

Data Collection and Preprocessing Phase

Date	15 Sept 2025
Team ID	SWUID20250185946
Project Title	Global Food Production Trends and Analysis:
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
Fasostat / Kaggle Csv	Missing Values in production.	Moderate	Missing values were addressed through interpolation or by replacing them with averages, ensuring consistency across years and countries for reliable analysis.
Fasostat / Kaggle Csv	Duplicate rows for same country, crop, and year.	Moderate	Duplicate records were eliminated using Power Query's "Remove Duplicates" function to ensure data integrity and prevent redundancy in analysis.
Fasostat / Kaggle Csv	Data type mismatches (numeric columns stored as text).	High	Converted Year to Whole Number, Production/Area/Yield High to Decimal Number, and names to Text.

Fasostat / Kaggle Csv	Combined columns like “1961F” for year/type.	Low	Split columns using Power Query ‘Split Column’ by position to separate Year and Type.
-----------------------------	---	-----	---