



# Historical Trade Data Analysis

---

*Assignment*

Prepared by: Chetan Tiwari

9/26/2024

---

# HISTORICAL TRADE DATA ANALYSIS

---

## INDEX:

- About the data
- Data Visualization and story telling
- Financial Metrics and ranking
- List of top 20 accounts based on ranking
- Methodology Findings and assumptions



### ABOUT THE DATA:

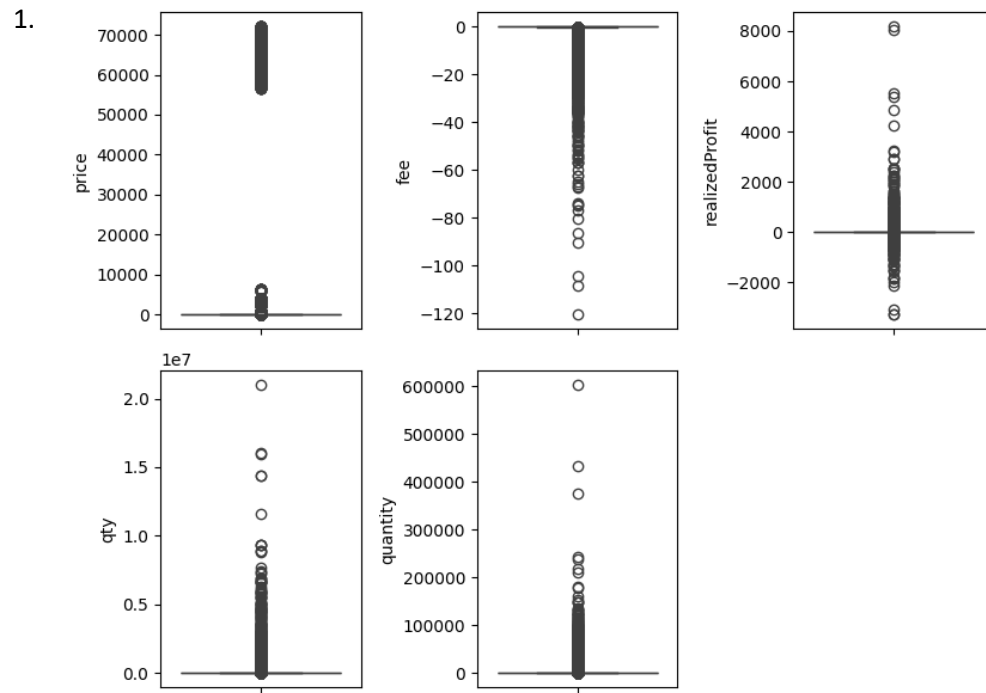
The dataset contains several columns that represent trades. Each row corresponds to a single trade and includes the following key columns:

1. Port\_IDs: A unique identifier for the portfolio or account executing the trade.
2. time: A timestamp indicating when the trade occurred.
3. symbol: The trading pair (e.g., SOLUSDT, DOGEUSDT), where the first part is the asset being traded and the second part is the currency (usually USDT, representing Tether).
4. side: Indicates whether the trade was a "BUY" or "SELL."
5. price: The price at which the asset was bought or sold.
6. fee: The transaction fee associated with the trade, which can be positive or negative, depending on whether it's a cost or rebate.
7. feeAsset: The currency in which the fee is charged (e.g., USDT).
8. quantity: The total monetary value of the trade (how much USDT was used for the trade).
9. quantityAsset: The amount of the asset being traded (e.g., how much of DOGE was bought/sold).
10. realizedProfit: The profit or loss realized from the trade, where a positive value indicates profit and a negative value indicates loss.
11. realizedProfitAsset: The asset in which the profit or loss is expressed (usually USDT).
12. baseAsset: The base asset being traded (e.g., SOL or DOGE).
13. qty: The quantity of the base asset in the trade.
14. positionSide: Indicates the position direction, such as "LONG" (buying) or "SHORT" (selling).
15. activeBuy: A boolean value indicating if the position is still open (True) or closed (False).

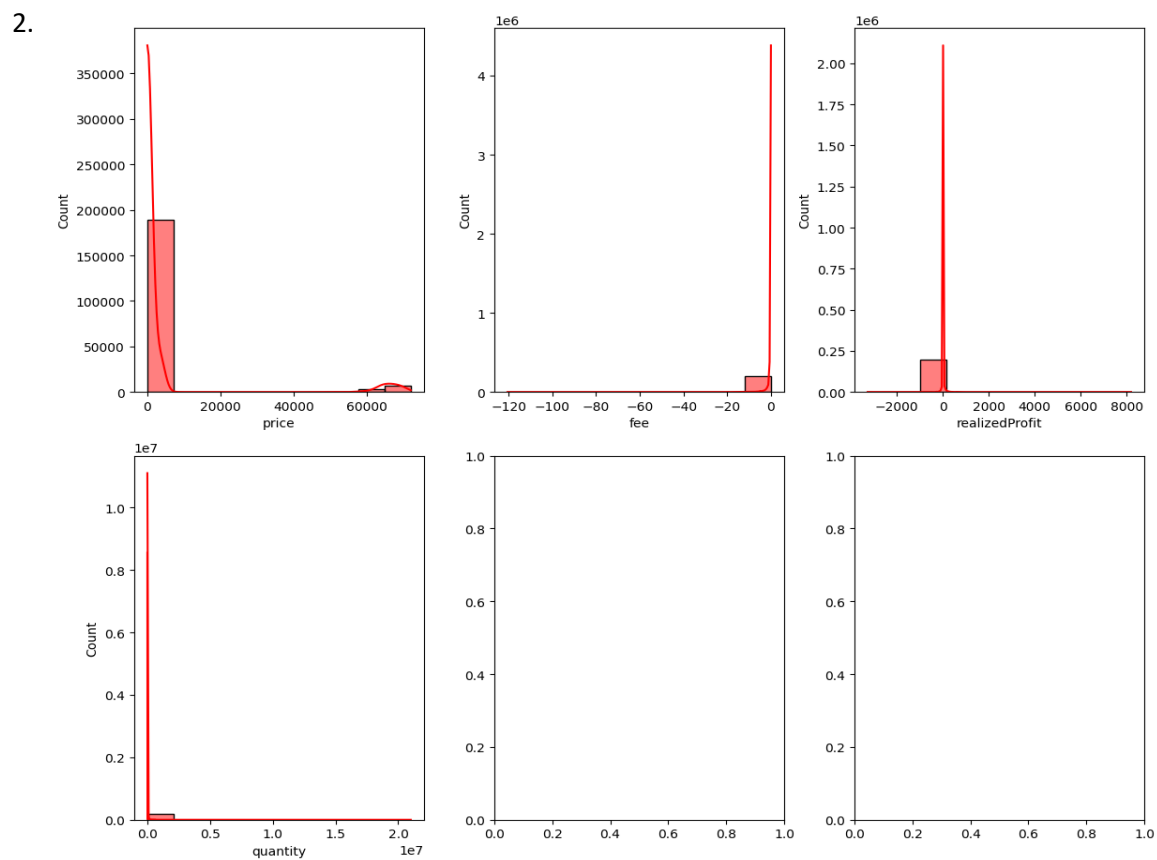


# HISTORICAL TRADE DATA ANALYSIS

## DATA VISUALIZATION AND STORY TELLING:



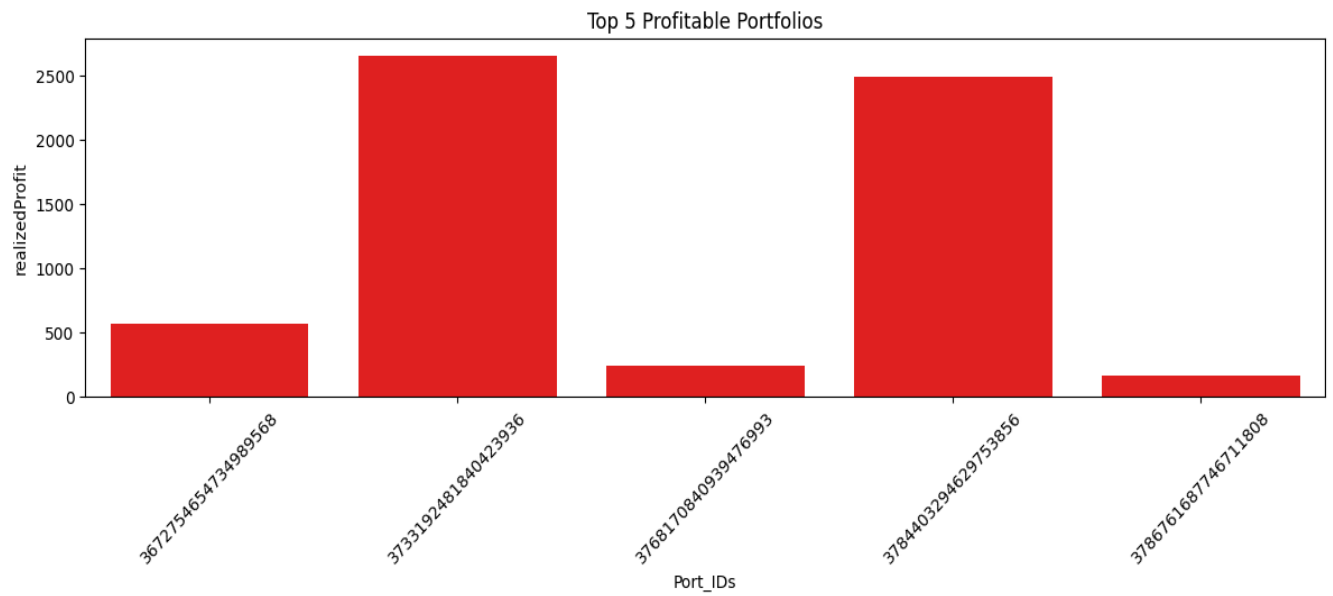
Box plot of all the numerical variables suggests the presence of massive outliers in the dataset



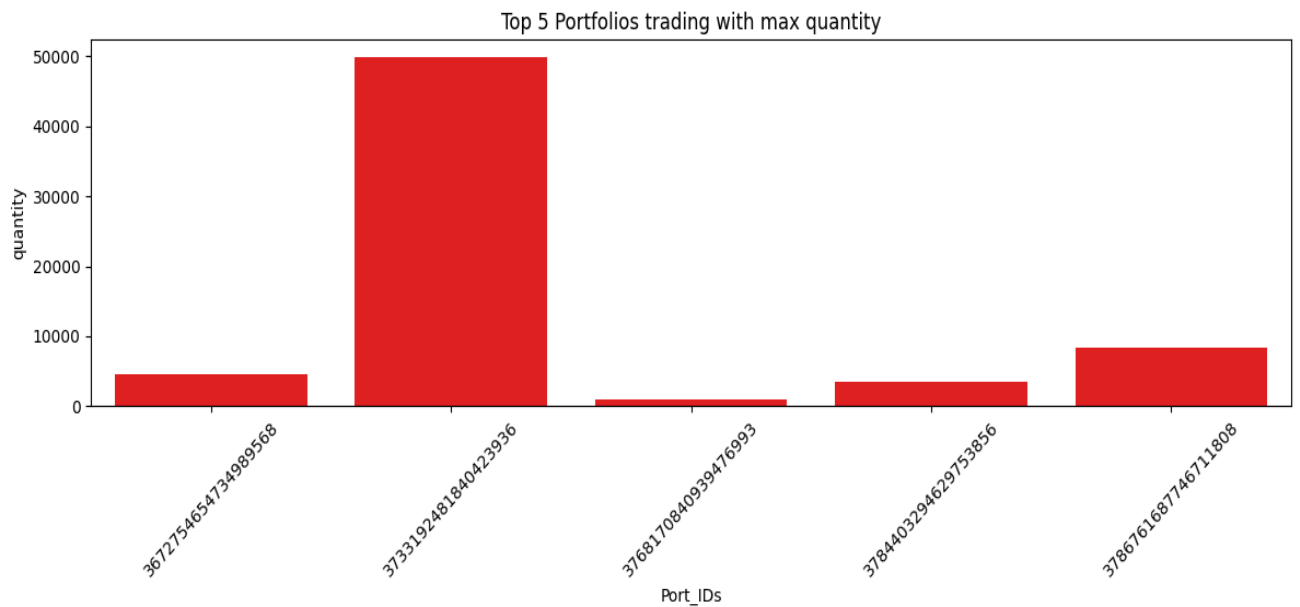
Histogram suggests that the data is skewed

## HISTORICAL TRADE DATA ANALYSIS

3.

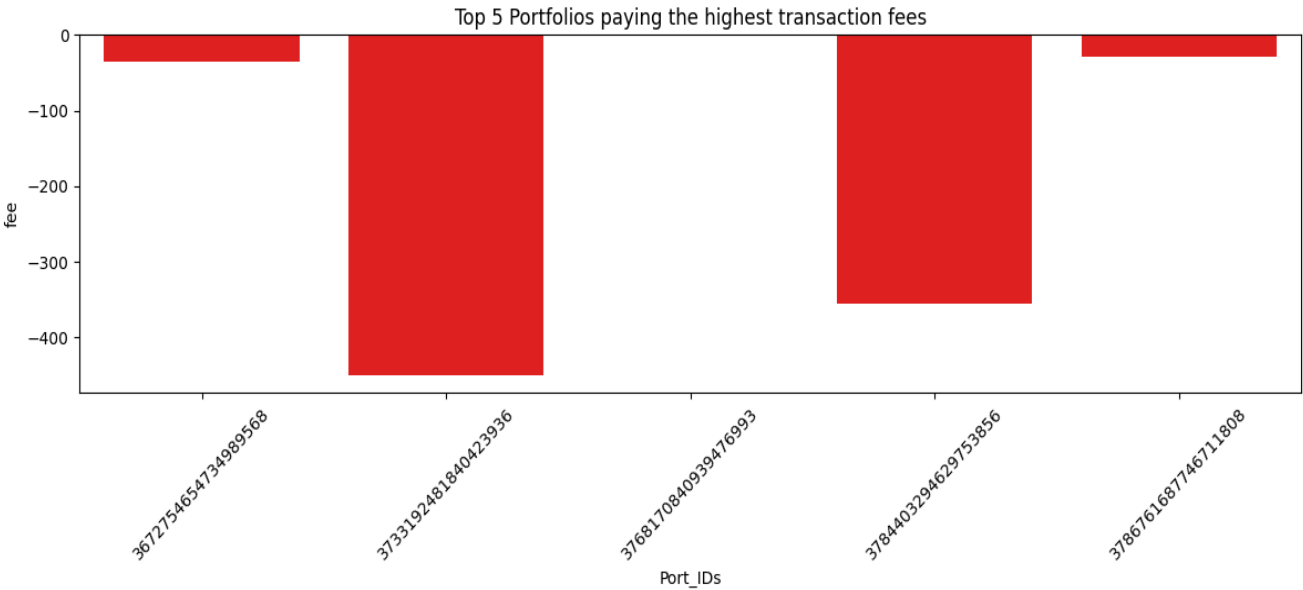


4.

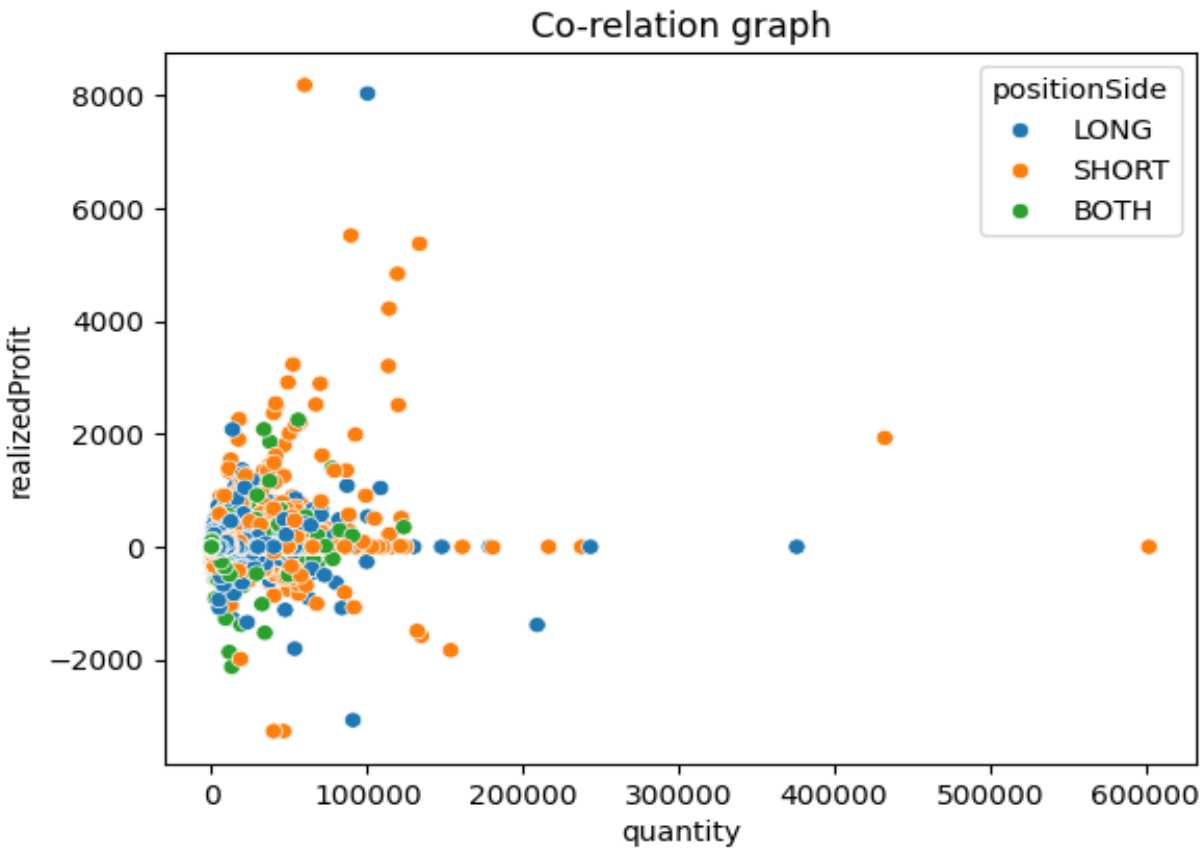


# HISTORICAL TRADE DATA ANALYSIS

5.

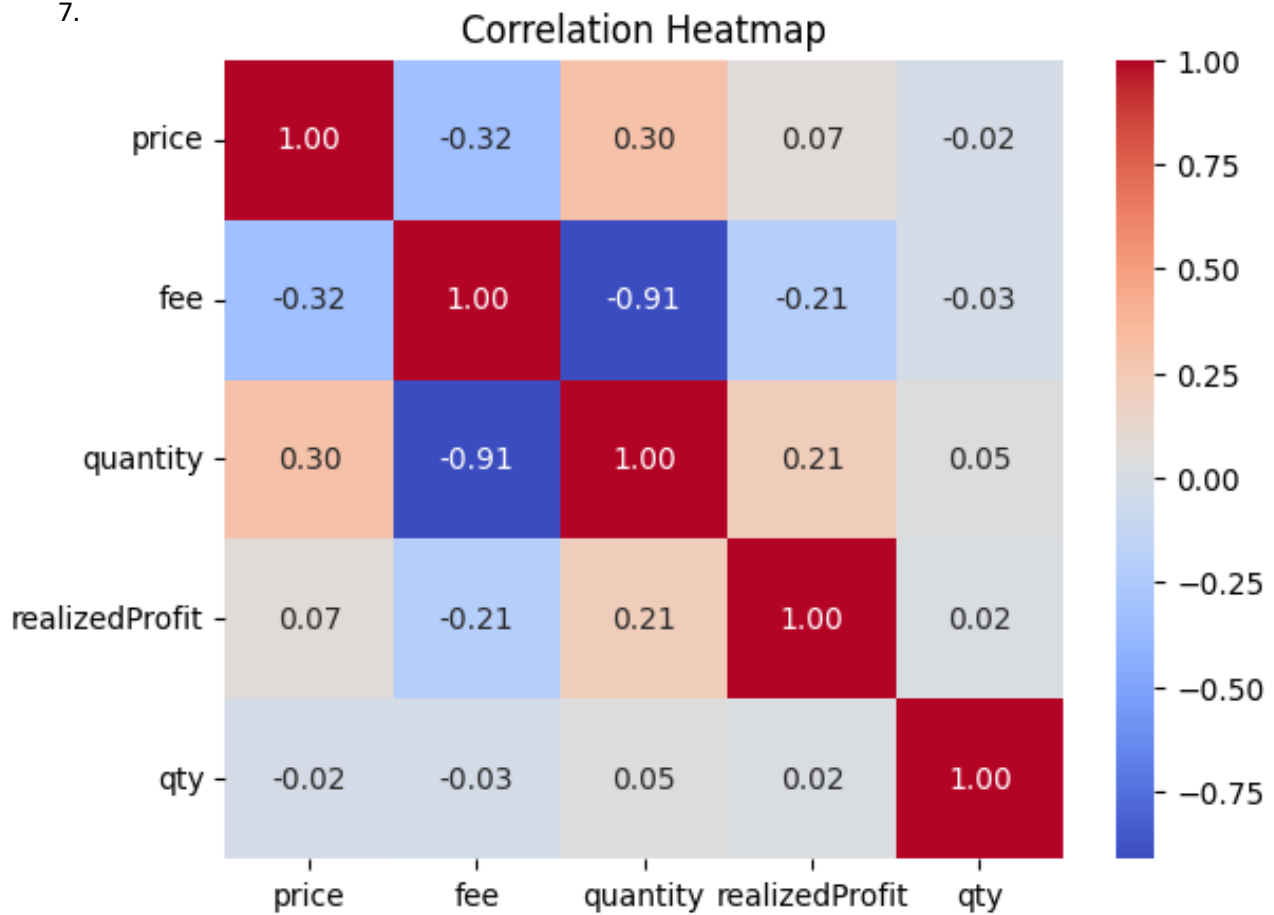


6.



## HISTORICAL TRADE DATA ANALYSIS

7.



There is co-relation between price and quantity.

There is co-relation between realised profit and quantity.



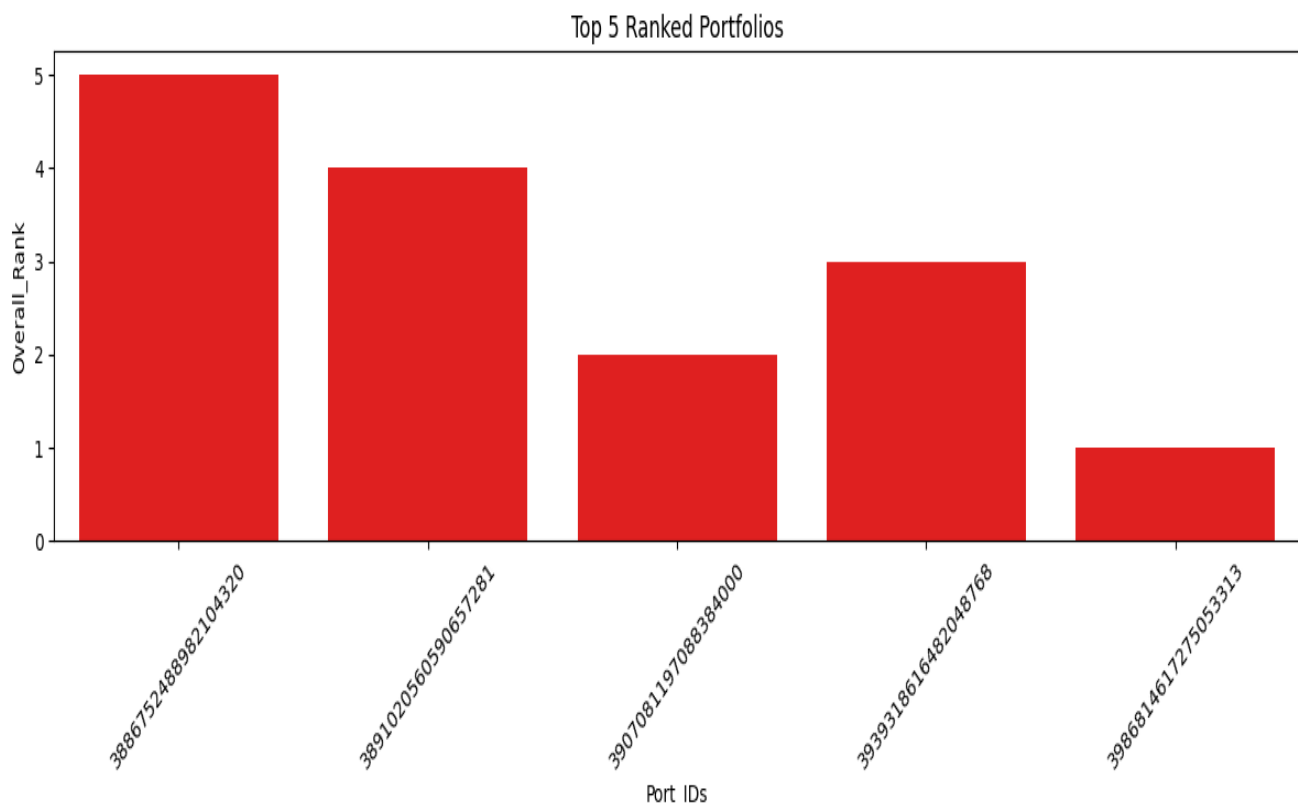
## HISTORICAL TRADE DATA ANALYSIS

### FINANCIAL METRICS AND RANKING:

Here's a breakdown of what the metrics represents:

1. Port\_IDs: Unique identifiers for each portfolio or position.
2. Total\_PnL: The total profit and loss for each Port\_ID.
3. ROI: Return on Investment for each portfolio. This shows how much return was earned relative to the investment.
4. Sharpe\_Ratio: Risk-adjusted return based on the portfolio's average return minus the risk-free rate divided by the standard deviation of returns. Higher values indicate better risk-adjusted returns.
5. MDD: Maximum Drawdown. This indicates the largest peak-to-trough drop in cumulative returns for the portfolio, shown as negative values.
6. Win\_Rate: The percentage of winning positions (profitable trades) over total positions.
7. Win\_Positions: The total number of profitable positions.
8. Total\_Positions: The total number of positions taken.

An overall rank for each account was computed by averaging the ranks across all selected metrics. The accounts were then sorted based on the overall rank in ascending order. Result Presentation: The ranked accounts were displayed alongside their metrics for easy comparison.





## HISTORICAL TRADE DATA ANALYSIS

---

### LIST OF 20 ACCOUNTS BASED ON RANKING:

Port_IDs	Overall_Rank
3986814617275053313	1.0
3907081197088384000	2.0
3939318616482048768	3.0
3891020560590657281	4.0
3886752488982104320	5.0
4029749871687083265	6.0
3999240873283311617	7.5
3977234346014419201	7.5
3826087012661391104	9.0
3944088772635000577	10.0
3943533600390906881	11.5
4035430878731345664	11.5
4028701921959171840	13.0
3910983669796782848	14.0
3956048468100538880	15.0
3936410995029308417	16.5
4022565861939831809	16.5
4017323550554338817	18.0
3956076827719377409	19.0
4017110277719148289	20.0

### METHODOLOGY FINDINGS AND ASSUMPTIONS:

#### 1. Methodology

Rank Calculation: For each metric, a rank was calculated: Total PnL, ROI, Sharpe Ratio, and Win Rate: Ranked in descending order (higher values are better). Maximum Drawdown: Ranked in ascending order (lower values are better). The ranks were computed using Pandas' built-in ranking functionality.

Overall Rank Calculation: An overall rank for each account was computed by averaging the ranks across all selected metrics. The accounts were then sorted based on the overall rank in ascending order. Result Presentation: The ranked accounts were displayed alongside their metrics for easy comparison.

#### 2. Findings

The analysis resulted in the following insights:

The top-performing accounts exhibit higher Total PnL, ROI, and Sharpe Ratios, reflecting effective management and profitable trades.

Accounts with lower Maximum Drawdown indicate better risk management and fewer significant losses.

Win Rate percentages highlight the proportion of profitable trades, suggesting effective trading strategies.

#### 3. Assumptions

The analysis is based on several assumptions:

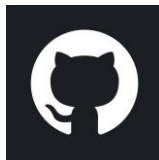
Metric Relevance: The selected metrics (Total PnL, ROI, Sharpe Ratio, MDD, Win Rate) are considered relevant indicators of account performance.

Equal Weighting: Each metric is assumed to have equal importance in determining the overall rank, which may not reflect the nuances of trading strategies.

Static Risk-Free Rate: The Sharpe Ratio calculation assumes a constant risk-free rate, which may not hold true in dynamic market conditions. 0.02 (or 2%)



: [www.linkedin.com/in/chetan-tiwari-](https://www.linkedin.com/in/chetan-tiwari-)



: <https://github.com/Chetan713205>

***-THANK YOU***