Data Science Report: Trader Behavior & Sentiment Analysis

This report summarizes the analysis performed in the project "Trader Behavior Insights: Fear & Greed Sentiment Analysis in Crypto Markets". The goal was to explore the relationship between trader profitability, trade volume, and risk-taking behavior with overall market sentiment (fear vs greed). Below is a step-by-step explanation of the methodology, data processing, analysis, and findings as documented in the accompanying Jupyter notebook.

This cell imports the pandas library to handle data manipulation. It loads two CSV files, 'historical_data.csv' and 'fear_greed_index.csv', into DataFrames. The key step is preparing and merging these datasets. It converts the date-time columns in both DataFrames into a standardized date format. This allows for a successful inner merge of the two datasets, combining crypto trading data with the corresponding daily sentiment scores. Finally, it prints the first five rows of the merged DataFrame to verify that the join was successful and the data is correctly structured for analysis.

This cell uses the value_counts() method on the 'Closed PnL' column to understand the distribution of profit and loss values. The output reveals that a large number of trades, 24,280 to be exact, resulted in a closed PnL of $0.00. This information is important for understanding the nature of the trading data, as many trades may have been opened and closed with no net gain or loss.

This cell performs a key step in the analysis by grouping the data by the classification column from the Fear & Greed Index. It then aggregates the data to calculate three important metrics for each sentiment category: Average_Closed_PnL: The mean profit/loss per trade. Average_Trade_Size_USD: The mean trade size in USD. Number_of_Trades: The total count of trades that occurred under each sentiment. This new sentiment_performance DataFrame is the basis for the core analysis in the following cells.

Using matplotlib and seaborn, this cell generates a bar plot to visually represent the Average_Closed_PnL for each sentiment classification. The visualization makes it easy to compare the profitability of trades across different market sentiments, showing which sentiment periods, on average, yielded the highest returns per trade.

This cell adds more calculated columns to the sentiment_performance DataFrame for a deeper analysis. Total_PnL: Calculated by multiplying the average PnL by the number of trades for each category. This metric reveals which sentiment periods were most profitable in aggregate. Average_PnL_per_USD: Calculated by dividing the average PnL by the average trade size in USD. This metric measures the efficiency of trades, showing which sentiment periods generated the most profit per dollar invested on average.

This cell creates three separate bar charts to visualize the newly calculated performance metrics. Average Closed PnL: A bar chart sorted to show that 'Extreme Fear' has the highest average PnL per trade, followed by 'Extreme Greed' and 'Fear'. Total PnL: This chart shows that the 'Fear' and 'Greed' categories generated the largest aggregate profits, likely due to the higher volume of trades during those periods. Average PnL per USD: This chart visualizes trade efficiency, with 'Extreme Fear' and 'Extreme Greed' showing the highest return on investment per dollar traded.