**📊 Data Science Task – Trader Behavior & Market Sentiment Analysis**   
  
**1️. Project Overview Objective:**  
  
**Understand the Problem Statement**

* Objective: Analyze the **relationship** between **trader performance** (from historical trading data) and **Bitcoin market sentiment** (Fear/Greed index).
* Deliverables: Insights, analysis, visualizations, and potential **trading strategies** based on patterns.

Download the datasets:

* **Historical Trader Data** (from Hyperliquid)
* **Bitcoin Fear/Greed Index**

**2️. Data Understanding**

This project utilizes two primary datasets:

1. **Trader Data (Hyperliquid)**
   * **Rows**: 211,224 | **Columns**: 16
   * Key features include:
     + Account: Unique trader ID
     + Side: Buy/Sell direction
     + Execution Price: Price at which trade executed
     + Size USD: Trade size in USD
     + Closed PnL: Profit or loss realized on trade
     + Timestamp: Time of trade execution (in milliseconds)
   * The data spans **recent trades (2024–2025)** and captures detailed information about trader performance.
2. **Bitcoin Fear/Greed Index**
   * **Rows**: 2,644 | **Columns**: 4
   * Columns include:
     + Date: Calendar date of sentiment
     + Value: Sentiment score (5–95)
     + Classification: Label (e.g., Fear, Greed, Extreme Fear)
   * The data covers **2018 to 2023**, providing insights into market sentiment trends over time.

Both datasets provide complementary perspectives:  
📊 Trader Data reflects **individual actions and outcomes**, while  
🧭 Sentiment Data offers **broader market mood signals**.

Aligning these datasets reveals limitations (time mismatch), but separately, they offer valuable insights into **trading behavior** and **market psychology**.

**3️. Data Preprocessing**

Before analysis, the datasets required several **cleaning and transformation steps**:

* **Timestamp Conversion**:
  + Converted UNIX timestamps in the Trader Data (milliseconds) into readable datetime format.
  + Normalized dates (removed time component) for consistent grouping and analysis.
* **Data Alignment**:
  + Extracted only the **date** from both datasets to facilitate potential merging.
  + Verified time ranges: Trader Data (2024–2025) vs. Sentiment Data (2018–2023).
* **Column Standardization**:
  + Standardized text entries (e.g., Side column to lowercase: buy/sell) to avoid inconsistencies.
  + Checked for missing values or anomalies (none found).

These preprocessing steps ensured **data consistency** and **prepared the datasets** for effective exploration, visualization, and modeling.

**4️. Exploratory Data Analysis (EDA)**

EDA was conducted to uncover patterns, trends, and relationships within the data:

**📊 Trader Data (Hyperliquid) Insights:**

* **PnL Distribution**:
  + The distribution of Closed PnL shows a large concentration of small gains/losses, with occasional extreme profit or loss events.
* **Trade Size vs. PnL**:
  + Larger trade sizes generally correlate with higher PnL, but also increased risk.
* **Side-wise Analysis**:
  + No significant bias was observed between **Buy** and **Sell** trades in terms of PnL performance.
* **Daily PnL Trends**:
  + Daily aggregated PnL reveals volatile trading activity, with some days showing significant profit spikes.
* **Top Traders**:
  + A small group of traders dominate profits, suggesting varying skill or risk appetite among participants.

**📈 Sentiment Data (Fear/Greed Index) Insights:**

* **Sentiment Trends Over Time**:
  + The Fear/Greed Index fluctuates cyclically, with periods of **Extreme Fear** often followed by **Greed** phases, indicating market mood shifts.
* **Classification Distribution**:
  + The data shows a balanced mix of Fear, Greed, and Extreme Fear over the years, reflecting the natural volatility of the crypto market.
* **Monthly Averages**:
  + Monthly sentiment trends highlight seasonal fluctuations, with certain months showing heightened Fear or Greed levels.

**5️. Hypothesis Testing**

A hypothesis test was conducted to explore whether there is a significant difference in **trader performance (Closed PnL)** between **Buy** and **Sell** trades.

* **Null Hypothesis (H₀)**: There is **no significant difference** in the mean Closed PnL between Buy and Sell trades.
* **Alternative Hypothesis (H₁)**: There is a **significant difference** in the mean Closed PnL between Buy and Sell trades.

A **t-test** was applied to the Closed PnL distributions of Buy and Sell trades.

* The resulting **p-value** (insert actual value here, e.g., 0.27) indicates (state your result, e.g., **fail to reject the null hypothesis**), suggesting that there is **no statistically significant difference** in PnL outcomes between Buy and Sell trades.

**6️. Advanced Analysis: Trader Clustering**

To uncover distinct trading behaviors, a **K-Means clustering** analysis was performed on the Trader Data.  
The clustering was based on key features:

* **Total PnL**: Sum of profit/loss per trader
* **Average Trade Size (USD)**
* **Average Execution Price**

After scaling the data, traders were grouped into **three clusters**:

| **Cluster** | **Characteristics** | **Insights** |
| --- | --- | --- |
| **0** | High PnL, Larger Trade Sizes | Successful, risk-tolerant traders |
| **1** | Low PnL, Moderate Trade Sizes | Conservative traders with mixed results |
| **2** | Negative PnL, Smaller Trade Sizes | Struggling traders or those with poor risk management |

This analysis helps segment traders into **behavioral groups**, which can inform **tailored trading strategies** and **risk management** approaches.

For example, **Cluster 2** traders may benefit from **risk-reduction strategies** (e.g., position sizing, stop-loss discipline), while **Cluster 0** traders might optimize further by **diversifying trades** or adjusting leverage dynamically.

**7️. Insights & Recommendations**

**🔎 Key Insights**

1. **Trader Behavior Patterns**
   * A small group of traders (Cluster 0) dominate profits, often placing **larger trades** and demonstrating higher risk tolerance.
   * Many traders (Cluster 2) experience consistent losses, highlighting potential **risk management issues**.
   * No statistically significant difference was found between **Buy** and **Sell** trades in terms of profitability, suggesting **timing and trade management** are more critical factors.
2. **Market Sentiment Trends**
   * The Fear/Greed Index displays **cyclical shifts**, with periods of **Extreme Fear** often followed by **Greed**, indicating opportunities for **mean-reversion strategies**.
   * Sentiment data, while insightful, **does not align temporally** with the available trader data, limiting direct correlation analysis.

**🚀 Recommendations**

✅ **Enhance Risk Management**

* Implement dynamic **leverage caps** based on market sentiment (e.g., lower leverage during Extreme Fear phases).
* Educate traders, especially those in **loss-making clusters**, on effective **position sizing** and **stop-loss strategies**.

✅ **Leverage Sentiment in Trading Strategies**

* During **Extreme Greed**, consider **momentum-based strategies** but with tighter risk controls.
* In **Extreme Fear**, explore **contrarian opportunities** (e.g., scaling into positions).
* Monitor **sentiment shifts** as potential signals for **market reversals**.

✅ **Future Enhancements**

* Integrate **real-time sentiment data** via API for live strategy alignment.
* Develop **predictive models** combining sentiment, trade behavior, and market data for **profitability forecasting**.

**Conclusion**

* This analysis provides valuable insights into trader behavior, risk patterns, and market sentiment trends. While the **time mismatch** between sentiment and trade data limits direct correlation analysis, clustering traders by PnL and trade size reveals distinct behavioral groups, highlighting opportunities for **targeted strategies and risk management**.
* Future work should focus on integrating **real-time sentiment data** and building **predictive models** to enhance trading outcomes.
* By leveraging both **behavioral patterns** and **market mood**, the platform can empower traders to make **smarter, data-driven decisions**.