

## Assessment Details

Unit Title	Programming For Data Analysis
Unit Code	B9DA108
Unit Leader	Amit Sharma
Level:	9
Assessment Title	CA
Assessment Number	2
Assessment Type:	Group of 3.
Assessment Weighting	40%
Issue Date:	Week of 7 <sup>th</sup> June 2023
Hand in Date:	20 <sup>th</sup> July 2023
Mode of Submission:	Online

## Assessment Task

### Instructions:

**[100 Marks]**

Primary objective: To design and develop a Data Acquisition, Pre-processing and Data Analysis Pipeline.

Details of assignment brief:

You are required to develop a Data Acquisition, Pre-processing and Data Analysis Pipeline of your choice, including data acquisition (API, Web scraping, DB Extract etc.), Extraction of features and Transformations as appropriate, followed by loading into an appropriate database. Finally, data analysis should be done which involves answering questions of (public or self) interest based on the data. The focus of the complexity of the pipeline is your choice. Python must be used for the assignment.

Use Git (e.g. GitHub / Colab) and commit/pin version regularly.

GitHub/GitLab/Colab MUST be used for this project to develop both the artefact and documentation.

If using Colab, it must be shared with your lecturer's @dbs.ie email address as editor, and documentation should be as text cells in line with the code. If using Git, documentation must be in the form of a text document (README.md / HTML / LaTeX) inside the git repository, and the repository must be public.

The documentation must contain a link to the repository itself, and the repository should be downloaded and submitted on the Moodle link.

Demonstrations will be in week 12.