

Data Collection and Preprocessing Phase

Date	15 March 2024
Team ID	SWUID20250213607
Project Title	Global Food Production Trends and Analysis: A Comprehensive Study from 1961 to 2023 Using Power BI
Maximum Marks	10 Marks

Data Exploration and Preprocessing Template

Identifies data sources, assesses quality issues like missing values and duplicates, and implements resolution plans to ensure accurate and reliable analysis.

Section	Description
Data Overview	Dataset contains global crop production data (rice, wheat, maize, coffee, tea, grapes, apples, bananas, oranges) from 1961 to 2023, sourced from FAO and World Bank databases.
Data Cleaning	Handled missing values for some years by interpolation.
Data Transformation	Used Power Query to filter relevant years and crops, sort data by year and entity, pivot data to wide format, and created calculated columns for production totals and growth percentages.
Data Type Conversion	Converted year columns to date format, production values to numeric types, and entity names as text fields for consistency in Power BI.
Column Splitting and Merging	Merged country and region columns where applicable for uniformity; split combined production fields into specific crop columns as needed.
Data Modeling	Define relationships between tables and create measures.

Save Processed Data	Saved the cleaned and transformed dataset as PBIX data model file and exported cleaned CSV for future use and reproducibility.
---------------------	--