Trader Behavior vs Market Sentiment Analysis

Candidate: *Praful Sasane*
Date: *22/9/2025*

---

1. Introduction

This project analyzes how trading behavior — specifically **profitability, risk, volume, and leverage** — aligns or diverges from the **Bitcoin Fear & Greed Index**.

By merging two datasets:

1. **Historical Trader Data (Hyperliquid)**
   - Trade-level details: account, symbol, execution price, trade size, side, timestamp, closed PnL, etc.
2. **Bitcoin Fear & Greed Index**
   - Market sentiment labels (Extreme Fear, Fear, Neutral, Greed, Extreme Greed).

The objective is to uncover **hidden patterns** that may help design **smarter trading strategies**.

---

2. Data Preprocessing

- Loaded both datasets (`historical_data.csv` and `fear_greed_index.csv`).
- Name given to them = trades(`historical_data.csv`) and sentiment(`fear_greed_index.csv`)
- Merged them together and created different columns for different insights.
- Cleaned missing values in `date`, `classification`, and `value` columns.
- Drop all the rows that do not contain value or null value in `date`, `classification`, and `value` columns.
- Converted `timestamp` columns to datetime for alignment.
- Merged the datasets using the common `timestamp` column.
- Kept **classification** as categorical (no numeric encoding).
- Made a Sentiment Code Column to make some Visualization by converting categorical data in classification into numerical data.
- Final Output File Name – merged_trades_sentiment.xlsx

Final Dataset Columns:
`account, coin, execution_price, size_usd, side, timestamp, closed_pnl, fee, classification, value , size_tokens , timestamp_ist , start_position , direction , transaction_hash , order_id , crossed , trade_id , date , profitable , sentiment_code .`

3. Exploratory Data Analysis (EDA)

3.1 Trade Volume vs Sentiment (Time-Series)

- Aggregated daily trade volume and plotted against sentiment score.
- **Insight:** [Insert your observation, e.g., "Trade volume spikes during Extreme Greed but sentiment lags during Fear."]

3.2 Trade Volume vs Sentiment (Categories)

- Compared trade size distributions across sentiment classes using boxplots.
- **Insight:** [E.g., "Extreme Greed shows larger but riskier trades, Extreme Fear has smaller but more frequent trades."]

3.3 Profitability vs Sentiment

- Analyzed average `closed_pnl` by sentiment class.
- **Insight:** [E.g., "Traders tend to lose more during Greed phases, while profits are steadier during Neutral/Fear."]

3.4 Risk Analysis

- Measured risk by variance in trade sizes across sentiment states.
- **Insight:** [E.g., "Extreme Greed phases show high variance (large, inconsistent trades), indicating higher risk-taking."]

3.5 Distribution of PnL by Market Sentiment

- Histogram or KDE plots of `closed_pnl` for each sentiment class.
- **Insight Example:** "PnL distribution is highly skewed during Extreme Greed (many losses, few big wins)."

3.6 Daily PnL vs Market Sentiment

- Aggregate `closed_pnl` by day and compare with sentiment.
- **Insight Example:** "Daily average PnL tends to decline in Extreme Greed phases."

3.7 Risk vs Sentiment Trend

- Compare **trade size variance / volatility** by sentiment class.
- **Insight Example:** "Risk-taking is highest in Extreme Greed, lowest in Neutral."

3.8 Trade Side (Long/Short) vs Sentiment

- Countplot or stacked barplot: `side` (Long/Short) vs sentiment.
- **Insight Example:** "Long positions dominate during Greed, Shorts increase during Fear."

3.9 Win Rate by Market Sentiment

- Define win = `closed_pnl > 0`.
- Calculate win percentage for each sentiment.
- **Insight Example:** "Win rate is highest in Neutral, lowest in Extreme Greed."

3.10 Top 10 Traders by PnL under Different Sentiments

- Rank top 10 accounts by total PnL for each sentiment.
- **Insight Example:** "Top traders consistently profitable across sentiments, but smaller traders overexpose themselves in Greed."

3.11 Trader Segmentation (PCA + KMeans)

- Features: avg trade size, avg pnl, win rate, risk level.
- Apply PCA → visualize clusters in 2D.
- Apply KMeans → segment traders into groups (e.g., "Risky but High Profit", "Conservative Steady", etc.).
- **Insight Example:** "Cluster analysis reveals 3 trader archetypes, with risk-seeking traders highly active in Greed phases."

3.12 Machine Learning: Predicting Trade Success

- Defined trade success = `1 if closed_pnl > 0 else 0`.
- Features used:
  - `execution_price`
  - `size_usd`
  - `side` (encoded)
  - `classification` (sentiment encoded or left categorical with one-hot encoding)
  - `value` (sentiment score)
  - `fee`
- Model: Logistic Regression or Random Forest.
- Evaluated with Accuracy, Precision, Recall, F1-score.
- **Insight Example:** "Model predicts trade success with ~72% accuracy. Sentiment and trade size were strong predictors."

3.13 RandomForestClassifier: Feature Importance in Predicting Trade Profitability

- Trained a `RandomForestClassifier` to predict `trade_success`.
- Extracted **feature importance** scores.
- **Key Findings:**
  - Sentiment (`classification`/`value`) is among top features.
  - Trade size and side (long/short) also highly influential.
  - Fees had minimal predictive power.

4. Key Findings

- **Volume:** Highest during Extreme Greed, but not always profitable.
- **Profitability:** Traders appear more disciplined and profitable during Neutral/Fear markets.
- **Risk:** Variability in trade size increases with Greed phases.
- **Leverage:** [If missing, state: "Leverage data not available in dataset, so risk was analyzed via trade size variance instead."]
- • Machine learning models confirm that **market sentiment significantly impacts trade success**.
- • Random Forest feature importance shows **trade size + sentiment** as top predictors.

5. Conclusion & Recommendations

- Market sentiment significantly influences trader behavior.
- **Extreme Greed**: High volume and high risk, often leading to losses.
- **Extreme Fear**: Lower volume but steadier profitability.
- **Recommendation:** Traders should exercise caution during Greed phases and consider opportunities in Fear phases.