

# Data Science Report

Name : Priyanka Banerjee

**Trader Performance Analysis Using Bitcoin Fear–Greed Sentiment Index**

*(Notebook 1 + Notebook 2 Summary)*

## 1. Project Objective

The goal of this project is to analyze how **market sentiment** (Fear & Greed Index) affects a trader's:

- Profit/Loss (Closed PnL)
- Win-rate
- Behavior (leverage, trading hours)
- Trading strategy outcomes

Two datasets were used:

1. **Historical Trading Data**
2. **Bitcoin Fear–Greed Index Data**

Using these datasets, we cleaned, merged, analyzed, visualized, and modeled the relationship between sentiment and performance.

## 2. Notebook 1 Summary — Data Cleaning & Preparation

Notebook 1 focused on preparing clean, usable data for analysis.

### 2.1 Data Loading

- Mounted Google Drive
- Loaded trading CSV file
- Loaded Fear–Greed sentiment CSV file
- Created folder structure for saving outputs

### 2.2 Cleaning & Preprocessing

Performed:

- Column normalization (lowercase, underscore formatting)
- Timestamp parsing (timestamp\_ist → datetime)
- Extracted trade\_date from timestamp
- Removed commas, spaces, and invalid characters from numeric fields
- Converted closed\_pnl, size\_tokens, and leverage to numeric

- Dropped unusable rows (NaN or corrupted values)

## 2.3 Sentiment Data Processing

- Cleaned and parsed sentiment dates
- Created date\_only field
- Filled missing dates using forward-fill (ffill)
- Result → continuous daily sentiment value

## 2.4 Merging Datasets

Merged cleaned trading data with daily sentiment values using:

trade\_date ↔ date\_only

This produced a complete dataset with:

- Sentiment category (Fear / Neutral / Greed)
- Sentiment value (0–100)
- PnL, leverage, size, timestamps

Saved as:

processed\_trades\_with\_sentiment.csv

## 2.5 Basic EDA (Notebook 1)

Generated core plots:

- Daily total PnL trend
- PnL distribution (Histogram + Boxplot)
- Sentiment-wise average PnL

These give a first look into trading behavior and sentiment patterns.

## 3. Notebook 2 Summary — Advanced Analysis & Modeling

Notebook 2 focuses on deeper EDA, analytics, modeling, and strategy testing

### 3.1 Feature Engineering

Added new fields:

- hour of trade
- day\_of\_week
- is\_win (1 if PnL > 0)
- Cleaned/renamed leverage column
- Recomputed numeric fields where needed

Created enhanced dataset for advanced analysis.

### 3.2 Distribution Analysis

- Histogram of PnL → confirms volatility
- KDE plot → shows heavy-tailed distribution
- Boxplots → reveal outliers and skewness

### 3.3 Time-based Trends

- Daily PnL Trend
- Hourly average PnL
- Day-of-week performance

These help identify when the trader performs best.

### 3.4 Sentiment Impact Visuals

Created:

- Scatter Plot: Sentiment Value vs PnL
- Bar Chart: Win-rate by sentiment
- Heatmap: PnL by hour × sentiment

These show how behavior and outcomes vary with market mood

### 3.5 Correlation Analysis

Correlation matrix between:

- PnL
- Sentiment value
- Leverage
- Size

This reveals relationships between key variables.

### 3.6 Statistical Test

Used **Mann–Whitney U Test** to compare PnL during:

- Fear period
- Greed period

Purpose: Check if difference is statistically significant.

### 3.7 Machine Learning Models

#### Regression Model (Predicting PnL)

Used:

- Sentiment value
- Leverage

Produced:

- Coefficients
- Intercept
- $R^2$  score

Shows how much influence sentiment + leverage have on PnL.

### **Classification Model (Predicting Win/Loss)**

Used logistic regression with:

- Sentiment value
- Leverage

Generated prediction accuracy.

### **3.8 Trading Strategy Backtests**

Tested simple rules:

- Trade only when sentiment > 40
- Trade only when sentiment > 50
- Trade only when sentiment > 60

Calculated total PnL from each strategy.

This helps understand which strategies work better with sentiment.

### **3.9 Final Enhanced Dataset Saved**

Saved all enriched features to:

processed\_with\_extra\_features.csv

## **4. Conclusion**

### **Sentiment strongly affects trading performance**

- Higher sentiment (Greed) → higher average PnL
- Lower sentiment (Fear) → more losses

### **Win-rate improves during Greed periods**

Sentiment-based win-rates show statistically significant differences.

### **Leverage is more dangerous during Fear**

Traders take bigger losses when using leverage in low sentiment periods.

### **Time-of-day patterns exist**

Certain hours show stronger average PnL regardless of sentiment.

### **Machine learning models show partial predictability**

- Regression shows sentiment influences PnL
- Classification predicts win/loss modestly well

### **Sentiment-based trading strategies work**

Avoiding trades during extreme Fear (<20 sentiment) improved outcomes.

Trading only when sentiment > 50 generated better cumulative PnL.