**Documentation**

Brief report summarizing your findings:

1. **Objective** : Analyze the dataset to calculate financial metrics for each account, rank them, and provide a top 20 list.
2. **Data Cleaning**: Handled missing data and parsed Trade\_History.

* Droped missing values as it is creating inconsistancing
* Parse the JSON-like Trade\_History Column. The Trade\_History column contains JSON-like objects. Converted these into a structured format for further analysis. After that extracted the trade details (e.g., symbol, price, quantity, etc.) from Trade\_History for each Port\_ID.

1. **Metrics**: metric calculated

* ROI (Return on Investment)
* PnL (Profit and Loss)
* Sharpe Ratio
* MDD (Maximum Drawdown)
* Win Rate
* Win Positions
* Total Positions

1. **Ranking Algorithm**: Mention how you weighted the metrics and ranked the accounts.

* Ranking Accounts - Combine all the metrics into one DataFrame and rank the accounts:

# Merge all metrics into one DataFrame

metrics\_df = roi\_df.merge(pnl\_df, on='Port\_ID') \

                   .merge(win\_rate\_df, on='Port\_ID') \

                   .merge(win\_positions\_df, on='Port\_ID') \

                   .merge(total\_positions\_df, on='Port\_ID') \

                   .merge(sharpe\_df, on='Port\_ID') \

                   .merge(mdd\_df, on='Port\_ID')

# Rank based on custom weighting (you can adjust weights)

metrics\_df['Score'] = (metrics\_df['ROI'] \* 0.2 +

                       metrics\_df['PnL'] \* 0.2 +

                       metrics\_df['Win\_Rate'] \* 0.2 +

                       metrics\_df['Sharpe\_Ratio'] \* 0.2 +

                       (1 - metrics\_df['MDD']) \* 0.2)

# Rank the accounts

metrics\_df['Rank'] = metrics\_df['Score'].rank(ascending=False)

# Sort by rank and select top 20

top\_20\_accounts = metrics\_df.sort\_values('Rank').head(20)

# Preview top 20 accounts

top\_20\_accounts[['Port\_ID', 'Score', 'Rank']]



Drive link : https://drive.google.com/drive/folders/14n4BclgqmZzrmvIyURqjTiWGu2CZy0Hl

Github link : https://github.com/SameerHussain128/Internshala\_DataScience\_Python\_Internship