R Analysis Project

**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. 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 | **MSFT** | **MMM** | **INTC** |
| Min | -12.10% | -7.08% | -9.54% |
| Max | 9.94% | 5.74% | 10.03% |
| Mean | 0.09% | 0.06% | 0.05% |
| Median | 0.05% | 0.08% | 0.07% |
| SD | 0.0145 | 0.0104 | 0.0148 |
| Skewness | -0.084 | -0.666 | -0.055 |
| Kurtosis | 11.24 | 8.32 | 8.12 |
| N | 1758 | 1758 | 1758 |

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 | **MSFT** | **MMM** | **INTC** |
| **MSFT** | 1 | 0.4738475 | 0.5441834 |
| **MMM** | 0.4738475 | 1 | 0.4900549 |
| **INTC** | 0.5441834 | 0.4900549 | 1 |

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 StockR 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:

,

where is the market excess return. Report the regression results for all the three stocks in tables (you can use only one table to summarize all the results or use three separate tables). [1.5 mark]

|  |
| --- |
| Stock MSFT  lm(formula = StockMSFT ~ MKTRF, data = ffstocklr)  Residuals:  Min 1Q Median 3Q Max  -0.120776 -0.005247 -0.000137 0.005126 0.098733  Coefficients:  Estimate Std. Error t value Pr(>|t|)  (Intercept) -0.0013111 0.0002652 -4.943 8.43e-07 \*\*\*  MKTRF 0.0117137 0.0003190 36.725 < 2e-16 \*\*\*  ---  Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1  Residual standard error: 0.0111 on 1756 degrees of freedom  Multiple R-squared: 0.4344, Adjusted R-squared: 0.4341  F-statistic: 1349 on 1 and 1756 DF, p-value: < 2.2e-16 |

|  |
| --- |
| Stock MMM  lm(formula = StockMMM ~ MKTRF, data = ffstocklr)  Residuals:  Min 1Q Median 3Q Max  -0.070824 -0.003346 0.000502 0.004046 0.053081  Coefficients:  Estimate Std. Error t value Pr(>|t|)  (Intercept) -0.0014784 0.0001838 -8.041 1.62e-15 \*\*\*  MKTRF 0.0092198 0.0002211 41.706 < 2e-16 \*\*\*  ---  Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1  Residual standard error: 0.007695 on 1756 degrees of freedom  Multiple R-squared: 0.4976, Adjusted R-squared: 0.4973  F-statistic: 1739 on 1 and 1756 DF, p-value: < 2.2e-16 |

|  |
| --- |
| Stock INTC  lm(formula = StockINTC ~ MKTRF, data = ffstocklr)  Residuals:  Min 1Q Median 3Q Max  -0.086934 -0.005501 0.000362 0.005965 0.086718  Coefficients:  Estimate Std. Error t value Pr(>|t|)  (Intercept) -0.0016500 0.0002839 -5.812 7.31e-09 \*\*\*  MKTRF 0.0112439 0.0003414 32.938 < 2e-16 \*\*\*  ---  Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1  Residual standard error: 0.01188 on 1756 degrees of freedom  Multiple R-squared: 0.3819, Adjusted R-squared: 0.3815  F-statistic: 1085 on 1 and 1756 DF, p-value: < 2.2e-16 |

1. Please 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). [1.5 mark]

|  |
| --- |
| StockMSFT  lm(formula = StockMSFT ~ MKTRF + SMB + HML, data = ffstocklr)  Residuals:  Min 1Q Median 3Q Max  -0.119932 -0.005537 -0.000272 0.004885 0.095255  Coefficients:  Estimate Std. Error t value Pr(>|t|)  (Intercept) -0.0013567 0.0002570 -5.278 1.47e-07 \*\*\*  MKTRF 0.0120516 0.0003151 38.248 < 2e-16 \*\*\*  SMB -0.0042517 0.0005384 -7.897 4.99e-15 \*\*\*  HML -0.0046318 0.0005453 -8.494 < 2e-16 \*\*\*  ---  Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1  Residual standard error: 0.01076 on 1754 degrees of freedom  Multiple R-squared: 0.4698, Adjusted R-squared: 0.4689  F-statistic: 518 on 3 and 1754 DF, p-value: < 2.2e-16 |

|  |
| --- |
| Stock MMM  lm(formula = StockMMM ~ MKTRF + SMB + HML, data = ffstocklr)  Residuals:  Min 1Q Median 3Q Max  -0.071312 -0.003185 0.000484 0.003914 0.052442  Coefficients:  Estimate Std. Error t value Pr(>|t|)  (Intercept) -0.0015012 0.0001817 -8.264 2.74e-16 \*\*\*  MKTRF 0.0095133 0.0002227 42.719 < 2e-16 \*\*\*  SMB -0.0024606 0.0003805 -6.467 1.29e-10 \*\*\*  HML 0.0003403 0.0003854 0.883 0.377  ---  Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1  Residual standard error: 0.007602 on 1754 degrees of freedom  Multiple R-squared: 0.5103, Adjusted R-squared: 0.5094  F-statistic: 609.2 on 3 and 1754 DF, p-value: < 2.2e-16 |

|  |
| --- |
| Stock INTC  lm(formula = StockINTC ~ MKTRF + SMB + HML, data = ffstocklr)  Residuals:  Min 1Q Median 3Q Max  -0.087409 -0.005584 0.000413 0.005989 0.086069  Coefficients:  Estimate Std. Error t value Pr(>|t|)  (Intercept) -0.0016711 0.0002827 -5.912 4.05e-09 \*\*\*  MKTRF 0.0114145 0.0003465 32.940 < 2e-16 \*\*\*  SMB -0.0020034 0.0005921 -3.384 0.000731 \*\*\*  HML -0.0018333 0.0005997 -3.057 0.002267 \*\*  ---  Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1  Residual standard error: 0.01183 on 1754 degrees of freedom  Multiple R-squared: 0.3882, Adjusted R-squared: 0.3871  F-statistic: 371 on 3 and 1754 DF, p-value: < 2.2e-16 |

1. Pls make your comments by comparing the results in *c* and *d.* [1 mark].

[Hint: which model can fit the data better?]

As we can see on the summary table of the models in c and d, it is suggested that adding SMB and HML to the model is better than using Mkt-Rf alone because the p-value of SMB and HML for all 3 stocks is very small 🡪 there is a significant improve in fitting the data compared to only using variable Mkt-Rf

1. If an investor would like to form a portfolio with a targeted expected return of 0.05% 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].

|  |  |  |
| --- | --- | --- |
| MSFT | MMM | INTC |
| -14.85% | 84.30% | 30.55% |

The Sharpe ratio of this optimal portfolio is 0.02817

1. If an investor would like to form a portfolio with a targeted expected return of 0.05% 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”]

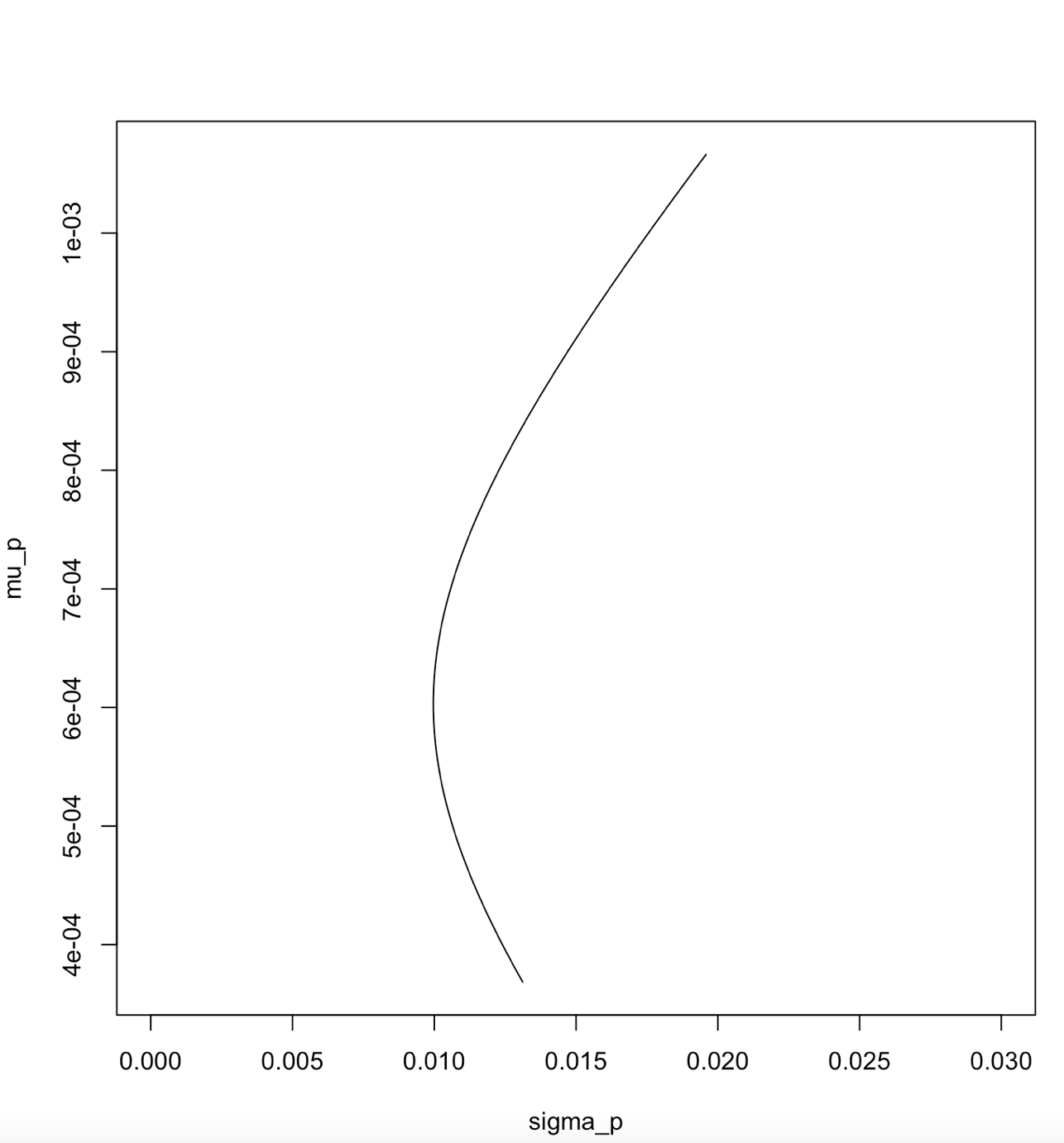
|  |  |  |
| --- | --- | --- |
| MSFT | MMM | INTC |
| 12.55% | 73.98% | 13.47% |

Sharpe ratio of this optimal portfolio is 0.03007

1. By comparing your answer in g) and h), is there any difference in their Sharpe ratios? Why that’s the case? [1 mark]

The Sharpe ratio of portfolio in h is higher than that of portfolio in g. The difference is that there is a short sale constraint 🡪 different weight of each stock in the two portfolio 🡪 different return and variance.

1. 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).



**Task 2: Loop and Conditional statement**

1. Write a program in R to print a pattern with n rows like [2 marks]:

1

2 3

6 5 4

7 8 9 10

15 14 13 12 11

…

1. Consider a Pokemon card collection game [3 marks]:

Assume there are totally 48 different types of Pokemon. And for each type there is an infinite number of cards for sale. Each time when you buy a Pokemon card, you can only get a random type (with an equal probability for any type). You want to know on average how many cards you need to buy to collect a full set of 48 Pokemon cards.

