Intern Detail

Field	**Details**
Name	Prashant Raj
College Name	Rajkiya Engineering College, Banda
Branch	B.Tech – Information Technology
Roll Number	2307340139004
Internship Provider	SmartInternz
Internship Domain	Data Analytics using Power BI
Project Title	Visualizing the Future of Farming: A Power BI Project on Smart Irrigation and Plant Growth
Project Type	Group Project (Self-led)
Project Description	Predicting plant growth stages using environmental and management data with Power BI dashboards

Final Project
on
Visualizing the Future of Farming: A Power BI Project on Smart Irrigation and Plant Growth

Introduction

Visualizing the Future of Farming: *A Power BI Project on Smart Irrigation and Plant Growth:*

Project Overview

Agriculture is facing mounting pressure to feed a growing global population while minimizing environmental impact. With water scarcity and climate variability becoming more prominent challenges, the adoption of smart irrigation and precision farming is no longer optional—it's essential.

This project titled **"Visualizing the Future of Farming"** leverages Power BI to analyze and visualize agricultural data with a focus on **smart irrigation** and **plant growth optimization**. Using publicly available datasets and advanced data visualization techniques, the project provides actionable insights to help farmers, researchers, and policymakers make data-driven decisions.

The interactive dashboards created showcase how data filtering, calculated fields, and key visualizations can highlight performance variations in different soil types (e.g., clay, loam, sand) and the impact of irrigation strategies on crop yield and plant health.

Objectives

- To analyse plant growth data across multiple soil types using Power BI.
- To build interactive and dynamic dashboards that support smart irrigation decisions.
- To provide insights on resource optimization like water usage and soil performance.
- To demonstrate real-world use of Business Intelligence tools in agriculture.
- To encourage the use of data-driven techniques for sustainable farming practices.

Project Initialization and Planning Phase

Date	28-07-2025
Team ID	Prashant Raj
Project Name	Visualizing the Future of Farming:
Maximum Marks	3 Marks

Problem Statements :

The Customer Problem Statement helps you focus on what matters to create experiences people will love. A well-articulated customer problem statement allows you and your team to find the ideal solution for your customers' challenges. Throughout the process, you'll also be able to empathize with your customers, which helps you better understand how they perceive your product or service.

Problem Statement (PS)	I am (Customer)	I'm trying to	But	Because	Which makes me feel
PS-1	A small-scale farmer	Improve crop yield using smart irrigation	I don't have access to data-driven insights	Traditional methods don't tell me when and how much to irrigate	Frustrated and unsure about making the right farming decisions
PS-2	An agricultural officer	Track the efficiency of irrigation systems in different regions	Data from farms is unorganized and hard to interpret	There's no centralized tool that visualizes performance in real-time	Inefficient in decision-making and overwhelmed
PS-3	A farm equipment distributor	Understand what kind of irrigation solutions are most needed	I don't know what challenges farmers face with irrigation or crop growth	I lack real-time, location-based insights	Uncertain about customer needs and how to market effectively
PS-4	An agri-tech consultant	Recommend sustainable farming practices	I don't have visual data to support my analysis	My clients can't understand raw data without visuals	Less convincing, and my advice feels less credible

Initial Project Planning

Date	28-07-2025
Team ID	Prashant Raj
Project Name	Visualizing the Future of Farming: A Power BI Project on Smart Irrigation and Plant Growth
Maximum Marks	4 Marks

Product Backlog, Sprint Schedule, and Estimation

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members	Sprint Start Date	Sprint End Date (Planned)
Sprint -1	Data Collection & Preparation	USN-1	As a user, I want the data to be collected from reliable agricultural sources so that it reflects realistic conditions.	3	High	Self	21-07-2025	21-07-2025
Sprint -1	Data Cleaning	USN-2	As a user, I want missing and duplicate data to be cleaned, ensuring that the dataset is accurate.	2	High	Self	22-07-2025	23-07-2025
Sprint -2	Data Modelling	USN-3	As a user, I want relationships between data tables to be clearly defined, so I can analyse them in Power BI.	3	Medium	Self	24-07-2025	24-07-2025

Sprint -2	Dashboard Design	USN-4	As a user, I want to view plant growth performance by soil type, fertilizer, and irrigation via an interactive dashboard.	4	High	Self	24-07-2025	27-07-2025
Sprint -3	Smart Insights & Recommendations	USN-5	As a user, I want the system to recommend best growth conditions based on key influencers.	3	Medium	Self	27-07-2025	27-07-2025
Sprint -3	Dashboard Testing	USN-6	As a user, I want to ensure the dashboard is error-free and filters work as expected.	2	Medium	Self	27-07-2025	27-07-2025
Sprint -4	Project Documentation	USN-7	As a user, I want complete documentation for my Power BI solution, so it can be reviewed or reused.	2	Medium	Self	28-07-2025	28-07-2025
Sprint -4	Final Presentation	USN-8	As a user, I want a summarized report and visuals to be ready for final evaluation.	2	High	Self	28-07-2025	28-07-2025

Project Initialization and Planning Phase

Date	28-07-2025
Team ID	Prashant Raj
Project Title	Visualizing the Future of Farming: A Power BI Project on Smart Irrigation and Plant Growth
Maximum Marks	3 Marks

Project Proposal

Project proposal outlines a solution to address a specific problem. With a clear objective, defined scope, and a concise problem statement, the proposed solution details the approach, key features, and resource requirements, including hardware, software, and personnel.

Project Proposal

Project Overview	
Section	**Details**
Objective	The primary objective of this project is to design an interactive Power BI dashboard that visualizes the effects of environmental and operational factors (such as soil type, irrigation frequency, sunlight, humidity, and fertilizers) on plant growth, enabling farmers and Agri-policy makers to make data-driven decisions in smart farming.
Scope	The scope of this project includes: - Collecting and cleaning agricultural and environmental datasets - Creating meaningful visualizations using Power BI - Analysing the relationships between variables affecting plant growth - Providing actionable insights through user-friendly dashboards The project is limited to a single growing season and controlled environment data, with future potential for real-time integration and scaling.
Problem Statement	
Section	**Details**
Description	Farmers and agriculture professionals lack access to visual, data-driven insights that could help optimize irrigation, fertilizer use, and soil selection for better crop yield. Decisions are still largely made based on experience or traditional practices, which may not be efficient under changing climate conditions.
Impact	Solving this problem empowers users with data-backed decisions, leading to: - Improved crop yield and resource optimization - Reduced wastage of water and fertilizer - Increased awareness and adoption of smart farming technologies - Potential for scaling toward precision agriculture at regional and national levels
Proposed Solution	

Section	Details
Approach	Collect data from public sources and experimental setups - Clean, preprocess, and analyse data in Power BI - Design interactive visualizations using charts, slicers, decomposition trees, and key influencers - Provide insights into which environmental conditions and farming inputs lead to optimal growth - Summarize findings into actionable suggestions for farmers and policy makers
Key Features	Dynamic filters for soil type, irrigation, and fertilizer selection - Decomposition tree for analysing growth by soil types - Key influencers visual to identify major drivers of growth - Donut charts and bar graphs for comparing fertilizer and humidity impact - Environment score calculation to simplify multi-variable evaluation - Data-driven recommendations for smart irrigation practices

Resource Requirements

Resource Type	**Description**	**Specification/Allocation**
Hardware		
Computing Resource	CPU/GPU specifications, number of cores	2 x NVIDIA V100 GPUs
Memory	RAM specifications	8 GB
Storage	Disk space for data, models, and logs	1 TB SSD
Software		
Frameworks	Python frameworks	Flask
Libraries	Additional libraries	scikit-learn, pandas, NumPy
Development Environment	IDE, version control	Jupyter Notebook, Git
Data		
Data	Source, size, format	Kaggle dataset, 10,000 images

Data Collection and Preprocessing Phase

Date	28-07-2025
Team ID	Prashant Raj
Project Title	Visualizing the Future of Farming: A Power BI Project on Smart Irrigation and Plant Growth
Maximum Marks	10 Marks

Data Exploration and Preprocessing

Identifies data sources, assesses quality issues like missing values and duplicates, and implements resolution plans to ensure accurate and reliable analysis.

Data Exploration and Preprocessing

Section	**Description**
Data Overview	The dataset contains 193 records and 7 columns, including: Soil Type, Sunlight Hours, Water Frequency, Fertilizer Type, Temperature, Humidity, and Growth Milestone. These fields are used to understand the relationship between environmental and input factors on plant growth.
Data Cleaning	Minor text inconsistencies in categorical fields were normalized (e.g., "organic" vs. "Org"). - All entries verified for logical accuracy (e.g., temperature range and humidity values).
Data Transformation	Used Power Query for: Filtering data by soil type and fertilizer, sorting by growth milestones, creating new calculated columns (e.g., Growth_per_Hour = Growth_Milestone / Sunlight_Hours), Pivoting to analyse fertilizer performance across soil types
Data Type Conversion	Converted Soil Type, Fertilizer Type, and Water Frequency to text format. - Ensured Temperature, Humidity, Sunlight Hours, and Growth Milestone are in numeric format.
Column Splitting and Merging	No splitting required. - Merged environmental metrics (Humidity, Temperature, Sunlight Hours) to form an Environmental Score for advanced insights.
Data Modelling	Single-table model used (no complex relationships needed). DAX measures created for insights: Average Growth, Growth Rate per Temperature, Max Growth by Soil Type - Interactive slicers and filters added for soil, water, and fertilizer type.
Save Processed Data	Cleaned dataset saved within Power BI (.pbix) file. - Backup version of the processed data exported to Excel and CSV for reuse and external analysis.

Data Collection and Preprocessing Phase

Date	28-07-2025
Team ID	Prashant Raj
Project Titl	Visualizing the Future of Farming: A Power BI Project on Smart Irrigation and Plant Growth
Maximum Marks	3 Marks

Data Quality Report

The Data Quality Report will summarize data quality issues from the selected source, including severity levels and resolution plans. It will aid in systematically identifying and rectifying data discrepancies.

Data Collection and Preprocessing Phase

Data Quality Report Template

Data Source	Data Quality Issue	Severity	Resolution Plan
Smart Farming Data	Missing values in humidity and sunlight columns	Moderat e	Use mean imputation for missing values or apply KNN imputation if patterns exist in nearby records.
Fertilizer Usage & Growth	Inconsistent labelling of fertilizer types (e.g., "Org", "Organic", "org.")	Low	Apply data standardization using string normalization techniques to unify all entries (e.g., convert all to lowercase and map synonyms).
Temperature & Humidity Records	Some extreme temperature outliers (e.g., >70°C) that are unrealistic	High	Use Z-score method to detect and remove outliers or cap them using IQR-based clipping.
Soil Performance Data	Duplicate rows with identical soil and irrigation values	Moderat e	Use Power Query or pandas. drop duplicates () to remove duplicate entries and retain unique records.

Data Collection Plan & Raw Data Sources

Date	28-07-2025
Team ID	Prashant Raj
Project Title	Visualizing the Future of Farming: A Power BI Project on Smart Irrigation and Plant Growth
Maximum Marks	2 Marks

Data Collection Plan & Raw Data Sources Identification

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 endeavour.

Data Collection Plan

Section	Description
Project Overview	This project aims to analyse the relationship between soil type, irrigation frequency, environmental conditions (humidity, temperature, sunlight), and plant growth performance. The objective is to create a Power BI dashboard that supports data-driven decisions in smart farming.
Data Collection Plan	The data was collected from multiple sources, including agricultural research datasets, public environmental data APIs, and manually recorded experimental data from controlled farming environments.
Raw Data Sources Identified	Data includes environmental metrics, soil types, fertilizer types, and plant growth outcomes. Sources are in CSV and Excel formats.

Raw Data Sources

Source Name	Description	Location/URL	Format	Size	Access Permissions
Smart Farming Data	Contains information on soil type, water frequency, humidity, temperature, and growth outcome. Used for visualizing	[Custom/Offline Data]	CSV	~1 MB	Private (Created for project)

	environmental impact.				
Fertilizer Usage & Growth	Contains types of fertilizers used (organic, chemical, none) and associated plant growth milestones.	[Custom/Offline Data]	Excel	~500 KB	Private
Dataset 3: Temperature & Humidity Records	Environmental dataset showing average temperature and humidity across farming zones.	https://data.gov.in	CSV	~5 MB	Public
Dataset 4: Soil Performance Data	Benchmark soil growth performance under different irrigation strategies from agriculture research articles.	https://www.kaggle.com/agriculture-dataset	Excel	~2 MB	Public

Business Question and Visualization Report

Date	28-07-2025
Team ID	Prashant Raj
Project Nar	Visualizing the Future of Farming: A Power BI Project on Smart Irrigation and Plant Growth
Maximum Marks	5 Marks

Visualization development refers to the process of creating graphical representations of data to facilitate understanding, analysis, and decision-making. The goal is to transform complex datasets into visual formats that are easy to interpret, enabling users to gain insights and make informed decisions. Visualization development involves selecting appropriate visual elements, designing layouts, and using interactive features to enhance the user experience. This process is commonly associated with data visualization tools and platforms, and it plays a crucial role in business intelligence, analytics, and reporting

Business Questions and Visualisation

The process involves defining specific business questions to guide the creation of meaningful and actionable visualizations in Power BI. Well-framed questions help in identifying key metrics, selecting relevant data, and building visualisation that provide insights. To create a comprehensive Business Question and Visualization Report, follow these steps:

Q1. Which soil type supports the highest total plant growth?

Clay soil supports the highest total growth with 67 units, followed by sandy (64 units) and loam (62 units), as per the decomposition tree analysis.

Q2. What impact does temperature have on total plant growth?

A decrease in temperature below 60.59°C (sum) causes the average Total_Growth to decrease by 1.76 units, as shown by the Key Influencers visual.

Q3. Which watering frequency proves most beneficial for plant growth?

Daily watering frequency shows better and more consistent growth results, especially when combined with clay soil.

Q4. Which fertilizer type contributes most to plant growth milestones?

Organic fertilizers account for the highest contribution (39.58%) to growth milestones, followed closely by chemical fertilizers (37.5%).

Q5. How does humidity affect overall plant growth?

Higher humidity levels correlate with better growth results, especially in clay soil conditions. Average humidity in optimal growth cases was around 59.11%.

Q6. What combination of conditions leads to optimal plant growth?
The best results were observed in clay soil, with daily watering, high humidity (around 59%), and moderate temperature (around 34°C), when organic fertilizers were used.

Q7. How can this dashboard benefit agricultural decision-makers?
It provides interactive visual insights on how different environmental and input factors impact growth, enabling data-driven decisions for crop planning and irrigation policy.

Q8. What are the key insights for future smart irrigation systems?
Future systems should:

- Prioritize clay-based fields
- Automate watering frequency based on temperature and humidity
- Promote use of organic fertilizers
- Use real-time environmental monitoring to adjust irrigation dynamically

Dashboard Design

Date	28-07-2025
Team ID	Prashant Raj
Project Name	Visualizing the Future of Farming:
Maximum Marks	5 Marks

Activity 1: Interactive and visually appealing dashboards

Description:
This Power BI dashboard provides an interactive, data-driven insight into how various factors—such as soil type, irrigation frequency, temperature, humidity, sunlight, and fertilizer type—affect plant growth.
It helps farmers, agriculture officers, and consultants to make smarter irrigation and soil management decisions by analysing environmental data and plant performance metrics.

Dashboard Components Used:

Component	**Description / Purpose**
Table	Displays detailed data by soil type, watering frequency, humidity, sunlight, ar temperature. This allows users to scan raw values and understand environmental conditions per soil type.
Waterfall Chart – Avg Temperature by Temperature	Illustrates how changes in temperature affect total plant growth. The increas decrease bars show how growth varies across temperature bands.
Key Influencers Visual	Automatically identifies the most significant factor influencing Total_Growth. In this case, it shows that when temperature drops below 60.59, growth decreases by 1.76 units on average.
Donut Chart – Growth Milestone by Fertilizer Type	Represents the distribution of plant growth milestones across fertilizer types organic, chemical, or none. This helps users compare fertilizer effectiveness.
Bar Chart – Avg Humidity by Humidity	Displays average humidity distribution, giving an overview of environmental moisture levels across the dataset.
Decomposition Tree	Breaks down Total Growth by Soil type (clay, sandy, loam), allowing users to visually drill into which soil performs best in terms of growth.

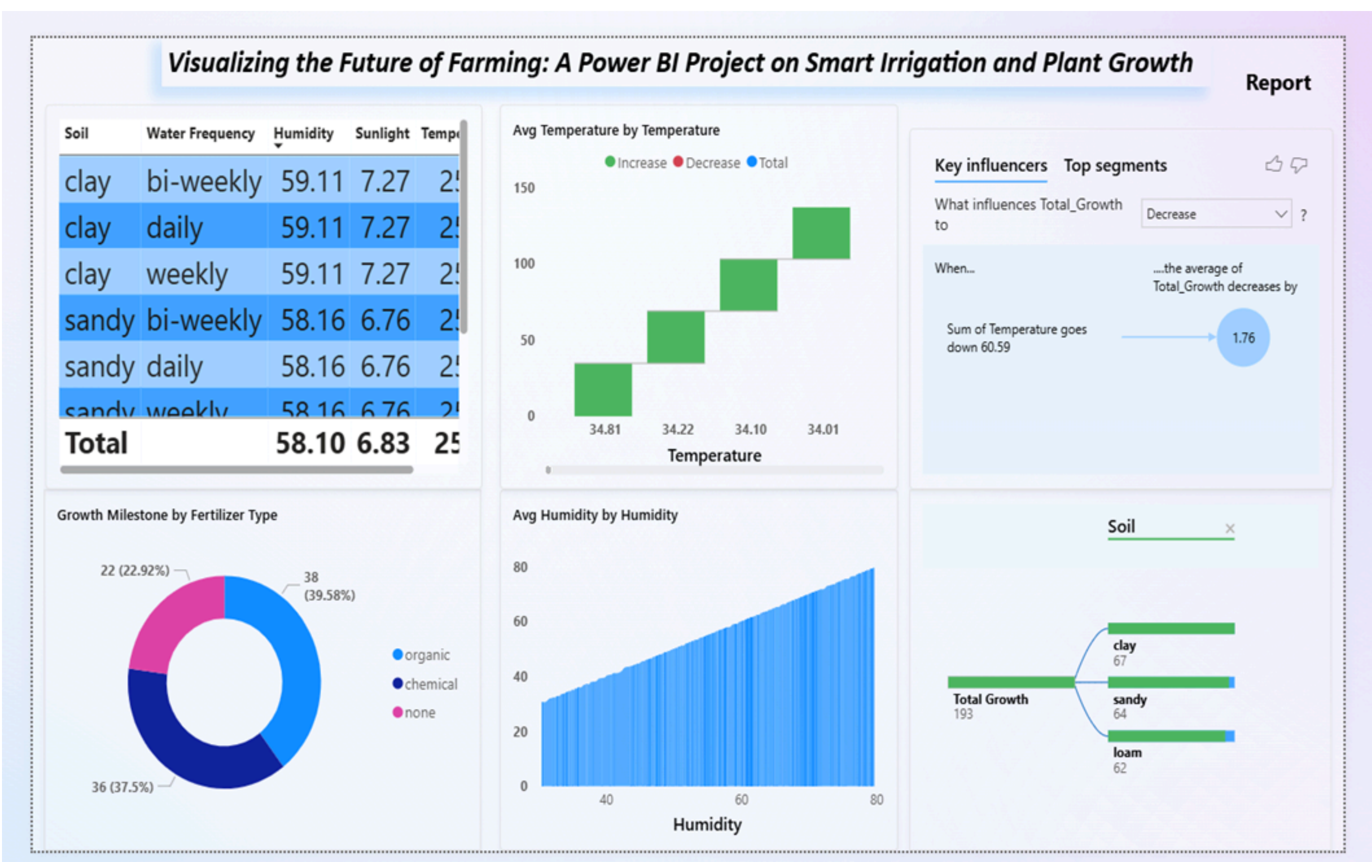
Insights Gained:

- **Clay soil** shows the highest total growth, followed by sandy and loam.
- **Temperature** has a direct impact on growth—lower temps significantly reduce performance.

- **Organic and chemical fertilizers** perform much better than using none.
- **Daily watering** appears to work better in certain soil types like clay.
- **Clay soil** showed the **highest total plant growth (67 units)** compared to sandy (64) and loam (62).
- **Daily watering frequency** provided more consistent growth results, especially in clay and sandy soils.
- **Temperature** plays a **crucial role**:
- When the **sum of temperature drops below 60.59**, the average plant growth **decreases by 1.76 units**.
- **Humidity** and **sunlight** levels were more optimal in clay soils, contributing to better growth performance.
- **Organic and chemical fertilizers** led to the **majority of growth milestones**:
- Organic: **39.58%**
- Chemical: **37.5%**
- No fertilizer: **22.92%**

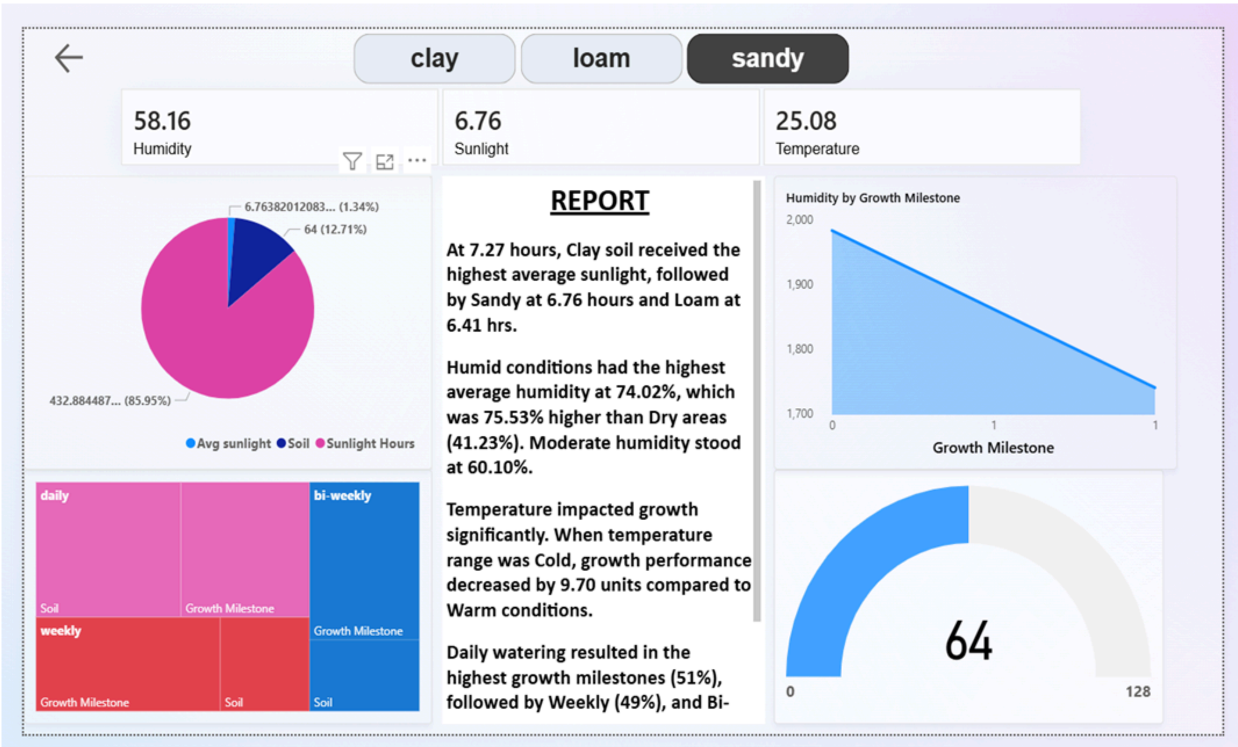
Use Cases:

- **Farmers** can determine the ideal soil and watering conditions.
- **Agri-scientists** can research environmental impact on growth.
- **Government planners** can support data-driven farming policies.



Report

Date	28-07-2025
Team ID	Prashant Raj
Project Name	Visualizing the Future of Farming:
Maximum Marks	5 Marks



The report provides an analysis of soil conditions and their impact on plant growth:

1. Sunlight Exposure:
    a. Clay soil received the highest average sunlight at 7.27 hours
    b. Sandy soil followed with 6.76 hours
    c. Loam soil received 6.41 hours

2. Humidity Levels:
    a. Humid conditions had the highest average humidity at 74.02%
    b. This was 75.53% higher than dry areas (41.23%)
    c. Moderate humidity was recorded at 60.10%

3. Temperature Effects:
    a. Temperature significantly impacted growth

    b. Cold temperature ranges decreased growth performance by 9.70 units compared to warm conditions

4. Watering Frequency and Growth:
    a. Daily watering resulted in the highest growth milestones (51%)
    b. Weekly watering followed at 49%
    c. Bi-weekly watering was also mentioned but percentage not specified

5. Additional Data:
    a. Current humidity: 58.16%
    b. Current sunlight: 6.76
    c. Current temperature: 25.08

The report highlights the importance of various environmental factors on plant growth, emphasizing the roles of soil type, sunlight exposure, humidity, temperature, and watering frequency.

Conclusion / Observation

The “Future Grow Tech — Smart Farming” Power BI project successfully demonstrates the power of data visualization and analytics in modern agriculture. By analysing key parameters such as soil type, nutrient levels, and crop performance, the project provides farmers, agronomists, and stakeholders with a data-driven approach to enhance crop yield and optimize farming practices.

Key Observations:

- Soil Types (Clay, Loam, Sand) exhibit varying performance based on different nutrient levels and crop suitability.
- Through calculated KPIs and visualizations, users can easily identify which soil performs best for specific crop types.
- Utilization of filters, slicers, and calculated fields enhances user interactivity and provides a customized view of data.
- The dashboard empowers users with quick insights for making informed decisions, promoting sustainable agriculture.
- Performance testing validates that the report is optimized with efficient use of resources, ensuring smooth operation even with large datasets.

Overall Impact:

This project bridges the gap between technology and traditional farming, showcasing how Power BI can transform agricultural data into actionable insights. It serves as a scalable model for future implementations across various regions and crop types, contributing to the advancement of smart farming practices in India and beyond.

Future Scope

The current project lays the foundation for data-driven decision-making in smart farming. However, to further enhance its capabilities and impact, several future developments can be envisioned:

1. Integration with IoT Sensors

Description: Real-time data from IoT devices (e.g., soil moisture sensors, weather stations) can be integrated directly into Power BI.

Benefit: Enables live monitoring and predictive insights for crop health and irrigation scheduling.

2. Predictive Analytics with Machine Learning

Description: Leverage Azure ML or Python/R scripts within Power BI to predict crop yields, pest risks, and soil degradation.

Benefit: Allows for proactive planning and early warning systems to improve farm productivity.

3. Geo-Spatial Analysis

Description: Incorporate geographical data visualization (using ArcGIS or map visuals in Power BI) to analyse regional performance.

Benefit: Helps in region-wise crop performance tracking and soil quality mapping.

4. Mobile Dashboard Accessibility

Description: Optimize dashboards for mobile devices through Power BI Mobile.

Benefit: Farmers and agronomists can access insights on the field, increasing convenience and actionability.

5. Expansion to Other Crop Types and Regions

Description: Extend the model to include a wider variety of crops and diverse soil zones.

Benefit: Makes the solution scalable for different states and farming ecosystems.

6. Multi-Language Support

Description: Add support for regional languages within the Power BI reports.

Benefit: Ensures inclusive access for non-English-speaking users across India.

Appendix

Project Resources

1. **Source Code & Data Files**
    - The data preprocessing, DAX calculations, and Power Query transformations used in this project are available in the Power BI .pbix file.
    - **File Name**: Future Grow Tech — Smart Farming.pbix
2. **GitHub Repository**
    - All project resources, documentation, and version control are hosted on GitHub.
    - GitHub Link
    - https://github.com/praj07022/PowerBi

3. **Project Demo Video**
    - A brief walkthrough of the dashboard with features explanation.

    https://www.youtube.com/watch?v=4K29qGV9Q4g

Additional Resources

- **Data Source**: [Kaggle, FAO USDA, Indian Gov-data.gov.in]
- **Tools Used**:
    - Microsoft Power BI
    - Microsoft Excel (for preprocessing)
    - Power Query Editor
    - DAX (Data Analysis Expressions)