Curtin University

ECOM 2001 Term Project Description

Due 30 September at 9:00AM AWST

Introduction

The aim of this project is to prepare, evaluate and analyse stock market data and to recommend an optimal portfolio consisting of two stocks. You have been assigned three stocks, all three must be included in the analysis which works towards your recommendation of a final optimal portfolio. The project requires a deep understanding of both the statistics and the mathematics components of this unit. It is recommended that you work on this on a weekly basis.

YOU MUST USE THE STOCKS ASSIGNED TO YOU. Any deviation from the assigned stocks will results in a grade of zero.

Refer to the rubric at the end of this document to understand how this assessment will be graded. In particular, note that all figures need to be numbered and labelled, and you need to include all the steps to involved with arriving at each of your answers.

Your final report should be a pdf document. An RMarkdown document to get you started is available on the unit Blackboard site. Show all of your coding by keeping `echo = TRUE`. Make sure to update your name and student ID in the YAML of the document.

You are **NOT ALLOWED** to engage any AI-assistive platforms to complete this assessments, unless you are told otherwise (in 2 questions below).

1 Import Data (2 points)

Import the adjusted stock prices for the three stocks which you have been assigned. See the Markdown file for hints.

2 The Analysis

2.1 Plot prices over time (4 points)

Plot the **prices** of each asset over time separately.

Succinctly describe in words the evolution of each asset over time. All axes and figures have to be properly labeled and described. (limit: 100 words for each time series).

2.2 Calculate returns and plot returns over time (4 points)

Calculate the **daily percentage returns** of each asset using the following formula:

$$r_t = 100 * \ln\left(\frac{P_t}{P_{t-1}}\right)$$

Where P_t is the asset price at time t. Then plot the **returns** for each asset over time.

2.3 Histogram of returns (6 points)

Create a **histogram** for **each of the returns series**.

You have to explain your choice of bins.

You will need to carefully label all axes and figures.

You are expected to: - Write a short paragraph to describe the trend of each time series; - Discuss the formula to calculate the bins.

(Hint: Discuss the formula you use to calculate the bins)

2.4 Summary table of returns (5 points)

Report the descriptive statistics **in a single table** which includes the **mean, median, variance, standard deviation, skewness and kurtosis** for each series. All tables need to be correctly labeled.

What conclusions can you draw from these descriptive statistics?

2.5 Are average returns significantly different from zero? (6 points)

Under the assumption that the **returns of each asset** are drawn from an **independently and identically distributed normal distribution**, are the expected returns of each asset statistically different from zero at the 1% level of significance?

Part 1: Provide details for **all 5 steps to conduct a hypothesis test**, including **the equation for the test statistic**. All steps have to be shown and this part has to be repeated for each hypothesis test. (1 points)

Part 2: Calculate and report all the relevant values for your conclusion and be sure to provide an interpretation of the results. (Hint: you will need to repeat the test for expected returns of each asset) (3 points - one for each stock)

Part 3: If you would have done this question using Chat-GPT, what answer will you get? (hints: you will need to describe how you **prompt** the question in Chat-GPT to guide the answer (1 point), would expect your answer to be different or similar to your answer above and provide your rationale? (1 point))

2.6 Are average returns different from each other? (7 points)

Assume the **returns of each asset** are **independent from each other**. With this assumption, are the mean returns statistically different from each other at the 1% level of significance?

Provide **details for all 5 steps to conduct each of the hypothesis tests** using what your have learned in the unit. All steps have to be shown and this part has to be repeated for each hypothesis test. (2 points)

Calculate and report all the relevant values for your conclusion and be sure to provide and interpretation of the results. (*Hint: You need to discuss the equality of variances to determine which type of test to use.*) (3 points)

If you have a chance to engage Chat-GPT, how would you approach this question? That is, you need to **clearly lay out ALL STEPS that you would ask the question to Chat-GPT**. (1 points)

Now, compare your answer to Chat-GPT, **why do you think your answer is different or similar?** Please attach a picture of the screenshot of the answer you have got from Chat-GPT. **What do you learn from this exercise?** (1 points)

2.7 Correlations (2 points)

Calculate and present the **correlation matrix of the returns**.

Discuss the direction and strength of the correlations.

2.8 Testing the significance of correlations (2 points)

Is the assumption of independence of stock returns realistic?

Provide evidence (the hypothesis test including **all 5 steps of the hypothesis test** and **the equation for the test statistic**) and a rationale to support your conclusion. All steps have to be shown and this part has to be repeated for each hypothesis test.

2.9 Advising an investor (12 points)

Suppose that an investor has asked you to assist them in choosing **two** of these three stocks to include in their portfolio. The portfolio is defined by

$$r = w_1 r_1 + w_2 r_2$$

Where r_1 and r_2 represent the returns from the first and second stock, respectively, and w_1 and w_2 represent the proportion of the investment placed in each stock. The entire investment is allocated between the two stocks, so $w_1 + w_2 = 1$.

The investor favours the combination of stocks that provides the highest return, but dislikes risk. Thus the investor's happiness is a function of the portfolio, r:

$$h(r) = \mathbb{E}(r) - \mathbb{V}ar(r)$$

Where $\mathbb{E}(r)$ is the expected return of the portfolio, and $\mathbb{V}ar(r)$ is the variance of the portfolio.[1]

Given your values for $\mathbb{E}(r_1)$, $\mathbb{E}(r_2)$, $\mathbb{V}ar(r_1)$, $\mathbb{V}ar(r_2)$ and $\mathbb{C}ov(r_1, r_2)$ which portfolio would you recommend to the investor? What is the expected return to this portfolio?

Provide evidence to support your answer, including all the steps undertaken to arrive at the result. You will need to solve the optinmisation problem using pen and paper, and you need to typeset your answer. You can then scan as picture to attach here as your answer. You can show the summary statistics using the coding learned in class, but the optimisation problem has to be solved by hand.

You will need to get your instructors to validate your identity of your work by asking them to sign your work when you complete it. Without their validation, you will automatically get a zero for this question.

Note: You will need to typeset your answer. Then, you need to put your name and student ID number on every page (and side) of your work. You will have the instructor to validate your information by signing your answer sheet. Then, you can scan the answer as picture(s) and embed it here as your answer.

Submission

1. Submit the pdf output of your completed project to the Turnitin.com link on the BlackBoard site for our unit.
    i. Keep the sections as they are in this document
    ii. Ensure that all Figures and Tables are numbered, and have appropriate captions.

[1] Note that $\mathbb{E}(r) = w_1 E(r_1) + w_2 \mathbb{E}(r_2)$, and $\mathbb{V}ar(r) = w_1^2 \mathbb{V}ar(r_1) + w_2^2 \mathbb{V}ar(r_2) + 2 w_1 w_2 \mathbb{C}ov(r_1, r_2)$

iii. **All** your calculations and steps used to produce the results should be included. So include any mathematical calculations and set `echo=TRUE` in all of your code chunk headers, including those used to generate figures.

2. Additional details

   - All results (numbers) should be accurate to 3 decimal places.
   - Proof-read your report - do not include spelling or grammatical errors.

Rubric

The submission is worth 50 Points in total and will be worth 50% of your final grade.

Table 1: Rubric

Question	(Maximum Score)	Fail (<25)	Pass (25)	Meets Expectations (25-40)	Above Expectations (40-50]
1	2	(0/2) The data are not imported into R, or the incorrect stock symbols were imported.	(1/2) The data were imported but the code or assigned symbols are not clear.	(2/2) The assigned stocks are correctly imported and identified in the report title.	
2.1	4	(0/4) A time series plot of the prices of each stock is missing.	(2/4) Plots are present but with omitted details or formatted poorly. Explanations are unclear and/or have spelling and grammatical errors. Coding is omitted.	(3/4) Plots are clear, but missing components such as captions or numbers. Explanations are present but may have spelling, grammatical or other minor errors. Coding is present but inadequately commented.	(4/4) Plot axes are labelled and the plot has an appropriate caption and Figure number. Explanations are clear, concise and free of spelling and grammatical errors. Plot coding is clear and commented.
2.2	4	(0/4) The calculation of returns is not present, no time series plot of the returns of each stock is included.	(2/4) Calculation of the returns is present, but may include errors. A time series plot is included by is missing details or poorly formatted. Coding is not commented.	(3/4) Calculation of returns is present and correct. A time series plot of the returns is present but details such as a caption and Figure number are missing. Code may be partially commented.	(4/4) Returns are correctly calculated and coding is clearly commented. A well-formatted time series plot of the returns to each stock is present, axes are labelled, the figure has a caption and a Figure number.
2.3	6	(0/6) A histograms of the returns are missing.	(3/6) A histogram of the returns is present for each stock, but is missing key details such as axis labels, caption or figure number. The selection of the number of bins to include is not discussed. Code may be uncommented.	(5/6) A histogram of returns is present for each stock, but some details may be missing. The selection of the number of bins to include is discussed, but may be poorly motivated or contain spelling or grammatical errors. Code is partially commented.	(6/6) Histograms of the returns to each stock are present, clearly formatted and include figure captions and numbers. A discussion of the method used to select the number of bins in the histogram is included. Code is commented. No spelling or grammatical errors are present.
2.4	5	(0/5) A summary table of the returns to each stock is missing.	(2.5/5) A summary table of the returns to each stock is included but omits key summary statistics (or included summary statistics which were not requested), poorly formatted and/or missing a table caption and number. Code is absent or uncommented. No conclusiosn or erroneous conclusions are drawn.	(4/5) A summary table of the returns to each stock is included and is formatted appropriately. Not more than 1 summary statistic is absent and no additional summary measures are included. Coding is commented but is missing some details. An explanation is provided but may have minor errors and/or spelling and grammatical errors.	(5/5) A well formatted summary table is presented which includes the requested summary statistics for each stock.The table has an appropriate caption and Table number. Code is well commented. Explanations are clear, concise and without spelling or grammatical errors.
2.5	6	(0/6) An appropriate statistical test is missing.	(3/6) Appropriate statistical tests are present for each stock but the results are poorly formatted. The steps involved in the hypothesis test procedure are present with errors. The equation of the test statistic contains errors. The results are partially interpreted. Coding is not commented.	(5/6) Appropriate statistical tests are present for all three stocks and the results are appropriately formatted. All five steps of the hypothesis testing procedure are present. The equation for the test statistic is present. Inference is drawn. Code is at least partially commented.	(6/6) An appropriate statistical test is present for each stock and the results appropriately formatted. All five steps of the hypothesis testing procedure are present and correct. The equation for the test statistic is present and correct. Inference is correct and well presented. Coding is clearly commented.

Table 1: Rubric *(continued)*

Question	(Maximum Score)	Fail (<25)	Pass (25)	Meets Expectations (25-40)	Above Expectations (40-50]
2.6	7	(0/7) No statistical tests for differences between average stock returns are present.	(3.5/7) Statistical tests for differences between average stock returns are presented but missing key details. The testing procedure is partially explained and equations for test statistics may be present but contain errors. Results are poorly formatted. Coding is incompletely commented.	(6/7) Appropriate statistical tests are presented and explained but poorly formatted. Hypothesis testing procedures are included and coding is at least partially commented. Inferences are drawn but may contain minor errors.	(7/7) Appropriate statistical tests are presented and well formatted. All five steps of the hypothesis testing procedures are clearly explained. Inference drawn from the test(s) is correct and informs any follow up tests. Code is fully commented.
2.7	2	(0/2) Correlations between stocks are missing.	(1/2) Correlations are present but formatted poorly. The direction and strength of correlations are discussed but may contain errors in interpretation.	(1.5/2) A correlation matrix is present and appropriately formatted. Direction and strength of correlation is discussed but may contain minor errors in interpretation, spelling or grammar. Code is present and at least partially commented.	(2/2) A well formatted correlation matrix is presented and discussed. Interpretations are accurate and free of spelling and grammatical errors. Code is well commented.
2.8	2	(0/2) No test for the significance of correlations is present.	(1/2) A test for the significance of correlation is present but poorly formatted and or incompletely explained (steps to the hypothesis testing procedure and an equation for the test statistic are incomplete). Inference may be incorrect.	(1.5/2) Tests for the significance of correlations is present and appropriately formatted. All five steps of the hypothesis testing procedure are present as well as the equation for the test statistic. Inference drawn may contain minor error including spelling and/or grammatical errors.	(2/2) A complete set of tests for significance of correlations between stock returns is present, all 5 steps to the hypothesis testing procedure are documented and include the test statistic. The results are well formatted, code is well commented and inferences drawn are correct and free of spelling and grammatical errors.
2.9	12	(0/12) No response is present.	(6/12) Some attempt to explain the mathematical process for arriving at the optimal portfolio is presented but may be poorly formatted or contain errors. A reccomendation is made, but may be accompanied by a weak rationale. Results of all portfolios are present but may be poorly formatted. Coding may be present but incompletely commented.	(8/12) The mathmatical process for arriving at the optimal portfolio is provided but may contain minor errors or lack appropriate supporting tests. Results for all portfolios are presented. A recommendation is provided and supported by a rationale. The answer may contain spelling and grammatical errors. Code is partially commented.	(12/12) A complete explanation of the process for arriving at the optimal portfolio is explained, and the result is supported by appropriate follow up tests. Coding is well commented and intelligently applied. Results are presented in a well formatted table that includes the optimal proportion to invest in each stock, as well as the expected return for each possible portfolio. A recommendation for the optimal portfolio is given and supported by a clear rational. No spelling or grammatical errors are present.
Total	50	-	-	-	-