**SCHOOL OF COMPUTING (SOC)**

**IT8701 Introduction to Programming for Data Science**

**Self Reflection (CA2)**

|  |
| --- |
| **Instructions:**   1. Submit this together with your other deliverables at Polymall “Assignments->CA2” folder 2. Name your file “YourStudentID-YourName-YourLecturer.docx” |

|  |  |
| --- | --- |
| **Your Lecturer’s Name** | Chee Seng Chong |
| **Your Name** | Li Hanwen Marcus |
| **Your Student ID** | P747480 |
| **Your Class** | IT8701 |

# QUESTION 1: RATE THE EFFORTS AND COMPETENCY THAT IS DEMONSTRATED IN THIS ASSIGNMENT

Tick in the column that best describes the efforts, technical competency and depth of data analysis that is demonstrated in this assignment.

Justify your rating in the second and third questions below

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **WAY Above Average** | **Above Average** | **Average** | **Below Average** | **Way Below Average** |
| Coding |  | x |  |  |  |
| Analysis |  | x |  |  |  |

# QUESTION 2: JUSTIFICATION FOR RATING GIVEN FOR CODING

Please provide evidence that you have met the requirements (AVERAGE) or if you think your submission is above average or even above average, state details of what you have done here so that your lecturer does not miss out the efforts you have put in for this assignment. For CA2, the basic requirements are to produce 4 different graphs with at least 3 datasets.

|  |
| --- |
| I think seaborn has definitely made the coding for exploratory data and for managing the csv files much more easier than just using numpy. So instead of something more intensive like numpy, this time it was just going through forums to see how I could make a graph work. Having said that I did quite enjoy making graphs using facetgrid as it was able to make a data set with multiple variables in different columns easily translated into 1 graph |

# QUESTION 3: JUSTIFICATION FOR RATING GIVEN FOR DATA ANALYSIS

Please provide evidence that you have met the requirements (AVERAGE) or if you think your submission is above average or even above average, state details of what you have done here so that your lecturer does not miss out the efforts you have put in for this assignment.

|  |
| --- |
| I think exploratory data wise it has been good, however now knowing that there are many types of tests in statistics, I’m wondering how we can apply those tests in our data analysis using python. Because of this lack of statistical analysis, it feels like the data analysis is incomplete. |

# QUESTION 4: YOUR FUTURE PLANS

How do you rate your programming competency with data analysis tasks after completing this assignment? Give yourself a rating from 0 to 10.

|  |
| --- |
| 7 |

After finishing the PDC1 of your Specialist Diploma, which do you think you prefer or is stronger at? The Statistics or Programming portion? How has this realisation affected your mindset of a Data Science job? Do you enjoy a Data Science role that mainly involves application of lots of statistical concepts (improving predictive algorithms for instance) or one that requires a lot of programming (e.g. code to acquire or clean data) or perhaps both equally excite you? 😊

|  |
| --- |
| I think the statistics part is much more difficult to comprehend, but I also understand that it is the analysis part of data analysis, whereas programming is used to help clean/write code for graphs. I’m not to excited about having to clean data but I do understand that both has to come hand in hand. But I think I wouldn’t be pursing a career like a data engineer. |

Are there any useful skills that you gained from this module? Share how you think the skills you learnt from this module can be applied in your current job or in a future career / job change.

|  |
| --- |
| I was hoping that I was able to learn mysql and nosql hands on but the module was that there were very little help with installing those programs and using them on a mac. I understand they were supplementary classes but I think that was one thing that I really wanted to learn the most. |

What was not taught in this module, but you wish to learn? How do you plan to learn these missing skills?

|  |
| --- |
| I think what is missing the most is using python for statistical tests such as anova, chi-square, t-tests etc. |

**-- End of Self-Reflection --**