Trader Behavior Insights Report

Summary of Findings

Summary of Findings

1. Dataset Coverage:

- Aggregated trader data covers the period from **March 28, 2023 to June 15, 2025**.
- Market sentiment data spans from **February 1, 2018 to May 2, 2025**.
- The merged analysis uses overlapping periods for robust comparison.

2. Profitability Over Time by Market Sentiment:
The line plot of average profitability (`avg_pnl`) over time, segmented by market sentiment categories (`Fear`, `Greed`, `Extreme Greed`, and `Neutral`), indicates:

- Noticeable variability in profitability during different sentiment phases.
- Some sentiments, such as **Greed**, show sustained higher average PnL during certain periods, suggesting trader optimism aligns with better profitability.
- However, the pattern is not uniform and warrants deeper exploration.

3. Trading Volume by Market Sentiment:
The boxplot visualization shows total trading volume distributions across different sentiment categories:

- Volume varies widely under all sentiments but is noticeably higher during **Fear** and **Greed** periods compared to **Neutral** and **Extreme Greed**.
- This suggests market activity intensifies during more emotional market states, possibly reflecting increased trader engagement or volatility.

4. Regression Model Analysis:
A linear regression model was developed to predict average profitability (`avg_pnl`) using:

- Market sentiment encoded as numeric (`sentiment_num`)
- Total trade volume (`total_volume`)

Evaluation Metrics:

- Mean Squared Error (MSE): 4299.71
- R^2 Score: -776,697.12 (negative, indicating poor model fit)

Model Coefficients:

Feature	Coefficient
sentiment_num	192.28
total_volume	0.0000002529839

Interpretation:

- Despite positive coefficients, the model's very low and negative R^2 indicates that a simple linear model with these features does not effectively explain variability in profitability.
- This suggests non-linear relationships, missing important predictors (e.g., leverage, risk metrics), or noisy data.
- Further modeling with additional features or more advanced techniques may improve predictive power.

Conclusions & Recommendations

- **Market sentiment and trading volume show some intuitive patterns in relation to profitability**, but complexity and noise limit straightforward modeling.
- The negative regression R^2 flags that basic linear approaches are insufficient alone—exploration of non-linear models, feature engineering, or additional data (like leverage, risk measures) is needed.
- **Trading volume spikes during emotional market states (Fear, Greed), highlighting heightened trader activity** which can be a useful signal for dynamic strategy adjustments.
- Reported findings should be contextualized with dataset limitations and encourage iterative refinement.

This analysis was performed using Python (pandas, seaborn, scikit-learn) in Google Colab. Plots are saved in the `outputs/` *directory.*