Develop a tool (ML or statistics-based approach) to predict market events (significant changes in a stock) before they happen.

Predictive tool modelling

Given the historical market data (best price and trade datasets) of 5 stocks containing volatility events (economic/news releases):

- **Step 1**: Design an interactive graph/visualization that can display the selected stock's data for a day or time range.
    - Using the graph, identify the timing of those events and observe how they impact the stock prices, and which stock usually lead the market movement.
    - Following these observations, can we predict the movement of a specific stock using the data from the other stocks? Additionally, how can we anticipate a sharp rise or drop in this stock's value?

  Bonus:

  1. Provide an option to plot the 30 sec and 60 sec standard deviation measures on the graph.
  2. Provide as many options as possible to make the graph/chart as useful as possible, for example:
     a. Provide options to select the values to display (bid, ask, mid-price, volumes)
     b. Highlight the low, high of the day, trades on the chart

- **Step 2**: Develop a tool to test and validate your analysis from step 1 (ML or statistics-based approach). Ideally, the tool should be able to generate signals predicting significant changes in the stock.
    - The data provided can be divided into training/validation/test sets at your preference.

  Bonus:

  1. Design a trading strategy using the signals generated by your tool and show the PNL of such a strategy (assume a starting fund of 1,000,000). ***Explaining the idea behind the strategy is key to securing more points***
     An example of strategies could be:
     - buy (or sell) a stock whenever a signal predicts a sharp rise and sell it once the price is higher.
     - Find a statistical relationship/ratio among the stocks and use the signal as a trigger to buy/sell one and sell/buy the other

A test dataset will be provided to you 5 hours before the deadline to validate your solution. The presentation should be based off that dataset.