

## Data Collection and Preprocessing Phase

Date	11 November 2025
Team ID	VIP-C3
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	A longitudinal dataset covering global production volumes of staple crops (Wheat, Rice, Maize) and fruits (Grapes, Apples, etc.) from 1961 to 2023.
Data Cleaning	Identification and removal of duplicate entries for specific years; handling null values in regional columns to ensure "Africa" coffee production totals are accurate.
Data Transformation	Using <b>Power Query</b> to filter for specific commodities and create a "Total Grain" calculated column combining Wheat, Rice, and Maize
Data Type Conversion	Converting "Year" to a Date format and "Production (tonnes)" to a Decimal Number to enable mathematical aggregations.
Column Splitting and Merging	Splitting "Entity" columns if they contain both Country and Continent data to allow for better regional filtering.
Data Modeling	Creating relationships between the Production table and a separate Date table; developing <b>DAX measures</b> for "Year-over-Year Growth %".

Save Processed Data	Saving the cleaned dataset as a Power BI Template (.pbix) or exporting to a clean CSV for documentation purposes.
---------------------	---