

SCHOOL OF COMPUTING

Programming for Data Science
Self-Reflection (CA1)**Instructions:**

1. Submit this at Blackboard “Assignments->CA1->Self-Reflection” folder
2. Name your file “YourModuleClass-YourStudentID-YourName.docx”

Name	Lim Chuan Hao
Student ID	1922264
Module Class	DIT/FT/1B/11

1 QUESTION 1: CHALLENGES - SELF-REFLECTION FOR CA1

Provide a brief reflection of the challenges you have faced in this assignment.

I faced challenges in both the technical aspects and in dealing with the data. For starters, I had to look at a lot of datasets from LTA to see which dataset seem interesting. Afterwards, I had to do some preliminary analysis on most of them before choosing which datasets to use.

I also faced challenges in trying to use numpy to do data manipulation. As for some operations that I wanted to do, I had to make my own utility functions. Some examples include merging, grouping and converting a column to DateTime.

As for matplotlib, a lot of plots had to be made manually. Such as adding labels, titles and legends.

Lastly, I also had trouble thinking of what plots to make and what kind of relationship I should be looking for. The conclusions I could draw from the plot also required some research and sometimes lead to more questions.

2 QUESTION 2: ACHIEVEMENTS - SELF-REFLECTION FOR CA1

Provide a brief reflection of what you think you have personally achieved in this assignment or the knowledge or skills you have found satisfaction in learning / acquiring.

I think I learned to work within the constraint of numpy, making my own functions to manipulate the data. In the end, I was able to get the data into the form that I wanted to plot the plots I wanted.

I also think I managed to get some meaningful and insightful conclusion from the data and plots I made. They also seem to correspond to what happened in the real world. An example would be seeing the trend in COE quotas, match up articles that Strait time would post on it.

Some of the plots I made also took quite some time, such as the 100% stacked area plot and dual-axis plot. I think those plots made the conclusion I wanted to make more apparent and served their purpose.

At the end of the day, I think taking each plot one at a time, and formulating questions to answer was what lead me to some of my conclusions. Understanding what conclusion could be drawn from each plot and asking more questions, got me to my final conclusions.

Personally, this assignment made me open my eyes to how hard it can be to do data science on real-world data. However, with enough patience and taking thing step-by-step, it can be possible to analyse real-world data.

-- End of Self-Reflection --