

# LDSCI5247 Foundations of Data Science AE1

## Assessment Details

<b>Course Leader:</b>	Anees Baqir
<b>First, Second, or Third Sitting:</b>	First
<b>Issue Date:</b>	8 September 2025
<b>Assessment Type:</b>	Code & Report
<b>Assessment Title:</b>	Global Student Migration Analytics
<b>Restrictions on Time/Length:</b>	1500 Words
<b>Assessment Weighting:</b>	50%
<b>Hand-in Deadline:</b>	7 November 2025
<b>Planned Feedback Deadline:</b>	28 calendar days after hand-in deadline or last presentation date
<b>File Format Accepted:</b>	.html export of Python Notebook
<b>Mode of Submission:</b>	Online (Canvas)
<b>Anonymous Submission:</b>	Yes
<b>Marking Scheme:</b>	Categorical



**Prohibited**

Use of AI is prohibited on this assignment.

## Assessment Task

### Introduction

Global education is undergoing significant transformation as student migration patterns evolve in response to geopolitical, economic, and social factors. International student mobility reflects not only educational aspirations but also broader global trends such as labour market demands, regional stability, and government policy. In recent years, events like the COVID-19 pandemic, shifts in visa regulations, and the growing importance of knowledge economies have reshaped how and where students pursue higher education abroad.

## Question 1 Data Exploration & Cleaning (20 marks)

Begin by loading the dataset and describe its structure, i.e., number of records, columns, and data types. Identify any missing or inconsistent entries (such as missing years, countries, or totals). Clean the dataset as needed: handle missing values (e.g. fill, drop), standardize country names, and correct inconsistent types. Provide summary statistics for numeric fields like student counts by year, origin, and destination countries. Explain your steps with Markdown, justifying the chosen cleaning strategies.

## Question 2 Temporal and Geographical Trends (20 marks)

Investigate changes in student migration from 2019 to 2023. Aggregate data to create insightful views, such as total outbound students per year or inbound per destination country. Use visualizations, such as, line charts for year-over-year trends, bar charts for top 5 sending/receiving countries, and comment on patterns observed: growth, decline, regional shifts, or anomalies (e.g. COVID effects). Provide at least two different visualizations and interpret their significance.

## Question 3 Feature Engineering, Group Analysis, & Visualisation (25 marks)

Create new, informative features, such as region-based groupings (e.g. Europe, Asia), per-capita flows relative to country population (if necessary, using external static values), or categorizing migration magnitude (e.g., high/medium/low flow). Use these features to compare groups, for example, continents or income brackets, across time or student flow volumes. Apply groupby operations and visualise results (e.g. box plots or grouped bar charts). Discuss how these derived features help reveal deeper insights beyond raw counts.

## Question 4 Hypothesis Testing on Migration Patterns (25 marks)

Formulate and test a hypothesis comparing international student flows between two groups, for instance, comparing the average annual student flows from Asia vs Europe between 2019 and 2023. Clearly state the null and alternative hypotheses. Aggregate data per group per year, check needed test assumptions (e.g. normality, equal variance), and perform an appropriate test, such as independent two-sample t-test or non-parametric alternative. Report test statistics, p-values, and interpret whether your data supports a statistically significant difference. Reflect on practical significance and any limitations.

**Note:** 10 marks are awarded for the English proficiency of your report.

The dataset required for this assessment will be provided by the instructor and made available exclusively through the course Canvas page. Please do not download external versions or alternative sources of the dataset.

## Assessment Criteria

The assessment should be completed in a Python notebook using Markdown cells for the written components of the task. You must remove any unnecessary code or markdown cells from the notebook before submitting the assignment.

## Assessment Criteria

70 or higher	There was evidence of the ability to perform all tasks correctly. The demonstration of the methods was excellent, coherent, well documented, and clearly explained.
60-69	There was evidence of ability to perform some tasks correctly. The demonstration of the methods is good, coherent, and reasonably detailed and explained.
50-59	There was evidence of ability to perform some tasks correctly, but the demonstration of the methods was limited, incoherent, not adequately documented and vaguely explained.
40-49	There was limited evidence of ability to perform tasks. The demonstration of the methods involved significant omissions and produced substantial inaccuracies.
39 or less	Failure to solve the tasks in the assignment. Methods were completely incorrect or absent.

## Submitting Assessments

For this assignment:

- Prepare a single Jupyter Notebook file that includes all the code, text, printed outputs, and data visualisations. Alternatively, prepare a Jupyter Notebook file that contains the code and prepare a .html file with the text and data visualisations.
- The notebook should be developed using Python 3 within a Jupyter environment, with well-commented code and appropriately structured text cells resembling a blog format.

- Do not include any identifying information such as your name.
- Remember to restart the kernel, thoroughly check the notebook for errors, and provide a working link to the dataset used in the analysis.
- Ensure that all code runs smoothly and create a compelling narrative by integrating analysis, visualisations, and reflections.
- Your submission may include auxiliary Python files (\*.py), where you packaged, e.g., some common functions and imported them in your notebook. In this case, submit a .zip file with all documents included.
- Do not include any dataset files in your submission.

You have three submission attempts, but only the last submission will be graded. If your last submission attempt is late, you will receive the late penalty even if you have a previous submission that was on time. Please make sure to avoid multiple submissions for assessments with multiple components, as only the last attempt will be graded. Upload several files in one submission attempt using the 'add files' function instead.

If you encounter issues with submission:

- Check the assessment details table at the top of the assessment brief to be sure you are submitting a permitted file format. Avoid zip files (unless explicitly required) and scanned PDFs. Use the 'add files' function to submit multiple files instead of a zip file.
- Make sure you have ticked the agreement box at the bottom of your Canvas submission page (scroll down if you don't see it). This will enable you to select 'Submit Assessment.'
- Try changing web browsers.

If you still cannot submit, e-mail a copy of your assignment before the deadline to [student.assessments@nulondon.ac.uk](mailto:student.assessments@nulondon.ac.uk) along with screenshots of the problem on Canvas, showing a timestamp.

If your assessment requires anonymous submission (see the assessment details table at the top of the assessment brief), be sure you have left your name off of your submission and out of the submission file name.

Please review the submitted file to ensure that everything is in order. To turn on notifications for submission confirmation emails in your Canvas settings: Account > Notifications > Turn on the bell for 'All submissions.' In the app this is via Settings > Email Notifications > All submissions.

## Marking

The University uses two assessment marking schemes – one for undergraduate and one for postgraduate – to mark all taught programmes leading to an award of the University.

More detailed information on the assessment marking scheme and the criteria can be found in the Course Syllabus, available on the University's VLE.

## Learning Outcomes

This assessment will enable students to demonstrate in full or in part the learning outcomes identified in the Course Descriptor.

On successful completion of this assessment, students should be able to:

### Knowledge and Understanding

K1b	Demonstrate knowledge and critical of well-established data representation and transformation concepts for multi-dimensional data.
K2b	Demonstrate ability to identify feasible operations and transformation on data, and their relationships in a data processing pipeline.
K3b	Demonstrate knowledge and critical understanding of plotting and visualising data.

### Subject-Specific Skills

S2b	Apply the data science theory learnt in class (e.g., well-established data transformation techniques) in an appropriate manner to a given dataset.
S3b	Identify the correct choice of appropriate data transformation techniques.

### Transferable Skills

T2b	Identify, transform, evaluate, and plot accordingly from a dataset.
T3b	Demonstrate a sound technical proficiency in written English and skill in selecting vocabulary so as to communicate effectively to specialist and non-specialist audiences.

## Accessing Feedback

Students can expect to receive feedback on all summative coursework within 28 calendar days of the submission deadline or, if applicable, the last oral assessment date, whichever later. The 28 calendar day deadline does not apply to work submitted late. Feedback can be accessed through the assessment link on the Canvas course page.

## Late Submissions

Please ensure that you submit your assignment well before the deadline to avoid any late penalties, as a submission made exactly on the deadline will be considered

late. Please keep in mind that there may be differences between your computer's clock and the server time, which can cause discrepancies, and that Canvas may take some time to process your submission.

Your Canvas submission portal displays two due dates: one is the deadline for your assignment, and the second is the latest possible date by which your assignment can be submitted late. Please make sure you submit by the assessment deadline in order to avoid late penalties.

If assessments are submitted late without approved Extenuating Circumstances, there are penalties:

- For assessment elements submitted up to one day late, any passing mark will receive 10 marks deducted or a threshold pass (40% for undergraduate students, 50% for postgraduate students), whichever is higher. Any mark below 40% for undergraduate students and below 50% for postgraduate students will stand.
- Students who do not submit their assessment within one day of the deadline, and have no approved Extenuating Circumstances, are deemed not to have submitted and to have failed that assessment element. The mark recorded will be 0%.
- For assessment subelements, late submission will result in non-submission penalties deducted according to the marking criteria above.

For further information, please refer to [AQF7 Part C in the Academic Handbook](#).

## Extenuating Circumstances

The University's Extenuating Circumstances (ECs) procedure is in place if there are genuine circumstances that may prevent a student from submitting an assessment. If the EC application is successful, there will be no academic penalty for missing the published submission deadline.

Students are normally expected to apply for ECs in advance of the assessment deadline. Students may apply for consideration of ECs retrospectively if they can provide evidence that they could not have done so in advance of the deadline. All applications for ECs must be supported by independent evidence.

Successful EC applications for live oral assessments, including vivas, will result in a deferral of the oral to be organized by faculty, students, and Timetabling for a date as close as possible to the original presentation date. The deadline for supplementary materials, if assigned, will be carried forward by the length of the oral assessment extension.

Missing an oral assessment, including a compulsory viva, without an approved EC will result in a non-submission for the entire assessment and, accordingly, a recorded mark of 0%.

Students are reminded that the ECs procedure covers only short-term issues (within 21 days leading to the submission deadline) and that if they experience longer-term matters that impact on learning then they must contact [Student Support](#) for advice.

Under the Extenuating Circumstances Policy, students may defer an assessed element on only one occasion and may request an extension on a maximum of two occasions.

For further information, please refer to the [Extenuating Circumstances Policy](#) in the Academic Handbook.

## Academic Misconduct

You must submit work for assessment purposes that is your own and meets good academic practice. Assessments must be completed strictly in accordance with the instructions outlined in the assessment brief. This includes ensuring that your work is appropriately referenced.

It is important to understand if artificial intelligence (AI) is permitted to be used or not, and if yes, the conditions for its use. For clarity, the assessment brief expressly states if AI can be used, and the parameters of its permitted usage. You are not permitted to use AI (in any form) if the assessment brief does not expressly state that usage is permitted.

Any concerns with the academic integrity of a submission will be addressed in accordance with the [Academic Misconduct Policy](#). This policy gives details of the different types of misconduct and the steps the University may take when a concern arises. Action may result in academic penalties being applied directly to you. You are advised to review this policy in full before completing and submitting any work for marking purposes. The University may also take steps to ascertain the authenticity of a submitted piece of work in cases of relevant concern, such as by holding a viva.

## Version History

Title: Assessment Brief Template					
Approved by: The Quality Team					
Version number	Date approved	Date published	Owner	Location	Proposed next review date
4.0	March 2023	March 2023	Registrar	VLE/ Faculty Resources Page	March 2024
3.0	August 2022	August 2022	Registrar	VLE, Faculty Resources Page	July 2023
2.3	December 2021	December 2021	Registrar	VLE	August 2022
2.2	August 2021	August 2021	Registrar	VLE	August 2022

2.1	September 2020	September 2020	Registrar	VLE	August 2021
2.0	September 2020	September 2020	Registrar	VLE	August 2021
1.0	August 2019	August 2019	Registrar	VLE	August 2020
Referenced documents	AQF7 Academic Regulations for Taught Awards; Extenuating Circumstances Policy; Academic Misconduct Policy; Course Syllabus				
External Reference Point(s)	UK Quality Code Theme: Assessment				