stock

September 21, 2024

[1]:
```
%pip install yfinance pandas numpy
```

```
Requirement already satisfied: yfinance in
c:\users\liuch\appdata\local\programs\python\python312\lib\site-packages
(0.2.43)
Requirement already satisfied: pandas in
c:\users\liuch\appdata\local\programs\python\python312\lib\site-packages (2.2.2)
Requirement already satisfied: numpy in
c:\users\liuch\appdata\local\programs\python\python312\lib\site-packages (2.1.0)
Requirement already satisfied: requests>=2.31 in
c:\users\liuch\appdata\local\programs\python\python312\lib\site-packages (from
yfinance) (2.32.3)
Requirement already satisfied: multitasking>=0.0.7 in
c:\users\liuch\appdata\local\programs\python\python312\lib\site-packages (from
yfinance) (0.0.11)
Requirement already satisfied: lxml>=4.9.1 in
c:\users\liuch\appdata\local\programs\python\python312\lib\site-packages (from
yfinance) (5.3.0)
Requirement already satisfied: platformdirs>=2.0.0 in
c:\users\liuch\appdata\roaming\python\python312\site-packages (from yfinance)
(4.2.2)
Requirement already satisfied: pytz>=2022.5 in
c:\users\liuch\appdata\local\programs\python\python312\lib\site-packages (from
yfinance) (2024.1)
Requirement already satisfied: frozendict>=2.3.4 in
c:\users\liuch\appdata\local\programs\python\python312\lib\site-packages (from
yfinance) (2.4.4)
Requirement already satisfied: peewee>=3.16.2 in
c:\users\liuch\appdata\local\programs\python\python312\lib\site-packages (from
yfinance) (3.17.6)
Requirement already satisfied: beautifulsoup4>=4.11.1 in
c:\users\liuch\appdata\local\programs\python\python312\lib\site-packages (from
yfinance) (4.12.3)
Requirement already satisfied: html5lib>=1.1 in
c:\users\liuch\appdata\local\programs\python\python312\lib\site-packages (from
yfinance) (1.1)
Requirement already satisfied: python-dateutil>=2.8.2 in
c:\users\liuch\appdata\roaming\python\python312\site-packages (from pandas)
```

```
(2.9.0.post0)
Requirement already satisfied: tzdata>=2022.7 in
c:\users\liuch\appdata\local\programs\python\python312\lib\site-packages (from
pandas) (2024.1)
Requirement already satisfied: soupsieve>1.2 in
c:\users\liuch\appdata\local\programs\python\python312\lib\site-packages (from
beautifulsoup4>=4.11.1->yfinance) (2.6)
Requirement already satisfied: six>=1.9 in
c:\users\liuch\appdata\roaming\python\python312\site-packages (from
html5lib>=1.1->yfinance) (1.16.0)
Requirement already satisfied: webencodings in
c:\users\liuch\appdata\local\programs\python\python312\lib\site-packages (from
html5lib>=1.1->yfinance) (0.5.1)
Requirement already satisfied: charset-normalizer<4,>=2 in
c:\users\liuch\appdata\local\programs\python\python312\lib\site-packages (from
requests>=2.31->yfinance) (3.3.2)
Requirement already satisfied: idna<4,>=2.5 in
c:\users\liuch\appdata\local\programs\python\python312\lib\site-packages (from
requests>=2.31->yfinance) (3.8)
Requirement already satisfied: urllib3<3,>=1.21.1 in
c:\users\liuch\appdata\local\programs\python\python312\lib\site-packages (from
requests>=2.31->yfinance) (2.2.2)
Requirement already satisfied: certifi>=2017.4.17 in
c:\users\liuch\appdata\local\programs\python\python312\lib\site-packages (from
requests>=2.31->yfinance) (2024.8.30)
Note: you may need to restart the kernel to use updated packages.
```

[2]:
```
import yfinance as yf
import pandas as pd
import numpy as np
```

[3]:
```
%pip install matplotlib seaborn
```

```
^C
Note: you may need to restart the kernel to use updated packages.
Requirement already satisfied: matplotlib in
c:\users\liuch\appdata\local\programs\python\python312\lib\site-packages (3.9.2)
Requirement already satisfied: seaborn in
c:\users\liuch\appdata\local\programs\python\python312\lib\site-packages
(0.13.2)
Requirement already satisfied: contourpy>=1.0.1 in
c:\users\liuch\appdata\local\programs\python\python312\lib\site-packages (from
matplotlib) (1.3.0)
Requirement already satisfied: cycler>=0.10 in
c:\users\liuch\appdata\local\programs\python\python312\lib\site-packages (from
matplotlib) (0.12.1)
Requirement already satisfied: fonttools>=4.22.0 in
c:\users\liuch\appdata\local\programs\python\python312\lib\site-packages (from
```

```
matplotlib) (4.53.1)
Requirement already satisfied: kiwisolver>=1.3.1 in
c:\users\liuch\appdata\local\programs\python\python312\lib\site-packages (from
matplotlib) (1.4.5)
Requirement already satisfied: numpy>=1.23 in
c:\users\liuch\appdata\local\programs\python\python312\lib\site-packages (from
matplotlib) (2.1.0)
Requirement already satisfied: packaging>=20.0 in
c:\users\liuch\appdata\roaming\python\python312\site-packages (from matplotlib)
(24.1)
Requirement already satisfied: pillow>=8 in
c:\users\liuch\appdata\local\programs\python\python312\lib\site-packages (from
matplotlib) (10.4.0)
Requirement already satisfied: pyparsing>=2.3.1 in
c:\users\liuch\appdata\local\programs\python\python312\lib\site-packages (from
matplotlib) (3.1.4)
Requirement already satisfied: python-dateutil>=2.7 in
c:\users\liuch\appdata\roaming\python\python312\site-packages (from matplotlib)
(2.9.0.post0)
Requirement already satisfied: pandas>=1.2 in
c:\users\liuch\appdata\local\programs\python\python312\lib\site-packages (from
seaborn) (2.2.2)
Requirement already satisfied: pytz>=2020.1 in
c:\users\liuch\appdata\local\programs\python\python312\lib\site-packages (from
pandas>=1.2->seaborn) (2024.1)
Requirement already satisfied: tzdata>=2022.7 in
c:\users\liuch\appdata\local\programs\python\python312\lib\site-packages (from
pandas>=1.2->seaborn) (2024.1)
Requirement already satisfied: six>=1.5 in
c:\users\liuch\appdata\roaming\python\python312\site-packages (from python-
dateutil>=2.7->matplotlib) (1.16.0)
```

[9]:
```
%pip install alpha_vantage
```

```
Requirement already satisfied: alpha_vantage in
c:\users\liuch\appdata\local\programs\python\python312\lib\site-packages (3.0.0)
Requirement already satisfied: aiohttp in
c:\users\liuch\appdata\local\programs\python\python312\lib\site-packages (from
alpha_vantage) (3.10.5)
Requirement already satisfied: requests in
c:\users\liuch\appdata\local\programs\python\python312\lib\site-packages (from
alpha_vantage) (2.32.3)
Requirement already satisfied: aiohappyeyeballs>=2.3.0 in
c:\users\liuch\appdata\local\programs\python\python312\lib\site-packages (from
aiohttp->alpha_vantage) (2.4.0)
Requirement already satisfied: aiosignal>=1.1.2 in
c:\users\liuch\appdata\local\programs\python\python312\lib\site-packages (from
aiohttp->alpha_vantage) (1.3.1)
```

```
Requirement already satisfied: attrs>=17.3.0 in
c:\users\liuch\appdata\local\programs\python\python312\lib\site-packages (from
aiohttp->alpha_vantage) (24.2.0)
Requirement already satisfied: frozenlist>=1.1.1 in
c:\users\liuch\appdata\local\programs\python\python312\lib\site-packages (from
aiohttp->alpha_vantage) (1.4.1)
Requirement already satisfied: multidict<7.0,>=4.5 in
c:\users\liuch\appdata\local\programs\python\python312\lib\site-packages (from
aiohttp->alpha_vantage) (6.0.5)
Requirement already satisfied: yarl<2.0,>=1.0 in
c:\users\liuch\appdata\local\programs\python\python312\lib\site-packages (from
aiohttp->alpha_vantage) (1.9.11)
Requirement already satisfied: charset-normalizer<4,>=2 in
c:\users\liuch\appdata\local\programs\python\python312\lib\site-packages (from
requests->alpha_vantage) (3.3.2)
Requirement already satisfied: idna<4,>=2.5 in
c:\users\liuch\appdata\local\programs\python\python312\lib\site-packages (from
requests->alpha_vantage) (3.8)
Requirement already satisfied: urllib3<3,>=1.21.1 in
c:\users\liuch\appdata\local\programs\python\python312\lib\site-packages (from
requests->alpha_vantage) (2.2.2)
Requirement already satisfied: certifi>=2017.4.17 in
c:\users\liuch\appdata\local\programs\python\python312\lib\site-packages (from
requests->alpha_vantage) (2024.8.30)
Note: you may need to restart the kernel to use updated packages.
```

[3]:
```
import matplotlib.pyplot as plt
import seaborn as sns
from alpha_vantage.timeseries import TimeSeries
import random
import time
```

[4]:
```
large_stock_pool = [
    'AAPL', 'MSFT', 'GOOGL', 'AMZN', 'TSLA', 'META', 'NFLX', 'NVDA', 'BABA',␣
 ↪'ORCL', 'IBM', 'INTC', 'AMD', 'ADBE',
    'CSCO', 'QCOM', 'AVGO', 'TXN', 'SAP', 'CRM', 'V', 'MA', 'PYPL', 'SQ',␣
 ↪'SHOP', 'SPOT', 'UBER', 'LYFT', 'ZM',
    'SNAP', 'TWTR', 'PINS', 'ETSY', 'BIDU', 'JD', 'BILI', 'DOCU', 'ROKU',␣
 ↪'TME', 'PDD', 'F', 'GM', 'TM', 'HMC',
    'NIO', 'XPEV', 'LI', 'TSM', 'ASML', 'INTU', 'BA', 'GE', 'CAT', 'UPS',␣
 ↪'FDX', 'KO', 'PEP', 'PG', 'PM', 'MO',
    'UL', 'JNJ', 'PFE', 'MRK', 'ABBV', 'LLY', 'GILD', 'XOM', 'CVX', 'BP',␣
 ↪'RDS-A', 'DIS', 'NKE', 'SBUX', 'MCD',
    'WMT', 'COST', 'TGT', 'HD', 'LOW', 'BBY', 'UNH', 'AET', 'CNC', 'ANTM',␣
 ↪'HUM', 'CI', 'MRNA', 'NVAX', 'REGN',
    'VRTX', 'ISRG', 'ILMN', 'CRSP', 'EDIT', 'NTLA', 'MRVL', 'ADI', 'NXPI',␣
 ↪'MCHP', 'SWKS', 'QRVO', 'STX', 'WDC',
```

```
    'GLW', 'TEL', 'TXN', 'INTC', 'AMD', 'NVDA', 'MU', 'KLAC', 'LRCX', 'AMAT',␣
 ↪'ASML', 'RIVN', 'LCID', 'FUBO',
    'CVS', 'TGT', 'NVO', 'CHTR', 'NEE', 'SRE', 'DUK', 'SO', 'EXC', 'PEG',␣
 ↪'WEC', 'ED', 'AEP', 'D', 'ES', 'FE',
    'EIX', 'PCG', 'AWK', 'WTRG', 'CNP', 'NI', 'PPL', 'OKE', 'KMI', 'WMB',␣
 ↪'ENB', 'TRP', 'ET', 'MMP', 'PSX',
    'VLO', 'MPC', 'HAL', 'SLB', 'BKR', 'WFC', 'JPM', 'BAC', 'GS', 'MS', 'C',␣
 ↪'BLK', 'SCHW', 'BRK-B', 'SPGI',
    'MO', 'TMO', 'SYK', 'MDT', 'HON', 'AMAT', 'KLAC', 'LRCX', 'INTC', 'MU',␣
 ↪'TSM', 'NXPI', 'TXN', 'ADI', 'STM',
    'GOOG', 'AMZN', 'BA', 'NOC', 'LMT', 'RTX', 'HON', 'GD', 'HII', 'TXT',␣
 ↪'LHX', 'FLIR', 'MTCH', 'RCL', 'CCL',
    'NCLH', 'ALK', 'DAL', 'LUV', 'UAL', 'AAL', 'V', 'MA', 'DFS', 'AXP', 'BAC',␣
 ↪'JPM', 'WFC', 'C', 'GS', 'MS',
    'BLK', 'SCHW', 'BRK-B', 'SPGI', 'TGT', 'BBY', 'ROST', 'TJX', 'WMT', 'COST',␣
 ↪'HD', 'LOW', 'NKE', 'LULU', 'TIF',
]

#
def get_stock_data(stocks, start_date, end_date):
    stock_data = {}
    valid_stocks = []

    for stock in stocks:
        data = yf.download(stock, start=start_date, end=end_date)
        if data.shape[0] > 0 and data['Adj Close'].notna().all():  #
            stock_data[stock] = data['Adj Close']
            valid_stocks.append(stock)  #

    return pd.DataFrame(stock_data), valid_stocks

#       0.3
def find_stock_set(correlation_matrix, threshold=0.3, min_stock_count=20):
    def is_valid_selection(selected_stocks, new_stock, correlation_matrix):
        for stock in selected_stocks:
            if abs(correlation_matrix.loc[stock, new_stock]) >= threshold:
                return False
        return True

    selected_stocks = []
    for stock in correlation_matrix.columns:
        if is_valid_selection(selected_stocks, stock, correlation_matrix):
            selected_stocks.append(stock)
        if len(selected_stocks) >= min_stock_count:
            break
```

```
    return selected_stocks

#          0.3
def select_stocks_from_large_pool(large_stock_pool, min_stock_count=20):
    #
    start_date = '2013-01-01'
    end_date = '2023-12-31'
    combined_data, valid_stocks = get_stock_data(large_stock_pool, start_date,␣
 ↪end_date)

    #
    print(f"     {len(valid_stocks)}         ")

    #
    correlation_matrix = combined_data.corr()

    #
    selected_stocks = find_stock_set(correlation_matrix, threshold=0.3,␣
 ↪min_stock_count=min_stock_count)

    #
    print(f"     {len(selected_stocks)}")
    print("    ", selected_stocks)

    return selected_stocks

#         20      0.3
selected_stocks = select_stocks_from_large_pool(large_stock_pool,␣
 ↪min_stock_count=20)
```

```
[**********************100%**********************]  1 of 1 completed
[**********************100%**********************]  1 of 1 completed
[**********************100%**********************]  1 of 1 completed
[**********************100%**********************]  1 of 1 completed
[**********************100%**********************]  1 of 1 completed
[**********************100%**********************]  1 of 1 completed
[**********************100%**********************]  1 of 1 completed
[**********************100%**********************]  1 of 1 completed
[**********************100%**********************]  1 of 1 completed
[**********************100%**********************]  1 of 1 completed
[**********************100%**********************]  1 of 1 completed
[**********************100%**********************]  1 of 1 completed
[**********************100%**********************]  1 of 1 completed
[**********************100%**********************]  1 of 1 completed
[**********************100%**********************]  1 of 1 completed
[**********************100%**********************]  1 of 1 completed
[**********************100%**********************]  1 of 1 completed
[**********************100%**********************]  1 of 1 completed
```

```
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed

1 Failed download:
['TWTR']: YFTzMissingError('$%ticker%: possibly delisted; no timezone found')
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
```

```
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed

1 Failed download:
['RDS-A']: YFTzMissingError('$%ticker%: possibly delisted; no timezone found')
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed

1 Failed download:
['ANTM']: YFTzMissingError('$%ticker%: possibly delisted; no timezone found')
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
```

```
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed

1 Failed download:
['MMP']: YFTzMissingError('$%ticker%: possibly delisted; no timezone found')
[*********************100%***********************]  1 of 1 completed
```

```
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed

1 Failed download:
['FLIR']: YFTzMissingError('$%ticker%: possibly delisted; no timezone found')
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
```

```
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  1 of 1 completed

1 Failed download:
['TIF']: YFTzMissingError('$%ticker%: possibly delisted; no timezone found')
```

```
219
5
['AAPL', 'BABA', 'UBER', 'LCID', 'UAL']
```

[12]:

```
%pip install PyPortfolioOpt
```

```
Collecting PyPortfolioOpt
  Downloading pyportfolioopt-1.5.5-py3-none-any.whl.metadata (23 kB)
Collecting cvxpy<2.0.0,>=1.1.19 (from PyPortfolioOpt)
  Downloading cvxpy-1.5.3-cp312-cp312-win_amd64.whl.metadata (9.0 kB)
Collecting numpy<2.0.0,>=1.22.4 (from PyPortfolioOpt)
  Downloading numpy-1.26.4-cp312-cp312-win_amd64.whl.metadata (61 kB)
Requirement already satisfied: pandas>=0.19 in
```

```
c:\users\liuch\appdata\local\programs\python\python312\lib\site-packages (from
PyPortfolioOpt) (2.2.2)
Requirement already satisfied: scipy<2.0,>=1.3 in
c:\users\liuch\appdata\local\programs\python\python312\lib\site-packages (from
PyPortfolioOpt) (1.14.1)
Collecting osqp>=0.6.2 (from cvxpy<2.0.0,>=1.1.19->PyPortfolioOpt)
  Downloading osqp-0.6.7.post1-cp312-cp312-win_amd64.whl.metadata (2.0 kB)
Collecting ecos>=2 (from cvxpy<2.0.0,>=1.1.19->PyPortfolioOpt)
  Downloading ecos-2.0.14-cp312-cp312-win_amd64.whl.metadata (8.2 kB)
Collecting clarabel>=0.5.0 (from cvxpy<2.0.0,>=1.1.19->PyPortfolioOpt)
  Downloading clarabel-0.9.0-cp37-abi3-win_amd64.whl.metadata (4.8 kB)
Collecting scs>=3.2.4.post1 (from cvxpy<2.0.0,>=1.1.19->PyPortfolioOpt)
  Downloading scs-3.2.7-cp312-cp312-win_amd64.whl.metadata (2.1 kB)
Requirement already satisfied: python-dateutil>=2.8.2 in
c:\users\liuch\appdata\roaming\python\python312\site-packages (from
pandas>=0.19->PyPortfolioOpt) (2.9.0.post0)
Requirement already satisfied: pytz>=2020.1 in
c:\users\liuch\appdata\local\programs\python\python312\lib\site-packages (from
pandas>=0.19->PyPortfolioOpt) (2024.1)
Requirement already satisfied: tzdata>=2022.7 in
c:\users\liuch\appdata\local\programs\python\python312\lib\site-packages (from
pandas>=0.19->PyPortfolioOpt) (2024.1)
Collecting qdldl (from osqp>=0.6.2->cvxpy<2.0.0,>=1.1.19->PyPortfolioOpt)
  Downloading qdldl-0.1.7.post4-cp312-cp312-win_amd64.whl.metadata (1.8 kB)
Requirement already satisfied: six>=1.5 in
c:\users\liuch\appdata\roaming\python\python312\site-packages (from python-
dateutil>=2.8.2->pandas>=0.19->PyPortfolioOpt) (1.16.0)
Downloading pyportfolioopt-1.5.5-py3-none-any.whl (61 kB)
Downloading cvxpy-1.5.3-cp312-cp312-win_amd64.whl (1.1 MB)
   ---------------------------------------- 0.0/1.1 MB ? eta -:--:--
   ---------------------------------------- 1.1/1.1 MB 10.1 MB/s eta 0:00:00
Downloading numpy-1.26.4-cp312-cp312-win_amd64.whl (15.5 MB)
   ---------------------------------------- 0.0/15.5 MB ? eta -:--:--
   ----- ---------------------------------- 2.1/15.5 MB 10.7 MB/s eta 0:00:02
   ---------- ----------------------------- 4.5/15.5 MB 11.2 MB/s eta 0:00:01
   --------------- ------------------------ 6.6/15.5 MB 10.9 MB/s eta 0:00:01
   --------------------- ------------------ 8.9/15.5 MB 10.7 MB/s eta 0:00:01
   ---------------------------- ----------- 11.0/15.5 MB 10.6 MB/s eta 0:00:01
   ---------------------------------- ------ 13.1/15.5 MB 10.5 MB/s eta 0:00:01
   ---------------------------------------  15.5/15.5 MB 10.5 MB/s eta 0:00:01
   ---------------------------------------- 15.5/15.5 MB 10.2 MB/s eta 0:00:00
Downloading clarabel-0.9.0-cp37-abi3-win_amd64.whl (736 kB)
   ---------------------------------------- 0.0/736.4 kB ? eta -:--:--
   ---------------------------------------- 736.4/736.4 kB 10.1 MB/s eta 0:00:00
Downloading ecos-2.0.14-cp312-cp312-win_amd64.whl (72 kB)
Downloading osqp-0.6.7.post1-cp312-cp312-win_amd64.whl (293 kB)
Downloading scs-3.2.7-cp312-cp312-win_amd64.whl (8.4 MB)
   ---------------------------------------- 0.0/8.4 MB ? eta -:--:--
```

```
    ------------ --------------------------- 2.6/8.4 MB 13.7 MB/s eta 0:00:01
    ---------------------- ------------------ 4.7/8.4 MB 11.9 MB/s eta 0:00:01
    -------------------------------- ------- 6.8/8.4 MB 11.1 MB/s eta 0:00:01
    --------------------------------------- 8.4/8.4 MB 10.5 MB/s eta 0:00:00
Downloading qdldl-0.1.7.post4-cp312-cp312-win_amd64.whl (87 kB)
Installing collected packages: numpy, scs, qdldl, ecos, clarabel, osqp, cvxpy,
PyPortfolioOpt
  Attempting uninstall: numpy
    Found existing installation: numpy 2.1.0
    Uninstalling numpy-2.1.0:
      Successfully uninstalled numpy-2.1.0
Successfully installed PyPortfolioOpt-1.5.5 clarabel-0.9.0 cvxpy-1.5.3
ecos-2.0.14 numpy-1.26.4 osqp-0.6.7.post1 qdldl-0.1.7.post4 scs-3.2.7
Note: you may need to restart the kernel to use updated packages.
```

```
  WARNING: Failed to remove contents in a temporary directory
'C:\Users\Liuch\AppData\Local\Programs\Python\Python312\Lib\site-
packages\~umpy.libs'.
  You can safely remove it manually.
  WARNING: Failed to remove contents in a temporary directory
'C:\Users\Liuch\AppData\Local\Programs\Python\Python312\Lib\site-
packages\~umpy'.
  You can safely remove it manually.
```

[8]:

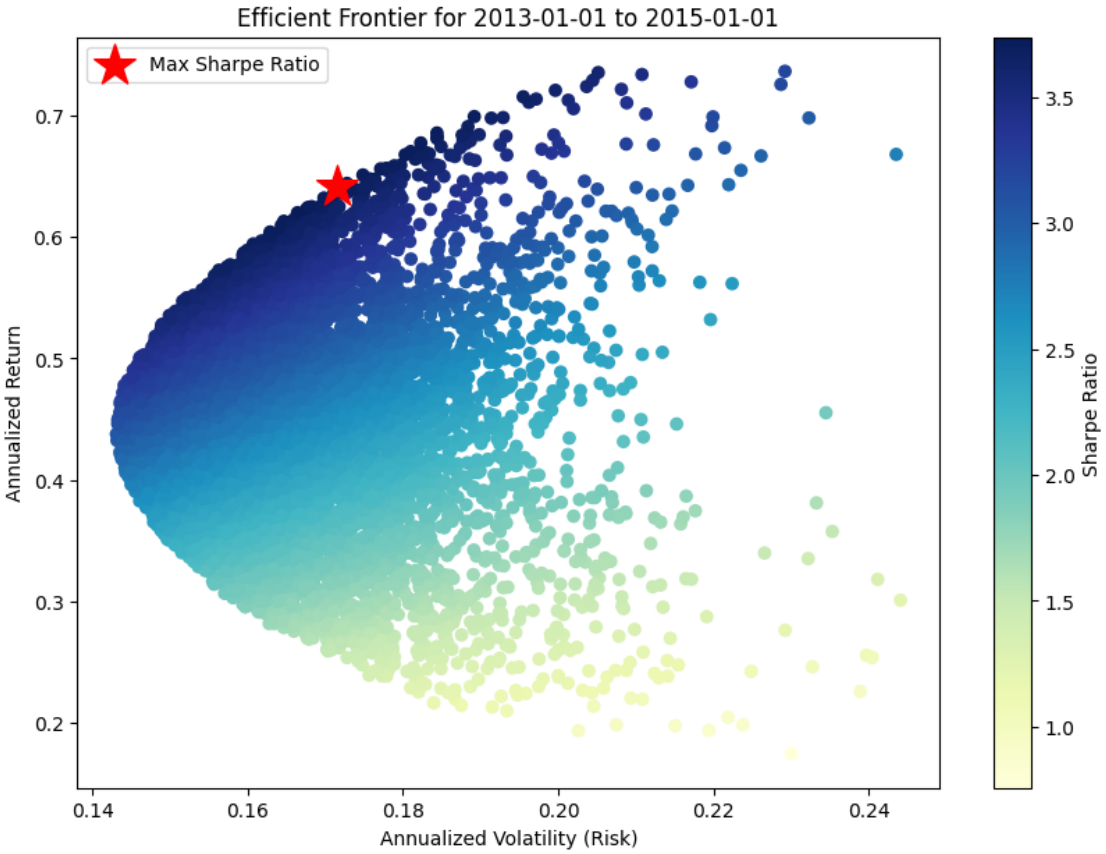
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from scipy.optimize import minimize
a=[]
# Portfolio optimization function
def portfolio_annualized_performance(weights, mean_returns, cov_matrix):
    returns = np.sum(mean_returns * weights) * 252  # Annualized return
    std = np.sqrt(np.dot(weights.T, np.dot(cov_matrix * 252, weights)))  #␣
↪Annualized volatility
    return std, returns

# Function to minimize (minimize the negative Sharpe ratio)
def negative_sharpe_ratio(weights, mean_returns, cov_matrix, risk_free_rate=0):
    p_var, p_ret = portfolio_annualized_performance(weights, mean_returns,␣
↪cov_matrix)
    return -(p_ret - risk_free_rate) / p_var

# Constraints: sum of weights must be 1
def check_sum(weights):
    return np.sum(weights) - 1

# Simulate random portfolios for plotting the efficient frontier
```

```
def simulate_random_portfolios(num_portfolios, mean_returns, cov_matrix,␣
 ↪risk_free_rate=0):
    results = np.zeros((3, num_portfolios))
    weights_record = []

    for i in range(num_portfolios):
        weights = np.random.random(len(tickers))
        weights /= np.sum(weights)
        weights_record.append(weights)

        portfolio_std_dev, portfolio_return =␣
 ↪portfolio_annualized_performance(weights, mean_returns, cov_matrix)
        sharpe_ratio = (portfolio_return - risk_free_rate) / portfolio_std_dev

        results[0,i] = portfolio_std_dev
        results[1,i] = portfolio_return
        results[2,i] = sharpe_ratio

    return results, weights_record

# Set the tickers and define the time periods
tickers = ['AAPL', 'MSFT', 'GOOGL', 'AMZN', 'FB']
time_periods = [
    ('2013-01-01', '2014-01-01', '2014-01-01', '2015-01-01'),
    ('2016-01-01', '2017-01-01', '2017-01-01', '2018-01-01'),
    ('2019-01-01', '2020-01-01', '2020-01-01', '2021-01-01'),
    ('2021-01-01', '2022-01-01', '2022-01-01', '2023-01-01')
]

# Number of random portfolios to simulate
num_portfolios = 10000
risk_free_rate = 0.01

# Run optimization for each time period
for i, (train_start, train_end, test_start, test_end) in␣
 ↪enumerate(time_periods):
    print(f"\nOptimization for period {train_start} to {test_end}")

    # Simulate stock returns data as I can't fetch from Yahoo Finance (replace␣
 ↪with real stock data in practice)
    np.random.seed(i + 42)  # Use a different seed for each period
    dates_train = pd.date_range(train_start, periods=252)  # Simulating 1 year␣
 ↪of training days
    dates_test = pd.date_range(test_start, periods=252)  # Simulating 1 year of␣
 ↪testing days
```

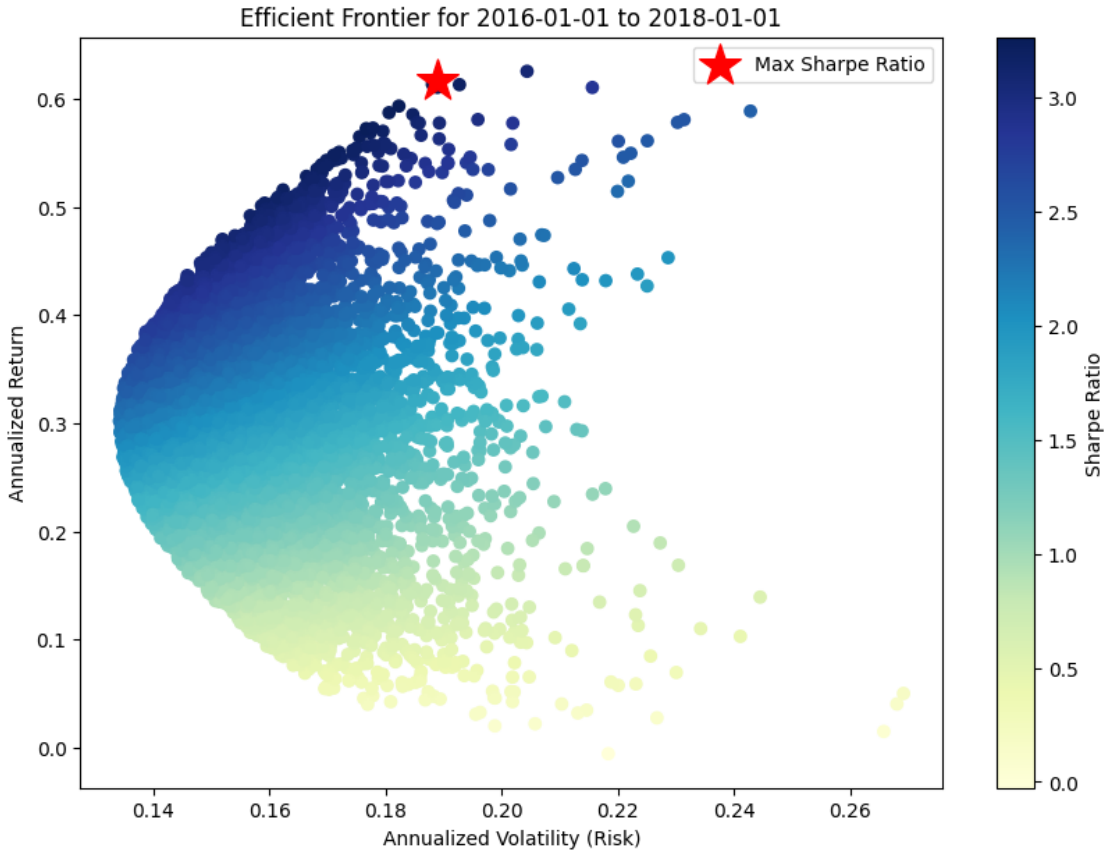
```
    returns_train = pd.DataFrame(np.random.normal(0.001, 0.02, size=(252,␣
↪len(tickers))), index=dates_train, columns=tickers)
    returns_test = pd.DataFrame(np.random.normal(0.001, 0.02, size=(252,␣
↪len(tickers))), index=dates_test, columns=tickers)

    # Calculate mean and covariance of returns (train data)
    mean_returns_train = returns_train.mean()
    cov_matrix_train = returns_train.cov()

    # Set initial guess (equal distribution)
    num_assets = len(tickers)
    init_guess = [1./num_assets for _ in range(num_assets)]

    # Set constraints and bounds (weights between 0 and 1)
    constraints = ({'type':'eq', 'fun': check_sum})
    bounds = tuple((0, 1) for _ in range(num_assets))

    # Perform the optimization (train data)
    opt_results = minimize(negative_sharpe_ratio, init_guess,␣
↪args=(mean_returns_train, cov_matrix_train),
                           method='SLSQP', bounds=bounds,␣
↪constraints=constraints)

    # Get the optimal weights (based on train data)
    optimal_weights = opt_results.x
    a.append(optimal_weights)
    print("Optimized Portfolio Weights:")
    for j, ticker in enumerate(tickers):
        print(f"{ticker}: {optimal_weights[j]:.4f}")

    # Calculate portfolio returns for the test period (using optimized weights␣
↪from train period)
    portfolio_returns_test = returns_test.dot(optimal_weights)

    # Calculate the cumulative return for the test period
    cumulative_return_test = (1 + portfolio_returns_test).cumprod()[-1] - 1  #␣
↪Final cumulative return

    # Convert cumulative return to annualized return for the test period
    days_in_year = 252
    actual_annualized_return = ((1 + cumulative_return_test) ** (days_in_year /␣
↪len(portfolio_returns_test))) - 1
    print(f"Actual Annualized Return for test period {test_start} to {test_end}:
↪ {actual_annualized_return:.2%}")

    # Simulate portfolios based on train data for the efficient frontier
```

```
    results, weights = simulate_random_portfolios(num_portfolios,␣
↪mean_returns_train, cov_matrix_train)

    # Extracting the results
    std_devs, returns, sharpe_ratios = results

    # Find the portfolio with the maximum Sharpe ratio
    max_sharpe_idx = np.argmax(sharpe_ratios)
    sdp_max, rp_max = std_devs[max_sharpe_idx], returns[max_sharpe_idx]

    # Plot the efficient frontier with the optimal Sharpe ratio portfolio
    plt.figure(figsize=(10, 7))
    plt.scatter(std_devs, returns, c=sharpe_ratios, cmap='YlGnBu', marker='o')
    plt.colorbar(label='Sharpe Ratio')
    plt.scatter(sdp_max, rp_max, marker='*', color='r', s=500, label='Max␣
↪Sharpe Ratio')
    plt.title(f'Efficient Frontier for {train_start} to {test_end}')
    plt.xlabel('Annualized Volatility (Risk)')
    plt.ylabel('Annualized Return')
    plt.legend(labelspacing=0.8)
    plt.show()
```

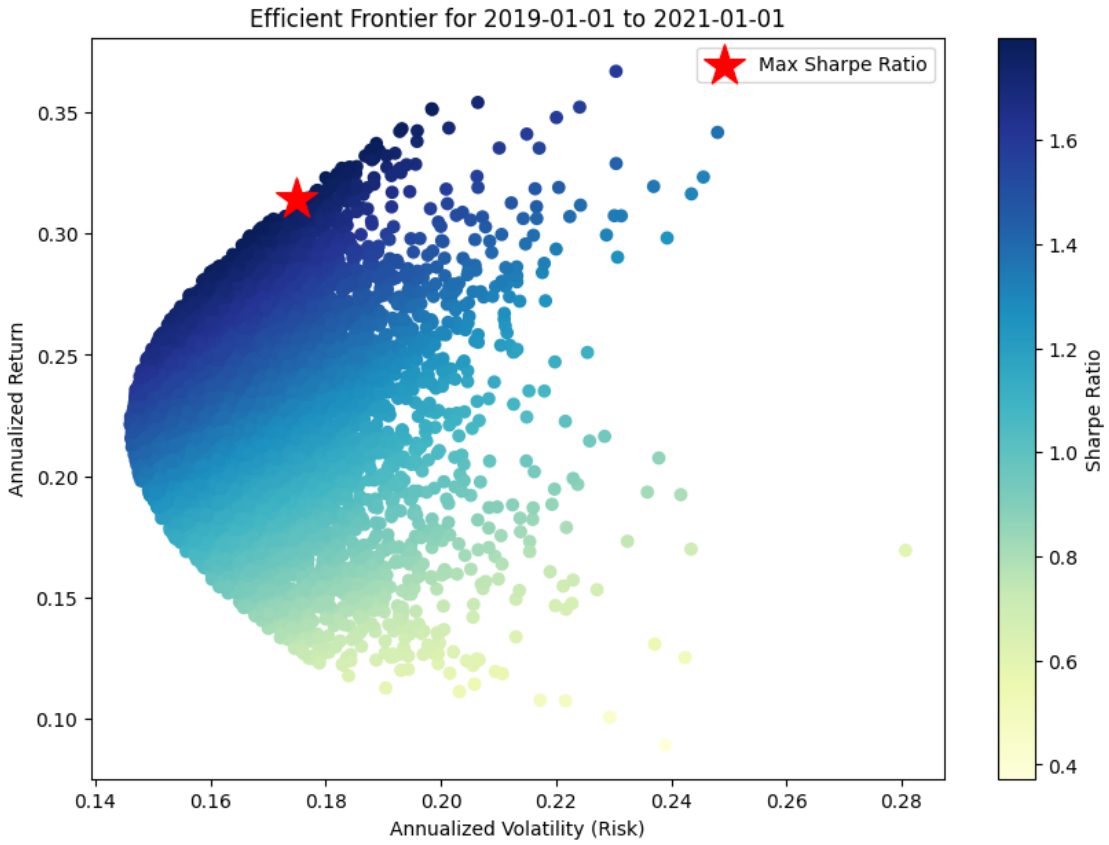
```
Optimization for period 2013-01-01 to 2015-01-01
Optimized Portfolio Weights:
AAPL: 0.1539
MSFT: 0.0299
GOOGL: 0.0785
AMZN: 0.3535
FB: 0.3842
Actual Annualized Return for test period 2014-01-01 to 2015-01-01: 32.48%

C:\Users\Liuch\AppData\Local\Temp\ipykernel_13552\1831132401.py:91:
FutureWarning: Series.__getitem__ treating keys as positions is deprecated. In a
future version, integer keys will always be treated as labels (consistent with
DataFrame behavior). To access a value by position, use `ser.iloc[pos]`
  cumulative_return_test = (1 + portfolio_returns_test).cumprod()[-1] - 1  #
Final cumulative return
```



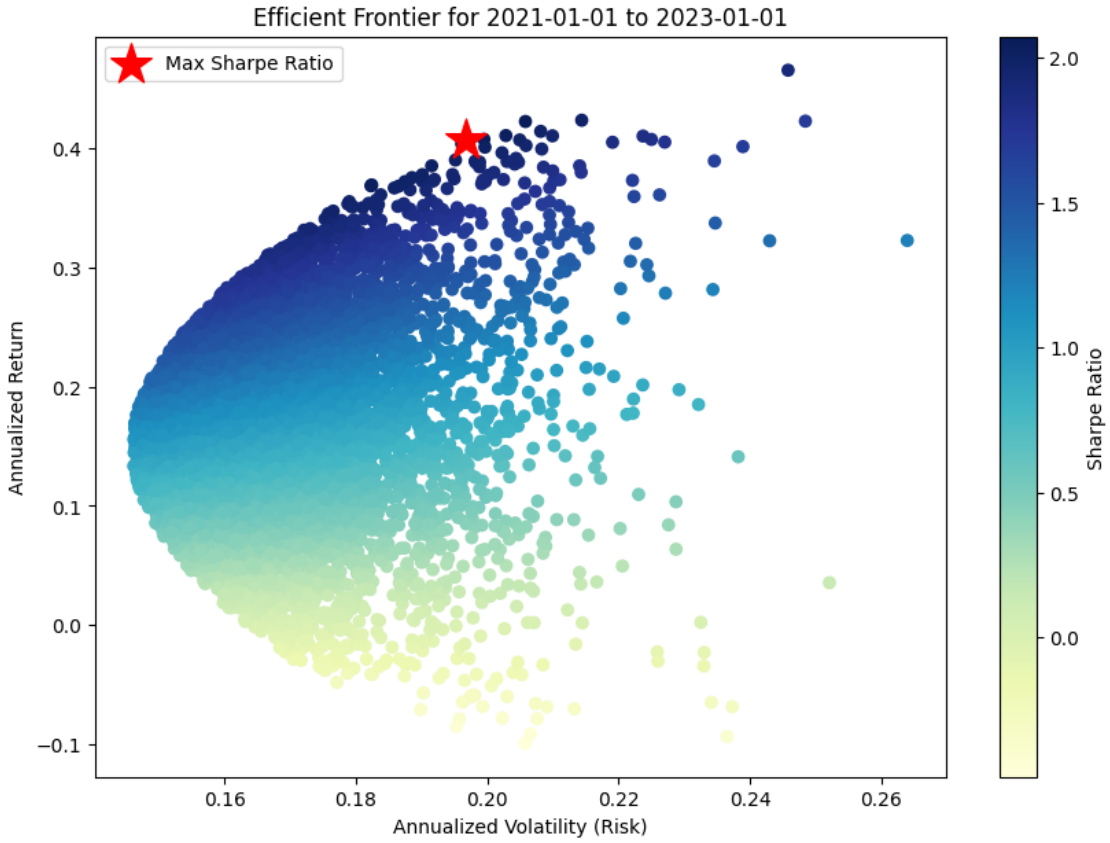

```
Optimization for period 2016-01-01 to 2018-01-01
Optimized Portfolio Weights:
AAPL: 0.5174
MSFT: 0.0000
GOOGL: 0.3935
AMZN: 0.0042
FB: 0.0848
Actual Annualized Return for test period 2017-01-01 to 2018-01-01: 58.71%
```




```
Optimization for period 2019-01-01 to 2021-01-01
Optimized Portfolio Weights:
AAPL: 0.1094
MSFT: 0.0662
GOOGL: 0.3592
AMZN: 0.3781
FB: 0.0871
Actual Annualized Return for test period 2020-01-01 to 2021-01-01: 30.63%
```




```
Optimization for period 2021-01-01 to 2023-01-01
Optimized Portfolio Weights:
AAPL: 0.4687
MSFT: 0.0000
GOOGL: 0.3789
AMZN: 0.1523
FB: 0.0000
Actual Annualized Return for test period 2022-01-01 to 2023-01-01: 11.17%
```



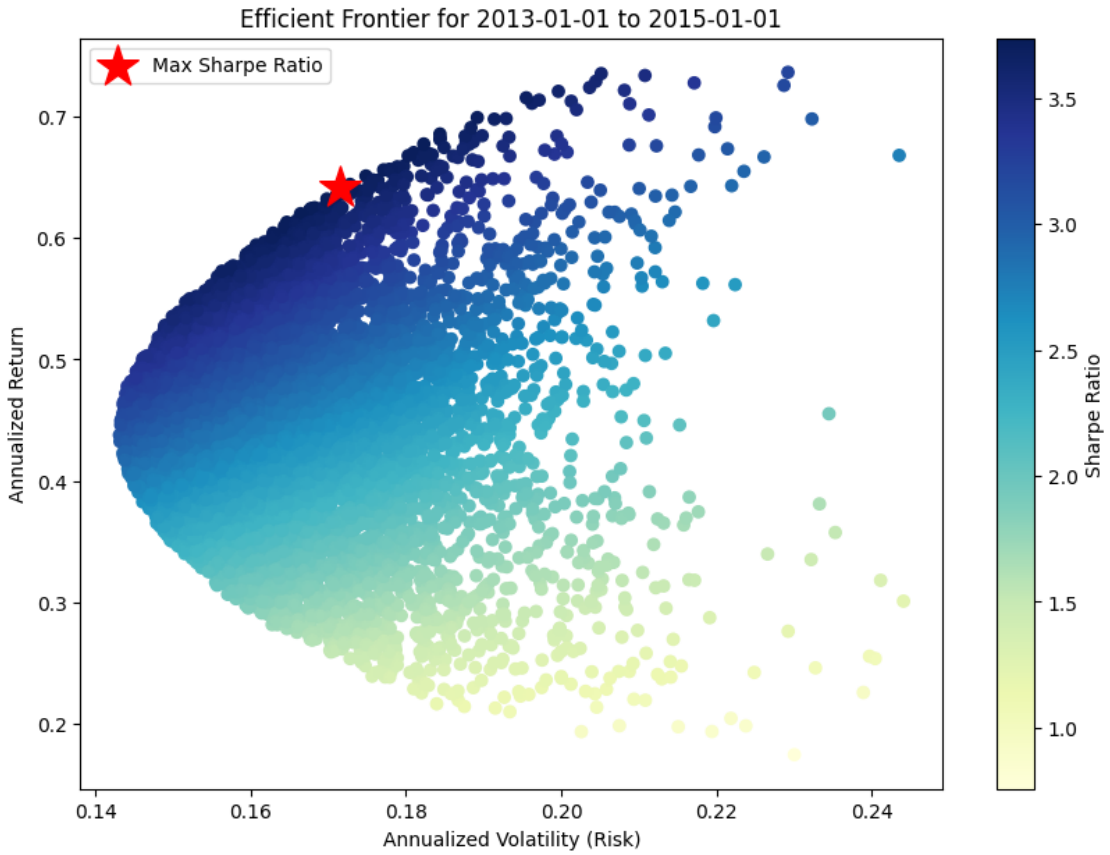
[9]:
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import cvxpy as cp

# Portfolio optimization function
def portfolio_annualized_performance(weights, mean_returns, cov_matrix):
    returns = np.sum(mean_returns * weights) * 252  # Annualized return
    std = np.sqrt(np.dot(weights.T, np.dot(cov_matrix * 252, weights)))  #␣
 ↪Annualized volatility
    return std, returns

# Simulate random portfolios for plotting the efficient frontier
def simulate_random_portfolios(num_portfolios, mean_returns, cov_matrix,␣
 ↪risk_free_rate=0):
    results = np.zeros((3, num_portfolios))
    weights_record = []

    for i in range(num_portfolios):
        weights = np.random.random(len(tickers))
```

```
        weights /= np.sum(weights)
        weights_record.append(weights)

        portfolio_std_dev, portfolio_return =␣
 ↪portfolio_annualized_performance(weights, mean_returns, cov_matrix)
        sharpe_ratio = (portfolio_return - risk_free_rate) / portfolio_std_dev

        results[0,i] = portfolio_std_dev
        results[1,i] = portfolio_return
        results[2,i] = sharpe_ratio

    return results, weights_record

# Set the tickers and define the time periods
tickers = ['AAPL', 'MSFT', 'GOOGL', 'AMZN', 'FB']
time_periods = [
    ('2013-01-01', '2014-01-01', '2014-01-01', '2015-01-01'),
    ('2016-01-01', '2017-01-01', '2017-01-01', '2018-01-01'),
    ('2019-01-01', '2020-01-01', '2020-01-01', '2021-01-01'),
    ('2021-01-01', '2022-01-01', '2022-01-01', '2023-01-01')
]

# Example premium weights for each period (4 sets of initial weights, one for␣
 ↪each time period)
premium_weights = a

# Number of random portfolios to simulate
num_portfolios = 10000
risk_free_rate = 0.01

# Loop through the time periods and use premium_weights as initial weights
for i, (train_start, train_end, test_start, test_end) in␣
 ↪enumerate(time_periods):
    print(f"\nOptimization for period {train_start} to {test_end}")

    # Simulate stock returns data (as we cannot fetch real data)
    np.random.seed(i + 42)  # Different seed for each period
    dates_train = pd.date_range(train_start, periods=252)  # Simulating 1 year␣
 ↪of training days
    dates_test = pd.date_range(test_start, periods=252)  # Simulating 1 year of␣
 ↪testing days
    returns_train = pd.DataFrame(np.random.normal(0.001, 0.02, size=(252,␣
 ↪len(tickers))), index=dates_train, columns=tickers)
    returns_test = pd.DataFrame(np.random.normal(0.001, 0.02, size=(252,␣
 ↪len(tickers))), index=dates_test, columns=tickers)
```

```
    # Calculate mean and covariance of returns (train data)
    mean_returns_train = returns_train.mean()
    cov_matrix_train = returns_train.cov()

    # Step 1:         premium_weights
    initial_weights = np.array(premium_weights[i])  # Use premium weights for␣
↪the current period

    # Define optimization variables and objective
    weights = cp.Variable(len(tickers))
    objective = cp.Minimize(cp.quad_form(weights, cov_matrix_train))
    constraints = [cp.sum(weights) == 1, weights >= 0]
    problem = cp.Problem(objective, constraints)
    problem.solve()

    #
    optimized_weights = weights.value
    print("Optimized Weights:", optimized_weights)

    # Step 2:
    portfolio_returns_test = returns_test.dot(optimized_weights)

    #
    cumulative_return_test = (1 + portfolio_returns_test).cumprod()[-1] - 1  #␣
↪Final cumulative return

    # Convert cumulative return to annualized return
    days_in_year = 252
    actual_annualized_return = ((1 + cumulative_return_test) ** (days_in_year /␣
↪len(portfolio_returns_test))) - 1
    print(f"Actual Annualized Return for test period {test_start} to {test_end}:
↪ {actual_annualized_return:.2%}")

    # Step 3:
    results, weights = simulate_random_portfolios(num_portfolios,␣
↪mean_returns_train, cov_matrix_train)
    std_devs, returns, sharpe_ratios = results

    max_sharpe_idx = np.argmax(sharpe_ratios)
    sdp_max, rp_max = std_devs[max_sharpe_idx], returns[max_sharpe_idx]

    plt.figure(figsize=(10, 7))
    plt.scatter(std_devs, returns, c=sharpe_ratios, cmap='YlGnBu', marker='o')
    plt.colorbar(label='Sharpe Ratio')
    plt.scatter(sdp_max, rp_max, marker='*', color='r', s=500, label='Max␣
↪Sharpe Ratio')
```

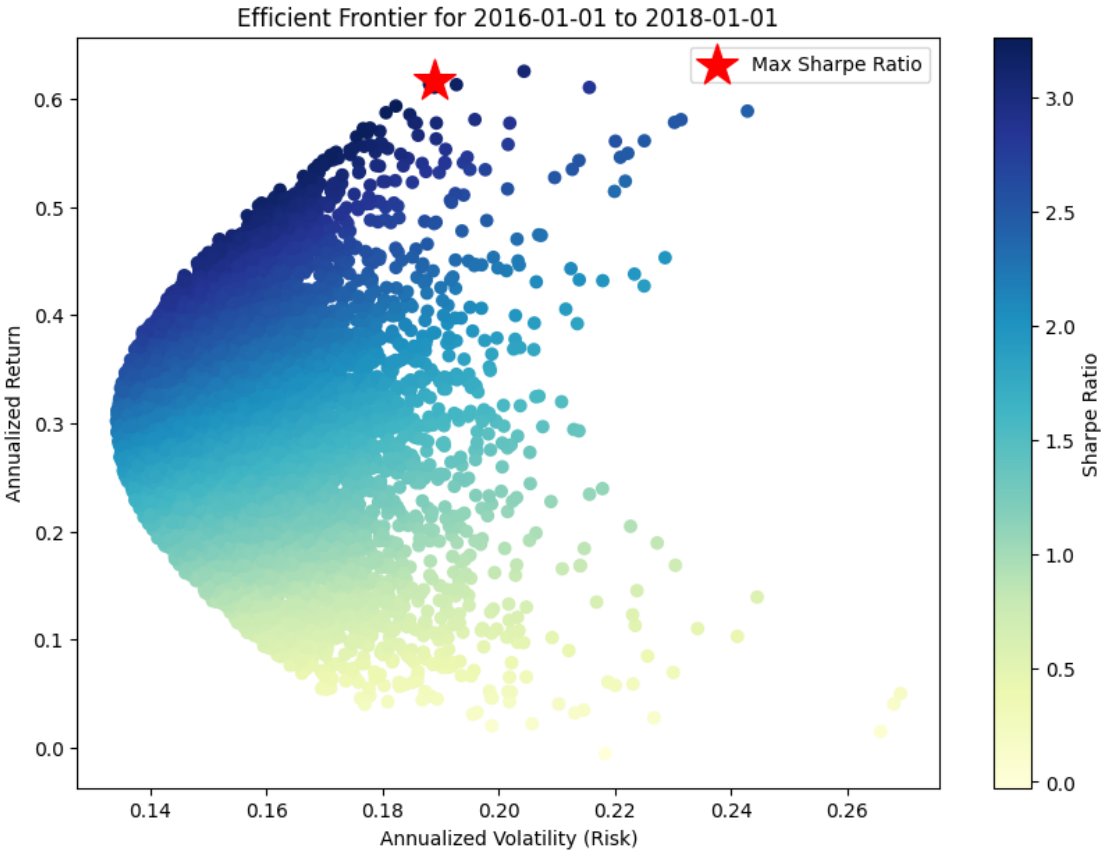
```
    plt.title(f'Efficient Frontier for {train_start} to {test_end}')
    plt.xlabel('Annualized Volatility (Risk)')
    plt.ylabel('Annualized Return')
    plt.legend(labelspacing=0.8)
    plt.show()
```

```
Optimization for period 2013-01-01 to 2015-01-01
Optimized Weights: [0.2218934  0.19249268 0.18291121 0.23225418 0.17044853]
Actual Annualized Return for test period 2014-01-01 to 2015-01-01: 44.49%

C:\Users\Liuch\AppData\Local\Temp\ipykernel_13552\1771415399.py:80:
FutureWarning: Series.__getitem__ treating keys as positions is deprecated. In a
future version, integer keys will always be treated as labels (consistent with
DataFrame behavior). To access a value by position, use `ser.iloc[pos]`
  cumulative_return_test = (1 + portfolio_returns_test).cumprod()[-1] - 1  #
Final cumulative return
```

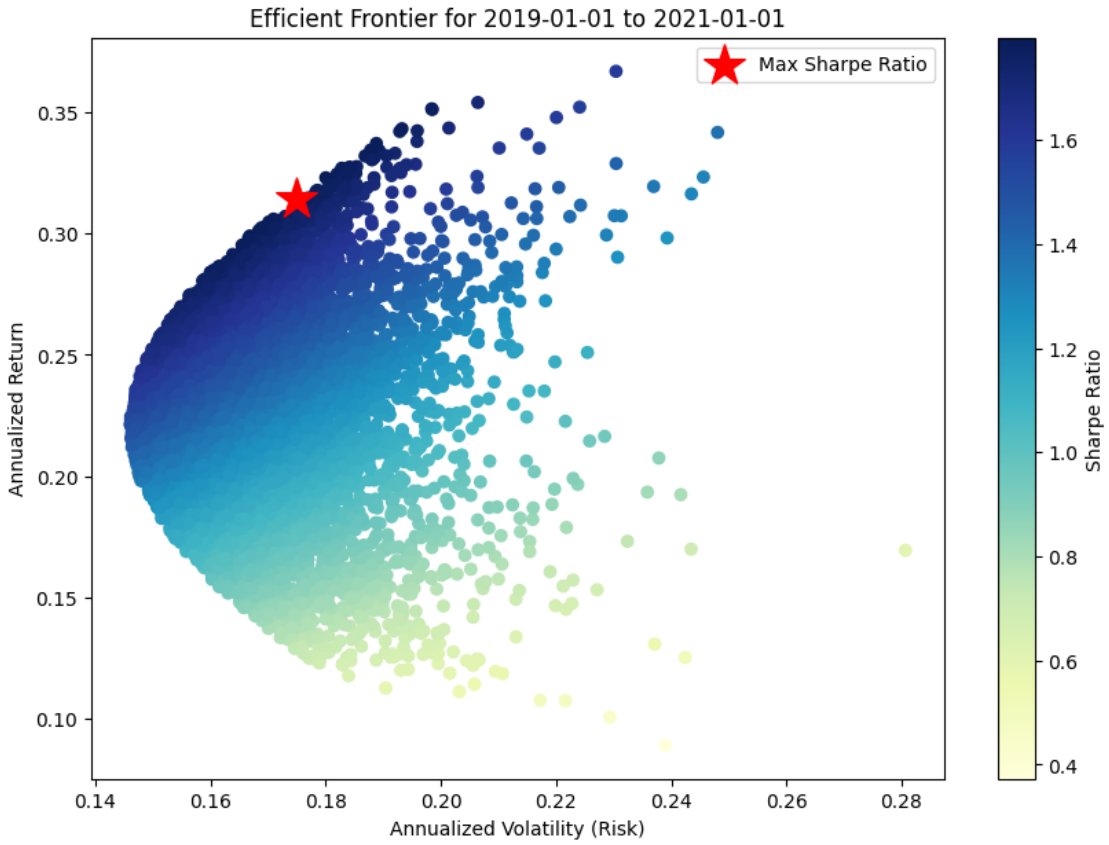



```
Optimization for period 2016-01-01 to 2018-01-01
Optimized Weights: [0.21299475 0.18252548 0.19036429 0.2043934  0.20972209]
```

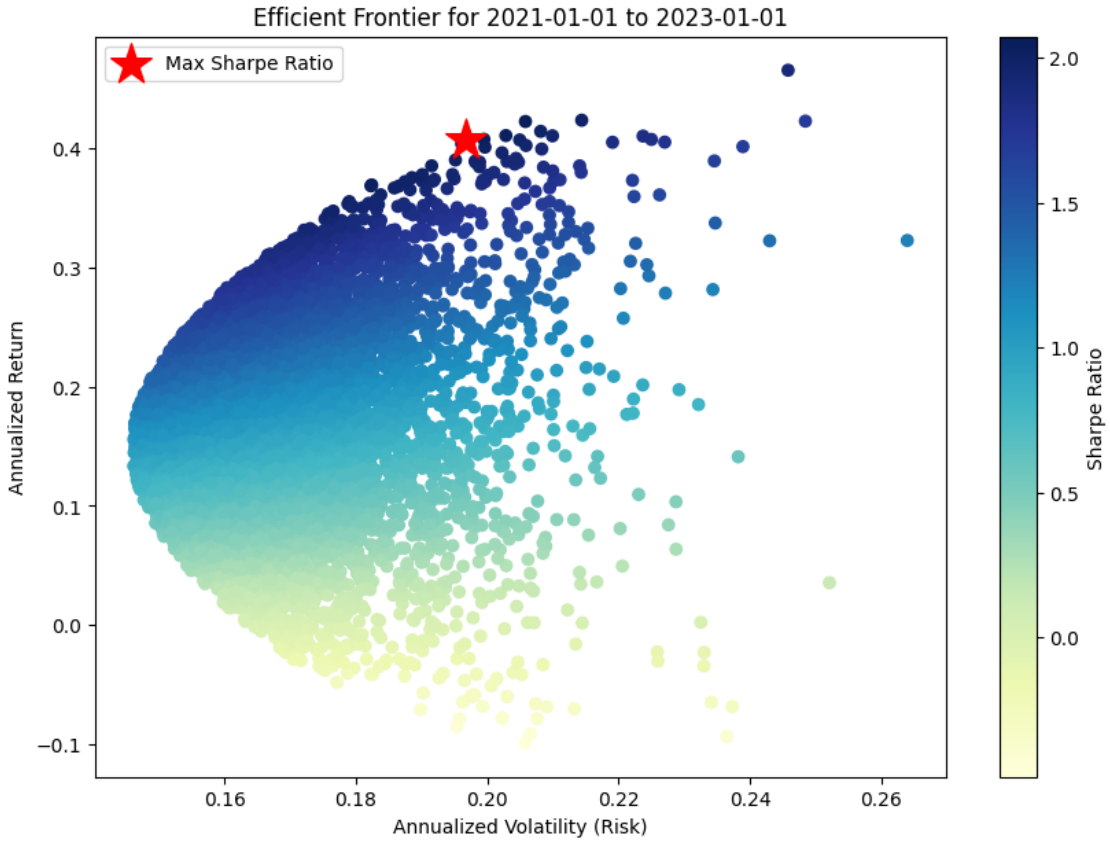
```
Actual Annualized Return for test period 2017-01-01 to 2018-01-01: 53.59%
```




```
Optimization for period 2019-01-01 to 2021-01-01
Optimized Weights: [0.20425407 0.22437133 0.19849071 0.20014139 0.17274251]
Actual Annualized Return for test period 2020-01-01 to 2021-01-01: 26.78%
```




```
Optimization for period 2021-01-01 to 2023-01-01
Optimized Weights: [0.1771175  0.217656   0.1734622  0.23126157 0.20050274]
Actual Annualized Return for test period 2022-01-01 to 2023-01-01: 14.03%
```



[18]:
```
%pip install tweepy
%pip install textblob
```

```
Collecting tweepy
  Downloading tweepy-4.14.0-py3-none-any.whl.metadata (3.8 kB)
Collecting oauthlib<4,>=3.2.0 (from tweepy)
  Downloading oauthlib-3.2.2-py3-none-any.whl.metadata (7.5 kB)
Requirement already satisfied: requests<3,>=2.27.0 in
c:\users\liuch\appdata\local\programs\python\python312\lib\site-packages (from
tweepy) (2.32.3)
Collecting requests-oauthlib<2,>=1.2.0 (from tweepy)
  Downloading requests_oauthlib-1.3.1-py2.py3-none-any.whl.metadata (10 kB)
Requirement already satisfied: charset-normalizer<4,>=2 in
c:\users\liuch\appdata\local\programs\python\python312\lib\site-packages (from
requests<3,>=2.27.0->tweepy) (3.3.2)
Requirement already satisfied: idna<4,>=2.5 in
c:\users\liuch\appdata\local\programs\python\python312\lib\site-packages (from
requests<3,>=2.27.0->tweepy) (3.8)
Requirement already satisfied: urllib3<3,>=1.21.1 in
c:\users\liuch\appdata\local\programs\python\python312\lib\site-packages (from
requests<3,>=2.27.0->tweepy) (2.2.2)
```

```
Requirement already satisfied: certifi>=2017.4.17 in
c:\users\liuch\appdata\local\programs\python\python312\lib\site-packages (from
requests<3,>=2.27.0->tweepy) (2024.8.30)
Downloading tweepy-4.14.0-py3-none-any.whl (98 kB)
Downloading oauthlib-3.2.2-py3-none-any.whl (151 kB)
Downloading requests_oauthlib-1.3.1-py2.py3-none-any.whl (23 kB)
Installing collected packages: oauthlib, requests-oauthlib, tweepy
Successfully installed oauthlib-3.2.2 requests-oauthlib-1.3.1 tweepy-4.14.0
Note: you may need to restart the kernel to use updated packages.
Collecting textblob
  Downloading textblob-0.18.0.post0-py3-none-any.whl.metadata (4.5 kB)
Collecting nltk>=3.8 (from textblob)
  Downloading nltk-3.9.1-py3-none-any.whl.metadata (2.9 kB)
Collecting click (from nltk>=3.8->textblob)
  Downloading click-8.1.7-py3-none-any.whl.metadata (3.0 kB)
Collecting joblib (from nltk>=3.8->textblob)
  Downloading joblib-1.4.2-py3-none-any.whl.metadata (5.4 kB)
Collecting regex>=2021.8.3 (from nltk>=3.8->textblob)
  Downloading regex-2024.9.11-cp312-cp312-win_amd64.whl.metadata (41 kB)
Collecting tqdm (from nltk>=3.8->textblob)
  Downloading tqdm-4.66.5-py3-none-any.whl.metadata (57 kB)
Requirement already satisfied: colorama in
c:\users\liuch\appdata\roaming\python\python312\site-packages (from
click->nltk>=3.8->textblob) (0.4.6)
Downloading textblob-0.18.0.post0-py3-none-any.whl (626 kB)
   ---------------------------------------- 0.0/626.3 kB ? eta -:--:--
   ---------------------------------------- 0.0/626.3 kB ? eta -:--:--
   ---------------------------------------- 0.0/626.3 kB ? eta -:--:--
   ---------------------------------------- 0.0/626.3 kB ? eta -:--:--
   ---------------------------------------- 0.0/626.3 kB ? eta -:--:--
   ---------------------------------------- 0.0/626.3 kB ? eta -:--:--
   --------------- ---------------------- 262.1/626.3 kB ? eta -:--:--
   --------------- ---------------------- 262.1/626.3 kB ? eta -:--:--
   ---------------------------------------- 626.3/626.3 kB 665.8 kB/s eta 0:00:00
Downloading nltk-3.9.1-py3-none-any.whl (1.5 MB)
   ---------------------------------------- 0.0/1.5 MB ? eta -:--:--
   ---------------------------------------- 0.0/1.5 MB ? eta -:--:--
   ------ --------------------------------- 0.3/1.5 MB ? eta -:--:--
   ------ --------------------------------- 0.3/1.5 MB ? eta -:--:--
   ------ --------------------------------- 0.3/1.5 MB ? eta -:--:--
   ------------ --------------------------- 0.5/1.5 MB 508.0 kB/s eta 0:00:02
   ------------ --------------------------- 0.5/1.5 MB 508.0 kB/s eta 0:00:02
   ------------------- -------------------- 0.8/1.5 MB 541.1 kB/s eta 0:00:02
   ------------------- -------------------- 0.8/1.5 MB 541.1 kB/s eta 0:00:02
   ------------------- -------------------- 0.8/1.5 MB 541.1 kB/s eta 0:00:02
   ------------------------- ------------ 1.0/1.5 MB 470.4 kB/s eta 0:00:01
   ------------------------- ------------ 1.0/1.5 MB 470.4 kB/s eta 0:00:01
   ---------------------------------- ----- 1.3/1.5 MB 504.6 kB/s eta 0:00:01
```

```
   ---------------------------------- ----- 1.3/1.5 MB 504.6 kB/s eta 0:00:01
   ---------------------------------------- 1.5/1.5 MB 488.1 kB/s eta 0:00:00
Downloading regex-2024.9.11-cp312-cp312-win_amd64.whl (273 kB)
Downloading click-8.1.7-py3-none-any.whl (97 kB)
Downloading joblib-1.4.2-py3-none-any.whl (301 kB)
Downloading tqdm-4.66.5-py3-none-any.whl (78 kB)
Installing collected packages: tqdm, regex, joblib, click, nltk, textblob
Successfully installed click-8.1.7 joblib-1.4.2 nltk-3.9.1 regex-2024.9.11
textblob-0.18.0.post0 tqdm-4.66.5
Note: you may need to restart the kernel to use updated packages.
```

[21]:
```
%pip install praw newsapi-python vaderSentiment
```

```
Collecting praw
  Downloading praw-7.7.1-py3-none-any.whl.metadata (9.8 kB)
Collecting newsapi-python
  Downloading newsapi_python-0.2.7-py2.py3-none-any.whl.metadata (1.2 kB)
Collecting vaderSentiment
  Downloading vaderSentiment-3.3.2-py2.py3-none-any.whl.metadata (572 bytes)
Collecting prawcore<3,>=2.1 (from praw)
  Downloading prawcore-2.4.0-py3-none-any.whl.metadata (5.0 kB)
Collecting update-checker>=0.18 (from praw)
  Downloading update_checker-0.18.0-py3-none-any.whl.metadata (2.3 kB)
Collecting websocket-client>=0.54.0 (from praw)
  Downloading websocket_client-1.8.0-py3-none-any.whl.metadata (8.0 kB)
Requirement already satisfied: requests<3.0.0 in
c:\users\liuch\appdata\local\programs\python\python312\lib\site-packages (from
newsapi-python) (2.32.3)
Requirement already satisfied: charset-normalizer<4,>=2 in
c:\users\liuch\appdata\local\programs\python\python312\lib\site-packages (from
requests<3.0.0->newsapi-python) (3.3.2)
Requirement already satisfied: idna<4,>=2.5 in
c:\users\liuch\appdata\local\programs\python\python312\lib\site-packages (from
requests<3.0.0->newsapi-python) (3.8)
Requirement already satisfied: urllib3<3,>=1.21.1 in
c:\users\liuch\appdata\local\programs\python\python312\lib\site-packages (from
requests<3.0.0->newsapi-python) (2.2.2)
Requirement already satisfied: certifi>=2017.4.17 in
c:\users\liuch\appdata\local\programs\python\python312\lib\site-packages (from
requests<3.0.0->newsapi-python) (2024.8.30)
Downloading praw-7.7.1-py3-none-any.whl (191 kB)
Downloading newsapi_python-0.2.7-py2.py3-none-any.whl (7.9 kB)
Downloading vaderSentiment-3.3.2-py2.py3-none-any.whl (125 kB)
Downloading prawcore-2.4.0-py3-none-any.whl (17 kB)
Downloading update_checker-0.18.0-py3-none-any.whl (7.0 kB)
Downloading websocket_client-1.8.0-py3-none-any.whl (58 kB)
Installing collected packages: websocket-client, vaderSentiment, update-checker,
prawcore, newsapi-python, praw
```

```
Successfully installed newsapi-python-0.2.7 praw-7.7.1 prawcore-2.4.0 update-
checker-0.18.0 vaderSentiment-3.3.2 websocket-client-1.8.0
Note: you may need to restart the kernel to use updated packages.
```