

FE 513 Practical Aspects of Database Design

Assignment I

Before starting this assignment, please go through the 'Google's R coding style guide' (available under Week 2 module on Canvas). Please submit R code as well as a pdf report file, containing results and the corresponding R code. Rmarkdown is highly recommended for generating the report. More guideline on using Rmarkdown please see help documentation under week 3 module.

1 Part I

1.1 Vector

1. Create 2 vector, each containing 10 random numbers.
2. Append the second vector to the first one.
3. Calculate the mean of the new combined vector.
4. For each number in the new combined vector, if it is larger than the mean then print out a 'True', otherwise print out a 'False'.

1.2 Matrix

1. Create a vector with 100 random numbers.
2. Transfer the above vector into a 10 by 10 matrix M .
3. Find the transposed matrix M^T . Print the value of element who is in the second row and the first column of M^T .
4. Write a nested loop to calculate the inner product between M^T and M . The result is also a matrix $N = \langle M^T, M \rangle$.
5. Calculate the same inner product using operator $\%*\%$. And compare two results.

1.3 Function

1. Load the given CSV file in R
2. Delete the columns containing NA(empty values).

3. Calculate daily log return for each stock. (**Hint.** log return is defined as $r_t = \ln\left(\frac{P_t}{P_{t-1}}\right) = \ln(P_t) - \ln(P_{t-1})$, where P_t is the stock price at time t .)
4. Calculate the mean and standard deviation of log return for each stock. Transfer the result into a 2 by N data frame (N is the number of stocks). (**Hint.** Function 'apply' is a better choice than 'for loop'.)
5. Build a graph with two sub-plots. In the first sub-plot, plot the first three stocks' daily prices. The y axis is stock price and x axis is date. In the second sub-plot, build a scatter plot of the statistical result you calculated above. In other words, the x-axis is the stocks' names and the y-axis is the statistical values. (**Notes.** Please include legend, title, and axis labels for each sub-plots.)

Bonus (10 pt)

Do problems in section 1.3 again using 'pipe'.

2 Part II

Try to use package 'quantmod' to download data, and search functions in 'quantmod' package to do following tasks.

1. Download Amazon daily stock price data from 2021-01-01 to 2021-12-31. And save the data to a csv file.
2. Calculate weekly log returns based on adjusted close price.
3. Calculate median, mean, standard deviation of log returns.
4. Plot the distribution of stock daily log returns (**Hints.** Histogram is a good choice to show the distribution. Controlling the number of bins in histogram can achieve a good resolution).
5. Count how many observation in this series whose log return is between 0.01 and 0.015.