COM 3590, Spring 2021
Assignment # 2

Due date: Tuesday, March 9th by 11:59 PM Total points: 20 points

In this assignment, our goal is to detect anomalous changes in the daily closing prices of various stocks. The input data BTC_GOLD_SP500.csv (available in Canvas under "Homework # 2" folder)contains the historical closing prices of stocks for currency, commodity and index (Bitcoin, gold, and SP500.)

(1) Compute the percentage of changes in the daily closing price of each stock. Perform the k-nearest neighbor approach (distance based approach) to calculate anomaly scores associated to the percentage of changes with $k = 2, 4, 15$. Plot the percentage of changes colored by the anomaly score.

(2) List the dates associated with the top-5 highest anomaly scores for each k. Are those dates associated to an important event? Explain.

You should submit:

- A jupyter notebook called "YOUR_NAME-Assignment2-data-cleaning.pynb" with codes outputs and explanations (marked-down mode in jupyter notebook.) Please submit a single notebok with code, answers and explanations.

Upload the file above to Canvas (under Assignment 2.) Please let me know if you are having troubles uploading this files to Canvas before the assignment deadline.