

## Data Collection and Preprocessing Phase

Date	24 June 2025
Team ID	XXXXXX
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	Short description about dataset The dataset used in this project is a comprehensive collection of global malnutrition estimates for children under five. It covers the years 1983 to 2019 and includes key indicators such as severe wasting, wasting, stunting, underweight, and overweight. The data is organized by country and classified by income level and special groups like Least Developed Countries (LDC), Low Income Food Deficit (LIFD), Landlocked Developing Countries (LLDC), and Small Island Developing States (SIDS). This rich dataset enables analysis of malnutrition trends over time and across different economic and geographic categories.
Data Cleaning	<ul style="list-style-type: none"> <li>• Handle missing values- deleting rows with empty cells or writing default values such as 0</li> <li>• Duplicates- deleting duplicate values by right clicking on the column and selecting remove duplicates</li> <li>• Remove errors- removing errors by right clicking on the column and selecting remove errors</li> </ul>
Data Transformation	Use of Power Query Editor to <ol style="list-style-type: none"> <li>1. remove null values (by clicking on the filter symbol beside the column name and deselecting the null values)</li> </ol>

	<ol style="list-style-type: none"> <li>change data types of many columns into correct type(Eg: column underweight datatype is changed from decimal number to fixed decimal number.)</li> <li>change the format of the data in the column( by using the column we can change data format ,Eg: changing U5 population format into percentage)</li> </ol>
Data Type Conversion	<p>When working with data in Power BI, it's important to make sure each column has the right type like numbers for malnutrition rates, dates for years, and text for country names. I use Power Query to check and change these types so everything works smoothly. Sometimes I clean up errors or fix tricky values before loading the data. Doing this makes sure the visuals and calculations in the report are accurate and easy to understand.</p>
Column Splitting and Merging	<p>Both actions are done easily in Power Query by selecting columns and choosing the split or merge options, helping you prepare data that's clean and well-structured for analysis. In this project I didn't split or merge columns.</p>
Data Modeling	<p>In Power BI, data modeling means organizing your tables so they work well together. First, you define relationships between tables for example, linking the malnutrition data table with the country classification table using a common field like "Country Name" or "Country Code." This lets you combine data from different tables smoothly.</p>
Save Processed Data	<p>Save the cleaned and processed data for future use.</p>