

## Data Collection and Preprocessing Phase

Date	14 September 2025
Team ID	SWUID20250172428
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 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
FAOSTAT / Kaggle CSV	Missing values in production, yield, or area harvested columns	High	Handled using interpolation or replacing with averages; ensured consistency across years and countries.
FAOSTAT / Kaggle CSV	Duplicate rows for same country, crop, and year	Moderate	Removed duplicate records using Power Query's "Remove Duplicates" feature.
FAOSTAT / Kaggle CSV	Inconsistent country and crop names (e.g., "U.S.A." vs "USA")	Moderate	Standardized names using Find & Replace and reference lookup tables in Power Query.

FAOSTAT / Kaggle CSV	Combined columns like “1961F” for year/type	Low	Split columns using Power Query “Split Column” by position to separate Year and Type.
FAOSTAT / Kaggle CSV	Data type mismatches (numeric columns stored as text)	High	Converted Year to Whole Number, Production/Area/Yield to Decimal Number, and names to Text.