Data Collection and Preprocessing Phase

Date	30 September 2025
Team ID	SWUID20250181744
Project Title	Global Food Production Trends and Analysis: A Comprehensive Study from 1961 to 2023 Using Power BI
Maximum Marks	3 Marks

Data Quality Report

The Data Quality Report 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 (Technical Solution)
Global Food Production Dataset (1961–2023)	Missing values in production data for certain years and commodities due to incomplete reporting by countries.	High	Used Power Query's "Replace Missing Values" and interpolation techniques to fill gaps based on historical averages or nearest available data. For non-critical gaps, replaced nulls with zeroes to maintain aggregation consistency.
Global Food Production Dataset (1961–2023)	Blank rows and metadata lines in raw CSV files.	Low	Removed unnecessary header/footer rows using Remove Top/Bottom Rows and Filter Non-Null Rows in Power Query during the data import stage.
Global Food Production Dataset (1961–2023)	Unbalanced data across regions (e.g., some countries missing multiple commodities).	Moderate	Identified coverage gaps using count and distinct measures in Power BI; documented limitations while ensuring consistent calculations across available data points.