Possible Dataset

Option 1 (<https://www.kaggle.com/datasets/cnic92/200-financial-indicators-of-us-stocks-20142018>)

* Stocks from 2014 to 2018
* 200 + features mostly found in financial reports
* 5 tables for each year (probably want to aggregate)
* Target column already made if the price stock when up over the year consider good stock 1 else 0 (binary classification)

Option 1 (<https://www.kaggle.com/datasets/pierrelouisdanieau/financial-data-sp500-companies>)

* S&P 500 4 annual reports per ticker dataframe size (2013,18)
* firm : company name
* Ticker : company ticker (the symbol)
* Research Development
* Income Before Tax
* Net Income
* Selling General
* Administrative
* Gross Profit
* Ebit
* Operating Income
* Interest Expense
* Income Tax Expense
* Total Revenue
* Total Operating Expenses
* Cost Of Revenue
* Total Other Income Expense Net
* Net Income From Continuing Ops
* Net Income Applicable To Common Shares

Option 2 (https://www.kaggle.com/datasets/camnugent/sandp500/code)

* S&P 500 daily information on stocks dataframe size (619041,7)
  + Date
  + Open
  + High
  + Low
  + Close
  + Volume
  + Ticker’s Name

<https://www.kaggle.com/datasets/hanseopark/sp-500-stocks-value-with-financial-statement>

* Lots of information but information is spread-out trough different csv file and sometimes represented in 1 column
* Price over 10 years
* Price open /close ..
* Current price
* Summary Statement by ticker
* Balance sheet
* Income statement
* Cash flow