**Data Collection and Preprocessing Phase**

|  |  |
| --- | --- |
| Date | 01 October 2025 |
| Team ID | SWUID20250215818 |
| Project Title | Global Food Production Trends and Analysis:  A Comprehensive Study from 1961 to 2023 Using Power BI |
| Maximum Marks | 2 Marks |

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

Elevate your data strategy with the Data Collection plan and the Raw Data Sources report, ensuring meticulous data curation and integrity for informed decision-making in every analysis and decision-making endeavor.

**Data Collection Plan Template**

|  |  |
| --- | --- |
| **Section** | **Description** |
| Project Overview | The project “Global Food Production Trends and Analysis (1961–2023)” aims to analyze agricultural production patterns across countries and regions to uncover global food production trends. The study focuses on identifying top-producing regions, crop-wise growth, and decade-wise production shifts using Power BI for interactive data visualization. |
| Data Collection Plan | The dataset was collected from Kaggle, a public platform for data science resources. The data represents global food and crop production from 1961 to 2023, including metrics such as production volume (tonnes), year, entity (country/region), and crop type. |
| Raw Data Sources Identified | The raw dataset, titled “FAO Global Food Production Dataset (1961–2023)”, was obtained from Kaggle. The data was originally compiled from the Food and Agriculture Organization (FAO). It contains over 1,80,000 records detailing production quantities across various crops, countries, and years. The dataset was downloaded in CSV format and later cleaned and transformed using Power Query in Power BI. |

**Raw Data Sources Template**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Source Name** | **Description** | **Location/URL** | **Format** | **Size** | **Access Permissions** |
| Dataset 1 | Description of the data in this source. | Link of Dataset 1 | CSV | XX GB | Public |
| Dataset 2 | Description of the data in this source. | Link of Dataset 2 | Excel | YY GB | Private (with access) |
| … | … | … | … | … | … |