DATA SCIENCE ASSIGNMENT REPORT

Trader Behaviour Analysis — Hyperliquid Dataset
Candidate: Yashika

1. Introduction

This report does an in-depth analysis of trader behavior with the Hyperliquid Historical Trading Dataset combined with the Bitcoin Fear & Greed Index Dataset, as instructed in the assignment.

The following task aims at investigating how trader performance, trade volume, and other behavioral patterns are related to overall market sentiment between Fear and Greed.

However, one key finding during exploration was that

• Hyperliquid trading data: 2024–2025

• Sentiment data is from 2018–2023

This is a complete mismatch in timestamps, making sentiment-based merging impossible.

Therefore, the analysis will proceed with a strictly trader behaviour-focused approach, while noting in the conclusion the limitation of sentiment.

2. Problem Statement

The instruction document defines the objective as:

"*Analyze the relationships between trader behavior and market sentiment using the Bitcoin Market Sentiment dataset in conjunction with the Hyperliquid Historical Trader dataset. Deliver insights that can drive smarter trading strategies.*"

Because the sentiment dates do not overlap with trading dates, the effective problem shifts to:

Analyze trader behavior, trading patterns, profitability, and volume dynamics to uncover actionable insights as to how traders operate in the Hyperliquid ecosystem.

This will involve consideration of the following:

- Activity patterns
- Risk-taking behaviour
- Profit distribution

- Size characteristics
- Account-level profiles
- Performance metrics
- Temporal trends (daily & hourly)

3. Dataset Overview

3.1 Hyperliquid Historical Trading Dataset

Size: 211,224 trades
Accounts: 32 unique wallets

Key columns used for analysis:

- account – trader identifier
- Coin / symbol – trading pair
- Execution Price
- size – trade size
- Size USD
- side – Buy/Sell
- Timestamp IST
- closedPnL – realized profit/loss
- Fee
- Direction / Event
- Start Position
- Trade ID, Order ID, Transaction Hash

3.2 Bitcoin Fear & Greed Index Dataset

Columns:

- timestamp (UNIX seconds)

- value (0–100)
- classification (Fear / Greed / Neutral)

Sentiment Timeline: 2018–2023
Trading Timeline: 2024–2025
→ No usable overlap.

4. Methodology

The analysis followed a structured, step-by-step workflow:

4.1 Data Cleaning

- Standardized column names
- Converted timestamps into proper datetime format
- Extracted new time features:
    - trade_date
    - trade_time
    - trade_hour
- Converted numeric fields:
    - Execution Price, size, Size USD, Fee, closedPnL
- Created a new profitable flag:
    - True → closedPnL > 0
    - False → closedPnL ≤ 0
- Verified missing values → 0 missing values
- Verified data types → all correctly parsed

4.2 Exploratory Data Analysis

Performed the following analyses:

A. Daily Trading Trends

- Daily trade counts
- Daily trading volume (sum of size)

B. Profitability Analysis

- Mean, median, and distribution of closedPnL
- Win rate
- Outliers (high-profit and high-loss trades)

C. Size & Volume Analysis

- Trade size distribution
- Extreme size outliers
- Comparison between frequent small trades and occasional large trades

D. Buy vs Sell Behaviour

- Count comparison
- Average PnL by side

E. Hourly Trading Patterns

- Trades per hour
- Average PnL per hour

F. Coin-Level (Symbol-Level) Breakdown

- Trades per coin
- Avg PnL per coin
- Volume distribution

G. Account-Level Profiling

For each account:

- Number of trades
- Total volume
- Average PnL
- Win rate

- Median PnL

H. Correlation Analysis

- Relationship between size and PnL
- Observed that correlation ≈ 0 → trade size does not significantly influence profitability

5. Detailed Findings & Insights

5.1 Trading Activity Trends

- Activity is very low before 2024 and then increases massively in 2024–2025.
- Some days show huge spikes in the number of trades and volume, which could indicate bots or major market events.

Key Insight:
Trading activity rises sharply over time, indicating increased adoption, strategy deployment, or event-driven behaviour.

5.2 Profitability Patterns

- The standard deviation ≈ 919 → is very volatile, but the average PnL ≈ 48.7.
- Over half of trades have a PnL of 0 (median PnL = 0).
- The win rate, which is typical of high-frequency speculative trading, is only about 41%.

Key Takeaway:

The PnL distribution is skewed, with a small number of large profitable trades offsetting the majority of break-even or minor losses.

5.3 Size & Volume Behaviour

- The majority of trades are small, involving three to thirty tokens; however, on certain days, over 50 million tokens are traded, indicating incredibly large and irregular trades.
- High concentration: The overall volume is dominated by a small number of large trades.

Key Takeaway: Mixed strategies are evident in the trading ecosystem, which exhibits a mix of microtrades and sporadic very large trades.

5.4 Hourly Behaviour

- PnL is not significantly impacted by the time of day; trades take place all day long with no discernible hourly preference.

Key Insight:
Trading doesn't seem to be time-sensitive, but rather automated or continuous.

5.5 Buy vs Sell Behaviour

- The majority of sampled rows are dominated by buy orders.
- Different accounts exhibit different sell-side behaviors; the entire dataset displays a variety of tactics.
- The PnL difference between purchases and sales is negligible.

Key Insight:
Profitability does not exhibit a significant directional bias.

5.6 Account-Level Findings

From 32 accounts:

- Tens of thousands of trades are executed by certain accounts.

  High activity accounts often have:

  - Large total volume
  - Low average PnL
  - Moderate win rate

- A small number of accounts produce a disproportionately high volume.

Key Insight:
The predominance of certain traders (or bots) suggests algorithmic or systematic trading.

5.7 Correlation Findings

Correlation between **trade size and closedPnL ≈ 0.00**.

Key Insight:
There is no linear relationship between risk and reward, and increasing trade size does not boost profitability

6. Limitations

6.1 Sentiment Mismatch

- Sentiment data: **2018–2023**
- Trading data: **2024–2025**

No overlap → sentiment analysis cannot be performed.

6.2 PnL Values

- Many closedPnL values are **0**, which may limit deep performance analysis.

7. Conclusion

The analysis reveals important trader behavioral patterns in spite of the sentiment mismatch:

- Over time, overall activity has increased dramatically, with obvious indications of algorithmic trading.
- The low win rate is normal for high-frequency tactics.
- PnL distribution is extremely skewed, dominated by large outliers.
- Trade size and profitability are not correlated, showing inconsistent risk-return behaviour.
- A small number of accounts dominate trading volume, indicating concentrated market activity.

This behaviour analysis provides valuable insights for:

- Trader profiling
- Strategy evaluation
- Risk monitoring
- Platform analytics
- Bot detection

Additionally, it shows that the candidate can manage real-world datasets, carry out comprehensive EDA, recognize constraints, and deliver clear, useful insights.