data science report – web3 trading team

candidate: prince kashav
project: trader behavior vs market sentiment analysis

files:

- **notebook_1.ipynb (google colab)**
- **csv_files/daily_trading_metrics.csv**
- **csv_files/merged_trading_sentiment.csv**
- **outputs/ (charts & visualizations)**

---

objective

the main goal of this project was to see how trader behavior (profitability, volume, risk etc) connects with the overall market mood (fear or greed).
i wanted to find patterns that show how people trade differently depending on what the market "feels" like – and how that can help in making smarter trading decisions on hyperliquid exchange.

---

datasets used

1. historical trader data (hyperliquid)

- **rows: ~211,000**
- **key columns: account, coin, execution_price, size_tokens, size_usd, side, timestamp_ist, start_position, direction, closed_pnl, leverage, fee**
- **it basically represents all trade-level activity for multiple traders across diff coins & times.**

2. bitcoin market sentiment (fear & greed index)

- **rows: ~2,600**
- **columns: timestamp, value, classification, date**
- **sentiment values range from 0–100, with labels like "fear", "greed", and "neutral".**

---

data cleaning & preprocessing

1. **column standardization – all column names were made lowercase and snake_case for uniformity.**

2. timestamp handling – converted timestamp_ist into proper datetime, and created a new date column for daily aggregation.
3. duplicates / missing data – removed duplicate trades and any rows missing timestamps.
4. numeric conversion – converted execution_price, size_usd, closed_pnl, fee into numeric using pd.to_numeric().
5. sentiment merge – merged both datasets on date, then forward-filled and back-filled missing sentiment days.

---

feature engineering & metrics

for each day, i calculated metrics like:

metric	description
total_trades	total trades per day
total_volume_usd	total trade volume (usd)
avg_execution_price	mean execution price
median_fee	median fee charged
unique_accounts	number of active traders
avg_closed_pnl	average profit/loss per trade
win_rate	% of profitable trades

then merged these with sentiment values (value, classification).

---

exploratory data analysis (eda)

1. avg daily volume by sentiment
trading volume was higher during *greed* phases than *fear*.
so, when the market is more positive, traders take bigger positions and trade more.

2. sentiment vs volume (time series)
a cyclical pattern appeared:

- during *greed*, both trading volume & pnl rise.
- during *fear*, both drop noticeably.

3. correlation insights

metric	corr w/ sentiment value
total volume (usd)	+0.42
average pnl	+0.35
win rate	+0.28
median fee	+0.05

so, higher sentiment = higher performance (roughly speaking).

---

statistical validation

i ran a mann-whitney u test comparing fear vs greed days.

metric	p-value	interpretation
total volume	< 0.001	significant difference
win rate	0.004	traders perform better in greed periods

hence, at 95% confidence, the differences are statistically valid.

---

machine learning (optional)

i also tried a quick kmeans clustering on scaled daily metrics.
it gave 2 clusters:

- cluster 0 (low activity) – low sentiment, low volume, low returns
- cluster 1 (high activity) – high sentiment, higher trade counts and profits

so the clustering supported what we saw in eda.

---

key insights

1. trader activity clearly follows sentiment – less active in fear, more active in greed.
2. profitability improves when market sentiment is higher.
3. when sentiment drops, traders reduce position sizes and trade more cautiously.
4. sentiment indicators can act like an *early warning system* for shifts in trading volume and performance.

---

conclusion

overall, market sentiment strongly influences trading behavior.
greed periods = more participation, higher profits, and larger trades.
using sentiment analysis inside trading models could really improve entry/exit timing and risk management.

---

deliverables summary

file	description
notebook_1.ipynb	complete colab notebook
csv_files/daily_trading_metrics.csv	aggregated trader stats
csv_files/merged_trading_sentiment.csv	final merged dataset
outputs/	eda charts, plots
ds_report.pdf	final report