PROJECT NAME: Cancer-Trends-Across-the-Globe---An-Analytical-Perspective-of-the-50-Most-Populated-Countries

Dataset Source: Kaggle  
Datasets Used: 3 (1 initially, 2 imported later)  
  
Dataset 1 (Initial Dataset):

Link to Dataset: <https://www.kaggle.com/datasets/ankushpanday1/cancer-datasettop-50-populated-countries>

Number of Rows: 160,000 rows  
Number of Columns: 23 columns  
Dataset Description:

**This dataset provides a detailed view of global cancer trends across the 50 most populated countries. With 160,000 records, it encompasses a wide range of variables including cancer types, risk factors, healthcare expenditure, and environmental factors. The data is designed to assist researchers, healthcare policymakers, and data scientists in identifying patterns, predicting future trends, and crafting effective cancer control strategies.**  
  
Dataset Variable Description:

Country, Age\_Group, Cancer\_Type, Risk\_Factors, Incidence, Mortality, Prevalence, Urban\_Population, Health\_Expenditure\_%GDP, Tobacco\_Use\_%, Alcohol\_Consumption\_Liters, Physical\_Activity\_%, Obesity\_%, Air\_Quality\_Index, UV\_Radiation, Family\_History\_%, Genetic\_Mutation\_%, Treatment\_Coverage\_%, GDP\_per\_Capita, Life\_Expectancy, Health\_Infrastructure\_Index, Education\_Index, Population\_Density.

Dataset 2:

Link to Dataset: <https://www.kaggle.com/datasets/folaraz/world-countries-and-continents-details?select=Countries+Longitude+and+Latitude.csv>

Number of Rows: 250  
Number of Columns: 27  
Dataset Description: World exploration for data scientist  
Dataset Variable Description:

name, official\_name\_en, official\_name\_fr, ISO3166-1-Alpha-2, ISO3166-1-Alpha-3, grid\_3x3M49, ITU, MARC, WMO, DS, Dial, FIFA, FIPS, grid\_3x3GAUL, IOC, ISO4217-currency\_alphabetic\_code, ISO4217-currency\_country\_name, grid\_3x3ISO4217-currency\_minor\_unit, ISO4217-currency\_name, grid\_3x3ISO4217-currency\_numeric\_code, is\_independent, Capital, Continent, TLD, Languages, vpn\_keyGeoname ID, EDGAR,

Dataset 3:

Link to Dataset: <https://www.kaggle.com/datasets/folaraz/world-countries-and-continents-details?select=Countries+Longitude+and+Latitude.csv>Number of Rows: 250  
Number of Columns: 4  
Dataset Description:

### **Context**

Can you tell geographical stories about the world using data science?

### **Content**

World countries with their corresponding continents , official english names, official french names, Dial,ITU,Languages and so on.

### **Acknowledgements**

This data was gotten from <https://old.datahub.io/>

### **Inspiration**

Exploration of the world countries:

* Can we graphically visualize countries that speak a particular language?
* We can also integrate this dataset into others to enhance our exploration.
* The dataset has now been updated to include longitude and latitudes of countries in the world.

Dataset Variable Description:

ID, Longitude, Latitude, Name

STEPS CONDUCTED: