

PLOT-A-THON

HEALTHY DIET DATASET

Submission Link: <https://forms.gle/2Nktv2STt2na6Mzz7>

Instructions:

- Merge the files and pre-process the database.
- Create visuals based on the following questions.
- The Visualization output should be in pdf format.
 - There should be creativity in the visuals.
 - The logical questions should be solved.
 - There should be an explanation for each visual.

Preprocessing Task:

The candidate will be required to:

1. Merge all four datasets based on appropriate keys (e.g., Entity, Code, and Year).
2. Handle missing values in all numerical columns (e.g., fill with mean or median).
3. Standardize the Entity column to ensure consistent naming of countries.
4. Generate derived metrics:
 - **Affordability Gap** = Share of the population who cannot afford a healthy diet
- Share of the population who cannot afford sufficient calories.
 - **Food Expenditure per Capita** = Total food expenditure / Food as share of total consumer expenditure.

Visualization Questions (5 Questions):

1. Plot the **distribution of countries** based on the percentage of the population unable to afford a healthy diet.
2. Show the **trend of total food expenditure** globally over a specific range of years.
3. Compare the **food expenditure per capita** across different countries for a specific year using a bar chart.
4. Visualize the **affordability gap** for the top 5 countries with the highest gap in a specific year.
5. Create a geographic heatmap of **food expenditure per capita** by country.