**Data Collection and Preprocessing Phase**

|  |  |
| --- | --- |
| Date | 6 October 2025 |
| Team ID | SWUID20250197150 |
| Project Title | Global Food Production Trends And Analysis (1961 – 2023) |
| Maximum Marks | 3 Marks |

**Data Quality Report Template**

The Data Quality Report Template will summarize data quality issues from the selected source, including severity levels and resolution plans. It will aid in systematically identifying and rectifying data discrepancies.

|  |  |  |  |
| --- | --- | --- | --- |
| **Data Source** | **Data Quality Issue** | **Severity** | **Resolution Plan** |
| Kaggle  Global Food Production  dataset | Missing values in the Value column for some countries and years. | High | Use Power Query in Power BI to handle missing values: either remove rows with insufficient data or use interpolation for continuous years. |
| Kaggle  Global Food Production  dataset | Duplicate rows based on Country, Year, and Indicator | Moderate | Remove duplicates using Power Query’s “Remove Duplicates” function. |
| Kaggle  Global Food Production  dataset | Outliers or incorrect malnutrition values (e.g., negative or extremely high values) | Moderate | Flag outliers using DAX measures and review for correction or removal. |
| Kaggle  Global Food Production  dataset | Data type mismatches (e.g., Year as text, Value as text) | Low | Convert columns to correct data  types in Power Query: Year →  Whole Number, Value → Decimal → Number. |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |