**SCHOOL OF DIGITAL MEDIA AND INFOCOMM TECHNOLOGY**

**IT8701 Introduction to Programming for Data Science**

**Self Reflection (CA1)**

|  |
| --- |
| **Instructions:**   1. Submit this at Polymall “Assignments->CA1->Self-Reflection” folder 2. Name your file “YourModuleLecturerName-YourStudentID-YourName.docx” |

|  |  |
| --- | --- |
| **Your Lecturer’s Name** | Dora Chua |
| **Your Name** | Tan Huimin Sherry |
| **Your Student ID** | P7319739 |
| **Your Class** | NSDAI |

# QUESTION 1: CHALLENGES - SELF-REFLECTION FOR CA1

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

|  |
| --- |
| 1. In the Singapore resident population data set, I wanted to visualize 2 variables of interest – population % change and population size for each town on the same chart. I found it difficult to order the data using only indexing as 3 different variables were involved (town, population change, and population). My solution was to use the *merge\_array* function from the *numpy recfunctions* library. This allowed me to create a new array allowing mixed data types and named columns 2. There was a lot of data for the resale transaction data set and it was a challenge deciding what to show to create a meaningful data narrative. Therefore, I focused on current data (e.g. I only used the latest data set). Analysis was then done on an aggregate level i.e. by comparing the highest resale price transactions and median values between towns to tease out interesting observations. Another option was to allow user-defined input to generate charts to visualize trends for different towns or flat types. 3. Plotting aesthetics for matplotlib was quite frustrating at times. As the charts would be used in a presentation, a lot of troubleshooting was needed to optimize padding between chart elements and font sizes. Finally, I used rcParams from matplotlib to set these plot attributes globally so that I didn’t have to keep bothering with them. Still, a lot of trial and error was required for some aesthetics that had to be tweaked for different charts. 4. I had to research unfamiliar python/numpy functions to achieve some of the aims for this assignment. A challenge was being able to recall them when the function was used more than once. |

# QUESTION 2: ACHIEVEMENTS - SELF-REFLECTION FOR CA1

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

|  |
| --- |
| 1. Learnt new functions not taught in class, such as *merge\_arrays from numpy.lib.recfunctions* and *rcParams* to set global plotting aesthetics, setting style sheets for plotting (though I eventually did not use them) 2. Obtained interesting insights by choosing appropriate visualizations and supplementing with other information sources (HDB’s and Stackedhome’s classification of geographical regions and estate maturity for each town) 3. Reinforced understanding and familiarity with functions taught in class by putting them to use in this assignment. |

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