REPORT

METHODOLOGY:

Data Preprocessing:

- Extracted trade details from a JSON-like nested structure for multiple portfolio IDs.
- Converted nested data into a tabular format with each trade represented as a row.
- Cleaned data by handling missing values and replacing infinite/large values.

Feature Engineering:

- Calculated critical metrics for each trade and portfolio:
    - **ROI (Return on Investment)**: (Realized Profit / Investment) × 100
    - **PnL (Profit and Loss)**: Cumulative realized profit for each portfolio.
    - **Winning Trades**: Flagged trades with positive realized profit.
    - **Sharpe Ratio**: Mean ROI / Standard Deviation of ROI for portfolio returns.
    - **Maximum Drawdown (MDD)**: Assessed peak-to-trough declines in portfolio performance.
    - **Total Positions**: (Total number of positions / trades of each portfolio id)
    - **Win Rate**: (Winning Trades / Total Positions) × 100

Machine Learning for Feature Importance:

- Attempted to use RandomForestRegressor to infer feature importance for the scoring model.

Ranking Algorithm:

- Developed a scoring system using a weighted combination of ROI, PnL, Sharpe Ratio, MDD, Win Rate, Win Positions, Total Positions

Findings:

Portfolio Insights:

- Portfolios with higher PnL scored better.
- Portfolios with low Sharpe ratios exhibited high volatility, indicating higher risk**.**

Trade Patterns:

- Winning trades contributed disproportionately to overall PnL.
- High investment volumes didn't always correlate with better performance due to poor risk management.

Assumptions:

Data Integrity:

- Assumed trade data provided was accurate and complete, with no missing trades for any portfolio.

Market Independence:

- Considered individual portfolio performance independent of broader market conditions.

Risk-Free Rate:

- Assumed a constant risk-free rate of 0% for Sharpe ratio calculations due to lack of contextual benchmarks.

Scoring Weights:

- Initial weights for the ranking system were heuristic and subject to optimization with more data.