

Assignment 1



NO PLAGIARISM!

In this assignment, we will analyze Bitcoin prices, returns, and volatility using a "Bitcoin_Data.csv" dataset. You will complete This assignment uses only (generic) Python. No additional Python modules like Pandas, Numpy, or SciPy are allowed. You will get 0 for this assignment if you use any of these modules. You are only allowed to use "math" module if needed.

Read about Bitcoin:

<https://en.wikipedia.org/wiki/Bitcoin>

Dataset Description: The dataset contains OHLC - Open, High, Low, Close (Price) daily prices for Bitcoin for about 10 years. There are 3,915 rows and 7 columns

1. **Date:** month/day/year

2. **Price**: last (the so-called "Close") price
3. **Open**: opening price for the day
4. **High**: max price of the day
5. **Low**: min price for the day
6. **Change%**: percent change in price from previous day (the so-called "return")

You will use a subset of this dataset (year Y) depending on the last digit L of your BUID as follows:

$$Y = 2014 + L$$

For example, if your BUID ends on 0, you will analyze data for 2014 only, if BUID ends on 1, you analyze 2015, ..., if your buid ends on 9, you analyze 2023.

Question 1:

1. load the "Bitcoin_Data" csv file as a list of lines using Python and construct a sublist for you group
2. how many entries are there?
3. compute the average price
4. compute the average return, max, min and standard deviation for all days in your year

Question 2: In this question, we will compare statistics for positive and negative days (by returns).

1. how many positive T^+ and negative days T^- are there?

2. compute average return, max, min, standard deviations for all days, for T^+ and for T^- days separately. How do they compare to each other?
3. find the three days with the highest positive and 3 days with the lowest negative returns
4. are these days consecutive?

Question 3: In this question, we will compare quantile statistics for positive and negative days (by returns).

1. compute median M , first quartile Q_1 , third quartile Q_3 and Interquartile range $IQR=(Q_3 - Q_1)$ of returns for T^+ and T^- days separately.
2. How do these compare with results from the previous question?
3. Is IQR higher or lower than the standard deviation?
4. consider the following non-parametric Pearson's skewness:

$$S = \frac{(\mu - M)}{\sigma}$$

5. compute S for your positive and negative return days
6. consider the following quantile (Galton's skew):

$$G = \frac{(Q_3 - M) - (M - Q_1)}{Q_3 - Q_1}$$

7. compute G for your positive and negative return days
8. compare the two measures of skewness S and G for your data and describe your observations

Question 4: For each day, we can compute "range" as the difference between the "High" price H and the "Low" price L .

1. What is the maximum range for negative, and what is the maximum range for positive return days
2. what are the median and mean values for these ranges
3. find the three days with the highest and lowest range
4. were these three days consecutive?

Question 5: There are a number of ways to measure volatility. One such method is called "True Range" or TR:

https://en.wikipedia.org/wiki/Average_true_range

It is computed as follows:

$$TR = \max \left[(H - L), |H - \text{PrevPrice}|, |L - \text{PrevPrice}| \right]$$

1. compute TR for your dataset and write the results into an output csv file
2. find three days with the highest TR. Are these days consecutive?

Question 6: What have you learned from this assignment?

