Solar Data Analysis Challenge – Week 0 Report

1. Project Overview

This Week 0 challenge focused on exploring solar radiation datasets from Benin, Togo, and Sierra Leone. The objective was to perform data profiling, cleaning, and exploratory data analysis (EDA) to identify high-potential regions for solar energy installation.

2. Data Sources

Three datasets were provided:

- benin-malanvile.csv
- togo-dapaong_qc.csv
- sierraleone-bunmuna.csv

Each dataset contains measurements including:

- Global Horizontal Irradiance (GHI)
- Direct Normal Irradiance (DNI)
- Diffuse Horizontal Irradiance (DHI)

Ambient temperature, humidity, wind speed, and more

3. Tools and Methods

The following tools and methods were applied:

- Python programming language
- Pandas for data cleaning and manipulation
- Seaborn and Matplotlib for visual analysis (plots summarized in text)
- Jupyter Notebook for coding and documenting workflow

Steps followed:

1. Load datasets individually and check structure.
2. Clean missing values and correct outliers.
3. Combine datasets into a single dataset with a country column.
4. Perform descriptive statistics and comparisons across countries.
5. Exploratory Data Analysis (EDA) Summary

6. Calculated summary statistics (mean, median, standard deviation) for key variables: GHI, DNI, DHI.
7. Identified missing values and handled them appropriately.
8. Compared solar irradiance patterns across countries.
9. Observed variations in solar potential, which can guide solar installation decisions.

Key observations:Differences in median GHI values between countries indicate regional variation in solar potential.

Some datasets had occasional outliers in module readings and wind speed, which were cleaned.

Relative humidity and ambient temperature showed expected correlations with solar irradiance.

5. Conclusion and Reflection

This Week 0 project helped develop skills in data cleaning, exploratory analysis, and cross-country comparison. The analysis provided insights into which regions may be better suited for solar energy investments.

Even with a tight timeline, completing this challenge strengthened practical skills in Python programming, data handling, and preparing a professional analysis report.