REAL WORLD PROBLEM (Three days to solve the problem)

Resources: Please find datasets linked
Gold Price Data:
https://drive.google.com/file/d/188xeEQtSoKliH5RyBNYXHU-BHfCJI1-g/view?usp=share_link
Stock Price Data:
https://drive.google.com/file/d/1T4qmYpgGh3tVb2T_L0v30yPXpZGWw474/view?usp=sharing

Problem Setup:

We are given NSE stock price data and gold price data over the past few years.

We need to analyse the impact of “extreme” gold price movement on stock prices.

Definitions:

- We are interested in intraday return (Close-Morning)/Morning for NSE stocks
- We will consider “extreme” gold movements on “previous” days -- business days right before we analyze stock price movement
- We define gold return as (Close - Close_previous)/(Close_previous)
- We define “extreme” gold movements as days when the gold return is in Top 95% of all gold returns upto that point or Bottom 5% of all gold returns upto that point

Your asks:

1. Since we need a combination of gold and stock data report the date range between which you will be able to run the experiments
2. For every year in the dataset Y, for all data upto the year end, find the Top Gold and Bottom Gold (95% return and 5% return respectively)
3. For every year Y, use the bottom gold and top gold value from the previous year Y-1 and use it to identify days in Y when the previous business day were Top Gold and Bottom Gold respectively
4. For the exact day following Top Gold and Bottom Gold, get intraday stock return for each company from NSE for that day
5. For the entire year, get the average stock return for each company on Top Gold following days of the year (and similarly separately for Bottom Gold following days of the year)
6. Now, for averages for each company for Top Gold, find a df.describe() to identify the Top 5% average return companies. Get the return value for Top 5%, 4%, 3%, 2% and 1% for each year. Do this for Bottom Gold and Top Gold (separately like always)
7. For every year for which we do step 6, list the 5% companies that were selected, 4%, 3%, 2% and 1%. Do this for Top Gold and Bottom Gold.

Please submit your solution in a Python Notebook.

Please reach out on Internshala for any questions