CAPSTONE PROJECT: SUSTAINABLE DEVELOPMENT GOALS (SDGS) DATA ANALYSIS

Submit your capstone project as a *.ipynb* file (Python Notebook).

Presentation Guideline

1. 10 minutes for each participant
2. The presentation is scheduled for Friday (28th June), from 5pm to 8.30pm.
3. Participants will be divided into two breakout rooms (15 pax per room), and we will have the presentations in both rooms simultaneously.

Instructions:

You are required to choose a dataset related to one of the Sustainable Development Goals (SDGs). Follow the steps below to complete the capstone project:

Step 1: Find a Dataset

1. **Question:**
   o Choose a dataset that aligns with one of the SDGs (e.g., No Poverty, Zero Hunger, Good Health and Well-being, Quality Education, etc.).
   o Describe the source of your dataset. Briefly describe the dataset, including its key features and variables.

Step 2: Create a Problem Statement

2. **Question:**
   o Formulate a clear and concise problem statement based on the dataset you have chosen. Your problem statement should reflect a specific aspect or challenge within the chosen SDG.
   o Explain why this problem is important to address and how solving it could contribute to achieving the SDG.

Step 3: Perform Data Cleaning and Transformation

3. **Questions:**
   o Describe the initial state of your dataset. Are there any missing values, duplicates, or outliers?
   o Outline the steps you took to clean the data. This may include handling missing values, removing duplicates, correcting data types, and addressing outliers.
   o Perform necessary data transformations/manipulations. Describe any new variables you created, data aggregations, or any other manipulation techniques applied to prepare the data for analysis.

Step 4: Perform Exploratory Data Analysis (EDA)

4. **Questions:**
   - o Conduct exploratory data analysis using the visualization techniques you have learned. This should include creating at least five different types of visualizations (e.g., histograms, bar charts, scatter plots, box plots, heatmaps).
   - o For each visualization:
     - ▪ Describe what the visualization is showing.
     - ▪ Explain the insights gained from each visualization in the context of your problem statement.

Step 5: Interpret the Visualizations

5. **Questions:**
   - o Connect your visualizations and findings back to your problem statement.
   - o Interpret the results and discuss how they address the problem you identified.
   - o What are the potential implications of your findings for the chosen SDG?
   - o Provide recommendations or next steps based on your analysis.