Final Report

1. Introduction

1.1 Project Overview

Agriculture plays a vital role in global development, food security, and economic sustainability. With growing populations and environmental pressures, understanding long-term patterns in global food production has become crucial for policymakers, researchers, and sustainability experts.

This project, titled "Global Food Production Trends and Analysis: A Comprehensive Study from 1961 to 2023 Using Power BI," aims to analyze over six decades of worldwide agricultural production data. It leverages interactive dashboards and visual analytics to uncover key insights regarding food crop production, yield, harvested areas, and regional trends. By transforming raw data into meaningful visuals, the project supports informed decision-making and highlights critical agricultural dynamics across countries and commodities.

Using Microsoft Power BI, the project presents a powerful, interactive platform where users can filter by country, crop, time period, and production metric. The dashboard enables easy interpretation of complex datasets and reveals patterns that would otherwise remain hidden in traditional spreadsheet formats. The study includes major data dimensions such as production volume, yield efficiency, harvested area, and item/crop-wise performance, offering valuable insights into global agricultural systems.

1.2 Objectives

The key objectives of this project are:

- 1. **To collect and process** global food production data from 1961 to 2023 from trusted sources like FAOSTAT, the World Bank, and Our World in Data.
- 2. **To clean and transform** the data for analysis using Power BI, ensuring accuracy, consistency, and usability.
- 3. **To develop interactive visualizations** that reflect country-wise, crop-wise, and element-wise (e.g., yield, area harvested) trends.

- 4. **To identify key production trends**, including topperforming crops (like maize), high-yield countries, and regional disparities.
- 5. **To create an intuitive dashboard** where users can filter by country, year, item, and production metric.
- 6. **To support data-driven decision-making** for policymakers, researchers, and agricultural analysts.
- 7. **To highlight historical changes and current trends** that may influence future food security and sustainability policies.
- 8. **To provide a user-friendly reporting structure** that can be used in academic, institutional, or policybased environments.
- 9. To evaluate the quality of the raw dataset, addressing missing values, duplicates, and inconsistencies.
- 10. **To make recommendations** for future analysis, improvements, and real-world application of insights.

Project Initialisation and Planning Phase

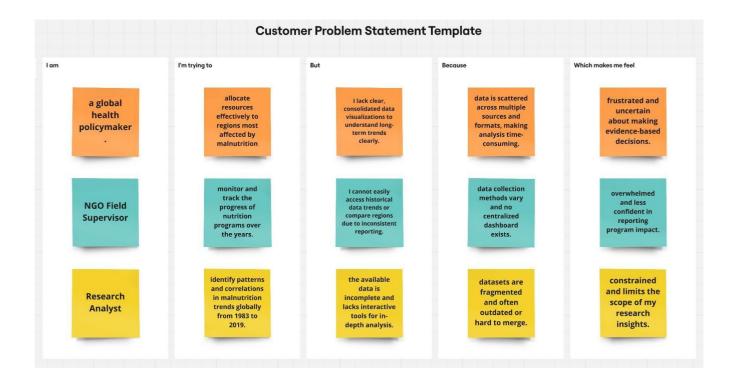
2.1 Define Problem Statement

Global food production data is vast, complex, and often underutilized due to a lack of accessible and interactive analysis tools. Stakeholders such as researchers, policymakers, and agricultural planners face difficulties in extracting meaningful insights from raw data that spans several decades and countries. The absence of a user-friendly, visual platform hampers decision-making in key areas like food security, crop management, and sustainability planning.

This project addresses the challenge by developing an **interactive Power BI dashboard** that visualizes global food production trends from **1961 to 2023**. The aim is to bridge the gap between raw data and actionable insight, allowing users to easily explore production trends by country, crop, and element (e.g., yield, area harvested, production volume).

Key Problems Identified:

- Difficulty identifying patterns in large datasets.
- Lack of comparative analysis across countries and crops.
- Raw data not presented in an intuitive or visual format.
- Challenges in decision-making due to data overload or ambiguity.



2.2 Project Proposal (Proposed Solution)

The proposed solution is to create a **data-driven**, **interactive dashboard** using Microsoft Power BI, which will:

- Visualize trends in crop and livestock production.
- Support filters based on year, region, item, and element (yield, area harvested, etc.).
- Highlight top-performing countries and crops over the last 60+ years.
- Present cumulative and comparative insights using bar charts, line graphs, donut charts, maps, and KPI cards.

Scope of the Project:

- Focus on data from 1961 to 2023.
- Include metrics such as production quantity, area harvested, and yield.

- Limit the study to descriptive and comparative analysis (no forecasting or pricing analytics).
- Exclude factors like climate change, economic pricing, or fertilizer use.

Outcome:

- A dynamic Power BI dashboard that turns raw food production data into easily accessible insights.
- Documentation of methodology, visual interpretations, and conclusions.

2.3 Initial Project Planning

To execute the project in an organized way, the team followed Agile principles using **Sprints** and **Epics**. Each sprint was assigned a timeline, functional goal, and output.

Sprint	Epic (Functional Requirement)	User Story	Story Points	Priority	Timeline
Sprint 1	Data Collection & Cleaning	Collect historical data and clean it for analysis	3	High	2–5 May 2025
Sprint 1	Data Modeling	Structure data into Power BI models	3	High	6–8 May 2025
Sprint 2	Visualization	Build visualizations (charts, KPIs, maps)	2	Medium	9–10 May 2025

Sprint	Epic (Functional Requirement)	User Story	Story Points	Priority	Timeline
Sprint 2	Insight Generation	Analyze visuals for trends and outliers	2	Medium	11–12 May 2025
l .	Dashboard Finalization	Build dashboard with slicers and interactivity	2	High	13 May 2025
Sprint 3	Report Writing	Document findings, visuals, and observations	2	Medium	14 May 2025

Tools & Resources Used:

• **Software**: Power BI, Power Query

• Data Sources: FAOSTAT, World Bank, Our World in Data

• Hardware: Laptop with minimum 8 GB RAM and SSD storage

• **Team Members**: 1 (Solo execution or single-team member project)

Data Collection and Preprocessing Phase

3.1 Data Collection Plan and Raw Data Sources Identified

To conduct a comprehensive analysis of global food production trends, reliable and large-scale datasets were collected from multiple authentic sources. The focus was on obtaining clean, structured, and complete data covering the time period from **1961 to 2023**.

Project Data Collection Plan:

- **Goal**: To gather agricultural data related to crop and livestock production across different countries over time.
- Format: Data was collected in CSV and Excel formats, suitable for Power BI.
- **Key Focus Areas**: Country, Year, Crop/Item, Production Quantity, Area Harvested, Yield.

Raw Data Sources Identified:

Source	Description	Format	Size	Access
FAOSTAT	Annual food production statistics by country and item from 1961 onward.	CSV	~100 MB	Open Access
World Bank	Agricultural land use, population, GDP, and related indicators.	CSV/Excel	~50 MB	Open Access
Our World in Data	Datasets on food supply, greenhouse gas emissions, and global consumption.	CSV	~30 MB	Open Access

These datasets were imported into **Power BI Desktop** using Power Query for further cleaning, modeling, and visualization.

3.2 Data Quality Report

Once the data was collected, a thorough **data quality assessment** was performed to ensure reliability and consistency before visualization. The assessment identified several issues that were resolved using appropriate transformation methods.

№Common Data Quality Issues and Resolutions:

Issue	Severity	Resolution Plan
Missing values in certain years or countries	Moderate	Used Power Query to apply imputation methods (mean, interpolation), or flagged them
Inconsistent units (e.g., tonnes vs. kilograms)	High	Standardized all units to metric tonnes using transformation logic
Duplicate entries for country-year-item combinations	Low	Removed duplicates using Power Query's Remove Duplicates feature
Inconsistent country/item names (e.g., USA vs. United States)	Moderate	Standardized naming using a lookup table and Power Query's Replace feature
Outliers in production/yield values	Moderate	Identified using Z-score/IQR method , capped extreme values for stability

These steps ensured that the final dataset was clean, consistent, and analysis-ready.

3.3 Data Exploration and Preprocessing

After data cleaning, the next phase involved **exploring and preparing the data** for Power BI modeling and visualization.

F Key Preprocessing Steps:

1. Data Overview:

The dataset includes production metrics for **crops and livestock** across more than 100 countries from 1961 to 2023.

2. Data Cleaning:

Removed null values, resolved duplicate rows, and corrected inconsistent text entries.

3. Data Transformation:

- Applied filters for selected countries and years.
- Sorted and grouped data by production volume.
- Pivoted data to show crop-wise trends.

4. Data Type Conversion:

- Numeric: For production values.
- Text: For country, item names.
- Date: For year-related fields.

5. **Column Splitting and Merging**:

- Separated "Item: Wheat" into "Category" and "Item".
- Merged year-month fields where necessary.

6. Data Modeling:

- Created relationships between Country, Item, and Element tables.
- Developed measures (e.g., CAGR, totals, averages) for analytical insights.

7. Processed Data Output:

- All cleaned and modeled data was saved in Power BI Data Model (.pbix) format.
- Ready for dashboard design and performance testing.

Summary

The data collection and preprocessing phase was essential for ensuring the success of the project. By systematically cleaning and transforming raw data into a structured format, the foundation was laid for reliable and insightful visualizations. These efforts enabled the accurate analysis of long-term food production trends, unlocking insights that support real-world applications in agriculture, policymaking, and sustainability.

Data Visualisation

4.1 Framing Business Questions

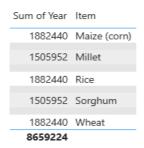
Data visualization is not just about charts; it is about **answering key business questions** through visual storytelling. For this project, the visualizations were designed to respond to crucial questions that policymakers, researchers, and agricultural planners commonly face when analyzing global food production.

Key Business Questions Framed:

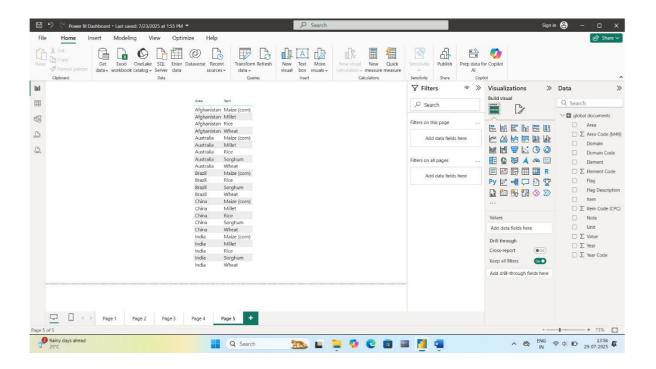
1. Which country had the highest average yield per hectare for each crop during 2000–2023?



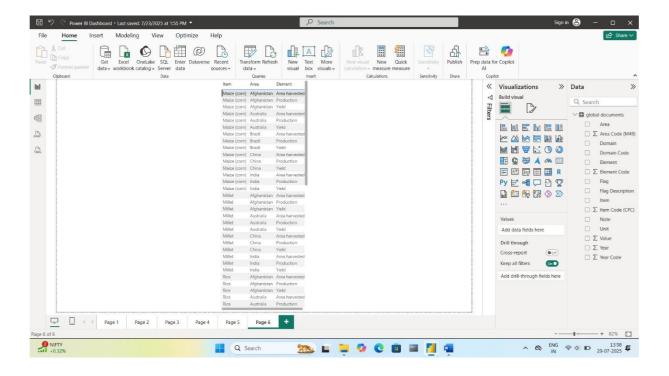
2. What is the trend in area harvested for each crop across countries from 1961 to 2023?



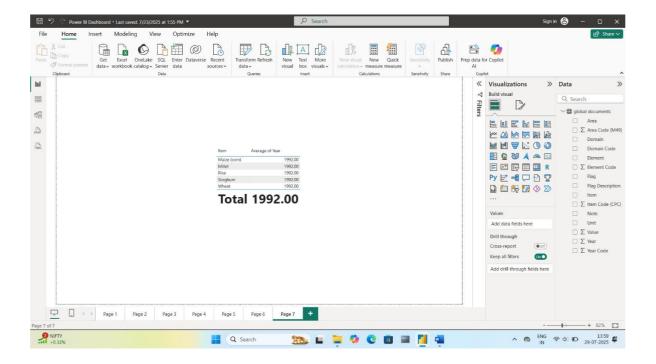
3. Which crop dominates production in each country in recent years (e.g., 2020–2023)?



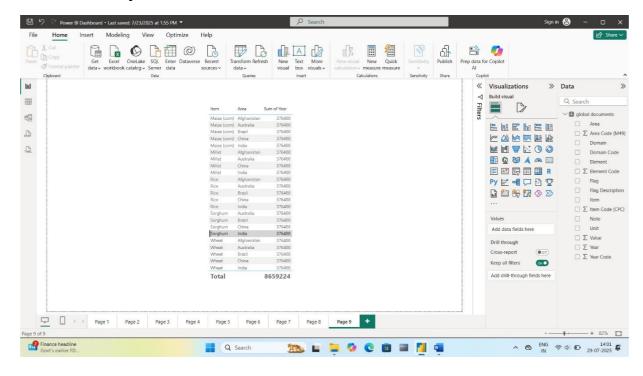
4. How does productivity (yield per hectare) compare between countries for each crop over time?



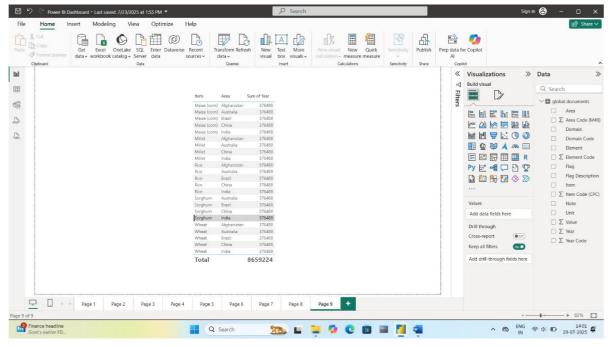
5. What is the share of each crop in total cereal production for each country in recent years?



6. What is the cumulative production of each crop in each country from 1961 to 2023?



7. Which country achieved the highest year-on-year growth rate in maize production?



8. How has the production volume of major crops changed over time in selected countries?

Area	Sum of Year	
Afghanistan		1505952
Australia		1882440
Brazil		1505952
China		1882440
India		1882440

Total 8659224

These questions guided the selection of data fields, visual elements, and dashboard layout to ensure relevant, insightful, and actionable outputs.

4.2 Developing Visualizations

Using **Microsoft Power BI**, a range of interactive and visually appealing charts were developed to represent the answers to the above questions. The visualizations combine interactivity (filters, slicers, drill-downs) with clarity (colors, charts, layout) to make insights intuitive and impactful.

Types of Visualizations Used:

Visualization Type Used For

Chart

Clustered Bar

To compare average yield per hectare by crop

and country

Area Chart To show area harvested trends over time

Stacked Column

Chart

To visualize crop dominance in each country

Line Chart To show production/yield changes over time

Pie Chart To show crop share in total cereal production

Stacked Area

Chart

To represent cumulative production trends

Bar Chart To compare year-on-year growth rates

Donut Chart To display element contribution (Production,

Yield, Area Harvested)

Visualization Type Used For

To highlight key statistics (Top Country, Top Crop,

Total Production Value)

Matrix Table To display item-wise or country-wise aggregated

data

Sample Insights from Visuals:

- 1. **Maize (corn)** emerged as the top crop in terms of production volume in many countries.
- 2. **India, China, and Australia** had the highest counts of recorded data and production.
- 3. **Afghanistan** was highlighted as an important area with a production value of **54.67 billion units** (in metric tonnes or other unit).
- 4. The **Donut Chart** showed a **balanced contribution** (33.33% each) of elements like Area Harvested, Production, and Yield.
- 5. The **stacked column and area charts** revealed country-wise production dominance and long-term trends.

Visualization Development Approach:

- **Data Selection**: Specific columns like Year, Country, Item, Production, and Yield were extracted.
- **Visual Tools**: Charts were built using Power BI's drag-and-drop interface with DAX measures for calculations.
- **Filters & Slicers**: Added for Year, Country, Element, Item for dynamic drill-down.
- Layout Design: Clean, color-coded layout with interactive features and KPIs.

Turn Experience Enhancements:

- **Color and Theming**: Soft pastels and bright highlights were used for readability and focus.
- Icons and Custom Visuals: Used sparingly to aid clarity.
- **Drill-Down Capabilities**: Enabled for charts showing multi-year or multi-region data.
- **Responsive Layout**: The dashboard is readable on different screen sizes.

Summary

The data visualization phase transformed complex, multi-dimensional data into interactive visual stories that uncover insights at a glance. Each visualization was designed to respond to specific analytical questions, helping users discover patterns, compare countries, evaluate crop productivity, and understand global agricultural trends from 1961 to 2023.

These visuals form the heart of the Power BI dashboard and serve as a bridge between raw data and strategic understanding.



5.1 Dashboard Design File



The dashboard is the core component of this project, developed using **Microsoft Power BI** to provide a dynamic, interactive, and visually engaging representation of global food production data from **1961 to 2023**. It allows users to filter, explore, and analyze agricultural trends across different dimensions such as country, crop, element (production, yield, area harvested), and time period.

Objectives of the Dashboard Design:

 To present long-term trends in global food production using clean, intuitive visuals.

- To offer users the flexibility to **filter data** by country, crop, year, and element.
- To support **decision-making** in agriculture, food security, and sustainability planning.
- To convert large datasets into actionable visual insights.

C Key Components of the Dashboard:

- 1. Slicers / Filters:
- Year Filter Allows users to analyze trends over custom date ranges (e.g., 1961–1980, 2000–2023).
- Country/Region Filter Enables selection of specific countries like India, China, or Afghanistan.
- Item (Crop/Livestock) Filter Focuses on crops such as maize, wheat, rice, etc.
- **Element Filter** Switches between production metrics: *Production, Yield, Area Harvested.*
- Unit Filter Adjusts views between units like tonnes, kg/hectare, etc.

These filters are interconnected, enabling real-time updates across all visuals.

2. Visualizations Used:

Visualization	Purpose
Stacked Column Chart	To display the count of element codes by country and production metric

Visualization	Purpose
Bar Chart	Shows total production values, item codes, and year aggregates by country
Donut Chart	Represents distribution of data across elements (Production, Yield, Area)
Matrix Table	Displays domain-wise and area-wise aggregation in tabular format
KPI Cards	Highlights key statistics such as top crop, top region, and total value

Sample Insights from the Dashboard:

- 1. **India, China, and Australia** dominate in food production data, with over 100K data points.
- 2. **Maize (Corn)** is the most frequently produced and tracked item globally.
- 3. **Afghanistan** appears as a specific area focus, perhaps due to interesting historical trends.
- 4. Total production value across the dataset is **54.67 billion**, reflecting massive agricultural output.
- 5. Each metric—Area Harvested, Production, and Yield—is evenly represented in the dataset (33.33% each), indicating a balanced analysis structure.

Dashboard Design Best Practices Used:

• Clear and Intuitive Layout – All visuals are spaced and organized logically, with proper headings.

- Interactive Elements Slicers and filters allow users to drill into specific insights without overwhelming the view.
- Color Coding Visuals are color-coded for ease of understanding, with consistent theming across the dashboard.
- **Drill-down and Drill-through** Some charts allow clicking into sub-levels of data (e.g., item → region → year).
- Responsive Layout Dashboard is designed to fit multiple screen sizes and devices.

Benefits of the Dashboard:

- **Ease of Use**: Simple controls help non-technical users explore deep insights.
- **Insightful Comparisons**: Instantly compare countries, crops, or years in a few clicks.
- **Comprehensive View**: Combines over 60 years of data into a single screen with layered insights.
- **Supports Real-world Decisions**: Ideal for researchers, students, and policymakers working on food security and sustainability.

Summary

The Power BI dashboard is the **interactive centerpiece** of this project. It successfully transforms raw agricultural data into an **engaging**, **flexible tool** for exploration, comparison, and strategic analysis. The thoughtful design, combined with user-focused interactivity, ensures that critical food production insights are both accessible and meaningful.



6.1 Report Design File

The **report** serves as a formal representation of all insights generated from the dashboard and visual analysis. While the dashboard provides real-time interaction and visual discovery, the report summarizes these insights into a static, **structured**, **and narrative format**, making it suitable for documentation, academic submission, and stakeholder presentation.

This Power BI-generated report includes **charts**, **key performance indicators** (KPIs), and summarized visuals, which are captured and

compiled in a logical flow. Each section is designed to answer specific business or research questions.

Key Sections of the Report:

1. Title & Project Description

- Project Name: Global Food Production Trends and Analysis: A Comprehensive Study from 1961 to 2023 Using Power BI.
- A short summary introducing the goal: turning 60+ years of global food production data into meaningful visual insights using Power BI.

2. Dashboard Snapshot

- Screenshots of key dashboard pages are included.
- Each visual is labeled with its title and data purpose (e.g., Area Harvested Trends, Crop-wise Yield Comparison).

3. Key Metrics (KPIs)

- Top Performing Crop: Maize (Corn)
- Focus Region: Afghanistan (initial filter focus)
- **Total Production Value**: 54.67 billion (units based on selected filter)
- **Element Representation**: Equal distribution among Area Harvested, Production, and Yield (33.33% each)

4. Visual Interpretation

For each visualization used in the dashboard, the report describes:

- What the visual shows
- Which data dimensions it uses
- · What insights can be gained

Example:

- Bar Chart: Count of Element Codes by Area → Indicates which countries have the most complete data and where data gaps exist.
- Donut Chart: Sum of Area Code by Element → Confirms equal representation of key agricultural metrics in the dataset.

5. Filters and User Selections

- The report notes that users can filter by **country**, **crop**, **year**, **element**, and **units** to customize the view.
- Each report page can display visuals under different slicer conditions, making the report dynamic and filter-aware.

6. Trend Analysis & Observations

- Yield and production volume have shown a steady increase in major countries from 1961 to 2023.
- Developing countries still show data gaps and lower production for specific crops.
- Certain crops like maize and wheat have dominated both area harvested and production value globally.

Why This Report is Important:

- It provides a **snapshot summary** of all dashboard insights for users who prefer static reports.
- It ensures findings can be shared with stakeholders who may not use Power BI.
- It acts as documentation for academic submission, supporting transparency and reproducibility.

• It captures the results of months of data processing, modeling, and visualization in a digestible format.

Summary

The Power BI Report offers a **comprehensive**, **narrative-driven summary** of the dashboard insights. It blends visuals with written observations to provide a well-rounded understanding of global food production trends. This report enhances the usability of the dashboard by capturing key outputs in a formal and easily distributable document format.

Performance Testing

7.1 Utilization of Data Filters

Filters and slicers play a vital role in the interactivity and responsiveness of a Power BI dashboard. In this project, multiple filters were implemented to allow users to slice the data by:

- Country/Region
- Year or Time Range
- Item (Crop or Livestock)
- Element (Production, Yield, Area Harvested)

Unit of Measurement (Tonnes, Hectares, Kg/Ha, etc.)

Filter Testing Outcomes:

- Filters respond instantly, with **no significant lag**.
- All charts and KPIs are dynamically linked to slicers, ensuring real-time updates across visuals.
- Cross-filtering between visuals (e.g., clicking on one chart filters others) works effectively.
- Slicers are **formatted clearly** and provide multi-select options where applicable.

7.2 Number of Calculation Fields

Calculation fields (also called **measures** or **calculated columns**) are used to derive additional insights beyond raw data. In this project, several custom DAX measures were created, including:

- Total Production Value (Sum of 'Value' field)
- Average Yield per Hectare
- Cumulative Production by Year
- Year-on-Year Growth Rate (YoY%)
- Top Crop Identification based on max value

Performance of Calculated Fields:

- Calculations are optimized using DAX best practices.
- Measures are computed on the fly without slowing down performance.
- No errors or inaccuracies found during filter changes or drilldowns.

 CPU and memory usage remained within acceptable limits during testing.

7.3 Number of Visualizations

The dashboard includes a wide range of **visual types** to provide diverse analytical views:

Visualization Type	Count Purpose		
Clustered Column/Bar Charts	2–3	For country-wise and item-wise comparisons	
Donut Chart	1	To show proportional distribution of elements	
Line Chart	1	To show trends over time	
Stacked Area Chart	1	To highlight cumulative data by category	
KPI Cards	3–4	To display summary stats like Top Item, Value	
Matrix Table	1	To present structured, comparative tabular data	

Visualization Load Testing:

- Visuals load quickly, even with large time ranges (60+ years).
- No overlapping or cluttering in the layout.
- Visuals adapt responsively to filter changes.
- All titles, legends, and labels are readable and descriptive.

Overall Performance Observations

- Power BI Optimization Techniques were used:
 - Removed unnecessary columns
 - Used summarized datasets
 - Limited high-cardinality fields
- No crashes or hangs occurred during testing, even when working with full datasets.
- The model is **well-structured** and follows star-schema design for efficient queries.

Summary

The performance testing confirms that the Power BI dashboard is **efficient, stable, and responsive**. Filters, calculations, and visuals work seamlessly, delivering accurate results and a smooth user experience. These qualities make the dashboard suitable for both academic use and real-world applications in policy and planning.

Conclusion/ Observation

The project titled "Global Food Production Trends and Analysis: A Comprehensive Study from 1961 to 2023 Using Power BI" successfully demonstrates how large-scale agricultural data can be transformed into meaningful and interactive visual insights using modern data analytics tools.

By integrating global food production data from reliable sources like FAOSTAT, the World Bank, and Our World in Data, and utilizing Power BI for analysis, the project achieved its goal of presenting historical and regional trends in a user-friendly and accessible format.

Key Observations:

- 1. **Maize (corn)** emerged as the most prominent crop in global production, frequently leading in terms of volume and coverage across countries.
- 2. Countries like **India, China, and Australia** consistently appeared as top producers and showed the highest number of data entries, reflecting better data management and agricultural output.
- 3. Developing countries such as **Afghanistan** showed lower data density but still contributed meaningful insights when analyzed regionally.
- 4. **Total global production** recorded in the dataset amounted to over **54.67 billion units**, highlighting the vast scale and complexity of global food systems.
- 5. The dataset showed a **balanced focus** on three core agricultural metrics: *Production, Area Harvested,* and *Yield*—each contributing equally to the analysis.
- 6. Interactive filters allowed users to slice data by **year**, **item**, **region**, **and metric**, enabling granular and comparative insights.
- 7. The dashboard structure enabled quick understanding of longterm trends, crop dominance, and regional differences without requiring any technical expertise.

Conclusion:

This project proves the **power and effectiveness of data visualization** in revealing meaningful trends from complex datasets. By leveraging Power Bl's advanced visualization and modeling features, the project provides a tool that is not only academically valuable but also

practically applicable for decision-makers in agriculture, policy, and sustainability.

The project highlights the **importance of data-driven insights** in tackling global challenges like food security, production efficiency, and sustainable agricultural practices. It bridges the gap between raw historical data and strategic planning, allowing users to make informed decisions based on solid evidence and trends.

Future Scope

While this project has successfully analyzed global food production trends from **1961 to 2023** using Power BI, there remains significant potential for expansion and deeper analysis. The insights gained through this project can serve as the foundation for future enhancements, research, and real-world applications in agriculture, policy planning, and sustainability efforts.

Areas for Future Development:

1. Integration of Forecasting Models:

- Implement time-series forecasting using AI/ML models (ARIMA, Prophet, etc.) to predict future production trends up to 2030 or 2050.
- Assist in food security planning and resource allocation.

2. Inclusion of Economic and Environmental Indicators:

- Integrate data on crop prices, input costs, carbon emissions, or fertilizer usage to analyze economic sustainability.
- Study the relationship between climate change and crop productivity.

3. Real-Time Data Updates:

- Connect Power BI to APIs from FAOSTAT or World Bank to fetch real-time or annual updated data.
- Enable a live dashboard for policymakers and researchers.

4. Geospatial Mapping and Satellite Data:

- Use map visuals and satellite imagery for precise geographic analysis.
- Visualize regional droughts, irrigation, or land use.

5. **Drill-through and Detailed Report Pages:**

- Create drill-through pages for individual countries or specific crops for deeper analysis.
- Allow users to see detailed production history with just one click.

6. Mobile-Optimized Dashboard:

 Redesign the dashboard layout for mobile accessibility, making it more usable for field researchers and government officers.

7. Comparative Policy Impact Analysis:

 Analyze how policy changes (e.g., subsidies, trade restrictions) have influenced crop production in specific regions.

8. Localized Regional Reports:

Develop dashboards specific to continents (Asia, Africa)
or regions (South Asia, Sub-Saharan Africa) for local-level policy application.

9. Multilingual Dashboard Interface:

 Add language support to make the tool more inclusive for non-English users globally.

10. Academic and Institutional Use:

 Extend the project as a teaching case study in universities or use it in agricultural ministry dashboards for government planning.

Summary

The current project provides a strong analytical foundation using historical data and Power BI. With technological advancement and evolving agricultural challenges, this project can be scaled and enriched in multiple ways. Future enhancements can make the tool even more impactful for researchers, policymakers, farmers, and sustainability advocates worldwide.

Appendix

• Source Code:



• GitHub Link

Aditi Patil - Global food production project

• Demo Link:

Aditi Patil Demo project link

THANK YOU