INFO 212: Data Science Programming I

Data Science Assignment 2: Economic Indicators

Objective: Develop a Python project to analyze and visualize economic indicators from JSON data files using Pandas and Seaborn or other visualization libraries.

Dataset: JSON files containing economic indicators such as GDP, unemployment rate, inflation rate, etc.

1. Data Collection (10 marks):
   - Download JSON data files containing economic indicators from reputable sources (e.g. such as the World Bank or IMF).
   - Ensure the data spans at least 5 years for multiple countries.

2. Data Preparation (10 marks):
   - If required, load the JSON files into Pandas DataFrames.
   - Clean the data by handling missing values, duplicates, and incorrect data types.

3. Data Analysis) (20 marks):
   - Perform exploratory data analysis (EDA) to understand the datasets.
   - Calculate summary statistics and identify trends over time for each indicator.

4. Data Visualization (20 marks):
   - Create visualizations to represent the trends and patterns discovered during EDA.
   - Use line plots, bar charts, and scatter plots to compare different economic indicators across countries.
   - Employ Seaborn's advanced features like facet grids or pair plots to visualize multi-dimensional data.

5. Interpretation (20 marks):
   - Interpret the visualizations to draw meaningful conclusions about the economic health of the countries.
   - Discuss any correlations or anomalies found between different economic indicators.

6. Presentation: (20 marks):
   - Compile the analysis and visualizations into a Jupyter Notebook.
   - Prepare a presentation summarizing the findings and insights from the project.

INFO 212: Data Science Programming I

Requirements

- Provide full and detailed answer to each question.
- Work in group of 4. University Academic Honesty Rules and Procedures will be adhered to strictly.
- There are 6 topics. The total marks are 100.

- Please name your assignment file as INFO212-assign1-group-number.zip, with the following **Deliverables:**
  - A Jupyter notebook containing the cleaned dataset, analysis code, and a summary of findings.
  - A brief pdf report outlining the approach taken and key observations.

- Submit your assignment file via moodle until the due **Due Date: 29-May-2024**
  - Assignments will not be accepted beyond that point. Missing work will earn a zero grade.

- Images must be clear and legible. Assignments will be judged on the basis of code, visual appearance, grammatical correctness, and quality of writing, as well as their contents. Please make sure that the text of your assignments is well-structured, using paragraphs, full sentences, and other features of well-written presentation. Text font size should be at least 11 points.