised, and statistically evaluated.

2. Data Description

- `historical_data.csv` — trade-level history with timestamps, positions, volume, PnL, and metadata.
 - `fear_greed_index.csv` — daily market sentiment classification (Fear / Greed) and index score.
-

3. Data Cleaning & Preprocessing

- Standardised column names (snake_case).
 - Fixed timestamp inconsistencies using the Timestamp IST field (%d-%m-%Y %H:%M).
 - Converted numeric fields: `execution_price`, `size_tokens`, `size_usd`, `closed_pnl`, `fee`.
 - Removed duplicate rows and filtered invalid entries.
 - Extracted date for daily-level aggregation.
-

4. Feature Engineering & Aggregations

Daily aggregates include:

- `total_trades`
- `total_volume_tokens`
- `total_volume_usd`
- `net_pnl`
- `avg_pnl`
- `pct_profitable`
- `buy_count / sell_count`
- `buy_sell_ratio`

These were merged with Fear/Greed classifications using a left join.

5. Exploratory Data Analysis

- Time-series analysis of net PnL over time.
 - Scatter plot with color-coded market sentiment (Fear = red, Greed = green, Missing = gray).
 - Patterns observed between profitability and sentiment-driven behaviour.
 - Visual outputs saved in outputs/.
-

6. Statistical Testing

- Conducted t-test and Mann–Whitney U test comparing net PnL on Fear vs Greed days.
- Insert your p-values here:

T-test p-value: 0.021709969966269897

Mann-Whitney p-value: 0.009470524637068167

Interpretation (example):

Low p-values indicate meaningful distribution differences in net PnL between Fear and Greed environments.

7. Limitations

- Missing sentiment for some dates.
 - Small dataset limits reliability of statistical conclusions.
 - Aggregations remove intra-day volatility information.
-

8. Recommendations

- Consider leveraging Fear days to adjust risk exposure.
 - Explore sentiment-aware position sizing.
 - Extend analysis to account-level behavioural clustering.
 - Expand dataset for stronger statistical power.
-

9. Acknowledgment of Assistance

Parts of the code structure, explanation formatting, and presentation improvements were created with guidance from AI-based tools (ChatGPT). All analysis, interpretation, debugging, and final decisions were performed manually by me.