

Data Collection and Preprocessing Phase

Date	20 June 2025
Team ID	Chaitra V
Project Title	Global Malnutrition Trends: A Power BI Analysis (1983-2019)
Maximum Marks	10 Marks

Data Exploration and Preprocessing Template

Identifies data sources, assesses quality issues like missing values and duplicates, and implements resolution plans to ensure accurate and reliable analysis.

Section	Description
Data Overview	The dataset was obtained from reliable sources including UNICEF, WHO, and the World Bank, containing detailed malnutrition statistics for children under the age of five between 1983 and 2019, covering indicators such as underweight, overweight, stunting, wasting, and income classifications globally.
Data Cleaning	Data cleaning was essential to ensure accuracy by removing duplicate records, correcting data entry errors, and handling missing values, thereby preparing the dataset for smooth analysis and avoiding misinterpretations that could arise from incomplete or incorrect information.
Data Transformation	Power Query was utilized extensively for transforming the data, including filtering out irrelevant entries, sorting the dataset based on income classifications and country, creating calculated columns for additional insights, and pivoting data structures to make trend identification and analysis easier.
Data Type Conversion	Data type conversion was carried out to ensure consistency across the dataset by rectifying mismatches, such as converting year columns to numerical format and appropriately formatting text, numeric, and categorical fields, which is crucial for accurate calculations and visualizations.

Column Splitting and Merging	Column splitting and merging were performed where necessary to reorganize the dataset, such as separating combined fields for clearer filtering or merging relevant columns to simplify analysis and enhance the usability of the dataset within Power BI.
Data Modeling	Data modeling was conducted by defining logical relationships between tables and fields, ensuring seamless connections between country, income group, malnutrition indicators, and survey data, along with creating calculated measures that enable accurate, dynamic visualizations and insights within the Power BI environment.
Save Processed Data	The cleaned, transformed, and modeled dataset was saved in multiple formats, including <code>.pbix</code> for Power BI dashboards and <code>.xlsx</code> for raw data reference, ensuring the data remains accessible, reusable, and ready for future analysis or project demonstrations.