ST446: DISTRIBUTED COMPUTING FOR BIG DATA - DATASET

ST446 GROUP PROJECT – WALL STREET MAVERICKS TEAM

50714   49775   49663   49872

TOPIC: "Dynamic Classification of S&P 500 Stocks Using Distributed Computing and Fundamental Ratios"

LINK : https://www.kaggle.com/datasets/laxmansudhan25002/st446-macroeconomic-and-fundamental-data/data

Note for Project Review:

1. We have created a Kaggle account and uploaded both datasets [macro-economic and fundamental dataset] there. The dataset is

titled "ST446 MACROECONOMIC AND FUNDAMENTAL DATA" and can be accessed via the shared Kaggle link. Please note that the dataset contains 146,587 entries, so it may take some time to download due to its large size.

2. We kindly request you to review the following Notebook file located in the "DATA EXTRACTION CODE" folder submitted within our project repository:

- ST446_DATA_EXTRACTION_CODE.ipynb

This file document our process of data extraction using APIs and various data engineering techniques.