# **Report on the Exploratory Data Analysis (EDA) Python Code**

## **Introduction**

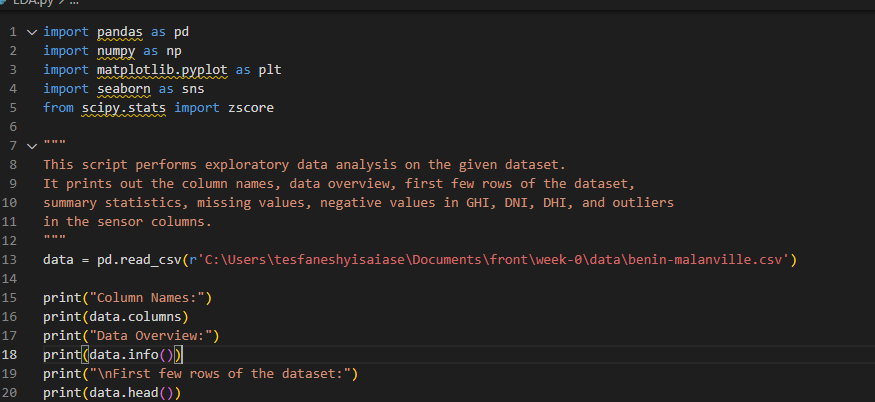
This report summarizes the Python code developed for Exploratory Data Analysis (EDA) on a dataset related to solar radiation, temperature, and other meteorological variables. The code employs various data analysis techniques to assess data quality, visualize trends, and explore relationships between different variables.

**Objectives**

The primary objectives of the EDA are:

1. **Data Overview**: Understand the structure and contents of the dataset.
2. **Data Cleaning**: Identify and handle missing values, outliers, and anomalies.
3. **Statistical Analysis**: Calculate summary statistics to understand data distributions.
4. **Visual Analysis**: Create visual representations to identify patterns and trends.
5. **Correlation Analysis**: Explore relationships between solar radiation and temperature measures.

### **1. Data Loading**

The code begins by importing necessary libraries and loading a CSV dataset using Pandas:

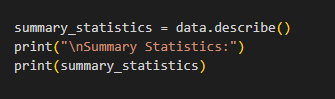
### **2. Data Overview**

Initial checks provide insights into the dataset:

* **Data Structure**: The dimensions and types of columns are printed.
* **First Few Rows**: A preview of the dataset is displayed to understand its contents.

### **3. Summary Statistics**

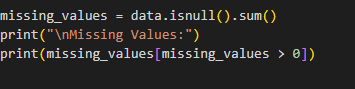
Summary statistics are calculated to provide an overview of key metrics:



This includes mean, median, standard deviation, min, max, and quartiles for each numeric column.

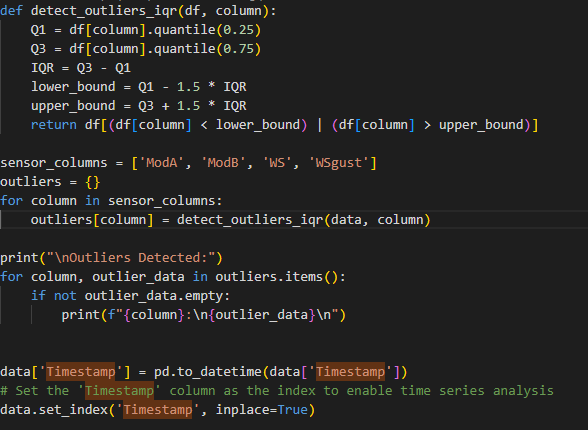
### **4. Missing Values and Negative Entries**

The code checks for missing values and identifies any negative entries in critical columns:



### **5. Outlier Detection and Time Series Analysis**

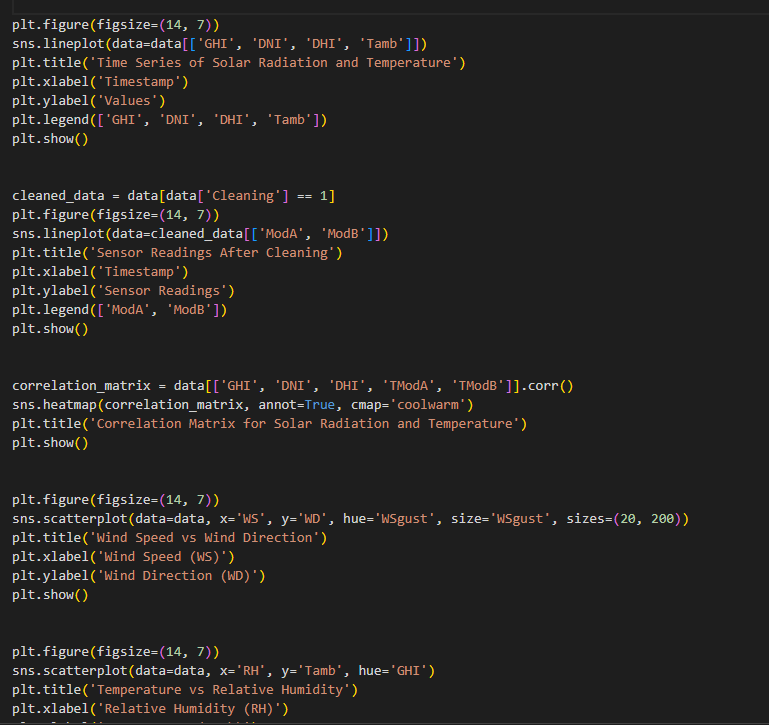
Outliers are detected using the Interquartile Range (IQR) method, specifically for sensor readings and wind speed: and The Timestamp column is converted to a datetime format, and various time series plots are generated to visualize trends in solar radiation and temperature over time:



### **6. Visualizing Relationships**

Visualizations such as scatter plots and histograms are used to analyze relationships between various variables:

* **Scatter Plots**: Used to assess the relationship between wind speed and direction.
* **Histograms**: Display the frequency distribution of key variables.

****

Name: Tesfanesh yisaiase