# School of Computing
# EP0302 Programming for Data Science PDS
# Assignment Brief: CA1 AY2526S2

Last updated 21/10/2025

## Assignment Rubrics

By completing this assignment, you will have been able to:

- Demonstrate basic competency in writing a Python program in a notebook environment
- Demonstrate basic competency in using the Python NumPy and Matplotlib packages for data visualization
- Demonstrate basic competency in applying the insights gained from the outputs of your Python programs to deliver a useful data visualization report for your stakeholders

## Assignment Instructions and Guidelines

1) This is an **individual** assignment which requires you to complete a Python notebook data science project that retrieves data from a CSV data source and perform basic operations on the dataset such as, but not limited to, transformation, cleansing and visualization.
2) The requirements and deliverables are outlined in the following section of this document.
3) The deadline of this assignment is on 7th December 2025, 11.59PM.
4) Submissions should be made via the Brightspace CA1 assignment submission link by the stipulated dateline.
5) Your deliverables should be contained in a zip file with the following convention for its name:
   CA1-[Elective Class]-[Admin Number]-[Name].zip
   Example: CA1-01-23120303-JohnDoe.zip
6) Your zip file should include the following deliverables:
   - One Python notebook environment that accomplishes the given tasks in the assignment brief below using the Python programming language.
   - One HTML exported version of the Jupyter notebook.
   - One PDF version of the report template for you to complete
   - One Declaration of Academic Integrity (SOC version).
7) As part of the assignment requirements, you will be interviewed based on the Jupyter notebook environment that you have submitted. Your module tutor may ask you to reproduce certain parts of your code during this interview session. You do not need to reproduce the same code in question, but the code should be able to perform the same task. Usage of Google will be allowed during the process. Usage of Gen AI tools and any other form of AI assisted tools will not be allowed during the interview process.
8) This assignment will account for 40% of the module grade.
9) No marks will be awarded upon meeting any of the following conditions:
   a. Your work is copied/plagiarized, or was allowed to have been copied/plagiarized.
   b. You are unable to reproduce code or are unable to answer most of the interview process.
   c. You use prohibited packages in your assignment.
   d. Your zip file is corrupted (Please double check by downloading your work after submitting it).
   e. You submit your assignment more than one day late.
10) 50% of the marks will be deducted for assignments that are received within one (1) calendar day after the submission dateline. No marks will be given thereafter.

# Assignment Brief and Scope

## Introduction and Brief

This individual assignment is a data science project to be completed in a Python notebook environment, as well as a written report. For this assignment, you are required to complete a Python notebook data science study on HDB housing prices in Singapore. Questions and problems will be posed throughout the notebook for you to solve. Upon the completion of the data science study, you will also be required to complete a report on the study that you have completed. Both the Python notebook data science study as well as the report template can be found in Brightspace.

## Scope

1) The scope of this assignment is limited to the project statement and problems posted in the Python notebook data science study around HDB housing.
2) You are required to complete the Python notebook by writing code that satisfies the output for each problem.
3) Your Python notebook should include the following:
    a. Your name and admin number at the beginning of the Python notebook
    b. Code that answers the questions posted in the Python notebook
    c. Filled in markdown cells that answers the questions posted in the Python notebook
    d. In the last section of the notebook, there will be a section where you will decide what data visualization technique to use to represent certain datasets.
        i) For each dataset, describe the insights you have gained from the data visualization and any conclusion that you might have made from the analysis **in the report**

# Marking Scheme

Marks will be awarded to each student based on the following marking scheme:

| Component | Weightage |
|---|---|
| **Base Assignment Requirements**<br>• Completion of the questions posted in the notebook<br>• Python code that shows competency in<br>  ○ Performing basic Python operations<br>  ○ Utilization and understanding of data structures in Python<br>  ○ Data manipulation and wrangling using the NumPy library<br>  ○ Data visualization using the Matplotlib library | 30% |
| **Notebook Quality**<br>• Python program technical competency<br>• Python program code quality and cleanliness | 15% |
| **Report Quality**<br>• Completeness in the analysis of the data<br>• Depth of the questions explored<br>• Quality of the answers you provide | 30% |
| **Interview**<br>• Technical competency<br>• Explanation of insights and thought process<br>• Flow of the interview process and content | 25% |
| **Total** | 100% |

-- End of Assignment Specifications --