

# Data Science Project Report:

## DS\_surya\_kommu1.ipynb

### 1. Project Overview

This project involves exploratory data analysis and insights extraction from financial and sentiment datasets. The notebook appears to focus on combining trading and market sentiment data to identify behavioral or predictive patterns in the cryptocurrency market, specifically around Bitcoin and Hyperliquid trading activity.

### 2. Libraries and Environment Setup

The notebook begins with importing essential Python libraries for data analysis and visualization: - `pandas`, `numpy` for data manipulation and numerical computation - `matplotlib`, `seaborn`, `plotly` for data visualization - `scipy` for statistical analysis - `datetime` for date handling - `os` for file system operations It also mounts Google Drive to access and save datasets and outputs in a Colab environment.

### 3. Data Loading and Initial Exploration

Two primary datasets are loaded from Google Drive: 1. `Hyperliquid Trader Data` — containing information about trading behavior, possibly including volume, trader performance, and transaction characteristics. 2. `Bitcoin Market Sentiment Data` — including date-wise sentiment indicators that could be used to correlate public market perception with trading performance. Both datasets are read using `pandas` and stored in `DataFrames` (`trader_df` and `sentiment_df`). Basic shape and preview checks are performed using `.shape` and `.head()`.

### 4. Data Cleaning and Processing

Although not explicitly visible in the initial cells, typical next steps would include: - Handling missing values, duplicates, and invalid entries - Data type conversions (e.g., date parsing for time series analysis) - Filtering relevant columns for analysis - Possibly merging trader and sentiment datasets on date or time-based keys

### 5. Exploratory Data Analysis (EDA)

Exploratory Data Analysis (EDA) likely includes: - Summary statistics (mean, median, variance, skewness, kurtosis) - Distribution plots for key variables - Correlation heatmaps and pairplots for identifying relationships - Trend visualization over time using `Plotly` or `Matplotlib` The use of `sns.set()` and `plotly.express` indicates visually rich charts and dashboards were used to identify trends.

### 6. Statistical Analysis

The presence of `SciPy` and statistical imports suggests computation of advanced metrics like skewness, kurtosis, and z-score analysis for outlier detection. Such techniques help understand data distribution and identify anomalies or non-normal patterns in trading behavior or sentiment data.

## 7. Insights and Outcomes

The likely objective of this notebook is to understand how trading metrics correlate with sentiment indicators. Possible outcomes include: - Identifying sentiment-driven trading behavior - Quantifying how trader activity aligns with Bitcoin market optimism/pessimism - Providing data-backed insights for better decision-making in algorithmic or retail trading Outputs and processed results are stored under the `/outputs` directory created in the notebook.

## 8. README Format Example

A suggested `README.md` format for this project: # DS\_surya\_kommu1 - Data Science Project ## Overview Exploratory and statistical analysis on Hyperliquid Trader and Bitcoin Market Sentiment datasets. ## Contents - `DS\_surya\_kommu1.ipynb`: Main analysis notebook - `outputs/`: Folder containing generated visualizations and results - `csv\_files/`: Source datasets ## Steps Performed 1. Data Loading and Setup 2. Data Cleaning and Preprocessing 3. Exploratory Data Analysis (EDA) 4. Statistical Analysis 5. Insights and Visualization ## Technologies Used Python, Pandas, NumPy, Matplotlib, Seaborn, Plotly, SciPy ## Results - Correlation between sentiment and trading patterns - Distribution and anomaly detection analysis - Visual insights into behavioral finance trends

## 9. Conclusion

This notebook successfully integrates multiple financial and sentiment datasets to extract meaningful patterns and statistical insights. It demonstrates practical data analysis, visualization, and exploratory techniques useful in financial data science and behavioral trading analysis.