Please refer to the following problems. If you describe the problem-solving process, you can get extra points even though your answer is not correct. Also, please kindly submit [.html] instead of [.ipynb] when you solve problems using jupyter-notebook.

P1. Regression

Attached .csv file has 1000 rows and four columns (from the left, A, B, C, and D). Please build a good model which explains Column D by a function of Column A, Column B and Column C.

P2. Growth – Sales/Earning

X(t) is a time series. Growth function is to capture relative change(increase or decrease) of *X(t)* where *t* means this quarter and *t-1* mean last quarter.

- Sales is the amount of money that a company actually receives during a specific period, including discounts and deductions for returned merchandise. Sales are positive number in most cases with few exceptions.
- Earnings are the amount of profit that a company produces during a specific period, which is usually defined as a quarter (three calendar months) or a year. Earnings can be both positive and negative values.

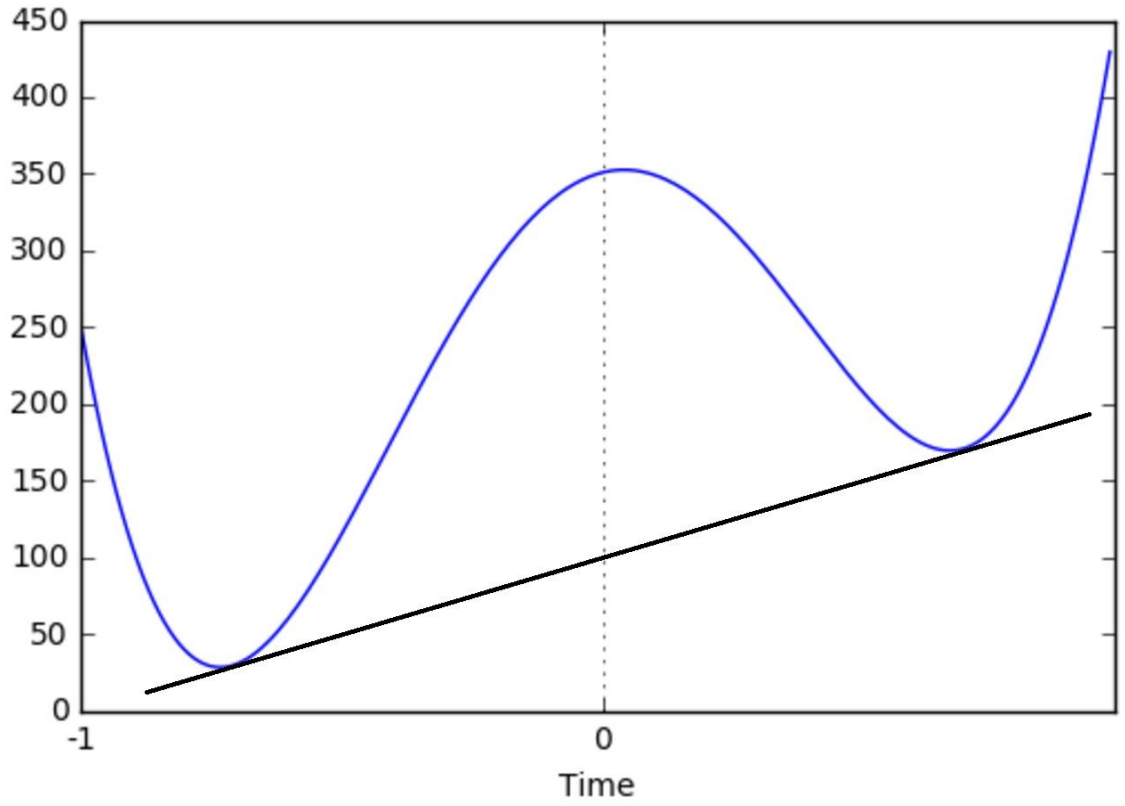
Considering above description, please improve below formula of Sales/Earning Growth to process *dirty* data values.

- Earnings Growth Formula: { Earnings(t)/Earnings(t-1) } – 1
- Sales Growth Formula: { Sales(t)/ Sales(t-1) } – 1

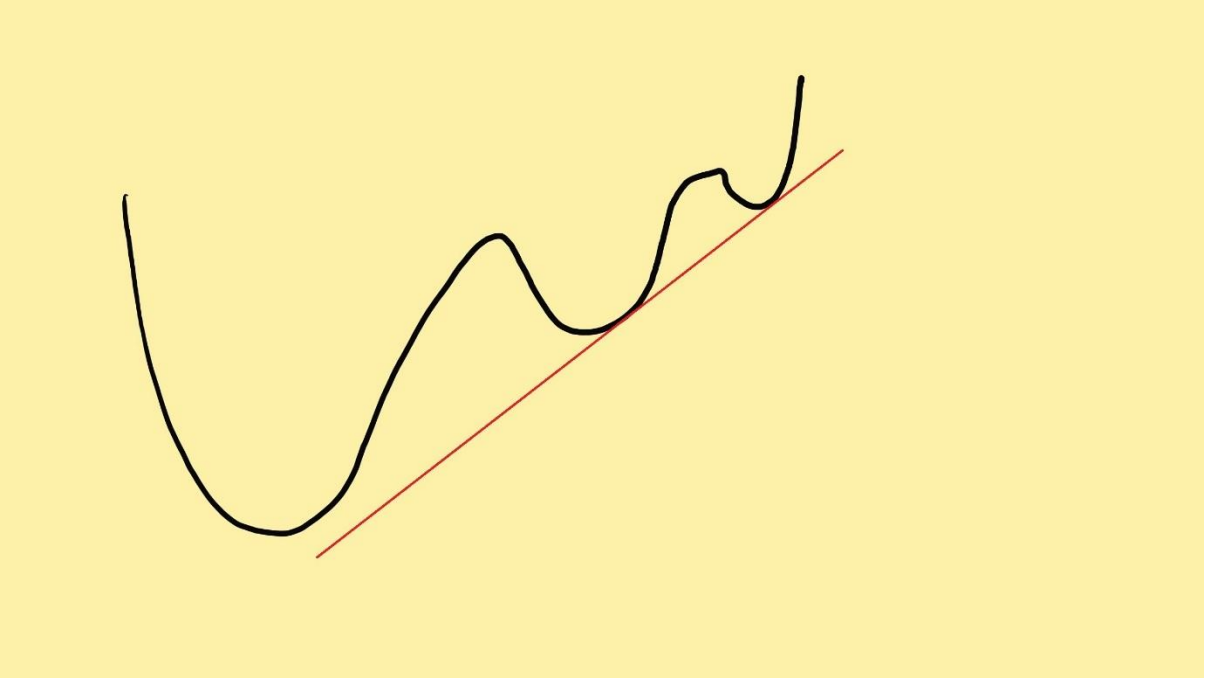
i.e. taking ratios between current quarter's value and last quarter's value and subtract by one.

P3. Technical indicator – trend-line

Below plot is daily price curve of 1,000 data-points of daily Price. In other word, Y-axis means Price and X-axis means day(t) – Price(t). For technical analysis, you need to draw a trend-line which defines a Supporting Line. This Supporting Line is a graph of linear function which is located below the daily price curve but intersects the daily price curve twice. Please note that intersections from below plots are not necessarily local minima.

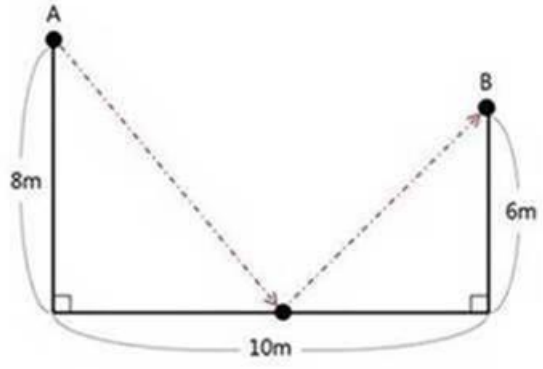


Now please describe (or write a pseudo-code) how to find the linear equation of Supporting Line with t-variable.



M1. Shortest Path

A traveler is departing point A toward point B with a constraint: the traveler must touch the horizontal line during the trip.
What is the distance of shortest path for the traveler?



M5. Poker Game

There are two sets of cards – blue and gold. Each set consists of ten cards numbered from 1 through 10. Now two players are playing “poker” with these two sets. Rules are simple

i. Shuffle 20 cards and each player randomly pick up two cards.
ii. A player with better combination of cards wins. If two players have the same cards regardless of their color, it is draw. Ex) Gold 9 – Blue 8 = Blue 9 – Gold 8
iii. Grade of card combination is as below. All Consecutive numbers are higher than Pairs. And all Pairs are higher than Any Number Cards.
iv. 1) Consecutive numbers: 10-1 > 9-10 > 8-9 > … > 2-3 > 1-2 > Pairs(10-10 > …)
2) Pairs: 10-10 > 9-9 > …. > 2-2 > 1-1 > High Number(10-8 > 10-7 > …)
3) Any card with high number: 8-1 > 7-4
4) If the high cards of both players are the same, the second cards determine the winner: 8-6 > 8-4

A. If you have 9-Gold at your hand, what is your odd of winning?

B. If you happen to have one extra card while the counterpart has only two cards, what is your odd of winning? In this case, you have three cards and you need to show the best combination out of three cards. For example, if you have Gold-1, Blue-2, and Blue-10, you may show Blue-10 and Gold-1

E1. Discounted Cash Flow problem

You have a condominium for rent. Your tenant will pay you yearly rent of $10,000.00 (ten thousand) If a rental contract is longer than two calendar years, yearly rent is adjusted every two year by inflation rate. For example, if annual inflation rate is 2%, rent will become $10,404 after two years. Now suppose that current inflation rate is 2% and interest rate will become 3% after 2 years. Then if your enant rent your condominium for 5 years, and wish to make a one-time payment in cash, what is the fair amount? (calculate using simple interest and rent is paid at the end of each year)

E2. Term Structure of Interest Rates

1) If Term structure of interest rates is upward sloping, fill this relationship (YTM, Spot Rate, Forward Rate)

2) I have a treasure bond that maturity is 3Y, face value 1000, coupon rate 5%, annual coupon

Calculate bond price
* 1Y par rate: 3%
* 2Y par rate: 4%
* 3Y par rate: 5%

Write efficient programs for the following questions using Python 2.7 or C++. Also, you will earn extra points if you refrain from using no built-in libraries such as numpy, scipy, and pandas. In other word, building your own libraries will earn you more points.

You are not expected to solve all the problems. Please submit whatever you could do by time out or request more time to take the test by email.

C1. Given a string, find the second most frequent vowel letter.

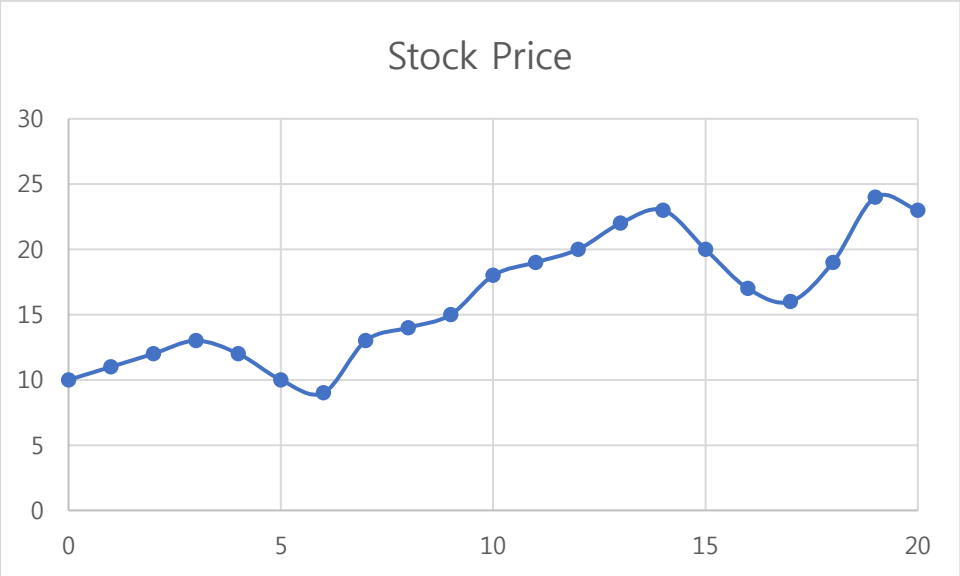
E.g. 'abcdeeeooauiea' -> a

C3. Given a sequence of daily stock price p_t, please write a fast code which returns k-largest maximum draw-downs(MDD) and t_r's where price recovers the previous maximum price.

Please return a list of tuples: [(mdd_size_1, t_mdd_recover_1), …, (mdd_size_k, t_mdd_recover_k)]

Use the input sequence(list) and two lists with length-k (to save MDDs) only – do not declare another list/array. A maximum drawdown (MDD) is the maximum loss from a peak to a trough of a portfolio, before a new peak is attained.

For below particular example, there were two MDDs. MDD#1 starts from t=3 and fully recovers at t=7; MDD#2 starts from t=14 and fully recovers at t=19. The size of MDDs are MDD#1 = 4 and MDD#2 = 7. So the latter is the largest MDD. Then the code is supposed return [(7,19), (4,7)] for k=2 or higher while the code should return [(7,19)] for k=1.



t	Price
0	10
1	11
2	12
3	13
4	12
5	10
6	9
7	13
8	14
9	15
10	18
11	19
12	20
13	22
14	23
15	20
16	17
17	16
18	19
19	24
20	23

C102. In the attached file, there are ten 200 by 200 matrices. Using the matrices, write a program as described. Each matrix is named A, B, C, D, ..., J respectively. When X(i,j) is the mean of A(i,j), B(i,j), C(i,j), D(i,j), ..., J(i,j) and Y(i,j) is the median of A(i,j), B(i,j), C(i,j), D(i,j), ..., J(i,j), you are required to calculate 200 by 200 matrices for mean(X) and median(Y). Notice when there are erroneous inputs, you need to correct the input values to be finite real numbers.

The attached file has the following format. Elements of each matrix are separated by "," Each row is separated by ";", and each matrix is separated by "|". For example, three 2 by 2 matrices are as follows.

1,2;3,4|0,1;1,0|5,3;3,1|

-END-