SOLUTION DESIGN DOCUMENT

MScFE Capstone Project

Acsadam Rex Eugenio – rexeugenio77@gmail.com

02/25/2023

1. Repository

Link to GitHub repository: https://github.com/rexlaboratory/adr-piotroski-f-score

2. Data Sources

2.1 ADR Stock Universe. The stock universe used in this research is from adr.com, a website owned by J.P. Morgan that houses information on depositary receipts including ADR stocks [2]. From this data, the list of ADR stocks and their ticker names are identified.

2.2 Stock Price and Financial Data. The ADR stock prices and financial information are retrieved from Yahoo Finance data.

3. Data Processing

3.1 Python Libraries. The following Python libraries are used in data processing.

1. Yfinance package is used to fetch the stock prices. Stock prices are used in calculating returns.
2. Yahoo Query package is used to fetch the financial information. Financial information is used in calculating the F-score.
3. Concurrent Futures package is used in parallel processing to aid in fetching the financial data. This helped gain a slight improvement in run time.
4. Pandas package is used in data merging, manipulation, transformation, and presenting the major results.

3.2 Functions. The following are the primary functions used in the Python code for analysis.

1. Calculate *Piotroski F-score* based on the financial data, using a nine-point scorecard [3]. The F-score is the basis for classifying stocks as 'expected winner' or 'expected loser'.
2. Calculate *equal-weighted returns* based on stock prices. The portfolio return rate is compared against the benchmark.
3. Calculate *precision* based on the number of 'actual winners' out of 'expected winners' and the number of 'actual losers' out of 'expected losers'.

4. Performance Metrics

4.1 Portfolio Returns and Benchmark. A benchmark index is created using equal-weighted stock returns by market group. The equal-weighted returns of high-F-score stocks, equal-weighted returns of low-F-score stocks, and the benchmark are then summarized and compared.

4.2 Accuracy Metric. The accuracy metrics are set up based on the definition of *precision*, a measure of specificity also known as *positive predictive value (PPV)* [1].

5. Pseudocode

1. Identify the list of foreign ADRs and extract it.
2. Retrieve the stock prices and financial information for each stock.
3. Transform data and compute Piotroski F-Scores.
   a. Transform data to prepare for F-score calculation.
   b. Calculate F-scores.
4. Measure the performance of the Piotroski F-score classifier in predicting 'winner' and 'loser' stocks.
   a. *Benchmark*. Set up an ADR stock index by market group (i.e., benchmark).
   b. *Return-based metrics*. Compare the return rates of High-F-score ADR stocks against the Benchmark and Low-F-score stocks.
   c. *Precision metrics*. Compute the precision of the F-score in predicting 'winner' stocks; compute the precision of the F-score in predicting 'loser' stocks.

Note: The visualization of the solution design and the data elements are provided in *Figure A1* and *Figure A2*, respectively, in the Appendix section.
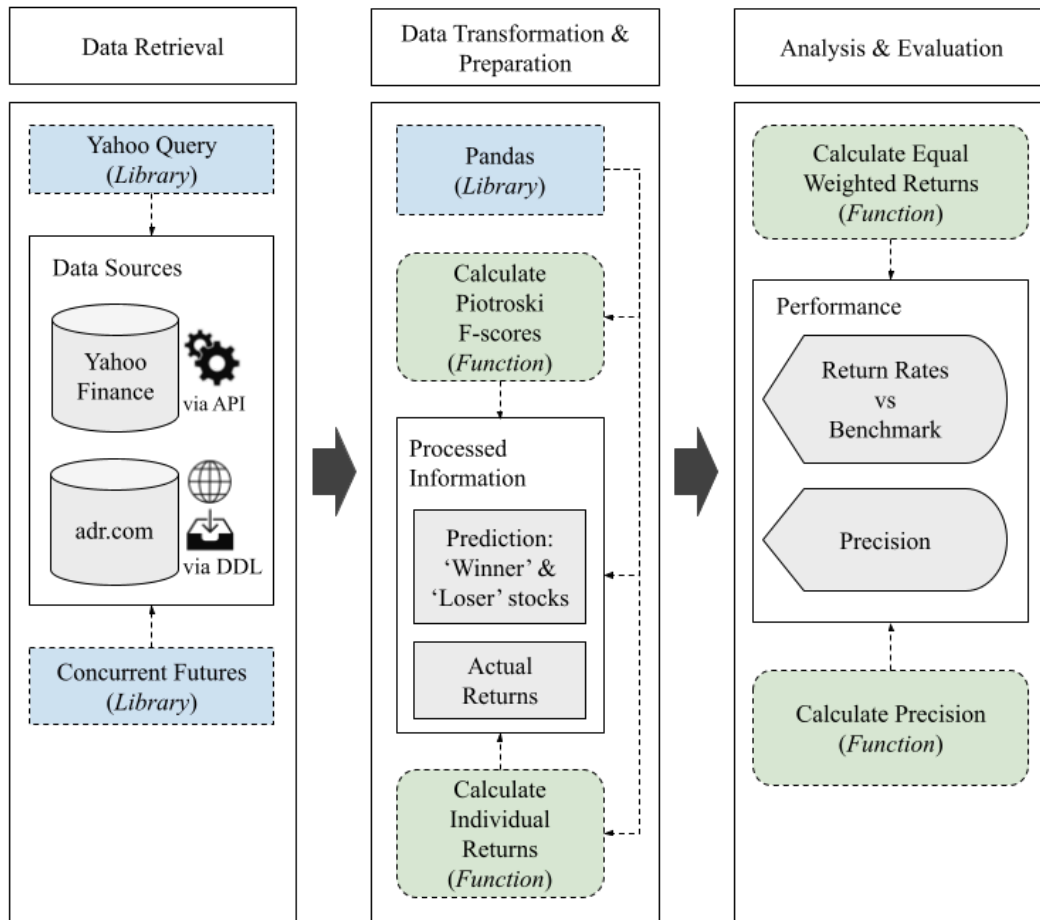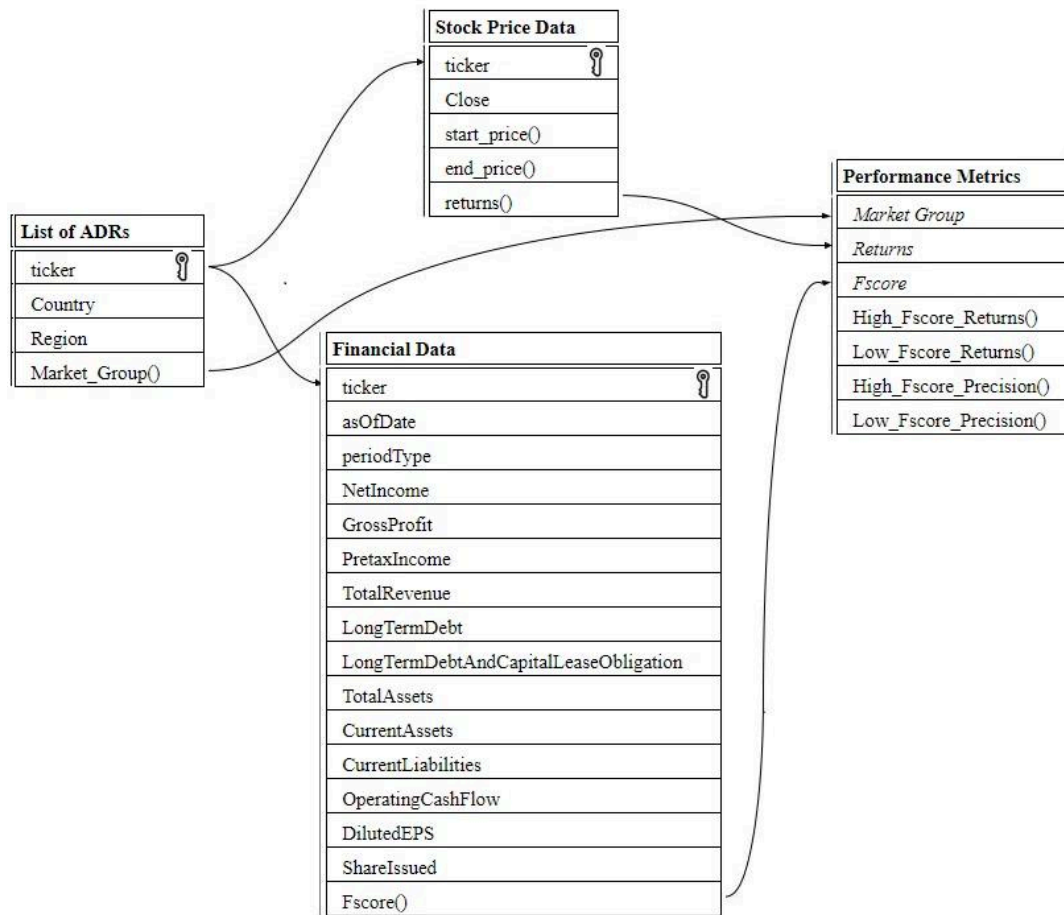
Appendix

Figure A1. Overview of Solution Design (Methodology)

Figure A2. Data Elements

Bibliography

[1] Irizarry, R. *Advanced Data Science - Statistics and Prediction Algorithms Through Case Studies*. Harvard University Dana-Farber Cancer Institute, 2023. Retrieved from https://rafalab.dfci.harvard.edu/dsbook-part-2/. Accessed 06 February 2023.

[2] J.P. Morgan Depositary Receipts. Depositary Receipt (DR) Universe. J.P. Morgan Chase & Co., 2024. Retrieved from https://adr.com/dr/drdirectory/drUniverse. Accessed 11 January 2024.

[3] Piotroski, J. D. *Value investing: The Use of Historical Financial Statement Information to Separate Winners from Losers*. Journal of Accounting Research, 2000. Retrieved from https://gm10b7le2-mp01-y-https-www-proquest-com.proxy.lirn.net/scholarly-journals/value-investing-use-historical-financial/docview/206723328/se-2. Accessed 10 January 2024.