

Project Report

Date	14 November 2022
Team ID	PNT2022TMID37430
Team Members	A. Suhail Basha Mohamed Faiz Ali. M Salman Farish A. J V. Vanchinathan
Project Name	Project- Estimate the crop yield using Data Analytics

1. INTRODUCTION

1.1 Project Overview

Our project focuses on estimating the crop yield in India using data analytics. The production of crops across the states of India is analyzed using complex IBM Cognos that is powered by an artificial intelligence that helps with analyzing and visualizing the data and makes the work easy for the user. We aim to create a web application to make the interaction easier for the user. Data Visualization such as Dashboard, Story, Exploratory Data Analysis and Report is embedded into the web application to access the information and insights about the given dataset.

1.2 Purpose

The purpose of this project is to work under the data analytics domain using the available technologies to achieve the goal which is to estimate the crop yield in India. We believe that it helps the client and people who are interested in agricultural studies to gain more insights. We work with the dataset that has data regarding the crop, production and area etc. and we use data analytic techniques to run various models and analysis on our dataset to gain more information which can help us understand the data.

2. LITERATURE SURVEY

2.1 Existing problem

- 1) Agriculture Data Analytics in Crop Yield Estimation
- 2) Prediction of Soil and Crop Yield by Big Data Analysis
- 3) Effective Data Analysis And Crop Yield Prediction Using Machine Learning Algorithm
- 4) Development of Crop Yield Estimation Model using Soil and Environmental Parameters
- 5) Comparison of Methods for Estimating Crop Yield at the County Level

2.2 References

[1]https://www.researchgate.net/publication/329467349_Agriculture_Data_Analytics_in_Crop_Yield_Estimation_A_Critical_Review

- [2] Prediction of Soil and Crop Yield by Big Data Analysis Venkata Chennareddy, S. Ramanayagam
- [3] EFFECTIVE DATA ANALYSIS AND CROP YIELD PREDICTION USING MACHINE LEARNING ALGORITHM, Vinayaka Manjunatha Purushan.
- [4] Development of Crop Yield Estimation Model using Soil and Environmental Parameters, Nisar Ahmed, H. M. Shahzad Asif, Gulshan Saleem, M. Usman Younus
- [5] Comparison of Methods for Estimating Crop Yield at the County Level, Michael E. Bellow (Research and Development Division Washington DC 20250).

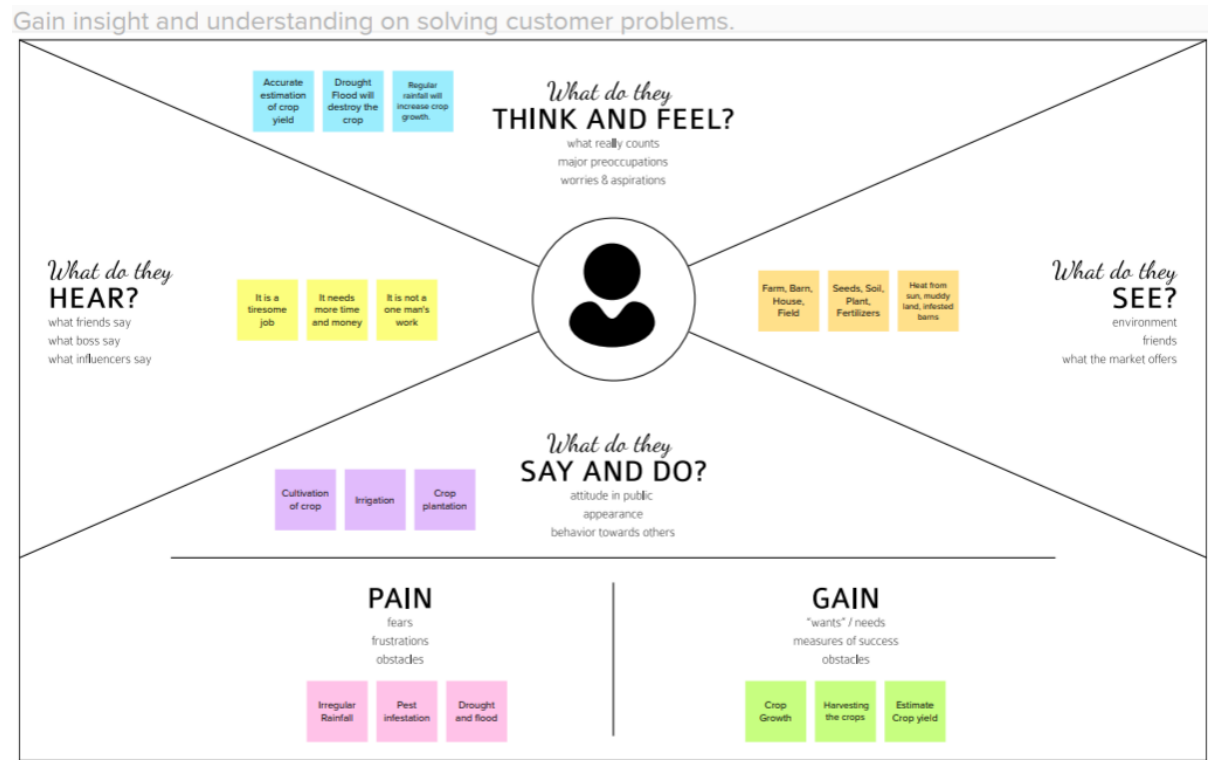
2.3 Problem Statement Definition

- 1)** Agriculture is important for human survival because it serves the basic need. It has become a challenging task to achieve desired targets in Agri based crop yield. Various factors are to be considered which have direct impact on the production, productivity of the crops. Crop yield prediction is one of the important factors in agriculture practices. Farmers need information regarding crop yield before sowing seeds in their fields to achieve enhanced crop yield.
- 2)** Optimization of agricultural practices for enhanced crop yield is considered to be an essential phenomena for the countries like India. In order to strengthen the economy and also to meet the food demand for the exponentially growing population, optimizing the agricultural practices has become a necessity. In India, weather and geographical conditions are highly variable and were thought to be the major bottleneck of agricultural practices to achieve improved crop yield.
- 3)** The agriculture sector in India is facing a tough problem to increase crop productivity. More than 60 per cent of the crop still depends on rainfall. Crop yield depends on many factors including soil, climate, rainfall, fertilizers and pesticides. Many factors have different effects on agriculture. Recent developments in information technology for agriculture have become an interesting research area for assessing crop yields. The yield estimation problem is a major problem that needs to be addressed based on available data.
- 4)** The main challenge using big data in agriculture is identification of impact and effectiveness of big data analytics. Efforts are going on to understand how big data analytics can be used to improve the productivity in agricultural practices. The analysis of data related to agriculture helps in crop yield prediction, crop health monitoring and other such related activities. In literature, there exist several studies related to the use of data analytics in the agriculture domain.
- 5)** The use of technology in agriculture has increased in recent years and data analytics is one

such trend that has penetrated into the agriculture field being used for management of crop yield and monitoring crop health. The recent trends in the domain of agriculture have made the people to understand the significance of data. The present study gives insights on various data analytics methods applied to crop yield prediction.

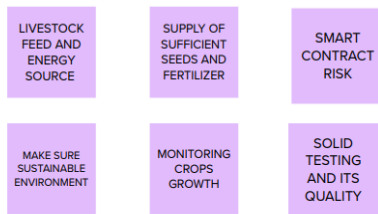
3. IDEATION AND PROPOSED SOLUTION

3.1 Empathy Map Canvas



3.2 Ideation and Brainstorming

SUHAIL BASHA



MOHAMMED FAIZ ALI



SALMAN FARISH



VANCHINATHAN



3.3 Proposed Solution

S. No.	Parameter	Description
1.	Problem Statement (Problem to be solved)	To aid farmers in making better decisions in order to have healthy crop production using data analytics.
2.	Idea / Solution description	To deliver a perfectly analysed Dashboard of historical agricultural production data from several Indian states so that farmers may forecast their crop yield.
3.	Novelty / Uniqueness	The dataset contains information about the crops in various Districts, States, Seasons, and Areas. Therefore, using all these facts, a thoroughly researched report will assist farmers in making the best crop choice for their region during a specific growing season to increase output
4.	Social Impact / Customer Satisfaction	The issues that farmers have with yield potential will all be resolved by this report. Therefore, this Dashboard will have a significant impact on farmers, and by adopting the advised crops, they can achieve enormous earnings.
5.	Business Model (Revenue Model)	Profit can be generated by marketing the solution as a freely accessible mobile application that anyone can use. Venture partnerships with the government may yield financial rewards.
6.	Scalability of the Solution	Regarding dataset storage and data gathering, there are no problems. As a result, the system may be readily scaled to manage rising user numbers, traffic, and requirements that must be met.

3.4 Problem Solution Fit

Define CS, fit into CC	1. CUSTOMER SEGMENT(S) CS <ul style="list-style-type: none"> Our target customer here is the farmer who mainly works with the crops in the field. 	6. CUSTOMER CONSTRAINTS CC <ul style="list-style-type: none"> Farmers don't make accurate predictions about the climate. They also stick to traditional methods of farming, therefore lacking modern irrigation, fertilization facilities. 	5. AVAILABLE SOLUTIONS AS <ul style="list-style-type: none"> They may have the dataset but a proper data report is not available to them. 	Explore AS, differentiate

Define CS, fit into CC	2. JOBS-TO-BE-DONE / PROBLEMS J&P <ul style="list-style-type: none"> Farmers need to have access to detailed reports and predictions to grow the right kind of crops and to take proper precautions. 	9. PROBLEM ROOT CAUSE RC <ul style="list-style-type: none"> Inaccurate predictions could lead to usage of wrong seeds, improper irrigation, and unpreparedness for drastic climate changes. 	7. BEHAVIOUR BE <ul style="list-style-type: none"> Farmers properly study and analyze their soil and decide what can be grown there. They also test crops with different weather conditions. 	Explore AS, differentiate

Identify strong TR & EM	3. TRIGGERS TR <ul style="list-style-type: none"> Destruction of crops because of climate change and growing competition in the market. 	10. YOUR SOLUTION SL <ul style="list-style-type: none"> It would help farmers a lot if crop yield predictions were made more accurately and the data is visualized and displayed on a dashboard for easier understanding. 	8. CHANNELS of BEHAVIOUR CH <ul style="list-style-type: none"> It may not be possible online as not every farmer has access to technology and the internet, but they can benefit from it offline from an agricultural office. 	Identify strong TR & EM
	4. EMOTIONS: BEFORE / AFTER EM <ul style="list-style-type: none"> Many farmers have faced huge losses in crop yield, which took months of hard work, leading them to commit suicide. When they are certain with the predictions and analysis, they are confident about making better decisions without much loss. 			

4. REQUIREMENT ANALYSIS

4.1 Functional Requirement

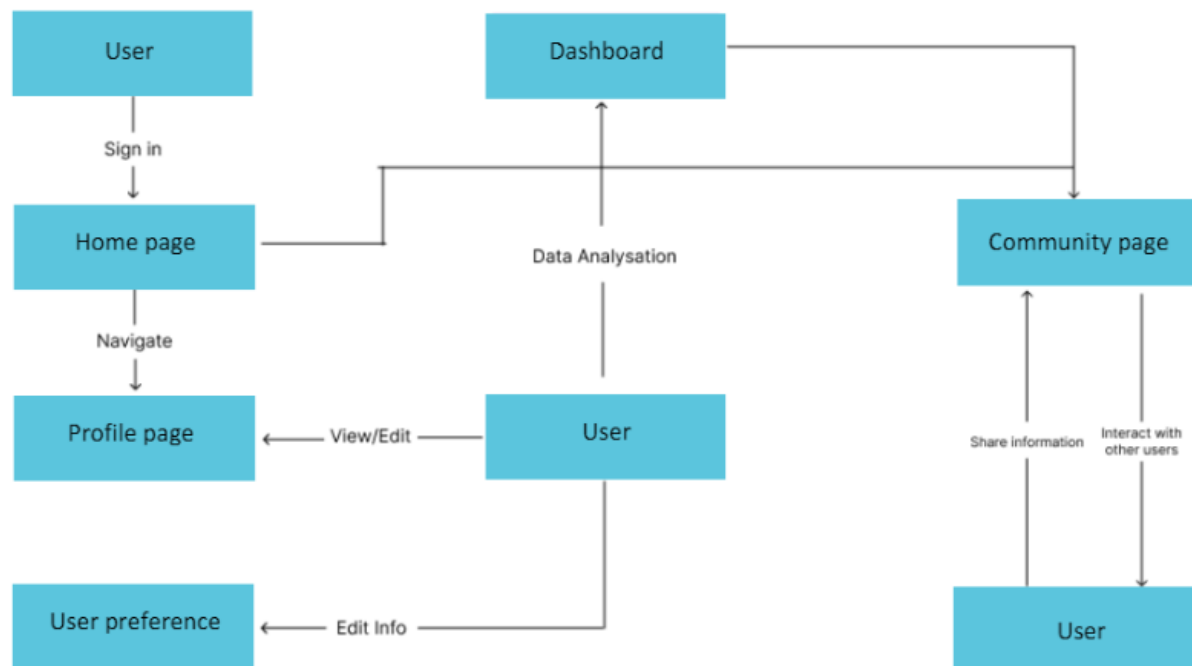
FR No.	Functional Requirement (Epic)	Sub Requirement (Story / Sub-Task)
FR-1	User Registration	Registration through Form Registration through Gmail Registration through LinkedIn
FR-2	User Confirmation	Confirmation via Email Confirmation via OTP
FR-3	Collects Data	Providing CSV file Authentic Datasets
FR-4	Cleans the given Data	Prepares data for EDA purpose
FR-5	Visualisation of Data	Identifying trends in given data Accurate visualisation of provided numbers
FR-6	Dashboard	Analysation of the datasets Key performance indicator

4.2 Non-Functional Requirement

FR No.	Non-Functional Requirement	Description
NFR-1	Usability	User friendly design with interactive UI/UX
NFR-2	Security	Users could create accounts that require authentication by sending OTP to their Email address. They could share their datasets in the app's community or make their account private.
NFR-3	Reliability	App could be run offline while server maintenance takes place. Server traffic wouldn't be an issue.
NFR-4	Performance	Requires minimum system requirements, hence could be accessible in many devices with faster loading time.
NFR-5	Availability	Server is online 24/7 hence users could use the app at any time. App will work offline as well.
NFR-6	Scalability	Scalability reflects the ability of the software to grow or change with the user's demands.

5. PROJECT DESIGN

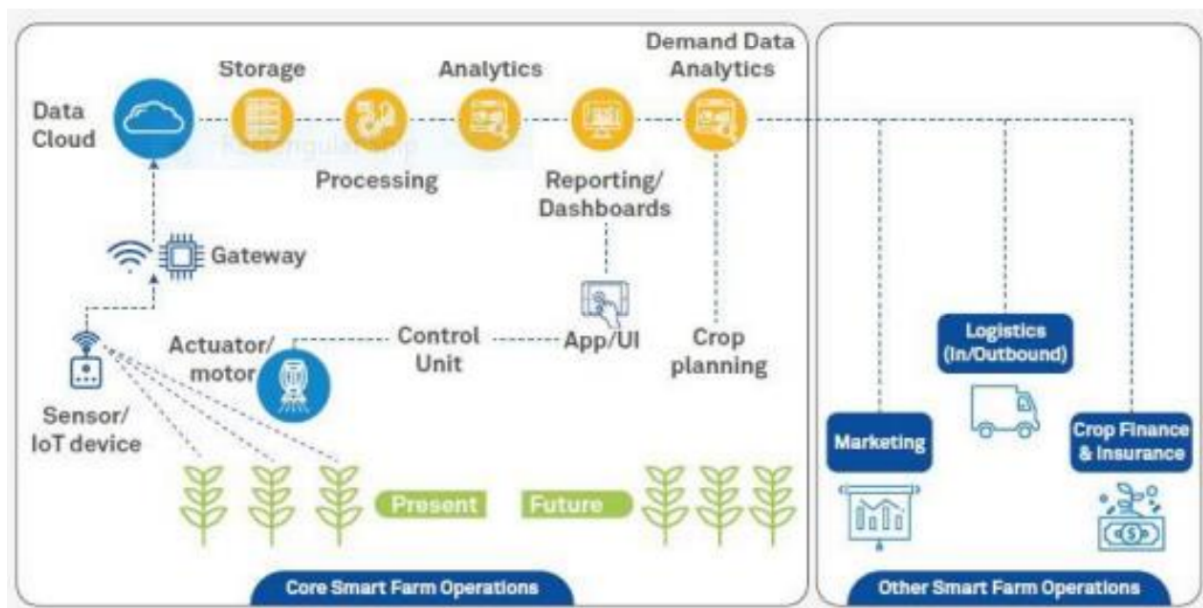
5.1 Data Flow Diagram



“

5.2 Solution & Technical Architecture

- This system is developed to help the farmers of Tamil Nadu region to improving their farm productivity.
- The agro user can get information which can be location based, check information about crop diseases, information about crop nutrient, information about crop storage, information about crop market price and get recommendations regarding crop production.
- It is used for the compare the models for the same datasets and find the best fitting model. The model with smallest AIC value is preferred.
- The farmers/users will be able to get current weather information of their locations through the system.
- Apart from that get past weather data to figure out patterns and weather conditions.
- Basically, the whole proposed system is divided in four modules. These modules and their functioning along with standards and specifications



5.3 User Stories

User Type	Functional Requirement (Epic)	User Story Number	User Story / Task	Acceptance criteria	Priority	Release
Customer (Mobile user)	Registration	USN-1	As a user, I can register for the application by entering my email, password, and confirming my password.	I can access my account / dashboard	High	Sprint-1
		USN-2	As a user, I will receive confirmation email once I have registered for the application	I can receive confirmation email & click confirm	High	Sprint-1
		USN-3	As a user, I can register for the application through Facebook	I can register & access the dashboard with Facebook Login	Low	Sprint-2
		USN-4	As a user, I can register for the application through Gmail		Medium	Sprint-1
	Login	USN-5	As a user, I can log into the application by entering email & password		High	Sprint-1
	Dashboard		As a user, I can enter data to conduct to conduct business analysis to make business decisions		High	Sprint-1
Customer (Web user)	Registration	USN-1	As a user, I can register for the application by entering my email, password, and confirming my password.	I can access my account / dashboard	High	Sprint-1
		USN-2	As a user, I will receive confirmation email once I have registered for the application	I can receive confirmation email & click confirm	High	Sprint-1
		USN-3	As a user, I can register for the application through Facebook	I can register & access the dashboard with Facebook Login	Low	Sprint-2
		USN-4	As a user, I can register for the application through Gmail		Medium	Sprint-1
	Login	USN-5	As a user, I can log into the application by entering email & password		High	Sprint-1
	Dashboard	USN-6	As a user, I can enter data to conduct to conduct business analysis to make business decisions		High	Sprint-1
Customer Care Executive			As a Customer Care Executive, I can answer users' queries		High	Sprint-1
Administrator			As an admin, I can make changes to the interface according the needs		High	Sprint-1

6. PROJECT PLANNING & SCHEDULING

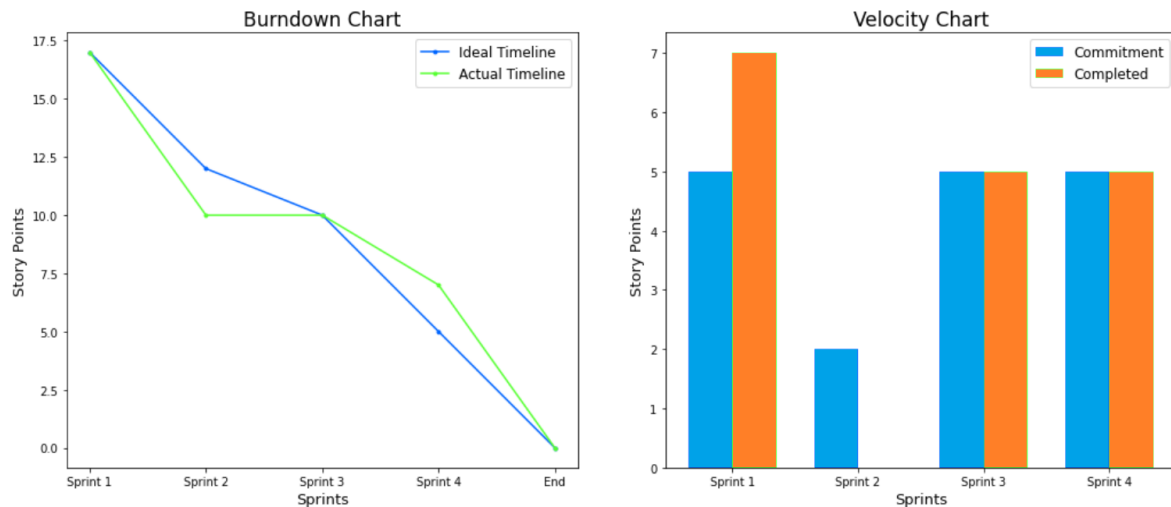
6.1 Sprint Planning & Estimation

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-1	Registration	USN-1	As a user, I can register for the application by entering my email, password, and confirming my password.	2	High	Suhail, Faiz Ali, Salman, Vanchi
Sprint-1	Registration	USN-2	As a user, I will receive confirmation email once I have registered for the application	3	High	Suhail, Faiz Ali, Salman, Vanchi
Sprint-1	Registration	USN-3	As a user, I can register for the application through Gmail	1	Low	Suhail, Faiz Ali, Salman, Vanchi
Sprint-1	Login	USN-4	As a user, I can log into the application by entering email & password	2	Medium	Suhail, Faiz Ali, Salman, Vanchi
Sprint-2	Dashboard	USN-5	As a user, I can enter my crop yield data to clean and prepare it for analysis	2	High	Suhail, Faiz Ali, Salman, Vanchi
Sprint-2	Dashboard	USN-6	As a user, I can identify trends in the data and make visualisations	2	High	Suhail, Faiz Ali, Salman, Vanchi
Sprint-2	Dashboard	USN-7	As a user, I can conduct crop yield analysis to make estimation	2	High	Suhail, Faiz Ali, Salman, Vanchi
Sprint-3	Customer Care	USN-8	As a Customer Care Executive, I can answer users' queries	2	Low	Suhail, Faiz Ali, Salman, Vanchi
Sprint-4	Administrator	USN-9	As an admin, I can make changes to the interface according to the needs	4	High	Suhail, Faiz Ali, Salman, Vanchi

6.2 Sprint Delivery Schedule

Sprint	Total Story Points	Duration	Sprint Start Date	Sprint End Date (Planned)	Story Points Completed (as on Planned End Date)	Sprint Release Date (Actual)
Sprint-1	8	6 Days	24 Oct 2022	29 Oct 2022	8	29 Oct 2022
Sprint-2	6	6 Days	31 Oct 2022	05 Nov 2022	6	05 Nov 2022
Sprint-3	2	6 Days	07 Nov 2022	12 Nov 2022	2	12 Nov 2022
Sprint-4	4	6 Days	14 Nov 2022	19 Nov 2022	4	19 Nov 2022

6.3 Reports from JIRA



7. CODING & SOLUTIONING

7.1 Feature 1

An interactive dashboard has been embedded into our web page.

Code:

<iframe

src="https://us1.ca.analytics.ibm.com/bi/?perspective=dashboard&pathRef=.my_folders%2FDashboard%2FDashboard&closeWindowOnLastView=true&ui_appbar=false&ui_navbar=false&shareMode=embedded&action=view&mode=dashboard&subView=mod

eI000001847fab81aa_00000000" width="500" height="900" frameborder="0" gesture="media" allow="encrypted-media" allowfullscreen=""></iframe>

7.2 Feature 2

An interactive report has been embedded into our web page.

Code:

```
<iframe
src="https://us1.ca.analytics.ibm.com/bi/?pathRef=.my_folders%2FReport%2FCrop%2BAnalysis%2BReport&closeWindowOnLastView=true&ui_appbar=false&ui_navbar=false&shareMode=embedded&action=run&format=HTML&prompt=false" width="500"
height="1500" frameborder="0" gesture="media" allow="encrypted-media"
allowfullscreen=""></iframe>
```

7.3 Database Schema

Data server connections

Name

Modified

CROPSDB2

11/12/2022 12:28 PM

Weather Company

3/31/2022 8:44 PM

CROPSDB2

Owner: Suhail ... f6937d

Created: 11/12/2022, 12:28 PM

Modified: 11/12/2022, 12:28 PM

Type: Data Server

General

Connections

Permissions

Name

Modified

CROPSDB2

11/12/2022 1:30 PM

CROPSDB2

Owner: Suhail ... f6937d

Created: 11/12/2022, 12:28 PM

Modified: 11/12/2022, 1:30 PM

Type: Connection

General

Settings

Schemas

Permissions

Status

Schema name

Tables loaded

☐

AUDIT

☐

DB2INST1

☐

ERRORSCHEMA

☒

TSK38118

4 / 4

Show system schemas

8. TESTING

8.1 Test Cases

Test case ID	Feature Type	Component	Test Scenario	Pre-Requlite	Steps To Execute	Expected Result	Actual Result	Status	Comments	TC for Automation(Y/N)	BUG ID	Executed By
Dashboard_TC_001	Functional	Home Page	Verify user is able to see the visualizations present in the dashboard	User should have good internet connectivity.	1. Enter URL and click go 2. Click on view dashboard button in the home page	UI elements of the home page are visible such as Navigation bar, header ,Dashboard ,Report etc	Working as expected	Pass	-	N	-	Suhail Basha , Mohamed Faiz Ali , Salman Farish ,
Dashboard_TC_002	UI	Home Page	Verify the user is able to interact with Dashboard using the Buttons and Widgets.	User should have good internet connectivity.	User should be able to click on functional buttons and view the animated illustration.	User could see the UI and interact with various buttons and elements in the web page.	Working as expected	Pass	-	N	-	Suhail Basha , Mohamed Faiz Ali , Salman Farish , Vanchinathan
Dashboard_TC_003	Functional	Home Page	Verify user is able to interact with different tabs present in the Dashboard and view other Visualisation.	User should have good internet connectivity.	1. Click on Services Section. 2. Click on view Dashboard Button. 3. Click on the respective Tabs to view other Visualisation	User should navigate to other tabs in the Dashboard	Working as expected	Pass	-	N	-	Suhail Basha , Mohamed Faiz Ali , Salman Farish , Vanchinathan
Report_TC_004	UI	Home Page	Verify user is able to see all the information in the report.	User should have good internet connectivity.	1. Click on Services Section. 2. Click on view Report Button.	User is able to view the report properly in the website.	Working as expected	Pass	-	N	-	Suhail Basha , Mohamed Faiz Ali , Salman Farish , Vanchinathan
Contact_TC_004	Functional	Home Page	Verify user is able to enter details in the Contact Us section.	User should have good internet connectivity.	1.Click on Contact Us section. 2.Enter Mail Id 3.Enter Message 4.Click on Send Button	User should be able to send messages through this form.	Working as expected	Pass	-	N	-	Suhail Basha , Mohamed Faiz Ali , Salman Farish , Vanchinathan
Story_TC_005	Functional	Home Page	Verify user is able to play the Story in the Website.	User should have good internet connectivity.	1.Click on Services 2.Click on View Story Button 3.Click on Play Button	User should be able to Play the Story in the Website Properly.	Working as expected	Pass	-	N	-	Suhail Basha , Mohamed Faiz Ali , Salman Farish , Vanchinathan

8.2 User Acceptance Testing

1. Purpose of Document

The purpose of this document is to briefly explain the test coverage and open issues of the Estimate the crop yield using Data Analytics project at the time of the release to User Acceptance Testing (UAT).

2. Defect Analysis

This report shows the number of resolved or closed bugs at each severity level, and how they were resolved

Resolution	Severity 1	Severity 2	Severity 3	Severity 4	Subtotal
By Design	3	1	1	1	6
Duplicate	1	0	0	0	1
External	0	2	0	0	2
Fixed	1	2	0	2	5
Not Reproduced	0	0	1	0	1
Skipped	0	0	0	0	0
Won't Fix	0	1	0	0	1
Totals	5	6	2	3	16

3. Test Case Analysis

This report shows the number of test cases that have passed, failed, and untested

Section	Total Cases	Not Tested	Fail	Pass
Home	4	0	0	4
Dashboard	2	0	0	2
Report	1	0	0	1
Story	4	0	0	4
Contact Us	4	0	0	4

9. RESULTS

9.1 Performance Metrics

S. No.	Parameter	Screenshot / Values
1.	Dashboard design	No of Visualizations / Graphs - 20
2.	Data Responsiveness	Very responsive
3.	Amount Data to Rendered (DB2 Metrics)	15 MB (crop_production.csv)
4.	Utilization of Data Filters	Utilized to full effectiveness
5.	Effective User Story	No of Scene Added - 4
6.	Descriptive Reports	No of Visualizations / Graphs - 4

10. ADVANTAGES &DISADVANTAGES

Advantages

- Risks can be measured when suitable mathematical and statistical model designs are applied on data related to soil, weather and past yield.
- The precise quantification of the rice productivity in various climatic conditions can help farmer to understand the optimum condition for the higher rice crop yield.

Disadvantages

- The main disadvantage of this strategy is that it does not allow for assessing the impact of intercropping on crop yields.
- This strategy is that crop areas cannot be aggregated at farm or higher levels, since intercropped fields would be double counted.

11. CONCLUSION

To conclude, we will say that we are providing an excellent website that can Estimate crop

yield using Data Analytics by using the latest Artificially powered tool called IBM Cognos.

12. FUTURE SCOPE

By using the data's, we can identify the output yield of the crops and forecast prices for the next few weeks will help the farmers to obtain maximum benefits.

13. APPENDIX

Source code - <https://github.com/IBM-EPBL/IBM-Project-46723-1660754993/tree/main/Final%20Deliverables/Source%20Code>

GitHub - <https://github.com/IBM-EPBL/IBM-Project-46723-1660754993>

Project Demo - <https://drive.google.com/file/d/1UPqqVOBY-unZ9ibMJhOT1tFE1qVVyV8W/view?usp=drivesdk>