Impact of Market Sentiment on Trading Behaviour

Candidate: Devendra Prasad
Date: 4^{th} January, 2026

<u>Abstract</u>:

The aim of the analysis is to examine how market sentiment (Fear vs Greed) influences **trading behaviour, PnL characteristics, and performance metrics** using real historical trade data. In this study we have tried to distinguish between **behavioural changes** and **profitability changes** across sentiment regimes. We have used the following datasets for analysis as provided in the assignment.

Daily **Fear & Greed Index** used to realize the sentiment throughout the timeline.

Historical trade dataset which was the main source for the statistical.

The period of time in the Fear Greed Index consisted of **2,647 Days** but the analysis was conducted over **731 trading days**, based on the h**istorical trade data** via date-based joining.

Methodology

- Statistical data analysis and data visualization of the given datasets.
- trade data resampled into daily features including trade count, participation, volume, PnL statistics, volatility, and win rate.
- Sentiment regimes grouped into **Fear (Fear + Extreme Fear)** and **Greed (Greed + Extreme Greed)**.
- Non-parametric statistical tests (Mann–Whitney U) applied to compare behavioural and risk metrics across regimes.
- Visual analysis complemented statistical results to validate regime-dependent patterns.

Key Insights

The dataset analysis showed that the dataset sentiment regimes were not scattered and usually lasted for a period of month before other one starts. The time periods for fear usually lasted longer than greed. The analysis of trades dataset showed that the mean of the PnL was ~$49 that signifies that the trades made very little or no profits with rare significant outcomes. The analysis of the datasets combined showed that fear regimes show **significantly higher trading activity**, with ~205% more trades, more participating accounts, and higher traded volume compared to greed regimes. PnL volatility during fear regimes is approximately **2× higher**, indicating elevated risk and outcome dispersion.

It was also observed that statistically no significant differences were observed in **mean PnL, median PnL, or win rate** across sentiment regimes.

These findings indicate that market sentiment primarily influences behavioural intensity and risk rather than predictive accuracy. The results support the use of sentiment as a regime-level risk and behaviour indicator rather than as a standalone predictive signal.

Introduction:

Financial markets are driven by information and fundamentals as well as by collective human behaviour. Periods of greed and fear are repeatedly observed during market cycles and are widely believed to influence trading decisions, capital allocation, and risk-taking behaviour. As a result, sentiment indicators such as the Fear and Greed Index are frequently used to justify timing decisions, and anticipate market turning points.

Sentiment indicators cannot be used solely as predictors of price direction or as general measures of market psychology. A change in trading behaviour such as increased activity, higher participation, or greater volatility does not necessarily imply improved predictability or profitability.

The objective of this study is to separate **behavioural effects** from **performance outcomes** by examining how sentiment regimes influence trading activity while evaluating their impact on profitability metrics. To achieve this, historical trade-level data are aggregated into daily behavioural metrics and aligned with a Fear & Greed sentiment index. Trading days are classified into fear and greed regimes, and non-parametric statistical methods are used to compare distributions of key metrics across regimes.

Data Description:

The initial stages of the study consist of analysis of given datasets. In further steps I have derived features from the given datasets that would help in better understanding the data. Next step integrates market sentiment data with historical trade-level records to examine the relationship between sentiment regimes and trading behaviour. Care was taken to ensure temporal alignment, data quality, and appropriate aggregation to support statistically valid comparisons.

Data Sources:

Two primary datasets were used:

Fear & Greed Index Dataset
This dataset provides a daily categorical data of overall market sentiment, classified into five states: *Extreme Fear, Fear, Neutral, Greed,* and *Extreme Greed*. The duration of the data is of 2,647 days spanning from 01/02/2018 to 02/05/2025. The dataset signifies continuous fear and greed regimes with slightly sudden transitions.

Historical Trades Dataset
The historical trades dataset contains transaction-level records marked by trade IDs and including trade timestamps (IST), trade side (buy/sell), trade size denominated in USD, account identifiers, and realized (closed) profit and loss values. Each record represents an executed trade and reflects actual trading activity rather than simulated outcomes. The duration of the data is of 732 days spanning from 01/05/2023 to 01/05/2025. Because trading activity is not present for all sentiment dates, the analysis window is limited to the trade dataset.

This asymmetry in coverage motivated the use of a **left join** during dataset integration, preserving all trading days while avoiding the inclusion of sentiment-only observations that would otherwise dilute behavioural analysis.

Temporal Processing and Data Validation

Trade timestamps were converted from string format into standardized datetime objects and resampled to a daily granularity. Each trade was assigned a trading date, which served as the primary key for aggregation and sentiment alignment.

Intraday timestamps were not used directly in the statistical analysis; instead, daily aggregation was chosen to:

- Reduce microstructural noise,
- Align with the daily resolution of sentiment data,
- Enable regime-based comparison at a consistent temporal scale.

Several data quality checks were performed prior to analysis:

- Summary-level date quality checks were conducted to verify continuity, coverage gaps, and temporal consistency.
- Invalid or unparsable timestamps were excluded.
- Closed PnL values were used only where explicitly available.
- Days with no closed trades were retained for behavioural metrics but excluded from performance-specific calculations.

These steps ensured that observed differences across regimes were not artifacts of missing data or preprocessing errors.

Feature Engineering and Dataset Construction

From the raw sources, multiple structured datasets were derived to support different stages of analysis through a multi-stage feature engineering pipeline. The objective of this process was to convert high-frequency, event-based records into interpretable daily behavioural and risk metrics that could be meaningfully aligned with sentiment regimes.

1. Account-Level Trading Frequency Dataset (accounts_trade_frequency.csv)

This dataset aggregates behaviour at the **account level** rather than the market level. This dataset helps to identify the number of dominant or high-frequency participants and the typical trading behaviour of each account.

Key features

- **Total trades per account**: Total trades made by given account
- **Trading days per account**: Total number of days the account made trades
- **Average trades per day**: Average daily trades of given account
- **Average daily volume**: The average trade volume of the account
- **Buy/sell volume ratios**: The ratio of buy a sell trades(to identify buying or selling in bulks)
- **Maximum daily volume concentration**

2. Daily Market Activity Dataset (daily_market_activity.csv)

This dataset represents the **first level of aggregation**, where intraday trade logs were resampled to daily granularity. For each day, the following metrics were derived:

- **Total Trades**: Count of executed trades per day
- **Unique Accounts**: Number of unique trading accounts active on that day

- **Total Volume (USD)**: Sum of traded notional value
- **Mean Closed PnL**: Average realized profit/loss of closed trades on that day
- **Median Closed PnL**: Median realized profit/loss of closed trades on that day
- **PnL Volatility**: Standard deviation of daily closed PnL
- **Win Rate**: Proportion of closed trades that resulted in profit

3. Trade Outcomes Distribution Dataset (trade_ outcomes_distribution.csv)

This dataset retains **trade-level granularity**, but only for relevant fields required for analysis. The following features were retained in this dataset: [‘Trade date’ , ’Closed PnL’ , ‘Side’ , ‘Trade size’]
This dataset was used to:

- Analyse PnL distributions and tail behaviour
- Compare gains vs losses
- Validate non-Gaussian characteristics of returns

4. Daily Trader Behaviour Dataset (daily_trader_behavior.csv)

This dataset performs daily aggregation by explicitly emphasizing **behavioural quality and risk**. A binary **activity flag** was introduced based on minimum thresholds for Trade count and Number of active accounts. This separates structurally meaningful trading days from less significant days. Beyond basic activity metrics, the following were incorporated:

- **Absolute Volume Normalization**: Ensures buy and sell volumes are comparable
- **Win Rate Adjustment**: Excludes zero-PnL trades from win-rate computation
- **PnL Volatility**: Used as a direct proxy for daily risk intensity

This dataset is the primarily the representation of **daily trader behaviour**, capturing intensity, participation, efficiency, and risk in a single table.

5. Daily Behaviour with Sentiment Dataset (daily_behavior_with_sentiment.csv)

This dataset represents the **final analytical table**, produced by merging daily trader behaviour with sentiment data. This dataset enables sentiment based behavioural comparison, statistical testing and visualization of sentiment-aligned activity patterns. It forms the basis for all behavioural analysis in the study.

Additional feature: Daily sentiment classification mapped from the Fear & Greed Index

6. Fear vs Greed Comparison Dataset (fear_vs_greed_comparison.csv)

This dataset captures **statistical summaries** of the sentiment differences. It compiles a report with the conclusions drawn from the data. This report can be checked and verified by others to clearly show the differences that depend on the specific sentiment being examined.

For each behavioural and risk metric we derived:

- Median values under Fear and Greed regimes
- Percentage change
- Mann–Whitney U test p-values
- Statistical significance indicators

The feature engineering process transforms raw transaction logs into a hierarchy of datasets that progressively capture **activity, behaviour, risk, and sentiment context**. This structured approach enables robust regime-based analysis while maintaining transparency between raw data and final conclusions.

EDA and Statistical Summary:

Exploratory Data Analysis (EDA):

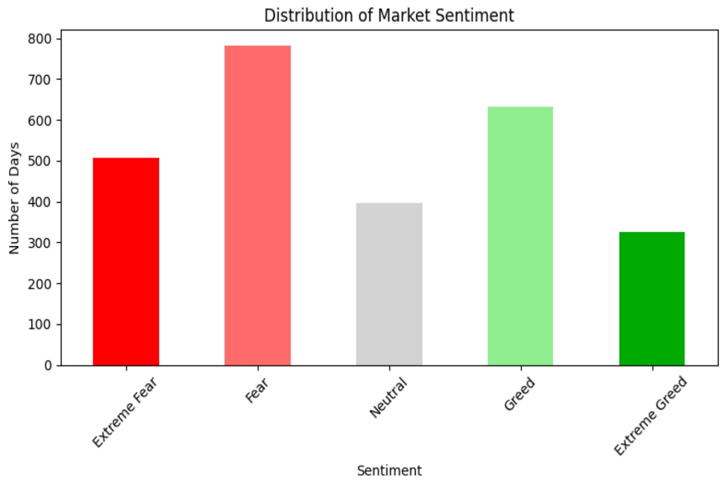
Exploratory Data Analysis was conducted to understand the structural properties of the data, identify regime-dependent patterns, and guide the choice of appropriate statistical methods.

i. Sentiment Distribution and Temporal Structure

Date Validation Checks:

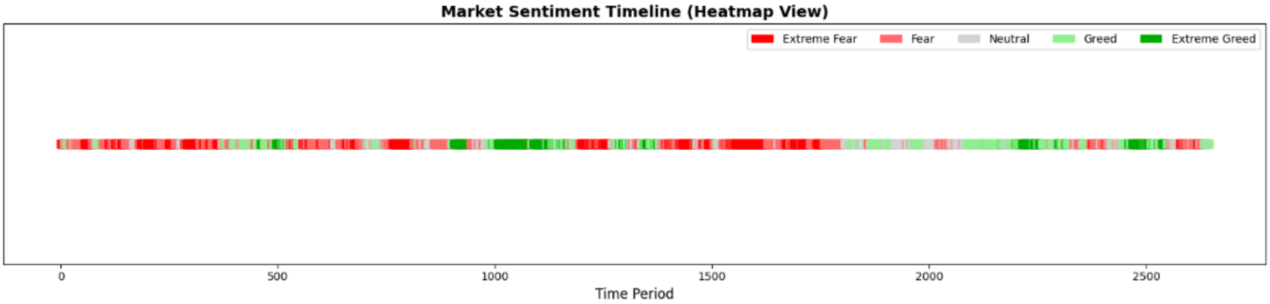
Dataset	Fear Greed Index
Earliest Date	2018-02-01
Latest Date	2025-05-02
Date Span Days	2647
Unique Days	2644
Total Records	2644
Duplicate Dates	0
Missing Dates	0

The Fear & Greed Index was first examined to understand the frequency and persistence of sentiment states over time. The distribution revealed an asymmetric structure. The market spends more time scared than optimistic while extreme greed is least frequent. This shows that the market is mostly in the state of fear and anxious state.



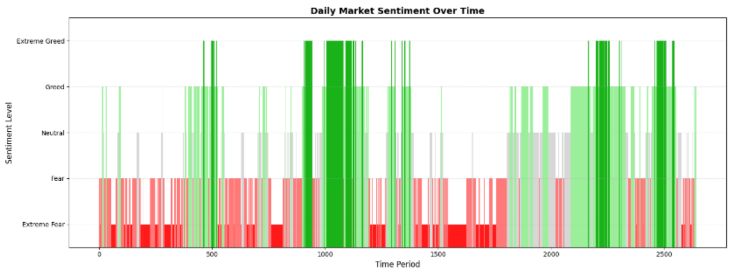
Classification	Count
Extreme Fear	508
Fear	781
Neutral	396
Greed	326
Extreme Greed	633

We also plotted the heatmap to understand the flow and regions of the given classifications. Sentiment clusters reveal distinct regions of fear and greed with regime transitions difficult to identify.



Timeline visualizations showed that sentiment transitions are abrupt rather than gradual, indicating regime-like behaviour instead of smooth sentiment drift. Fear states are more fragmented but frequent and Extreme Greed comes in **compact, dense clusters**.

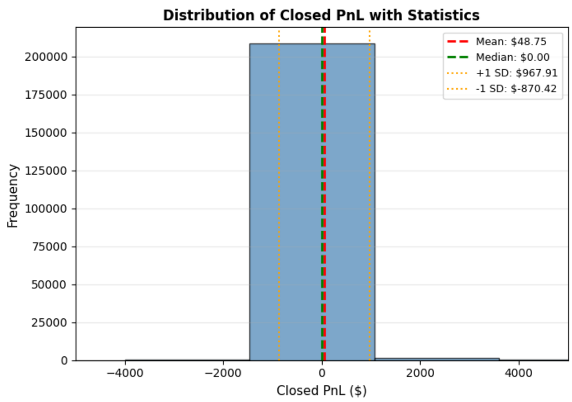
The monthly sentiment distribution shows that fear and greed states are not random and usually persist over consecutive weeks and sometimes months, reinforcing the interpretation of sentiment as a **market regime variable** rather than a short-lived signal.

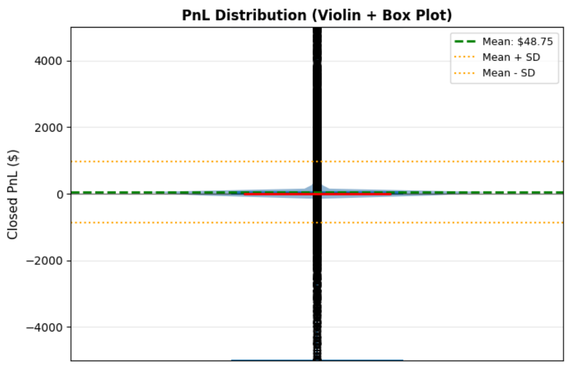


ii. PnL Distribution Analysis

Mean	$48.75
Median	$00.0
Std. Deviation	$919.16
Minimum	- $ 117,990
Maximum	$135,329

PnL distributions were found to be sharply peaked around zero with long tails on both the positive and negative sides. Boxplots and violin plots confirmed non-Gaussian characteristics, with outcome dispersion mainly affected by rare but extreme events.



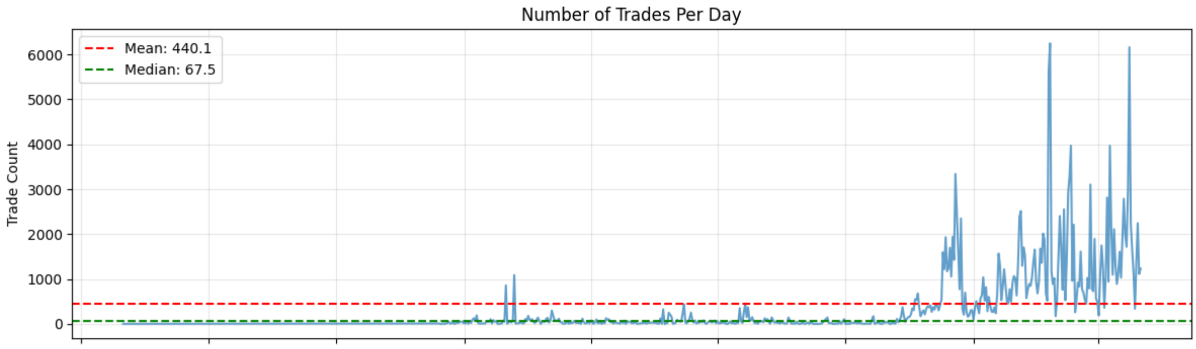


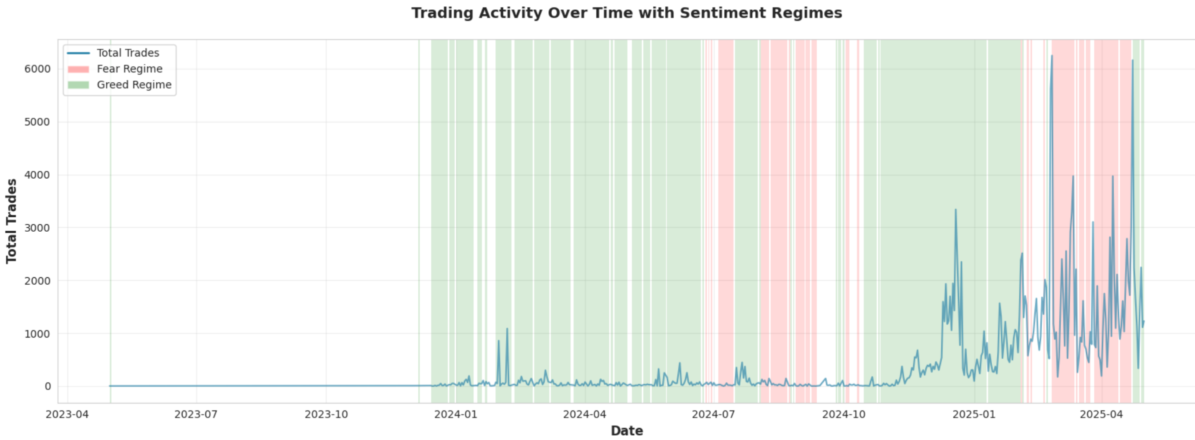
Gains (86869) slightly outnumber losses (17539), suggesting positive trade frequency bias. Despite this, the asymmetric loss tail suggests insignificant or small positive payoffs, making the strategy highly sensitive to avoid high loss magnitude.

iii. Trading Activity and Participation Patterns

Daily trading activity shows 480 active days with mean 440.1 trades/day, max 6246 trades/day. The extreme divergence between mean and median of trades confirms sentiment-dependent activity, where risk and PnL are concentrated in high-intensity trading periods.

Daily trading activity was analysed through total trade counts and unique account participation.





Overlaying sentiment regimes on trading activity revealed that high-activity days cluster disproportionately within fear regimes. Greed regimes, by contrast, showed lower and more stable levels of participation. This preliminary observation motivated a deeper investigation into whether sentiment systematically conditions trading intensity.

Key EDA Takeaways:

- The market sentiment is dominated by **fear**.
- Trading activity, participation and volume are **heavily clustered in time**.
- Fear regimes visually align with spikes in activity and dispersion.
- Profitability metrics show no obvious regime-dependence and cluster around zero.
- Data distributions violate normality assumptions, justifying non-parametric testing.

Statistical Methodology:

The statistical methodology was designed to statistically compare trading behaviour and risk characteristics in both sentiment regimes while accounting for the non-normally distributed trade data.

Trading days were classified into two primary sentiment regimes:

- **Fear Regime:** Fear + Extreme Fear
- **Greed Regime:** Greed + Extreme Greed

Neutral sentiment days were excluded from regime comparison to focus on the biased market sentiment states where behavioural effects are expected to be strongest.

Choice of Statistical Tests
Given the skewed distributions, presence of outliers, and unequal sample sizes across regimes, **non-parametric tests** were selected.

- **Mann–Whitney U Test**
  Used to compare median values of behavioural and risk metrics between fear and greed regimes assuming non-normality.
- **Kolmogorov–Smirnov (KS) Test**

  Used to assess whether entire distributions differ across regimes, not just central tendency.

These tests are applicable to heavy tails and are appropriate for observational financial data.

Metrics Evaluated
The following daily metrics were tested across regimes:

- Total trades
- Unique accounts
- Total traded volume (USD)
- Mean closed PnL
- Median closed PnL
- PnL volatility
- Win rate

Each metric was evaluated independently to avoid conflating activity, risk, and performance effects.

Significance Criteria

- A significance level of **α = 0.05** was used.
- P-values below this threshold were considered statistically significant.
- Percentage differences in medians were reported with p-values to show practical relevance.

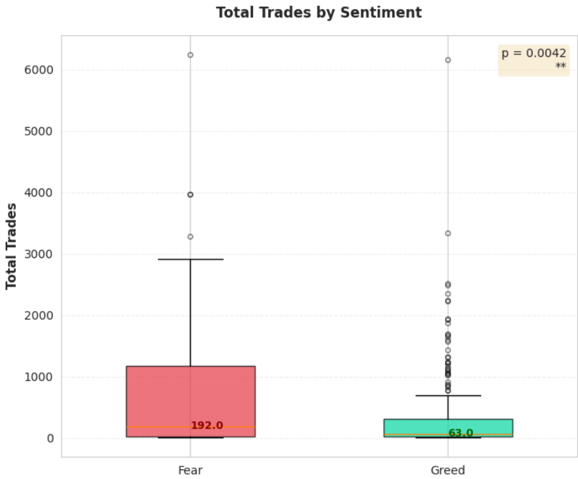
Fear vs Greed Regime Comparison:

This section presents a quantitative comparison of trading behaviour, trade volume, risk characteristics, and performance metrics across fear and greed sentiment regimes. The objective is to identify whether observed behavioural differences are statistically significant and give behavioural insights. [A significance level of α = 0.05 was used. P-values below this threshold were considered statistically significant.]

Trading days were classified into **Fear** (Fear + Extreme Fear) and **Greed** (Greed + Extreme Greed) regimes, and non-parametric statistical tests were applied to compare median values across regimes.

i. Trade Frequency

Trading frequency differs by a reasonable amount across sentiment regimes.

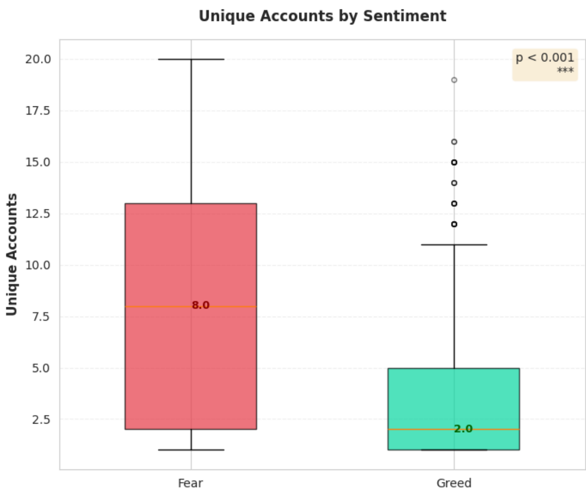


- **Median total trades (Fear):** 192 trades/day (days = 105)
- **Median total trades (Greed):** 63 trades/day (days = 307)
- **Relative increase:** +204.8%
- **Mann–Whitney U p-value:** 0.0042

The difference is statistically significant, indicating that fear regimes are characterized by higher trading intensity. This suggests reactive or panic behaviour, where market participants increase trade frequency in response to heightened uncertainty.

ii. Market Participation

Participant breadth expands significantly during fear regimes.

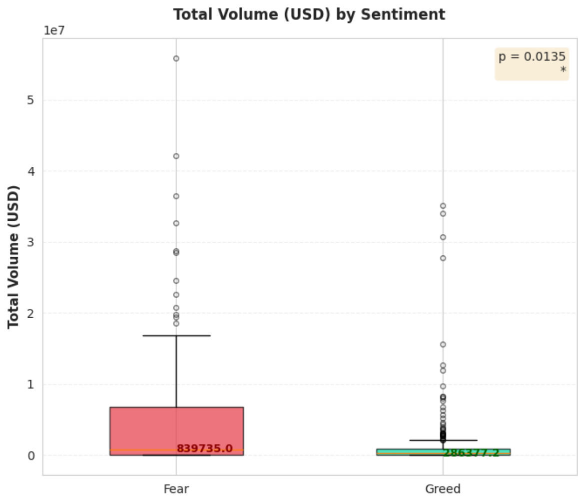


- **Median unique accounts (Fear):** 8 accounts/day
- **Median unique accounts (Greed):** 2 accounts/day
- **Relative increase:** +300.0%
- **Mann–Whitney U p-value:** < 0.001

The strong statistical significance indicates that fear induces broad participation across market participants, consistent with herd behaviour and forced activity during stressed conditions.

iii. Capital Deployment

Capital flow differs across regimes.

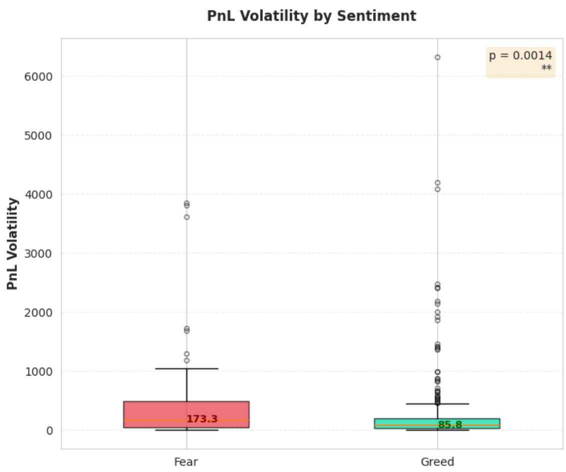


- **Median total volume (Fear):** $839,734.97/day
- **Median total volume (Greed):** $286,377.20/day
- **Relative increase:** +193.2%
- **Mann–Whitney U p-value:** 0.0135

Fear regimes exhibit significantly higher capital movement, suggesting large-scale position adjustments, liquidations, or defensive reallocations rather than opportunistic accumulation.

iv. Risk Characteristics

Outcome dispersion increases sharply during fear regimes.



- **Median PnL volatility (Fear):** 173.28
- **Median PnL volatility (Greed):** 85.85
- **Relative increase:** +101.8%
- **Mann–Whitney U p-value:** 0.00138

This result confirms that fear regimes are associated with significantly higher risk and instability, which is further shown by widened PnL distributions.

v. Performance Metrics

In contrast to activity and risk metrics, performance measures show no statistically meaningful differences across regimes.

- **Mean Closed PnL:** No significant difference
- **Median Closed PnL:** 0.0 in both regimes
- **Win Rate:** No significant difference

The absence of statistical significance indicates that sentiment regimes do not materially improve predictive accuracy of trade outcomes and expected per-trade profitability.

Thus, the results demonstrate that sentiment acts as a **behavioural and risk amplifier**, not as a performance enhancer. Fear changes *how much* and how violently the market trades, but not how well it predicts outcomes.

Results Summary:

Quantitative comparison across sentiment regimes reveals strong and statistically significant differences in trading behaviour:

- Trading activity is substantially higher during fear regimes, with median daily trade counts increasing by approximately **205%** relative to greed regimes.
- Market participation broadens significantly under fear, with the median number of active accounts increasing by **300%**.
- Capital deployment intensifies during fear regimes, with median traded volume rising by nearly **193%**.

Risk measures exhibit pronounced regime dependence:

- PnL volatility during fear regimes is approximately **2× higher** than during greed regimes.
- Elevated volatility coincides with periods of high trading activity and volume, suggesting concentration of risk within fear-driven intervals.

These findings confirm that fear regimes are structurally more unstable and are associated with heightened outcome dispersion.

In contrast to behavioural and risk metrics, performance-related measures show no statistically meaningful regime dependence:

- Mean closed PnL does not differ significantly between fear and greed regimes.
- Median closed PnL remains at zero across both regimes.
- Win rate exhibits no statistically significant variation across sentiment states.

This indicates that increased activity and risk during fear regimes do not translate into improved predictive accuracy or profitability.

Taken together, the results demonstrate a clear asymmetry between behavioural intensity and performance outcomes:

- Sentiment regimes materially influence how much, how broadly, and how violently markets trade.
- Sentiment regimes do not materially influence how well markets perform at the trade level.

Practical Implications:

The findings of this study have direct implications for how market sentiment should be interpreted and used in real-world trading, risk management, and analytical systems. Rather than being used as a predictive signal, sentiment is found to be a **modifier of behaviour and risk**.

The consistent increase in trading activity, participation, capital deployment, and volatility during fear regimes indicates that sentiment can be effectively used to **identify high-risk market conditions**.

Practical implementation:

- Use fear sentiment as a **risk signal**, not a trade trigger.
- Flag fear regimes as periods requiring heightened monitoring and stricter controls.
- Integrate sentiment into dashboards as a **market stress indicator**.

Since fear regimes are associated with approximately **2 times the PnL volatility**, exposure should be adjusted dynamically based on sentiment conditions.

- Reduce position sizes during fear regimes to maintain constant risk.
- Apply volatility-scaled sizing conditioned on sentiment.
- Avoid increasing leverage during fear-driven activity spikes.
- Prioritize trade quality over trade quantity when sentiment indicates stress.

Because risk is temporally concentrated in fear regimes, drawdowns are more likely to occur during these periods.

- Tighten stop-loss limits and drawdown thresholds during fear regimes.

The analysis demonstrates that market sentiment directly affects **how markets behave**, not the **outcomes**. Fear regimes concentrate participation, capital flow, and risk, while leaving profitability unchanged.

Therefore, sentiment is best deployed as a **risk-awareness and regime-conditioning tool**, guiding exposure control, execution discipline, and evaluation frameworks rather than serving as a standalone trading signal.