Midterm Project: Research on ETF intraday volatility vs volume

The goal

- Find the mathematical equation(s) to describe the relationship between intraday volatility vs volume.
- Hints:
  - Your equation(s) must have intraday volatility and volume. You may include anything else into your equation(s).
  - The possible equation(s) may be: linear equations, linear systems, non-linear systems, polynomials, ODE or PDE
  - Show that your data fits your equation(s).

The background knowledge

- An exchange traded fund (ETF) is a type of security that involves a collection of securities—such as stocks—that often tracks an underlying index, although they can invest in any number of industry sectors or use various strategies. ETFs are in many ways similar to mutual funds; however, they are listed on exchanges and ETF shares trade throughout the day just like ordinary stock.

- Some well-known example is the SPDR S&P 500 ETF (SPY), which tracks the S&P 500 Index. ETFs can contain many types of investments, including stocks, commodities, bonds, or a mixture of investment types. An exchange traded fund is a marketable security, meaning it has an associated price that allows it to be easily bought and sold.

Step 1: We will start from SPY in this project.

- import the useful packages

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
%matplotlib inline
plt.rcParams["figure.figsize"] = [20, 10]
```

- You can download the csv file from the page, or use the pandas_datareader in the sample code. For example, use the below URL to load the daily data of 2022 into dataframe

  https://query1.finance.yahoo.com/v7/finance/download/SPY?period1=1666471279&period2=1698007279&interval=1d&events=history&includeAdjustedClose=true (https://query1.finance.yahoo.com/v7/finance/download/SPY?period1=1666471279&period2=1698007279&interval=1d&events=history&includeAdjustedClose=true)

  Your dataframe will look like

Date	Open	High	Low	Close	Adj Close	Volume
1/3/2022	476.3	477.85	473.85	477.71	464.9233	72668200
1/4/2022	479.22	479.98	475.58	477.55	464.7676	71178700
1/5/2022	477.16	477.98	468.28	468.38	455.8431	104538900
1/6/2022	467.89	470.82	465.43	467.94	455.4149	86858900
1/7/2022	467.95	469.2	464.65	466.09	453.6144	85111600
1/10/2022	462.7	465.74	456.6	465.51	453.05	119362000
1/11/2022	465.23	469.85	462.05	469.75	457.1765	74303100
1/12/2022	471.59	473.2	468.94	471.02	458.4125	67605400

- We compute the intraday volatility as Volatility = (High-Low)/Close
- We then compute the daily volume change as Volume_delta = Volume_today - Volume_previousday
- We then compute the daily volatility change as Volatility_delta = Volatility_today - Volatility_previouday
- Plot Volume_delta and Volatility_delta.

  Use the sample code to generate graph

  ```
  your_data_frame.plot(x='Volume_delta', y='Volatility_delta', style='.')
  ```

- Based on your observations, can you find the mathematical equation(s) to describe the relationship between intraday volatility vs volume.
  - Hint: You can guess equation(s) and leave some coeffecients unknown, which you will find their values by numerical methods.
  - Then you can show how well your data fits your equation(s)
- Use the last 5 years data of SPY to test your equation(s).

Step 2: Continue your research on other ETFs

- You pick other ETFs and research the intraday volatility vs volume as Step 1. You can pick from the list

  https://etfdb.com/compare/volume/ (https://etfdb.com/compare/volume/)

- Do other ETFs fit your equation(s)? Better fit or worse fit? Please provide some explanation.

Step 3: Extend your research on stocks.

- Conduct the same research on the most active stocks

  https://www.barchart.com/stocks/most-active/daily-volume-leaders (https://www.barchart.com/stocks/most-active/daily-volume-leaders)

- Do stocks fit your equation(s)? Better fit or worse fit? Please provide some explanation.

Submission

- This is a group project. Each team may have up to 3 (three) members.
- Write a summary about what you find. Anything interesing and/or worth to mention will be good enough.
- Submit in ipynb and HTML format on canvas. Each group only need to submit one copy on canvas. Please include all members on the top of your jupyter notebook.