

# M.Sc. Software Design with Artificial Intelligence (AL\_KAIMF\_9\_1)

## Data Visualization

### Assignment 2

**Marks:** 35% of the Module.

**Deadline:** Sunday, 30<sup>th</sup> March 2025 (11:59 P.M.)

**Restrictions:** The visualization chosen **must not** endorse or forward political, racist, or sexist messages, furthermore it **must not** endorse hate speech aimed at any group in society. Failure to follow these instructions **may result in the forfeiture of the marks** for this assignment.

#### Overview:

This assignment will focus on using the Dash framework in Python to create an interactive data visualization dashboard. You will analyze and visualize data from the World Bank's World Development Indicators (WDI) for your assigned country. You will also demonstrate the importance of data cleaning and how it impacts analysis.

#### Task:

1. **Country Focus:** You will be assigned a specific country. You will be responsible for understanding and communicating your assigned country's development using World Bank Indicators:  
<https://databank.worldbank.org/reports.aspx?source=world-development-indicators>
2. **Data Selection & Analysis (10 Marks):**
  - a. Select at least three distinct topics relevant to your assigned country's development. Examples include:
    - i. Economic Growth and Indicators (GDP, GNI, trade, etc.)
    - ii. Health and Demographics (life expectancy, infant mortality, etc.)
    - iii. Education Levels and Access
  - b. For each topic, **download** relevant indicators and time periods from the WDI.
  - c. **Data Cleaning:** Identify and address at least three different data quality issues in your data (e.g., missing values, inconsistencies in format, outliers). Document your cleaning process and rationale in the report and the code comments.  
For example, identify data type inconsistencies and provide codes to address that.
3. **Interactive Dashboard with Dash (15 Marks):**
  - a. Create an interactive dashboard using Dash (Python library).
  - b. Your dashboard must include:
    - i. **At least three different visualizations**, each corresponding to one of the topics you chose. Use appropriate chart types (line chart, bar chart, scatter plot, or a combined chart).
    - ii. **Interactive elements:** at least one dropdown menu or slider to filter/change the view of the data.
    - iii. A clear and user-friendly layout.
    - iv. Clear labels and titles for each visualization.
    - v. A title at the top of the webpage.
    - vi. **Visualizations of your chosen country and the region it belongs to. Provide some visual comparisons.**
4. **Analysis & Explanation (5 Marks):**
  - a. In the report, for each of your visualizations, provide:
    - i. A clear heading (e.g., "Dashboard Chart 1: GDP Growth Over Time").
    - ii. The World Bank indicators and years used for this visualization.
    - iii. A brief explanation of the purpose of the visualization, what insights it conveys, and the rationale behind your choice of chart type.
    - iv. Explain how the data was cleaned before visualization.

- v. **Specifically address** what you learned from the data cleaning process and how this changed your understanding of the data/analysis.

**5. Learning Outcomes & Challenges (1 Mark):**

- a. Briefly reflect on what you learned about your assigned country and what challenges you faced while completing the assignment. (No more than one page).

**Deliverable:**

**1. Report (PDF) 3 Marks): A concise report (no more than 4 pages) containing:**

- a. Assignment Cover Sheet
- b. Table of Contents
- c. Report Body: Your analysis, explanations, and data cleaning steps, following the structure outlined above.
- d. Learning Outcomes and Challenges section.
- e. Reference List
- f. *Note: the visualizations are embedded in the dashboard and so do not have to be added to the report, simply discuss them in your report as you see on the dashboard*

**2. Code and Data Files (ZIP) (1 Mark): A zipped folder containing:**

- a. All original data files from the World Bank.
- b. Cleaned data files.
- c. Your Python code files for the Dash dashboard.
- d. Your report in PDF format.

**Formatting Requirements:**

- **Font:** Times New Roman, 12pt for the main body text.
- Justify all body paragraphs
- **Page Numbering:** Number all pages consecutively, starting with the first page after the Table of Contents.
- **References:** Use Vancouver style for referencing.

**Grading Emphasis:**

Your assignment will be graded according to the detailed rubric provided separately. The rubric outlines the specific criteria and mark allocation for each aspect of the assignment. Refer to the rubric for a complete breakdown of the assessment criteria and expectations.

**Submission:**

Submit the completed report and the zipped folder with the code and datasets using the assignment upload on Moodle.

**N.B.**

Do not use any of the visualizations contained in your Data Visualization notes as your chosen visualization. You can, however, refer back to them if you think it relevant.