**Project Initialization and Planning Phase**

|  |  |
| --- | --- |
| Date | 24 June 2025 |
| Team ID |  |
| Project Title | Global Food Production Trends and Analysis (1961–2023) using power bi |
| Maximum Marks | 3 Marks |

**Project Proposal (Proposed Solution) template**

This project analyzes global food production from 1961 to 2023, focusing on trends in crop yields, harvested areas, and production volumes across countries and regions. Using FAOSTAT and related datasets, the study identifies top producers, emerging contributors, and shifts in commodity patterns. Data will be cleaned, modeled, and visualized in Power BI dashboards to provide insights, comparisons, and simple forecasts. The outcomes will support policymakers, researchers, and agribusinesses in addressing food security and planning challenges.

|  |  |
| --- | --- |
| **Project Overview** | |
| Objective | The primary objective of this project is to examine global food production trends from 1961 to 2023 and derive meaningful insights for food security and agricultural planning. Specifically, the project aims to collect and preprocess production data from reliable sources, analyze long-term patterns in yields, harvested areas, and output across countries and commodities, and identify both leading and emerging producers. Additional objectives include evaluating per-capita production to understand population–food balance, creating interactive dashboards in Power BI for intuitive visualization, and generating simple forecasts. These objectives will support policymakers, researchers, and businesses in decision-making and planning. |
| Scope | This project focuses on analyzing global food production data from 1961 to 2023, covering major commodities such as cereals, pulses, oilseeds, sugar crops, fruits, vegetables, and roots & tubers. The analysis spans multiple countries and regions to capture both global patterns and regional variations. Key metrics include production quantity, harvested area, yield, and per-capita availability. The scope also involves building a structured data model, cleaning and preprocessing datasets, and creating interactive dashboards in Power BI for visualization. Price data and trade flows are excluded, ensuring the study remains focused on production trends and agricultural productivity. |
| **Problem Statement** | |
| Description | Global food production has significantly increased since 1961, yet challenges such as unequal regional distribution, population growth, climate change, and resource limitations continue to threaten food security. While some countries have achieved remarkable yield improvements, others still depend on expanding agricultural land, leading to sustainability concerns. Policymakers, researchers, and businesses lack an integrated, long-term analytical view of how production patterns have evolved across commodities and regions. Without such insights, it is difficult to plan effectively for future demand, assess vulnerabilities, or identify emerging opportunities. This project addresses the need for data-driven analysis of food production trends. |
| Impact | The absence of comprehensive analysis on global food production trends limits the ability of governments, researchers, and agribusinesses to make informed decisions. Without clarity on historical shifts, yield improvements, and emerging production hotspots, policymakers struggle to design effective food security strategies. This gap also hinders sustainable resource management, leaving vulnerable regions exposed to risks from climate change and population growth. Businesses may miss investment opportunities in high-potential commodities or regions. By not addressing these insights, stakeholders risk inefficient planning, reduced productivity, and widening inequalities in food availability and access worldwide. |
| **Proposed Solution** | |
| Approach | * **Data Collection:** Gather food production datasets (1961–2023) from FAOSTAT and other reliable sources.   + **Data Cleaning & Preprocessing:** Standardize country names, commodities, units, and handle missing values for consistency.   + **Data Modeling:** Organize data into a star schema with fact and dimension tables for efficient analysis.   + **Visualization:** Build interactive Power BI dashboards to showcase global, regional, and commodity-wise trends.   + **Advanced Analysis:** Compute CAGR, per-capita production, and global share for deeper insights.   + **Forecasting:** Apply simple forecasting methods to predict future production trends.   + **Decision Support:** Enable policymakers, researchers, and businesses to make data-driven decisions for food security. |
| Key Features | The proposed solution is designed to deliver clear, actionable insights into global food production trends. Its key features include comprehensive data coverage from 1961 to 2023, ensuring a long-term perspective across commodities and regions. The solution applies robust data cleaning and preprocessing techniques to maintain accuracy and reliability. A structured data model supports efficient analysis, while interactive Power BI dashboards make insights easily accessible through visual storytelling. Advanced metrics such as CAGR, per-capita production, and global share enrich the analysis. Additionally, simple forecasting highlights future trajectories, enabling proactive decision-making for policymakers, researchers, and agribusiness stakeholders. |

**Resource Requirements**

|  |  |  |
| --- | --- | --- |
| **Resource Type** | **Description** | **Specification/Allocation** |
| **Hardware** | | |
| Computing Resources | CPU/GPU specifications, number of cores | |  | | --- | | 4 vCPUs (cloud VM), GPU  not required; scalable cloud  compute on Azure/AWS |  |  | | --- | |  | |
| Memory | RAM specifications | 8–16 GB RAM per machine; cloud compute with 16 GB RAM for Power BI dashboard rendering |
| Storage | Disk space for data, models, and logs | e.g., 1 TB SSD or cloud storage (Google Drive, AWS S3, OneDrive) |
| **Software** | | |
| Frameworks | |  | | --- | | Backend/data frameworks |  |  | | --- | |  | | Power BI (for dashboard), Python (optional: Flask or Streamlit for custom tools) |
| Libraries | |  | | --- | | Analytical and data processing  libraries |  |  | | --- | |  | | |  | | --- | | pandas, numpy, scikit-learn,  matplotlib (for preprocessing or  modelling) |  |  | | --- | |  | |
| Development Environment | IDE, version control | |  | | --- | | Jupyter Notebook, Visual Studio  Code, Git & GitHub for version  management |  |  | | --- | |  | |
| **Data** | | |
| Data | Source, size, format | Kaggle dataset,  FAOSTAT datasets (1961–2023), World Bank/UN population data; CSV/Excel format; ~200–300 MB, expandable to cloud feed |