Data Collection and Preprocessing Phase

|  |  |
| --- | --- |
| Date | 15 June 2024 |
| Project Title | Global Malnutrition Trends: A Power BI Analysis (1983-2019) |
| Maximum Marks | 3 Marks |

Data Quality Report Template

The Data Quality Report summarizes the data quality issues identified in the two selected datasets, along with their severity levels and resolution plans. This ensures the reliability and consistency of data for analysis.

|  |  |  |  |
| --- | --- | --- | --- |
| **Data Source** | **Data Quality Issue** | **Severity** | **Resolution Plan** |
| **country-wise- average.csv** | Presence of missing values in some columns like "Overweight" and "Underweight" for certain countries. | Moderate | Impute missing values using mean/median of respective columns or remove  rows if too few. |
| **country-wise- average.csv** | Inconsistent country names (possible duplicates or mismatches across datasets). | High | Standardize country names using Power Query text transformations or reference  mappings. |
| **country-wise- average.csv** | Some numeric columns detected as text (data type mismatch). | Moderate | Convert data types (e.g., to decimal or whole number) using Power BI or  Excel Power Query. |
| **malnutrition-**  **estimates.csv** | Duplicate records for certain  country-year combinations. | High | Remove duplicates  based on country |

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  | and year using "Remove Duplicates" in  Power Query. |
| **malnutrition- estimates.csv** | Null or blank fields in critical measures (e.g., Stunting, Wasting). | High | Apply forward/backward fill or use statistical imputation methods; ensure critical fields have  no blanks. |
| **malnutrition- estimates.csv** | Incorrect data type for "Year" column (treated as text). | Low | Convert to integer or date format for time-based  analysis. |
| **Both Datasets** | Mismatch in country names between both datasets (join problem). | High | Create a country mapping table to align country names across both datasets for accurate  relationships. |