Assessment Brief

Module title	Data Science
Module code	COMP5122M
Assignment title	Coursework 1: Global Climate Attitudes and Eco-Anxiety Part II
Assignment type and description	Data analysis, predictive modelling, and clustering using survey data
Rationale	This coursework challenges you to analyse real-world climate survey data and communicate findings effectively. You will apply the core stages of the data science pipeline (Part II is focused on Predictive Modelling and Clustering and Unsupervised Learning Techniques) to investigate global climate attitudes and eco-anxiety. The task strengthens your ability to train supervised machine-learning models for prediction, evaluate their performance using appropriate metrics, and apply unsupervised learning techniques to cluster data and identify key predictors and country typologies related to support for strengthening climate commitments.
Word limit and guidance	Cover page with group name and group members' names, surnames and usernames (1 page), Report of up to four pages, single-spaced, 11pt font. References may be included in an additional page (max. 1 page).
Weighting	32%
Submission deadline	23:59, Friday 12th Dec 2025
Submission method	Via Minerva
Feedback provision	Marks and feedback returned via Minerva
Learning outcomes assessed	- Understand the work of a data scientist - Model data using Machine Learning approaches

	- Evaluate Machine Learning models - Visualize data effectively - Apply problem-solving skills to analyse data and communicate findings for a global societal issue
Module leader	Dr. Duygu Sarikaya

1. Assignment guidance

The objective is to assess your ability to analyse data and communicate your findings. You need to do this coursework in the groups that are defined on Minerva. A group comprises a maximum of four people (You can find your assigned group on Minerva).

Background and Motivation

Image from https://www.savethechildren.org/

The Peoples' Climate Vote 2024 is the world's largest survey of public opinion on climate change, conducted by the United Nations Development Programme (UNDP) in partnership with the University of Oxford. It covers 73 countries, representing over half of the global population, and gathers responses across different age groups, genders, and education levels.

The survey explores several themes: how often people think about climate change, how worried they are about future generations, whether they have experienced extreme weather events, and how much they trust their governments, businesses, and international organisations to act. It also asks people whether they believe countries should strengthen or weaken climate commitments, how quickly fossil fuels should be replaced by renewable energy, and how much nature should be protected.

The dataset is aggregated: each row records the percentage of respondents in a given country, age group, and education group who selected a particular response. This allows rich cross-country and demographic comparisons, highlighting patterns of support, scepticism, and trust.

Understanding these patterns matters because climate change is not only a scientific and technological problem, but also a social and political one. Public opinion influences the policies governments adopt, the urgency of action, and the willingness to collaborate internationally. Identifying where age gaps, education divides, or trust deficits exist can inform more targeted communication strategies and policymaking.

This coursework therefore challenges you to:
1. Analyse whether there is a consistent age gap across countries in support for renewable energy transition. By completing this analysis, you will not only practice the core stages of the data science pipeline (Part I is focused on Exploratory Data Analysis) but also contribute to a deeper understanding of how societies around the world perceive and respond to the climate crisis

2. Assessment tasks

You will analyse the UNDP Peoples' Climate Vote 2024 dataset, which covers 73 countries and records responses by age, education, and country. It includes public attitudes towards climate change, the energy transition, and institutional responsibility.

You can download the dataset here: https://peoplesclimate.vote/data-center

Your tasks for the Coursework Part II are as follows:

a) Predictive modelling (20 marks)

- Build a supervised machine learning model to predict whether respondents support strengthening climate commitments.
- Use demographics and survey attitudes (e.g., trust in institutions, extreme weather experience) as features.
- Compare at least two models (e.g., logistic regression vs. random forest/gradient boosting).
- Evaluate performance using appropriate metrics (e.g. accuracy, F1, balanced accuracy if imbalanced).
- Interpret outputs to identify whether demographics or attitudes matter more.

b) Clustering analysis (8 marks)
- Construct a typology of countries based on aggregated climate attitudes.
- Apply an unsupervised method (e.g., k-means, hierarchical clustering, PCA).
- Label and interpret clusters and compare clusters across world regions.

c) Conclusions (4 marks)
- Summarize key findings: strongest predictors, country typologies

3. General guidance and study support

- Handle preprocessing carefully (e.g., "All Education" / "All Ages" aggregates).
- Justify your choice of visualisation and modelling approaches.
- Clearly explain your evaluation metrics.
- Deep learning methods are out of scope for this coursework.

Links: Module Handbook, Skills@Library, and Minerva resources.

4. Assessment criteria and marking process

Part II
Predictive modelling: 20 marks
Clustering analysis: 8 marks
Conclusions: 4 marks

Total: 32 marks

5. Presentation and referencing

- Report format (PDF).
- Four pages maximum (excluding cover page and references).
- Professional structure, clear writing, appropriate figures/tables.
- Academic referencing in additional page (max. 1 page).
- The quality of written English will be assessed: paragraphs, flow, appropriate word choice, and correct grammar.

The quality of written English will be assessed in this work. As a minimum, you must ensure:
- Paragraphs are used
- There are links between and within paragraphs although these may be ineffective at times
- There are (at least) attempts at referencing
- Word choice and grammar do not seriously undermine the meaning and comprehensibility of the argument
- Word choice and grammar are generally appropriate to an academic text

These are pass/ fail criteria. So irrespective of marks awarded elsewhere, if you do not meet these criteria, you will fail overall.

6. Submission requirements

Submit electronically via Minerva before the deadline. File format: PDF.

Please also add a cover page that clearly states your group name, your group members' names, surnames, and usernames. It is preferable that you add your group's Pokemon as a logo on the cover page.

Your file should be named after your group (e.g. "Mewtwo") and be structured as yourgroupname_PartII_report (e.g. Mewtwo_PartI_report).

You will only submit your report, however, I reserve the rights to ask for the code or a document outlining your Tableau workflow along with queries you have used. Be prepared to submit these if asked.

7. Academic misconduct and plagiarism

Leeds students are part of an academic community that shares ideas and develops new ones.

You need to learn how to work with others, how to interpret and present other people's ideas, and how to produce your own independent academic work. It is essential that you can distinguish between other people's work and your own, and correctly acknowledge other people's work.

All students new to the University are expected to complete an online Academic Integrity tutorial and test, and all Leeds students should ensure that they are aware of the principles of Academic integrity.

When you submit work for assessment it is expected that it will meet the University's academic integrity standards.

If you do not understand what these standards are, or how they apply to your work, then please ask the module teaching staff for further guidance.

By submitting this assignment, you are confirming that the work is a true expression of your own work and ideas and that you have given credit to others where their work has contributed to yours.

8. Use of GenAI in this assessment

RED: AI tools cannot be used. The purpose and format of this assessment make it inappropriate for AI tools to be used.

9. Assessment/ marking criteria grid

Criteria	Marks	High (70%+)	Medium (50–69%) / Low (0–49%)
Predictive Modeling	20	Well-designed pipeline, clear comparison, an excellent selection of methods/tools well-suited to the task, justification of these methods, tools and metrics, proper metrics chosen, justification of metrics used, methods applied appropriately, correct findings, Whether the issues with dataset such as imbalance are considered, insightful interpretation of results	Medium: Adequate modelling, limited metrics, partial interpretation Low: Poorly executed, inadequate evaluation

Clustering Analysis:	8	Well-designed pipeline, insightful interpretation, an excellent selection of methods/tools well-suited to the task, justification of methods, methods applied appropriately, correct findings, insightful interpretation of results	Medium: Basic clustering, limited interpretation Low: No or poor clustering
Conclusion: strongest predictors, country typologies	4	Concise, clear, supported by analysis, clear take-aways, good storytelling, discussion of findings	Medium: Some conclusions, partial clarity / Low: Weak or missing, not supported