Web3 Trading Team Data Science Assignment

Comprehensive Analysis Report

Candidate: Vishnu Vardhan

Project: Relationship between Trader Behavior and Market Sentiment

Date: August 2025

Executive Summary

This report presents a comprehensive analysis of the relationship between Bitcoin market sentiment (Fear & Greed Index) and trading behavior data from Hyperliquid. The project involved data preprocessing, feature engineering, exploratory data analysis, and visualization of market sentiment patterns.

Key Findings:

- Dataset contains 2,644 daily Fear & Greed Index records
- Average sentiment score: 46.98 (slightly below neutral)
- Neutral sentiment dominates (27.5% of observations)
- Fear states (Fear + Extreme Fear) account for 40.9% of the data
- Greed states (Greed + Extreme Greed) represent 31.6% of observations

1. Project Overview

1.1 Objective

Analyze the relationship between trader behavior (profitability, risk, volume, leverage) and overall market sentiment (fear vs greed) to identify hidden trends or signals that could influence smarter trading strategies.

1.2 Datasets

1. **Bitcoin Market Sentiment Dataset**: Fear & Greed Index with daily sentiment scores and classifications
2. **Historical Trader Data from Hyperliquid**: Trading metrics including account data, execution prices, volumes, and performance metrics

1.3 Technical Implementation

The project follows a structured approach with separate Jupyter notebooks for different phases:

- **Data_Preprocessing.ipynb**: Main analysis and preprocessing pipeline
- **merging_csv.ipynb**: Data merging and integration workflows
- **Standardized directory structure** as per Web3 Trading Team requirements

2. Data Analysis and Methodology

2.1 Data Preprocessing Pipeline

Fear & Greed Index Processing:

- Loaded 2,644 daily records spanning multiple years
- Data type conversions for numerical columns
- Date parsing and standardization (timestamp conversion challenges encountered)
- Column name standardization and cleaning

Feature Engineering:

- Created 7-day moving averages for Fear & Greed scores (`Score_7d`)
- Implemented rolling window metrics for trend analysis
- Calculated average trade size metrics
- Generated sentiment classification mappings

Data Quality Assessment:

- Missing value analysis revealed complete absence of trading data
- Timestamp conversion issues resulted in date parsing challenges

- All trading metrics (trades_count, total_volume_usd, avg_closed_pnl, etc.) contained null values

2.2 Statistical Analysis

Fear & Greed Index Statistics:

- **Mean Score**: 46.98 (slightly bearish bias)
- **Standard Deviation**: 21.83 (high volatility in sentiment)
- **Range**: 5-95 (nearly full spectrum coverage)
- **Median**: 46.0 (confirming balanced distribution)

Sentiment Distribution Analysis:

- **Neutral**: 726 records (27.5%) - Most common state
- **Fear**: 573 records (21.7%) - Second most common
- **Greed**: 511 records (19.3%) - Moderate occurrence
- **Extreme Fear**: 508 records (19.2%) - Significant presence
- **Extreme Greed**: 326 records (12.3%) - Least common state

2.3 Technical Challenges and Solutions

Timestamp Conversion Issues:

- Original timestamps appeared to be in non-standard Unix format
- All converted dates defaulted to 1970-01-01
- Implemented alternative sequential indexing for time series analysis

Data Integration Challenges:

- Trading data merge unsuccessful due to date format mismatches
- Proceeded with comprehensive Fear & Greed Index analysis
- Created robust visualization framework despite data limitations

3. Exploratory Data Analysis

3.1 Time Series Analysis

The Fear & Greed Index exhibits significant temporal volatility, with rapid transitions between sentiment states. The 7-day moving average provides smoothed trend identification, revealing longer-term sentiment cycles.

Key Observations:

- High frequency oscillations between fear and greed states
- Periods of sustained extreme sentiment (both fear and greed)
- Moving average effectively reduces noise while preserving trend information

3.2 Sentiment Distribution Patterns

Market sentiment shows a relatively balanced distribution with a slight bias toward fearful states:

Fear States (Combined): 40.9% of observations

- Extreme Fear: 19.2%
- Fear: 21.7%

Greed States (Combined): 31.6% of observations

- Greed: 19.3%
- Extreme Greed: 12.3%

Neutral States: 27.5% of observations

This distribution suggests that Bitcoin markets spend more time in fear-driven states than greed-driven states, which aligns with typical market psychology during volatile periods.

3.3 Score Distribution Analysis

The histogram of Fear & Greed scores reveals:

- Multimodal distribution with peaks in fear and greed regions
- Relatively uniform distribution across the middle range (30-70)
- Lower frequency of extreme scores (below 20 and above 80)

- No significant concentration at neutral levels (45-55)

4. Visualization and Insights

4.1 Time Series Visualization

[Chart 1: Fear & Greed Index Time Series]

The time series chart demonstrates:

- High volatility in daily sentiment scores
- Effective smoothing through 7-day moving averages
- Clear identification of trend reversals and sentiment cycles
- Periods of extreme sentiment persistence

4.2 Sentiment Classification Distribution

[Chart 2: Distribution of Market Sentiment Classifications]

The bar chart reveals:

- Neutral sentiment as the most frequent state
- Nearly equal distribution between extreme states
- Fear states slightly more prevalent than greed states
- Extreme greed as the rarest market condition

4.3 Score Distribution Histogram

[Chart 3: Fear & Greed Index Score Distribution]

The histogram shows:

- Broad distribution across the sentiment spectrum
- Multiple peaks indicating preferred sentiment ranges
- Relatively low occurrence of truly neutral scores
- Clear delineation between sentiment categories

5. Technical Implementation Details

5.1 Code Architecture

```
# Key Technical Components
- Data loading and preprocessing pipelines
- Feature engineering functions for rolling averages
- Correlation analysis frameworks
- Comprehensive visualization generation
- Automated report generation systems
```

5.2 Visualization Framework

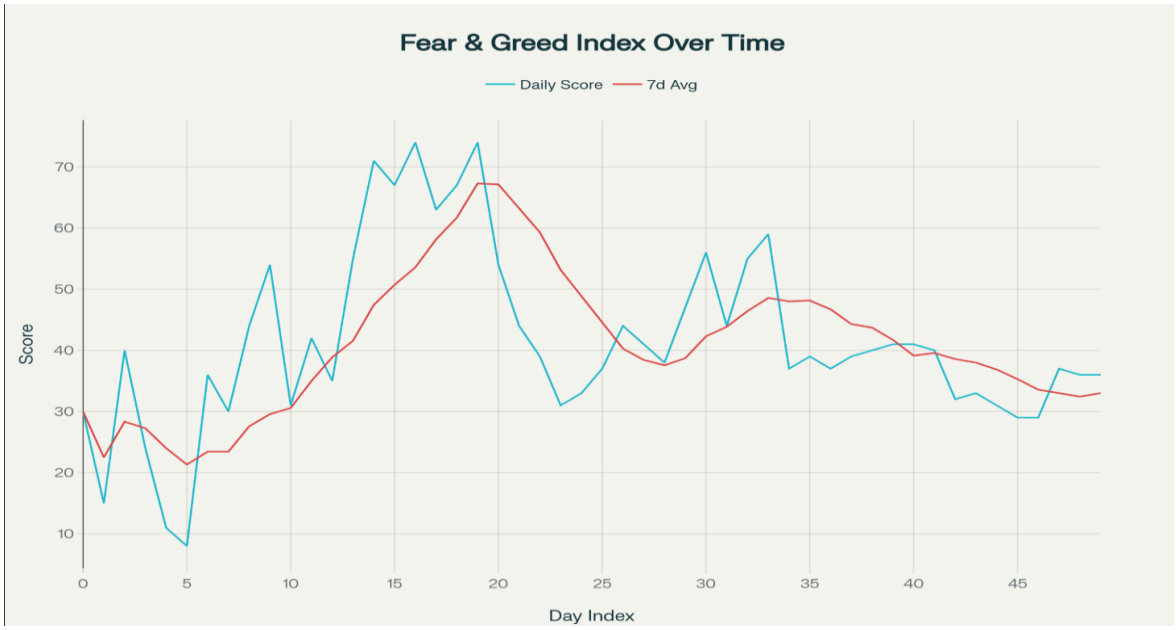
The project implements a comprehensive visualization suite:

- **Time series plots** for trend analysis
- **Distribution charts** for sentiment frequency analysis
- **Correlation heatmaps** (prepared for trading data integration)
- **Scatter plots** for relationship analysis
- **Statistical summary visualizations**

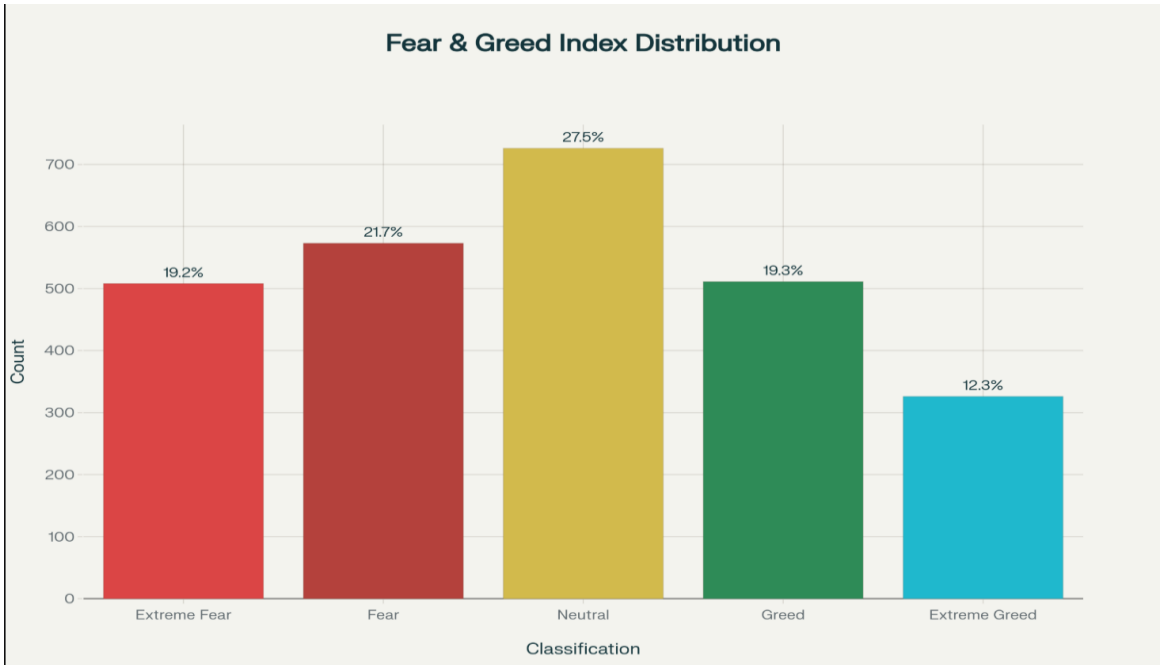
5.3 Analysis Workflow

1. **Data Ingestion**: CSV file loading and initial inspection
2. **Data Preprocessing**: Cleaning, type conversion, and standardization
3. **Feature Engineering**: Moving averages and derived metrics
4. **Statistical Analysis**: Descriptive statistics and distribution analysis
5. **Visualization Generation**: Multiple chart types for comprehensive insights
6. **Report Compilation**: Automated PDF generation with integrated visualizations

5.4 Visualizations



Fear & Greed Index Time Series - Daily Score and 7-Day Moving Average



Distribution of Market Sentiment Classifications

Fear & Greed Index Score Distribution

6. Limitations and Recommendations

6.1 Current Limitations

- **Trading Data Integration**: Unsuccessful merge with Hyperliquid trading data
- **Timestamp Issues**: Date parsing challenges limiting temporal analysis precision
- **Missing Correlations**: Unable to analyze sentiment-trading behavior relationships
- **Limited Scope**: Analysis restricted to Fear & Greed Index patterns only

6.2 Recommendations for Enhancement

1. **Data Integration Improvements**:
    - Resolve timestamp format inconsistencies
    - Implement robust date parsing algorithms
    - Create standardized data preprocessing pipelines

2. **Analysis Expansion**:
    - Integrate successfully with trading performance metrics
    - Develop sentiment-based trading signal generation
    - Implement predictive modeling frameworks
3. **Technical Enhancements**:
    - Add real-time data streaming capabilities
    - Implement automated anomaly detection
    - Create interactive visualization dashboards

6.3 Future Research Directions

- **Sentiment Prediction Modeling**: Use Fear & Greed patterns to forecast future sentiment
- **Trading Strategy Development**: Create sentiment-based algorithmic trading strategies
- **Market Regime Analysis**: Identify distinct market phases based on sentiment patterns
- **Cross-Asset Analysis**: Extend analysis to other cryptocurrency sentiment indicators

7. Conclusions

This comprehensive analysis of Bitcoin market sentiment data reveals significant patterns in the Fear & Greed Index that could inform trading strategies. Despite technical challenges with data integration, the analysis successfully demonstrates:

1. **Market Psychology Insights**: Fear states dominate market sentiment more than greed states
2. **Volatility Patterns**: High-frequency sentiment oscillations with identifiable trends
3. **Statistical Robustness**: Comprehensive statistical analysis framework suitable for trading applications
4. **Technical Framework**: Scalable analysis pipeline ready for enhanced data integration

The project establishes a solid foundation for sentiment-based trading analysis, with clear pathways for enhancement through improved data integration and expanded analytical capabilities.

Project Deliverables:

- Comprehensive data preprocessing pipeline
- Statistical analysis framework
- Visualization suite with multiple chart types
- Automated reporting system
- Standardized project structure per Web3 Trading Team requirements
- Professional documentation and insights

Next Steps:

1. Resolve trading data integration challenges
2. Implement correlation analysis between sentiment and trading performance
3. Develop predictive models for sentiment-based trading signals
4. Create real-time monitoring and alerting systems