Name: Dhruv Kasana Course B.Tech. C.S.E.

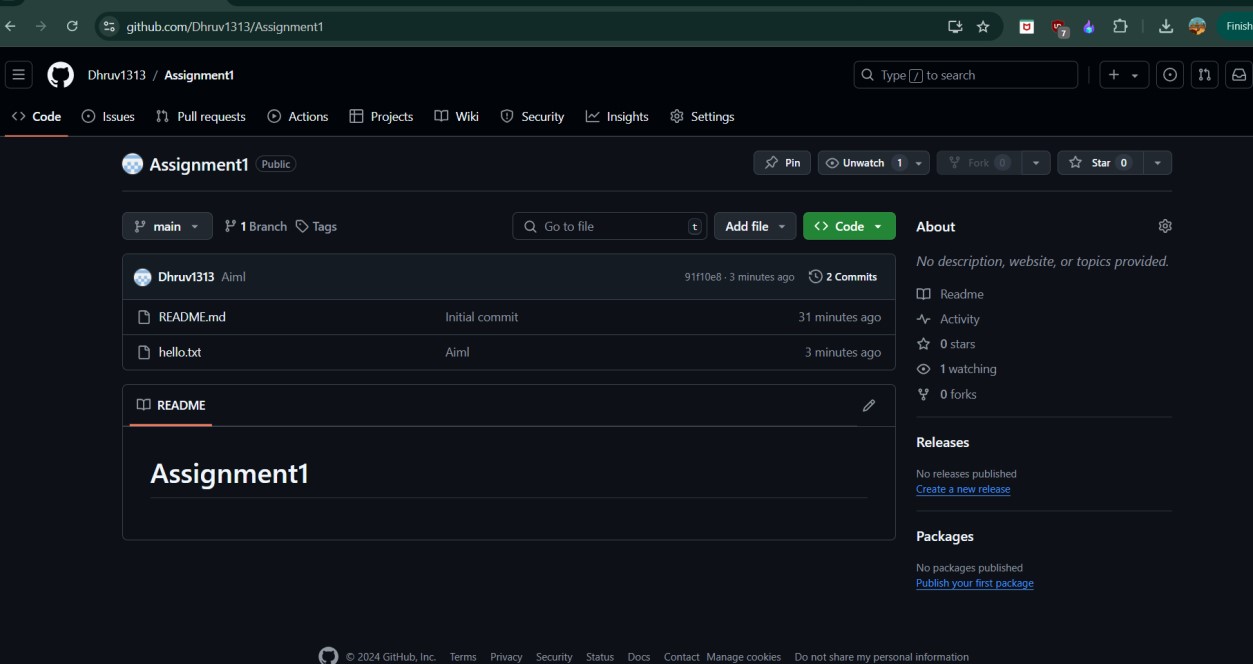


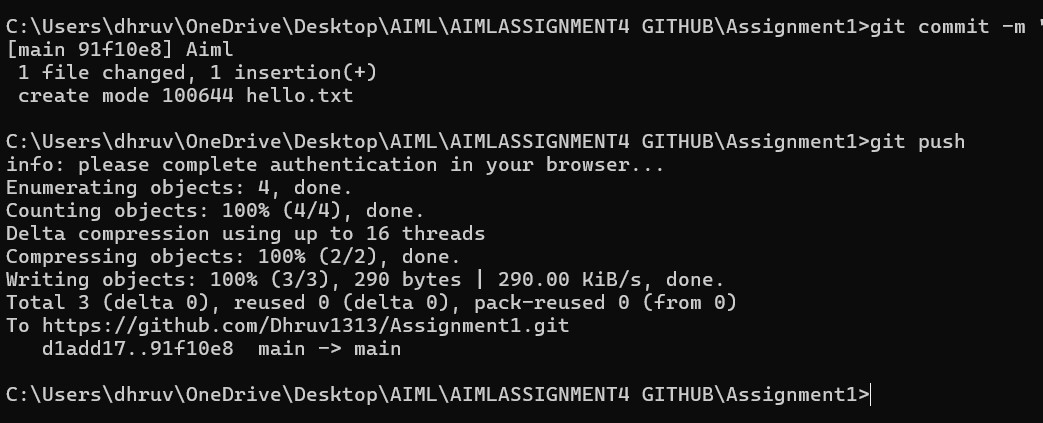
Batch: 11

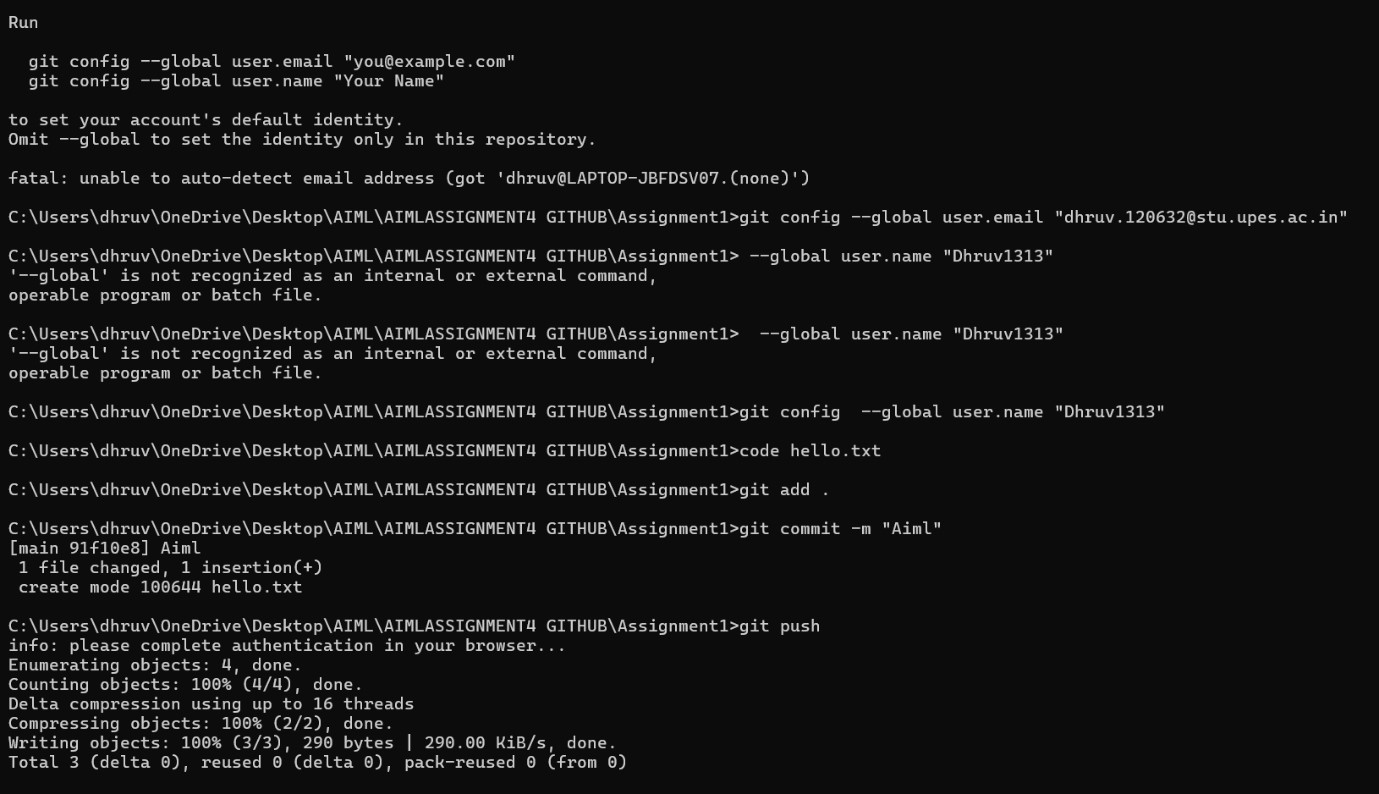
Sap.id.: 500120632

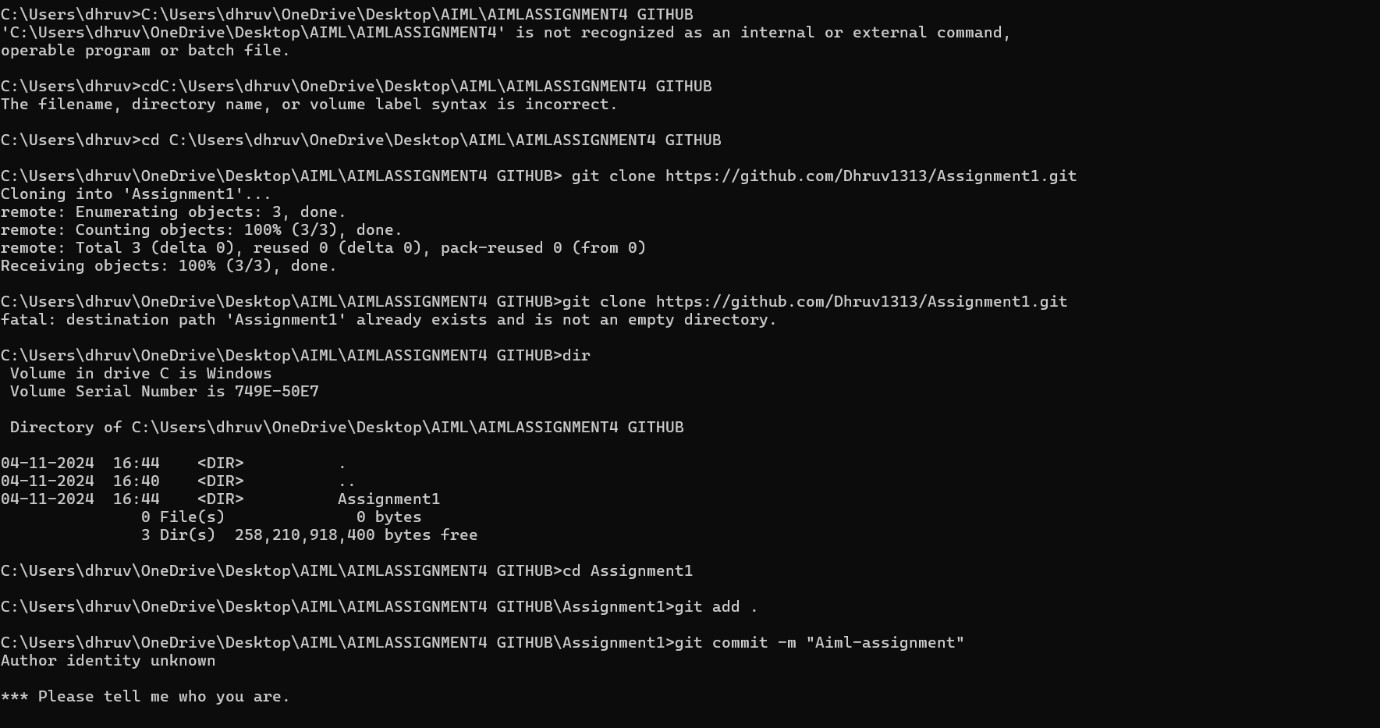
Enrolment no.: R2142230320

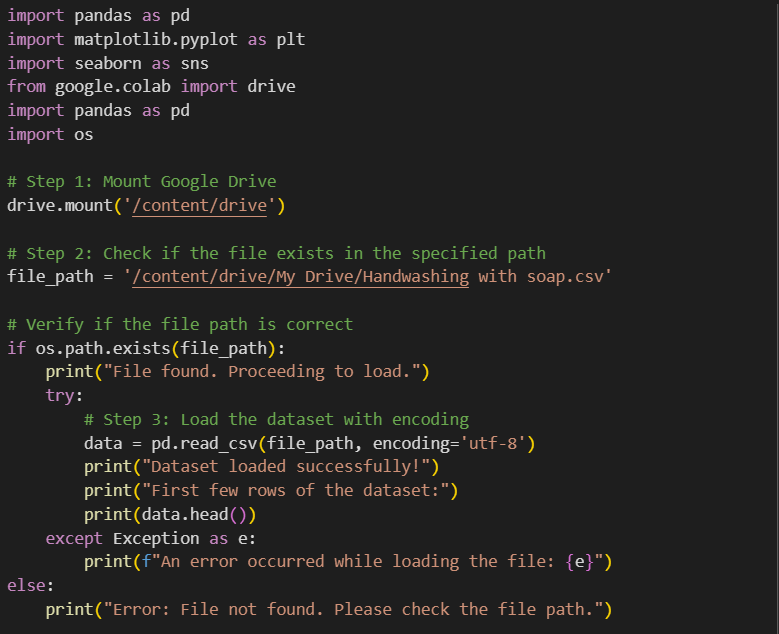
Instructor: Prof. C.M. Sherma

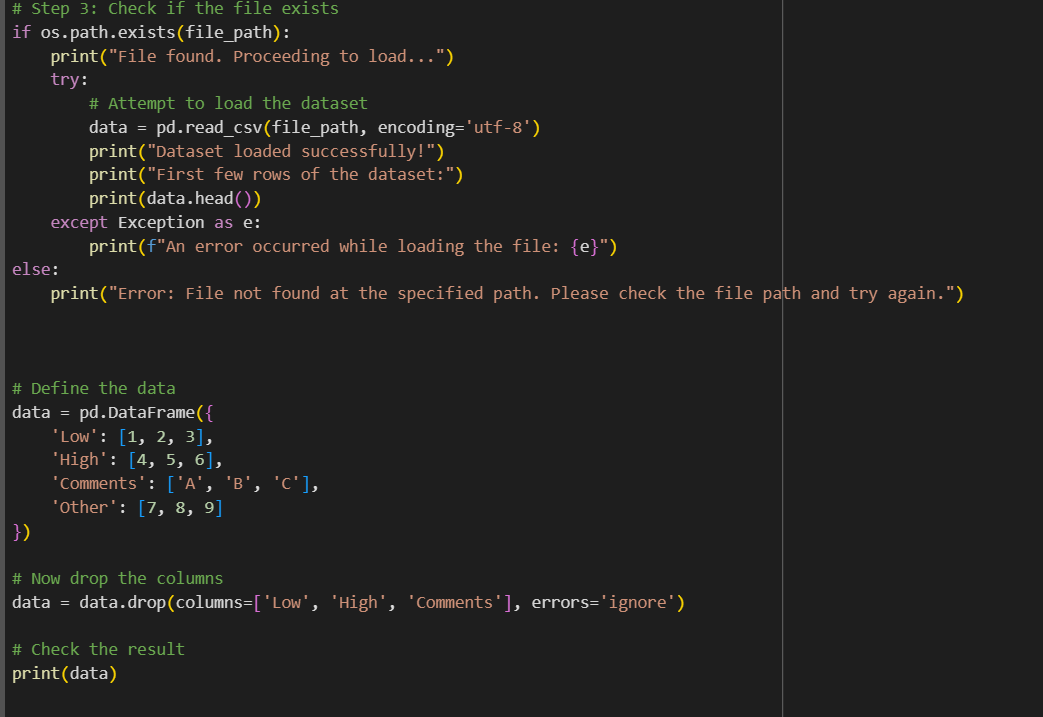


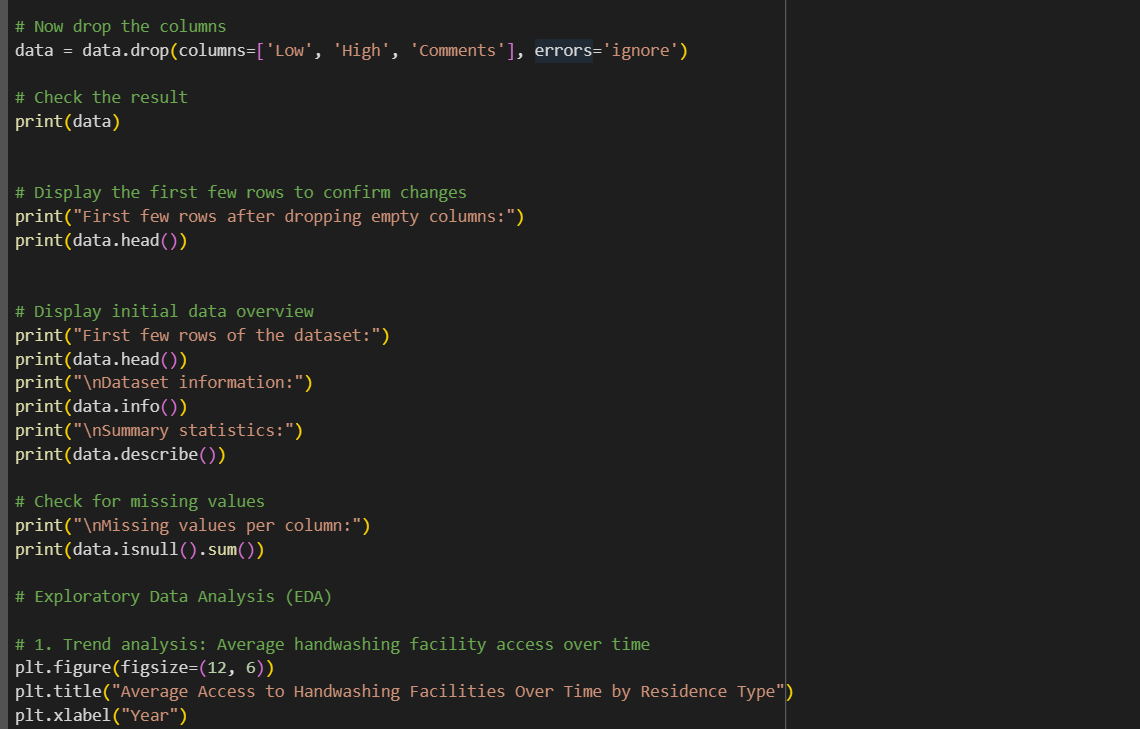


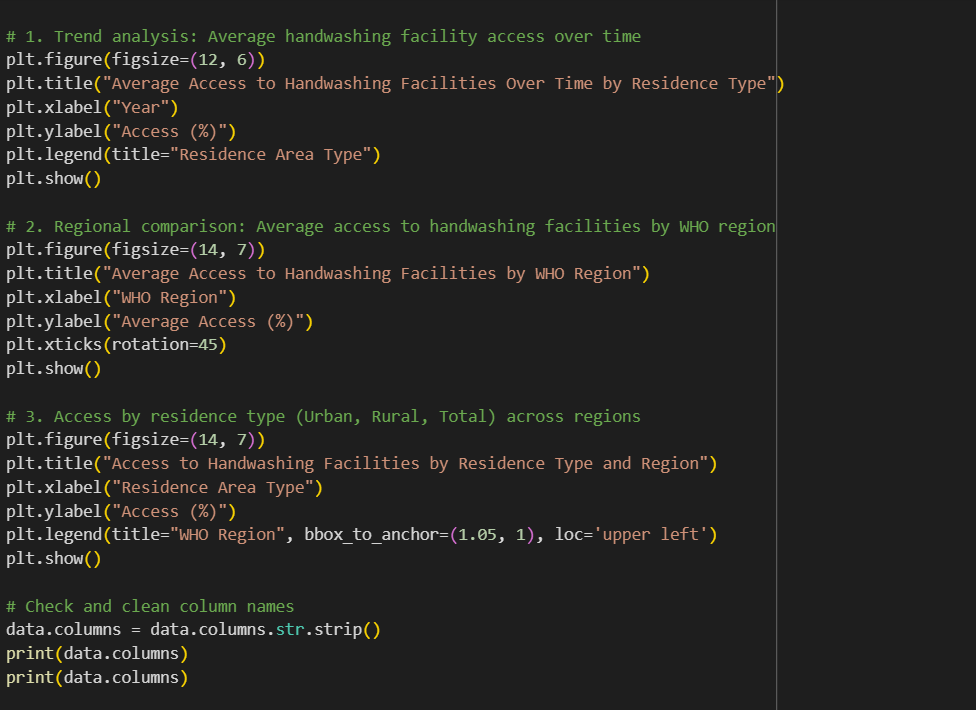


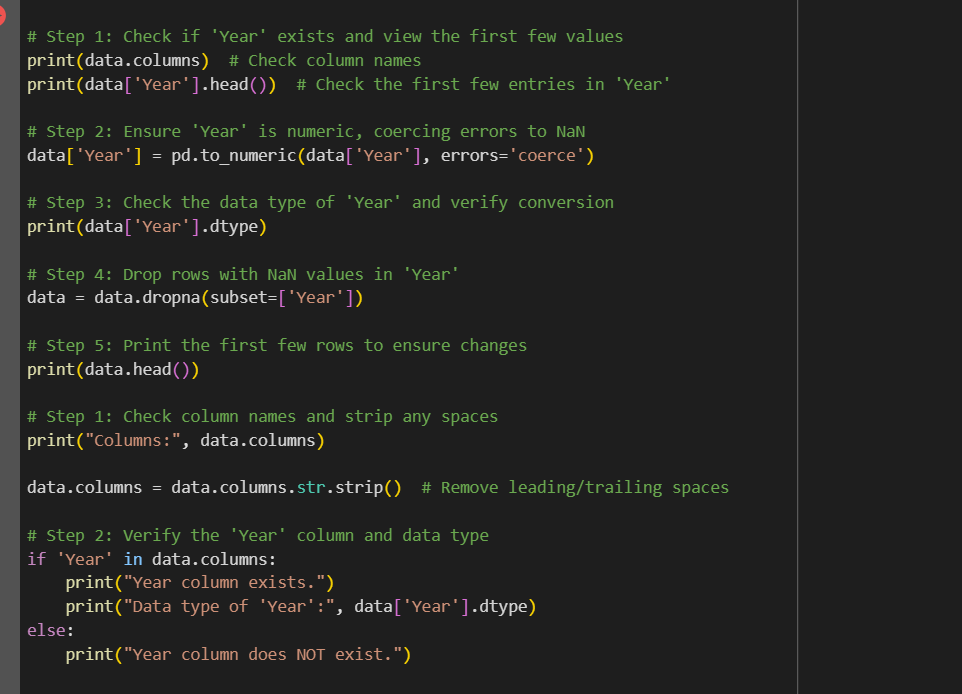


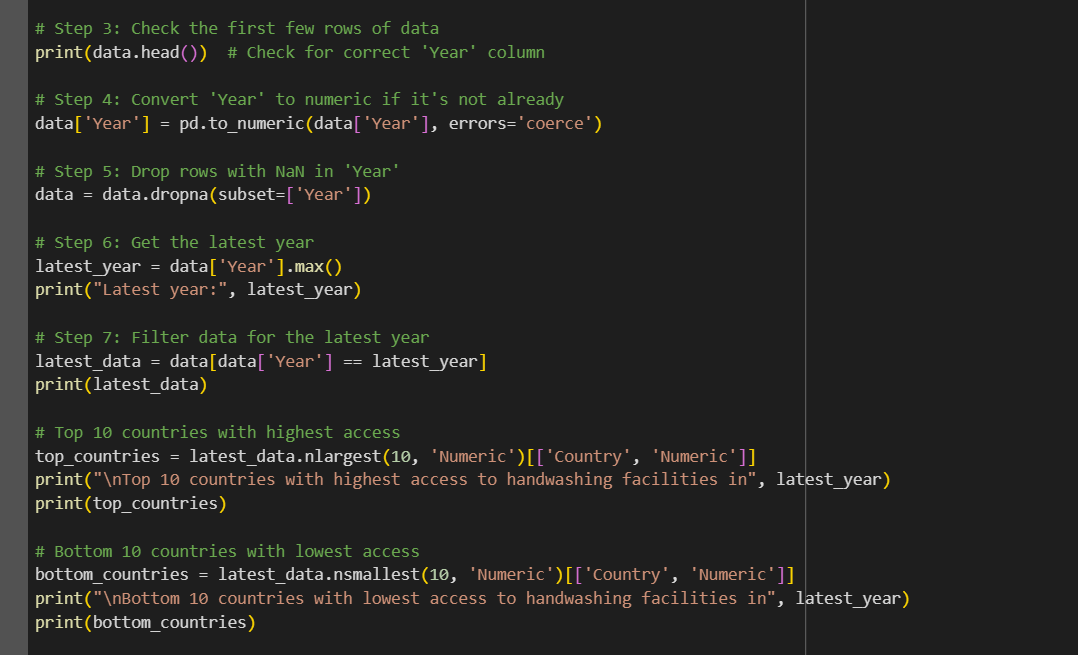












**Introduction:**

The **Clean Water and Sanitation Dataset** is a crucial dataset that provides insights into the global access to clean water and sanitation facilities. It is a part of the United Nations Sustainable Development Goals (SDGs), specifically focusing on **SDG 6**: Clean Water and Sanitation. This goal aims to ensure the availability and sustainable management of water and sanitation for all by 2030.

The dataset includes data on the percentage of populations that have access to safe drinking water and improved sanitation facilities across various countries and regions. It allows policymakers, researchers, and organizations to analyse the disparities in access to clean water and sanitation, track progress towards SDG 6, and plan interventions to improve public health globally.

**Purpose:**

The primary purpose of the Clean Water and Sanitation Dataset is to track and measure the progress made by countries in providing access to clean water and sanitation facilities. The key purposes of the dataset include:

* **Monitoring Progress**: Track global and regional progress towards ensuring universal access to clean water and sanitation.
* **Identify Gaps**: Identify regions or countries where access to clean water and sanitation is lacking, helping to prioritize interventions.
* **Aid Decision-Making**: Support governments, international organizations, and NGOs in making data-driven decisions to improve public health and water management.
* **Research and Analysis**: Enable academic researchers and policymakers to study the relationships between water access, health, and socio-economic factors.

**Dataset Used:**

The dataset typically includes the following variables:

* **Country/Region**: The geographic location for which the data is provided (e.g., country, region, or global).
* **Year**: The year for which the data was recorded (often annually).
* **Access to Safe Drinking Water (%)**: The percentage of the population with access to safe drinking water sources.
* **Access to Improved Sanitation (%)**: The percentage of the population with access to improved sanitation facilities (e.g., toilets, sewage systems).
* **Residence Area Type**: Data might be split by urban and rural populations to highlight differences in access.
* **Income Group**: Some datasets categorize countries by income group (low, middle, and high income) to show the relationship between income and access to water and sanitation.
* **Health Indicators**: Some datasets might include health-related variables (e.g., waterborne diseases, child mortality) linked to sanitation access.

**Libraries Used:**

To analyse the Clean Water and Sanitation dataset, several Python libraries might be used, including:

* **Pandas**: For data manipulation and cleaning (e.g., reading CSV files, filtering data, handling missing values).
* **NumPy**: For numerical operations (e.g., handling large arrays, performing statistical operations).
* **Matplotlib**: For data visualization (e.g., plotting graphs to visualize trends in access to clean water and sanitation).
* **Seaborn**: For statistical data visualization (e.g., heatmaps, box plots) to help identify patterns in the data.
* **SciPy**: For statistical tests and more advanced data analysis.
* **Plotly**: For interactive data visualizations.
* **Geopandas**: If the data involves spatial analysis or mapping.
* **Scikit-learn**: For performing machine learning and data modelling (if the analysis requires predictive modelling).
* SUMARY:
* The Clean Water and Sanitation Dataset provides a comprehensive view of global efforts to achieve universal access to clean water and sanitation. By analysing this dataset, we can understand how different regions and countries are performing in terms of providing clean drinking water and sanitation facilities to their populations. The dataset is vital for tracking progress toward **SDG 6**, which aims to ensure clean water and sanitation for all by 2030.
* The dataset includes variables related to the percentage of people with access to safe water and improved sanitation, and it may break this down by income group, geographic region, or year. This data is essential for policymakers, governments, and international organizations as they design strategies to address water and sanitation issues worldwide.

