**Data Collection and Preprocessing Phase**

|  |  |
| --- | --- |
| Date | 24 June 2025 |
| Team ID |  |
| Project Title | Global Food Production Trends and Analysis (1961–2023) using power bi |
| Maximum Marks | 2 Marks |

**Data Collection Plan & Raw Data Sources Identification Template**

Elevate your data strategy with a robust Data Collection Plan and a clearly defined Raw Data Sources report. For this project, datasets from FAOSTAT, World Bank, and UN repositories ensure historical coverage (1961–2023) across key commodities and regions. This meticulous approach to data curation, preprocessing, and validation safeguards integrity and accuracy, laying the foundation for reliable insights. By structuring data effectively and documenting every source, the project empowers policymakers, researchers, and agribusiness stakeholders to make informed, evidence-based decisions that strengthen food security and agricultural planning worldwide.

**Data Collection Plan Template**

|  |  |
| --- | --- |
| **Section** | **Description** |
| Project Overview | |  |  |  | | --- | --- | --- | | |  | | --- | | This project analyzes global food production trends from 1961 to 2023, focusing on crop yields, harvested area, and production volumes across regions and commodities. The objective is to identify leading producers, emerging contributors, and long-term shifts while providing interactive Power BI dashboards and insights to support policy-making, research, and agribusiness planning. |  |  | | --- | |  | |  |  | | --- | |  | |
| Data Collection Plan | Data is sourced from FAOSTAT, UN, and World Bank repositories. Key variables include production (tonnes), harvested area (hectares), and yield (tonnes per hectare) for major food commodities. Additional population data is used to compute per-capita production. Data is collected annually for each country and commodity from 1961 to 2023. |
| Raw Data Sources Identified | - **FAOSTAT Production Data:** Annual country-wise data for crops (1961–2023).  - **FAOSTAT Harvested Area & Yield Data:** Used to calculate productivity trends.  - **Population Data:** World Bank/UN datasets for per-capita analysis.  - **Commodity Classifications:** FAO commodity codes for grouping items.  - **Regional Mapping:** Country-to-region mapping for aggregation and comparison. |

**Raw Data Sources Template**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Source Name** | **Description** | **Location/URL** | **Format** | **Size** | **Access Permissions** |
| **World Food Production**  Different Food Production Data Since 1961-2023 | **About Dataset**  This dataset contains data of total production of different types of foods in per year in each country from 1961-2023 of all country. | <https://www.kaggle.com/datasets/rafsunahmad/world-food-production> | CSV | 856 KB | Public |