MATH 70076 Data Science
2022/23 - Assessed Coursework 1

Deadline: **Friday 17 Feb 2023 at 12:00** (Midday, UK).

QUESTION SHEET

You should submit two files in total for this assessment. These files should be submitted via the Imperial College VLE on Blackboard by the deadline stated above. Note that large files can take quite some time to upload. These files should be named:

1. `YOURCID-submission.pdf`

2. `YOURCID-project-1.zip`

Please ensure that you upload each document to the correct part of the learning space. Allowances will not be made for incorrect submission.

All submitted materials should be clearly presented and be understandable as stand-alone documents. This coursework is expected to take approximately **15 hours** of individual effort and will be marked out of a total of **100 marks**. This coursework counts for **30%** of your total mark for this course.

In submitting your work for this assessment you are agreeing to the following statement of academic integrity:

I certify that this assessed coursework is my own work, unless otherwise acknowledged, and includes no plagiarism. I have not discussed my coursework with anyone else except when seeking clarification with the module lecturer via email or on MS Teams. I have not shared any code underlying my coursework with anyone else prior to submission.

Plagiarism: As per the above, your submission should be your own work. Note that software tools are used for plagiarism detection.

---

1 Project Brief

1.1 Submission

For this assessment you will submit a piece of reproducible data journalism. This will be published in a new magazine, to be launched as a competitor to New Scientist, and should follow the Urban Institute style guidelines.

Your piece should be submitted as a single file named `YOURCID-submission.pdf` and may take one of two forms:

- a single, full page (20x24cm) infographic;

- a short article of 200-300 words with 2-3 data visualisations to accompany the text.

Note 1: Submitted articles should be clearly presented but final typesetting and arrangement will be controlled by the magazine's editorial board. To ensure high print quality, all figures should be provided as either vector graphics or with a resolution of least 300 dpi.

Note 2: Your submitted article or infographic should be based on a data set published on data is plural archive between 2022-10-01 and 2023-01-18. Each student must use a distinct data set; you should **register your your CID against your chosen dataset** in this spreadsheet.

1.2 Project Files

In addition to your main submission, you are required to submit `YOURCID-project-1.zip`. This is a zipped folder containing the code, data, documentation, and anything else needed to reproduce your submission.

This project folder should contain a file named `notes-to-editor.md`. This file should contain a reflective summary of how you have designed your submission and applied the style guide, addressing each of the following points in ~ 100 words (300 words total):

- Summarise the style guidelines that are relevant to your submission;

- Clearly state which guidelines are and are not met.

- Describe any alterations you would make to your submission with the assistance of the magazine's team of experienced data journalists.

2 Markscheme

This assessment will be graded with the award of up to 20 marks in each of the following categories:

1. Submitted infographic or article

2. Reflective summary of style guide

3. Project-level structure and documentation

4. Code-level structure and documentation

5. Reproducibility and accuracy of overall workflow (incl. data gathering and manipulation)

The number of marks allocated in each category will be guided by the following criteria:

Mark	Description
0	Missing of highly flawed
4	Substandard: major revisions required before approval
8	Acceptable: correct up to small errors, moderate revisions
12	Good: Requires only minor revisions before approval
16	Excellent: Could be approved as presented
20	Exceptional: Exceeds expectations or extends taught material