PROG1000 – Introduction to Programming

Semester assignment 2022

Handed out: 4 April 2022 at Canvas

Handed in: 26 April 2022, 15:00 in Inspera.

Lecturer: Erik Smith-Meyer

(Last update: 4 April 2022)

It is possible to hand in the files (pdf + py/ipynb) from a week before deadline.

You will produce two outputs:

- **Report**: A written report of 2000-3000 words, containing analysis and graphs. Upload this report as a .pdf file.
- **Code:** You have to upload the file with the code behind your analysis. This will either be a file created by Jupyter notebook (.ipynb) or a file created by Spyder (.py).
- Please write the group name and the names of all of the group members on the first page of the report (and not the candidate numbers from Studentweb)

You work as an analyst at Tryad Asset Management when your boss, Mr. Bing, asks you to analyse the effect of different business strategies and product strategies on firm performance. This will help our private equity branch to unlock value from re-structuring our privately owned businesses.

According to Porter (1980), a firm can make two decisions regarding product strategies, *differentiation*, or *cost leadership*. If a firm produces different products than other firms, then this is more expensive, but the firm might be able to charge a higher price for them. If a firm produces the same products as other firms, then the firm must focus on keeping costs lower than other firms for them to get a higher profit.

According to Zott and Amit (2007), a firm can make two decisions regarding the business model, *novelty,* or *efficiency*. If a firm tries to sell its products to the customers in a new way, hoping for increased prices, then the firm follows a novelty business model. If a firm tries to sell its products to the customers in the same way as before, but at a lower cost, then the firm is following an efficiency business model.

The data base administrator, Mrs. Green, has downloaded two data sets for you from the data base. These might help you reach some conclusions:

- The file "market_cap.csv" contains the following variables:
    - Symbol – This is the code for the company in the market. This is unique for each of the companies.
    - Name – The name of the company.
    - Sector – The business sector in which the company operates. The data contains firms from three sectors.
    - FirmV2017 – The value of the firm in April 2017.
    - FirmV2022 – The value of the firm in April 2022.
- The file "survey.csv" contains the following variables:
    - Symbol – This is the code for the company in the market. This is unique for each of the companies.
    - Name – The name of the company.
    - Differentiation – The *differentiation* score of the company.
    - Cost leadership – The *cost leadership* score of the company.
    - Efficiency – The *efficiency* score of the company.
    - Novelty – The *novelty* score of the company.

The scores in "survey.csv" comes from a survey of market experts and each firm receives a score between 0 and 100 on the four different strategy metrics.

1)

a) Import the file "sp500_marketvalues.csv".
b) Import the file "sp500_survey.csv"
c) Merge the two data sets. Use "Symbol" as a common column.
d) Drop any NaNs.
e) Create a new variable called "returns" which is the yearly stock return of each firm in the data from April 2017 to April 2022. The formula for the returns is:

$$returns = \left(\frac{FirmV2022}{FirmV2017}\right)^{\frac{1}{5}} - 1$$

2) Explore the data set.
   a) Present descriptive statistics for the data set.
   b) Create a figure with four sub-plots in two rows and two columns. "returns" should be on the y-axis and differentiation, cost leadership, efficiency, and novelty should be on the x-axis. Add the correlation coefficients in the subplot titles.

3)
   a) Create a bar plot to show the average yearly returns for each sector.
   b) Create a plot with three subplots, one for each of the sectors. The subplots should be a histogram of the yearly returns within each sector.

4)
   a) Estimate the regression model:

$$returns = \alpha + \beta_1 Differentiation + \beta_2 Cost\ leadership + \beta_3 Efficiency + \beta_4 Novelty + \epsilon$$

   How much of the variation in returns does the model explain?
   b) If a company in the financial sector has a score of:
   Differentiation: 54
   Cost leadership: 61
   Efficiency: 57
   Novelty: 42
   What would be the expected returns?
   c) If the company in 4) b) managed to change the way it sold the products to the customers and thereby increased novelty to 55. What would be the change in returns?

5)
   a) Split the data into three parts, one for each of the sectors (Financials, Industrials, Information Technology).
   b) Run a regression on each of the sub-sets of data. Does any of these regressions obtain a higher adjusted R^2 than the regression in 4) a) which used the whole data set?

6)

a) Create a function which takes in two arguments (y) and (x), where y is the dependent variable and x is the explanatory variables. The function should run a regression and return the models' adjusted R^2.

Write a report where you:

Present your investigation and analysis of how product and business strategy are related to business performance. The report must be between 2000 and 3000 words, and it should provide answers to all the tasks above. Beside the defined tasks, you are of course free to include what you have found out from the data sets that is not a part of the task descriptions.

The pdf report should not include any code. Instead, you should explain and summarize your results in a way that the reader is able to understand the report without any knowledge of Python or your Jupyter notebook. Remember to not just present the results from your Jupyter notebook, the report should also explain those results. For example, it is not sufficient to simply include graphs, but you must also explain them (e.g. this figure shows…, in this figure we have plotted… etc.).