# Project Brief: Age-Wise Global Literacy Analysis

## Problem Statement

You are tasked with analyzing global literacy patterns across different age groups using real-world data. The dataset is available in tabular form on the Wikipedia page: https://en.wikipedia.org/wiki/List\_of\_countries\_by\_literacy\_rate  
  
This table provides age-wise annual literacy rates (e.g., youth, adult, elderly, total) by country and year. Your job is to extract this data using web scraping techniques (requests + BeautifulSoup), organize it into a structured format (CSV), and conduct Exploratory Data Analysis (EDA) to draw insights.

## What You Have to Do

1. 1. Web Scraping

- Use `requests` and `BeautifulSoup` to scrape the relevant table from the Wikipedia page.  
- Extract relevant columns: Country, Year, Youth Literacy Rate, Adult Literacy Rate, Elderly Literacy Rate, Total Literacy Rate (as available).  
- Clean and preprocess the data and save it in a CSV file.

1. 2. Exploratory Data Analysis (EDA)

- Perform univariate and bivariate analysis, generate summary statistics, and create visualizations.  
- Analyze trends in literacy rates by age group.  
- Identify countries with high/low literacy and generational gaps.  
- Group data by region if applicable.

1. 3. Answer Key Questions

- You will find the list of questions inside the Jupyter notebook template as comments.  
- Each answer must be supported with proof through visualizations and/or clear data-driven analysis—not just plain text.  
- Use bar charts, boxplots, scatterplots, heatmaps, or other visual tools as suitable.

1. 4. Documentation & Explanation

- Interpret your findings in markdown cells.  
- Write clear explanations for your observations and insights drawn from visualizations.

## What to Submit

You must create a GitHub repository and upload the following:  
1. Final CSV file – cleaned, structured dataset  
2. Jupyter Notebook (.ipynb) – containing the code, analysis, visualizations, and answers to the questions (with proof)  
  
Then:  
- Create a text file named submission\_link.txt  
- Paste the link to your GitHub repository inside this file  
- Upload the text file to the Google Form (link will be provided)

## Important Notes

- Answers to the questions must be supported by code and visualizations.  
- Submissions without code-based justification will be considered incomplete.  
- Your notebook should reflect a proper data analysis pipeline—from data collection to insight generation.  
- Ensure your GitHub repository is public or shareable.