**SCHOOL OF COMPUTING**

**Programming for Data Science**

**Self-Reflection (CA2)**

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

|  |  |
| --- | --- |
| **Your Name** | Jurgen Tan Yu Teng |
| **Your Student ID** | 1715217 |
| **Your Class** | DBIT 3B01 |

# QUESTION 1: CHALLENGES - SELF-REFLECTION FOR CA2

Provide a brief reflection of the challenges you have faced in this assignment (CA1) with a comparison of those you faced in CA1. Did you manage to apply any learning experiences from the previous assignment to improve the way you did this assignment?

|  |
| --- |
| One of the challenges that I faced was choosing of the dataset, as the dataset had many values and different categories. This made it hard to drill down as to which category to choose from. The dataset originally had data from ITE all the way up to Universities. Hence, I decided to use only the Universities data as it contained the main chunk of the dataset and since most of the audience of the presentation are going to be enrolling into University soon.  Subsequent challenge that I faced was different date ranges for each dataset. The datasets that were found on the data.gov website as inconsistent in the data ranges. Some of the datasets had data from the 1980s while come only had data for a single year. This made it really hard to draw a connection and inference based on data that are from different years as there are many other factors that could affect in different years.  Another challenge which I faced was the lack of examples and documentation from Seaborn. This could be a personal peeve, as I like to read proper documentations and methods of usage before trying it out. Hence, trying to find information about the library was quite challenging. However, in the end I had come to realized that it is very similar to Matplotlib and many codes that were used in Matplotlib could be used in Seaborn.  I had managed to apply the things that I learnt in the previous assignment in this assignment. By trying to add in more statistics and adding more aesthetically appealing graphs. |

# QUESTION 2: ACHIEVEMENTS - SELF-REFLECTION FOR CA2

Provide a brief reflection of what you think you have personally achieved in this assignment. Do not be shy to highlight extra efforts you have put in for this assignment. This is the place for you to impress upon your tutor on your competency in this area.

|  |
| --- |
| I feel that I had managed to achieve the skill of being able to code out proper graphs using different libraries and doing basic cleansing of unclean and unstructured data using the Pandas library. While doing this assignment, I had also managed to stumble upon the Plotly library, which allowed the graphs that were produced to be interactive and show the other statistics or information about the graph. Which proved to be very informative for me when I was trying to draw an inference and conclusion. I had also created a GitHub repository which can be accessed via  <https://github.com/jurgen-tan-git/PDS-Assignments> |

# QUESTION 3: EVALUATE YOUR CA2

What grade do you think you deserve for CA2? Justify the grade you gave yourself with reasons.

|  |
| --- |
| What grade do you think you deserve for CA2?  A  State your justifications here  I feel that I have applied the features and concepts that were taught in this module and fulfilled the basic requirements. On top of that, I had applied the fundamentals of learning how to use different libraries for the data cleansing and data visualization before analyzing the graphs. |

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