

Data Collection and Preprocessing Phase

Date	25 March 2024
Team ID	44
Project Title	Global Food Production Trends and Analysis
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	The dataset contains 11913 rows and 24 columns. It includes 24 categorical column (Item), 2 numerical columns (1961-2020, Tonnes), and 40geographical column (Country).
Data Cleaning	Removed rows with missing values and eliminated duplicate entries to maintain data consistency.
Data Transformation	Used Power Query for filtering irrelevant data, sorting by year and value, and creating new calculated columns for year-wise growth.
Data Type Conversion	Converted text-based year and value columns to proper numeric types for accurate analysis.
Column Splitting and Merging	No column splitting or merging was performed.
Data Modeling	No relationships between tables were defined as only one primary dataset was used.
Save Processed Data	The cleaned and transformed data was saved as a new CSV file for further visualization and reporting.