University of Hertfordshire UH
School of Physics, Engineering and Computer Science

Weighting %:	60%	**Submission deadline (for students):**	23 Nov 2025, 2359 Class presentation: **24 Nov 2025**
Authorship:	Individual	**Target date for returning marked coursework:**	30 Nov 2025
Tutor setting the work:	WONG KIM SIONG	**Number of hours you are expected to work on this assignment:**	40-50

This Assignment assesses the following module Learning Outcomes (from Definitive Module Document):

1. Demonstrate knowledge of and understand a wide range of data formats and data types and the methods used to read them.
2. Demonstrate knowledge of and understand common techniques for handling and manipulating large data sets.
3. Demonstrate understanding of fundamental concepts and techniques in data visualisation.
4. Demonstrate how to use visualisation to explore data and communicate in an effective manner.
5. Demonstrate knowledge of different visualisation tool.

Assignment Tasks:

Final Visualization Report – Tell a story using data

This is an individual assignment that requires you to design and produce a visual report based on a dataset of **your choice**. The primary objective is to tell a compelling, evidence-based story using data, supported by well-designed visualizations and appropriate analytical commentary.

You are free to propose your own topic of interest for this project. The dataset you select must be:

- Clearly referenced with a source citation and URL
- Openly accessible to the public
- Sufficiently rich to support analysis and visualisation

Some example domains you might consider include (but are not limited to):

- Health and wellbeing
- Finance and economics
- Climate and environment
- Education, poverty, or inequality
- Public policy and government services
- Sports analytics
- Media and entertainment

Project Requirements:

- You must use Python (e.g., pandas, numpy, matplotlib, seaborn, plotly) to process and visualise the data.
- Visualizations should be customized beyond default settings and optimized for clarity and presentation.
- The report must reflect a structured analysis, showing evidence of intermediate to advanced data handling and exploratory insight.
- Each plot must reveal new insights (i.e., avoid repetition or redundant charts).
- All visual elements (plots, tables) should be well annotated with legible fonts, appropriately scaled axes, and consistent styling.
- The table should ideally present summary statistics or an overview of key variables, formatted clearly on one page.
- Provide a Link to your Github repository

Guidelines:

- Focus on presenting insightful summaries, uncovering trends, and answering relevant questions with your data.
- Review lecture materials on best practices in visualisation design.
- Avoid overly complex visuals unless they genuinely support the narrative.
- Ensure all figures are readable in print format (avoid clutter and overlapping labels).
- Proofread your report for clarity, conciseness, and coherence.

Submission Requirements (Deliverables):

1. A 15-minute presentation to explain your project
2. **A PDF report (**about **800** words±20%, excluding figure captions), which must include:
   - A descriptive title and your name/student ID
   - A brief overview of the dataset and its source
   - At least three professionally styled visualizations
   - At least one summary table with relevant statistics
   - An interpretive narrative that ties the data and visuals together. In other words:
     - 1. **Coherent writing** - You need to explicitly illustrate, in your written report, how the data blend in with the visuals to construct effective visual storytelling.
     - 2. **Coherent visuals** - At the same time, the quality and clarity of your visualisation should be such that the reader, without having to read your report, be able to understand the points that you are trying to convey.
   - Clear citations for all data sources
3. **Code and dataset files** - used to create your plots and tables. These should be provided in a readable and executable format, such as:
   - **Code file** - .py scripts or Jupyter Notebooks (.ipynb)
   - **Data files** - in .csv, .parquet, or equivalent formats

Grading Rubrics:

No.	Component	Weighing (%)	Description
1	Data Interpretation	20%	Demonstrates understanding of the dataset's structure and variables.
2	Plot Clarity and Quality	20%	Plots are clearly labeled, easy to read, and communicate key messages effectively. Proper axis scales, font sizes (≥14), and labeling conventions are used consistently.
3	Choice of Charts for Visualization	20%	Visualization types are appropriate for the data and insight. Includes a variety of chart types (e.g., scatter, histogram, boxplot) that effectively represent different aspects of the dataset.
4	Styling and Customisation	15%	Plots are visually polished with appropriate use of color, spacing, layout, and annotation. Goes beyond default settings to improve aesthetic and clarity.
5	Code Structure and Comments	15%	Code is well-structured, logically sequenced, and reproducible. Includes clear, informative comments to guide readers through the analysis.
6	Coherent visual story-telling, structure and flow	10%	The sequence of visualizations supports a clear narrative or insight progression. Markdown cells and annotations connect ideas cohesively and aid reader comprehension.

Type of Feedback to be given for this assignment:

Written feedback on the code and report.

Additional information:

- Regulations governing assessment offences including Plagiarism and Collusion are available from https://www.herts.ac.uk/__data/assets/pdf_file/0007/237625/AS14-Apx3-Academic-Misconduct.pdf (UPR AS14).
- Guidance on avoiding plagiarism can be found here: https://herts.instructure.com/courses/61421 (see the **Referencing** section)
- Adhere to UH's policy on use of GenAI:
    - "Unauthorised use of artificially generated material (AI) in researching or presenting material for an assessment is an academic misconduct offence if you use AI tools in producing your assessment, unless the use of AI tools is expressly permitted. However, even if expressly permitted, where you do not declare that you have used an artificial intelligence tool(s) in the production of your assessment, or you are dishonest about the extent to which such tools have been used, you will have committed academic misconduct."
- For **postgraduate modules**:
    - a score of 50% or above represents a pass mark.
    - late submission of any item of coursework for each day or part thereof (or for hard copy submission only, working day or part thereof) for up to five days after the published deadline, coursework relating to modules at Level 7 submitted late (including deferred coursework, but with the exception of referred coursework), will have the numeric grade reduced by 10 grade points until or unless the numeric grade reaches or is 50. Where the numeric grade awarded for the assessment is less than 50, no lateness penalty will be applied.