

Data Science Assignment Report: Trader Behavior vs. Market Sentiment

Sujal Gupta

Date: October 24, 2025

1. Introduction and Project Objectives

1.1. Context and Problem Statement

The cryptocurrency market, particularly Bitcoin, is notoriously volatile and driven significantly by emotional factors. The core challenge for profitable trading is distinguishing between emotional herd behavior and rational, high-conviction market participation. This report investigates the relationship between collective trading behavior on a major decentralized exchange (Hyperliquid) and the overall market mood, as quantified by the Fear & Greed Index (FGI).

The primary hypothesis is that successful trading activity (high PnL) is inversely related to market sentiment—meaning the most profit is captured when the majority of market participants are in a state of Fear or Extreme Fear.

1.2. Assignment Objectives

This analysis was structured to meet the following key objectives:

1. Data Integration: Synchronize high-frequency trade data with daily sentiment data.
2. Behavioral Metric Derivation: Aggregate raw trade data into daily, measurable indicators of Profitability, Volume, and Activity.
3. Divergence Analysis: Quantify how trading performance and volume align with or diverge from the five FGI classifications.
4. Signal Identification: Identify concrete, data-driven trends and signals that can be implemented into a smarter, contrarian Web3 trading strategy.

2. Data Sources and Methodology

2.1. Data Sources

Two distinct datasets were used for this analysis:

Dataset	Granularity	Key Columns Used	Purpose
Bitcoin Market Sentiment	Daily	date, classification, value	Provides the emotional state of the market (Extreme Fear to Extreme Greed).
Historical Trader Data	Per-Trade	Timestamp IST, Closed PnL, Size USD, Trade ID	Provides raw, high-frequency data on actual executed trades, profits, and size.

2.2. Data Preprocessing and Feature Engineering

The primary challenge in this analysis was merging the high-frequency trade data with the low-frequency daily sentiment data. The following steps were executed:

1. Time Synchronization: The Timestamp IST column from the historical data was converted to a uniform date format. The timestamp column in the FGI data was converted to its corresponding date.
2. Trader Data Aggregation (Feature Engineering): The historical_data.csv was grouped by the synchronized Date to calculate three critical daily metrics:
 - Total Closed PnL: Sum of all daily profits/losses (Closed PnL).
 - Total Trading Volume: Sum of daily traded size (Size USD).
 - Total Trade Activity: Count of all daily trades (Trade ID).
3. Final Merged Dataset: The daily aggregated trader metrics were joined with the FGI classification on the common date key, resulting in a single time series ready for comparative analysis.

2.3. Analytical Approach

The analysis focused on comparing the central tendency and dispersion of the three behavioral metrics against the five ordinal FGI classifications.

- Profitability Analysis: Box plots were used to visualize the distribution of Total Daily PnL across each sentiment category, revealing both the average performance and the volatility (risk/reward profile) within that market state.

3. Exploratory Data Analysis (EDA) & Visual Findings

3.1. Overview of Sentiment and Collective Performance

The summary statistics derived from the merged dataset immediately suggest a strong contrarian signal in the market, where high profits concentrate in phases of low sentiment.

FGI Classification	Mean Daily PnL (USD)	Median Daily PnL (USD)	Mean Daily Volume (USD)	Mean Daily Trades
Extreme Fear	\$52,793.59	\$22,561.74	\$8,177,447	1,528
Fear	\$36,891.82	\$1,412.31	\$5,311,261	679
Neutral	\$19,297.32	\$1,818.57	\$2,690,180	562
Greed	\$11,140.57	\$678.48	\$1,495,246	260
Extreme Greed	\$23,817.29	\$3,127.54	\$1,091,799	350

3.2. Profitability Distribution by Sentiment

The box plot visualization (refer to outputs/01_pnl_vs_sentiment.png) graphically confirmed the core finding:

- **Extreme Profit Skew:** Days classified as Extreme Fear consistently yielded the highest mean PnL, with the highest positive outliers, indicating highly successful, large-scale trades occurred during market panic.
- **Greed's Poor Performance:** Conversely, days categorized as Greed exhibited the lowest average profitability, suggesting that the "chasing" behavior common during bullish excitement rarely resulted in substantial collective gains. The average PnL during Extreme Fear was nearly 5x higher than during Greed.

3.3. Volume and Activity Analysis

Market activity, measured by both Volume and the number of Trades, closely follows the profitability pattern, creating a powerful behavioral footprint.

- **Fear-Driven Liquidity:** The Total Daily Trading Volume peaked dramatically during Extreme Fear (Mean Volume: \$8.1M).
- **Apathy in Greed:** Periods of Greed showed the lowest volume and trade count (Mean Volume: \$1.5M; Trades: 260). This behavioral drought suggests a lack of new liquidity entering the market to sustain the rally, making the existing price movement highly vulnerable to reversal.

4. In-Depth Behavioral Analysis and Strategic Signals

4.1. The Contrarian Principle

The analysis validates the famous trading axiom: "Be fearful when others are greedy, and greedy when others are fearful." The profitability data shows that the aggregate trader base on Hyperliquid, consisting of sophisticated participants, systematically makes its largest profits when the retail or less-informed segments of the market are most emotional.

- Fear Zone (≤ 50): High PnL is achieved by capitalizing on the liquidity provided by sellers in distress. High volume here is the signature of true conviction.
- Greed Zone (> 50): Low PnL is often a consequence of trades that enter late into a rally, leading to smaller average gains and higher chances of liquidation during local tops.

4.2. Identifying the Hidden Signal

To move beyond simple correlation, a predictive signal must incorporate both the emotional state (FGI) and the magnitude of collective commitment (Volume/Activity). The most critical hidden trend identified is the synergistic effect of Extreme Sentiment paired with Volume deviation.

The signal dictates filtering high-volume events based on extreme sentiment to pinpoint optimal entry and exit timing.

Trading Action	FGI Sentiment Filter	Behavior Filter (Hidden Signal)	Strategic Rationale
Optimal LONG Entry	Extreme Fear / Fear	Daily Volume ≥ 90 th Percentile of Historical Volume	A market classified as fearful <i>and</i> experiencing exceptionally high transactional flow indicates a major capitulation event is being bought up by large, highly profitable entities. This marks a low-risk, high-reward accumulation zone.
Optimal SHORT/Hedge	Greed / Extreme Greed	Daily Trades \leq Median Trade Count	A market classified as greedy <i>but</i> suffering from low transactional activity suggests the current price movement is running on fumes. Low activity signals a lack of new buyers, making the market susceptible to a sharp corrective move.

5. Conclusion and Future Work

5.1. Summary of Findings (Page 5/5)

The comprehensive analysis of Hyperliquid trade behavior and Bitcoin sentiment unequivocally demonstrates that profitable trading is a contrarian activity. The most significant collective profits and market-moving volume are generated when the Fear & Greed Index signals distress. Conversely, chasing momentum during periods of Greed proves to be the least lucrative and highest-risk strategy for the collective trader base. The highest concentration of "smart money" activity is found in the Extreme Fear category.

5.2. Strategic Recommendation

The most effective strategy is to leverage the FGI as an entry filter for high-conviction trades, specifically initiating long positions when the FGI is in the Fear zone and validated by a spike in volume.

5.3. Next Steps and Future Enhancements

To evolve this model into a production-ready Web3 trading tool, the following steps are recommended:

1. Account Segmentation: Isolate and track the PnL and activity of the top 5% of consistently profitable Account addresses. Analyzing their specific volume and side (BUY/SELL) during FGI extremes would provide a highly granular "smart money indicator."
2. Perpetual Contract Symbol Analysis: Extend the analysis to specific perpetual contract symbols (Coin) to determine if the FGI signal is uniform across all crypto assets or if it provides a stronger signal for major tokens (e.g., BTC/ETH) compared to smaller altcoins.
3. Liquidation Analysis: Incorporate liquidation data (if available) to directly measure the cost of emotional trading and correlate it with FGI scores, adding a direct metric for "pain" in the market.

This analysis provides a foundational, data-backed signal for enhancing trading efficacy by systematically exploiting the market's inherent emotional biases.