Data Analysis Project [15 marks]

Due date: 5pm, 29 April 2020

**Requirement: Pls finish the tasks according to the requirements. All the tasks need to be finished by using R. Pls present your results in the word file, copy all your R code in the end of this word file, and then submit your word file via the Turnitin link in iLearn.**

**Task 1: Stock Return and Portfolio Analysis**

In the file named as “Stock.csv”, you have been provided with the daily prices of three stocks from 2012 to 2018.

1. Plot and present the stock prices in time series with appropriate labels. [1 mark]
2. Calculate the log returns of all the three stocks and express them **in percentages**.

Pls report the descriptive statistics of log returns of three stocks in Table 1. Pls change the names in Table 1 to the stock names in your file. [1 mark]

Table 1: Summary Statistics

|  |  |  |  |
| --- | --- | --- | --- |
| Variable | **Stock 1** | **Stock 2** | **Stock 3** |
| Min |  |  |  |
| Max |  |  |  |
| Mean |  |  |  |
| Median |  |  |  |
| SD |  |  |  |
| Skewness |  |  |  |
| Kurtosis |  |  |  |
| N |  |  |  |

1. Pls report the **correlation** matrix of log returns of three stocks in Table 1. Pls change the names in Table 1 to the stock names in your file. [1 mark]

Table 2: Correlation Matrix

|  |  |  |  |
| --- | --- | --- | --- |
| Variable | **Stock 1** | **Stock 2** | **Stock 3** |
| **Stock 1** |  |  |  |
| **Stock 2** |  |  |  |
| **Stock 3** |  |  |  |

1. There is a file named as “FF3factors.csv” in your folder. Merge your stock data with the Fama French three factors data, which are all in **percentages**. Pls create a new variable Stock.Rf for all the three stocks, which equals to the difference between log return of each stock and RF (available in the FF3factors.csv), and run a regression of :

.

Report the regression results for all the three stocks in tables (you can use only one table to summarize all the results or three separate tables). [1 mark]

1. Pls run regressions of:

.

Report the regression results for all the three stocks in tables (you can use only one table to summarize all the results or three separate tables). [1 mark]

1. Pls make your comments by comparing the results in *d* and *e* [1 mark].
2. If an investor would like to form a portfolio with a targeted expected return of 0.08% and achieve the minimized standard deviation by investing in these three stocks in the “Stock.csv” file. If the daily risk free rate is 0.02%, what is the Sharpe ratio of this optimal portfolio, given there is **no short sale constraint**? [1 mark] [hint: can use the library of “quadprog”] [pls provide your R code used to form the optimal portfolio in the end of the word file].
3. If an investor would like to form a portfolio with a targeted expected return of 0.08% and achieve the minimized standard deviation by investing in these three stocks. If the daily risk free rate is 0.02%, what is the Sharpe ratio of this optimal portfolio, given there is **short sale constraint** (i.e., you cannot short sell the stocks)? [1 mark] [hint: can use the library of “quadprog”]
4. By comparing your answer in g) and h), is there any difference in their Sharpe ratios? Why that’s the case? [1 mark]
5. Given there is **no short sale constraint**, pls draw an efficient frontier that satisfies the following conditions: 1. The range of the mean return of the portfolio is from 0.75\*min of the mean return of three stocks to 1.25\*max of the mean return of three stocks; 2. Pls create 500 optimal portfolios in this range; 3 Then pls plot all the risk-return combination of the 500 optimal portfolios (i.e., the efficient frontier). [1 mark] [Pls provide your R code in the end of word file]

**Task 2: Stock Market Reaction to FOMC Announcement**

US monetary policy can be measured by changes in the Federal funds rate[[1]](#footnote-1), which are announced by the Federal Open Market Committee (FOMC) meetings. Given the importance of monetary policy to the whole economy, investors pay special attention to the FOMC announcements. In the attached dataset (“Event Study.csv”), we have collected data on FOMC announcements from 1989 to 2007 and the aggregate stock market reactions to it.[[2]](#footnote-2)

Here is a list of variables:

* Date: FOMC announcement date
* Return: S&P 500 index return on the announcement date (in percentage)
* Total change: change of Federal fund rate (in bps)
* Expected: investors’ expected change of Federal fund rate, which is estimated based on the futures contracts written on the Federal fund rate (in bps)
* Surprise: the difference between Total change and Expected, which measures the change that is out of expectation of investors (in bps)
* Scheduled: equals to 1 if the FOMC is scheduled, and 0 if not.

1. **Run three regressions and report the estimation results according.**

[Note: please report the coefficient and t-statistics, indicate significance with difference levels, and the adjusted R-squared of the regression.] **(1.5 marks) [pls provide your R code in the end of word file]**

In the first regression, Y is Return, X is Total change;

In the second regression, Y is Return, X is Expected;

In the third regression, Y is Return, X is Surprise.

1. **Comparing the three regression results, which X variable explains Y the best? (0.5 mark)**

**Why that is the case?** [Note: you may link it with market efficiency and event study] **(1 mark)**

1. **Create a variable which equals to Scheduled\*Surprise, and then run a regression in**

**which Y is Return, the first X variable is Surprise, and the second X variable is Scheduled\*Surprise. Report the regression results. (1 mark) [pls provide your R code]**

1. **Based on the results, is the effect of Surprise on Return dependent on whether the FOMC meeting is scheduled or not? Please explain. (1 mark)**

**R code:**

1. <https://en.wikipedia.org/wiki/Federal_funds_rate> [↑](#footnote-ref-1)
2. Bernanke and Kuttner (2005) attached is a good reference. [↑](#footnote-ref-2)