# Yahoo Stock Data Scraper - Project Documentation

## 1. Introduction

This project is a Python-based web scraper designed to extract and analyze the most active stock data from Yahoo Finance. The scraper uses Selenium to navigate through web pages, extract stock information, clean the data, and export it to an Excel file. The script has been compiled into a standalone executable (.exe) file for ease of use.

## 2. Features

- Navigates Yahoo Finance to fetch the most active stock listings.

- Scrapes relevant data like stock symbol, name, price, change, change percentage, volume, market cap, and PE ratio.

- Cleans and formats the data using pandas.

- Automatically saves the data to an Excel file in the Downloads folder.

- Packaged as a .exe file so it can be run without needing Python installed.

## 3. How It Works

1. Opens the Yahoo Finance website using Selenium WebDriver.

2. Navigates through the 'Markets' menu to the 'Most Active Stocks' section.

3. Iteratively extracts data from all pages of stock listings.

4. Cleans and converts the raw strings into appropriate numerical formats using pandas.

5. Saves the cleaned data as an Excel file named 'yahoo\_finance-stocks.xlsx' in the Downloads folder.

## 4. Technologies and Libraries Used

- Python 3.x

- Selenium: for web automation and scraping

- pandas and numpy: for data cleaning and processing

- openpyxl: for writing Excel files

- pyinstaller: for creating a standalone .exe file

## 6. Usage Instructions

1. Double-click on 'Executable\_Script.exe'.

2. A Chrome browser window will open and navigate to Yahoo Finance.

3. The script will automatically extract the stock data and save the Excel file.

4. Check your Downloads folder for 'yahoo\_finance-stocks.xlsx'.