10 Academy - Week 0: Solar Data Discovery Challenge

Challenge Overview

This project focuses on analyzing solar radiation and environmental data from three African countries — Benin, Sierra Leone, and Togo — to identify high-potential regions for solar farm installations. The aim is to support MoonLight Energy Solutions in making data-driven investment decisions that enhance operational efficiency and long-term sustainability. The tasks involved setting up a collaborative Git environment, performing Exploratory Data Analysis (EDA), and comparing key metrics across the three countries.

Task 1: Git & Environment Setup

The development environment was prepared using Python (venv) and version control with GitHub. A repository named *solar-challenge-week0* was initialized and structured as per 10 Academy's recommended folder format. Dependencies such as pandas, numpy, matplotlib, seaborn, scipy, and streamlit were listed in requirements.txt. Continuous Integration (CI) was implemented through a GitHub Actions workflow that verified environment setup and package installation.

Task 2: Data Profiling, Cleaning & EDA

Each country's dataset underwent profiling, cleaning, and exploratory data analysis to reveal solar radiation patterns and relationships among environmental variables. The process included handling missing values, detecting outliers, and generating descriptive statistics. Key steps included:

1. Computed summary statistics and missing-value reports.
2. Handled missing data using median imputation.
3. Removed outliers using Z-scores ($|Z| > 3$).
4. Analyzed temporal patterns of GHI, DNI, and DHI using line charts.
5. Investigated correlations among irradiance, temperature, and humidity.
6. Assessed the effect of cleaning events on sensor readings (ModA and ModB).

Findings revealed that solar irradiance levels (GHI, DNI, DHI) fluctuated throughout the day, peaking around midday. Relative humidity was inversely related to ambient temperature, and cleaning events generally improved module performance readings.

Task 3: Cross-Country Comparison

Cleaned datasets for Benin, Sierra Leone, and Togo were compared to assess overall solar potential. Statistical summaries and visualizations (boxplots and bar charts) were used to compare the GHI, DNI, and DHI metrics.

- Benin displayed the highest mean Global Horizontal Irradiance (GHI), suggesting strong solar potential.
- Togo showed moderate irradiance levels with stable daily patterns and fewer outliers.
- Sierra Leone had the lowest median GHI and higher variability, likely influenced by higher humidity and precipitation.

A one-way ANOVA confirmed statistically significant differences ($p < 0.05$) between the three countries' GHI values, indicating that solar potential varies meaningfully by region. Based on the analysis, Benin emerges as the most promising location for large-scale solar investment.

Bonus: Streamlit Dashboard

A Streamlit dashboard prototype was developed to visualize the cleaned datasets interactively. Users can select a country, explore irradiance trends, and view summary statistics in real time. The dashboard provides intuitive visual comparisons and supports data-driven decisions for solar site selection.

Conclusion

This Week 0 challenge provided a practical introduction to the data analysis lifecycle — from environment setup and data cleaning to visualization and interpretation. By combining analytical rigor with effective visualization techniques, this project highlights how solar data can inform sustainable energy strategies across Africa. Benin shows the most favorable conditions for solar energy deployment, followed by Togo, while Sierra Leone's variability suggests the need for further investigation.

Submitted by: Rahel Abera

Date: November 9, 2025