

# PROJECT REPORT

## Estimate The Crop Yield Using Data Analytics

### 1.INTRODUCTION:

Crop production in India is one of the most important sources of income and India is one of the top countries to produce crops. In this project, we will be analyzing some important visualization, creating a dashboard and by going through these we will get most of the insights of Crop production in India.

#### 1.1 Project Overview:

Agriculture is important for human survival because it serves the basic need. Due to variations in climatic conditions, there exist bottlenecks for increasing the crop production in India. 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. The use of technology in agriculture has increased in recent year and data analytics is one such trend that has penetrated into the agriculture field.

#### 1.2.Purpose:

Farming depends on various factors like climate and economic factors like temperature, irrigation, cultivation, soil, rain fall, pesticide and fertilizers. Historical information regarding crop yield provides major input for companies engaged in this domain. Crop yield prediction helps the farmers in various ways by providing the record of previous crop yield. Technology is the emerging trend in the research in this area in recent days. One of the areas being explored is the problem of yield prediction which is a major concern. Data mining techniques are being widely used as a part of solution for crop yield prediction. data analytics techniques use artificial intelligence, statistics, machine learning and database system. This survey focuses on various methods being used for crop yield prediction.

## **2.LITERATURE SURVEY:**

### **2.1.Existing Problem:**

At present, we are at the immense need of another Green revolution to supply the food demand of growing population. With the decrease of cultivable land globally and the decreased cultivable water resources, it is almost impossible to report higher crop yield. Agricultural based data analytics is one approach, believed to have a significant role and positive impact on the increase of crop yield by providing the optimum condition for the plant growth and decreasing the yield gaps and the crop damage and wastage.

### **2.2.References:**

1. Dhivya B H, Manjula R, Siva Bharathi S, Madhumathi R.  
A Survey on Crop Yield Prediction based on Agricultural Data, International Journal of Innovative Research in Science, Engineering and Technology, 2017.
2. Jharna Majumdar, Sneha Naraseeyappa, Shilpa Ankalaki.  
Analysis of agriculture data using data mining technique and application of big data, 2017.

### **2.3.Problem Statement Definition:**

The accurate prediction of crop yield certainly benefits the farmers in choosing the right method to reduce the crop damage and gets best prices for their crops. It is done with an objective of accurate prediction of crop yield through data analytics to assess various crop yield factors. The yield maps developed shall provide a unique opportunity to overcome both spatial and temporal based scaling up challenges and thus improve the ideology of crop yield prediction.

## **3.IDEATION & PROPOSED SOLUTION:**

### **3.1.Empathy Map Canvas:**

In this empathy map, we have stated the users' needs, pain points and also satisfies the customer needs.



2

## Brainstorm

Write down any ideas that come to mind that address your problem statement.

🕒 10 minutes

### TIP

You can select a sticky note and hit the pencil [switch to sketch] icon to start drawing!

### Deevija T

We must use natural fertilizers.	Woodings should be removed periodically.	Soil must be checked in agri labs.
Water should be fed to the crops regularly.	Irrigation methods should be followed on the basis of type of crops.	crops should not get damaged in case of natural disasters like.
agro chemicals can be used	necessary sunlight should be provided to crops.	pH level should be maintained

### Kavishankari S P

Intensive farming can be used to increase the crop production.	Survey from farmers can be done and can be used based on survey.	Choosing yied potential zones
Modern equipments can be used	Organic fertilizers can be used like cow dung.	Using high yielding variety seeds.
Disease resistance can be increased	Climatic changes should not affect crops	conservation tilage can be done

### Kowsalya A

Choosing right seeds for planting	crop protection methods should be taken.	Crop rotation so as to increase the fertility of the soil.
Modern irrigation methods can be used	pesticides should be such that it should not affect both soil and crops.	High nitrogen fertilizers can be used for high yield of crops.
Regular Scouting can be done	Seasonal croppings can be done	Alluvial soil is preferable

### Nivethitha M

Automation strategies reduce the overall production cos.	Sand and clay are not suitable for growing crops.	Companion Planting can be done
There should be required gaps between crops.	Increased temperature and humidity can be used to increase the growth of crops.	Organic fertilizers should be used after each collection.
Seeds quality should be checked before sowing	Harvesting should be done at run time	Plants need appropriate the nutrients

3

## Group ideas

Take turns sharing your ideas while cluster. In the last 10 minutes, give each cluster a s than six sticky notes, try and see if you and

🕒 20 minutes

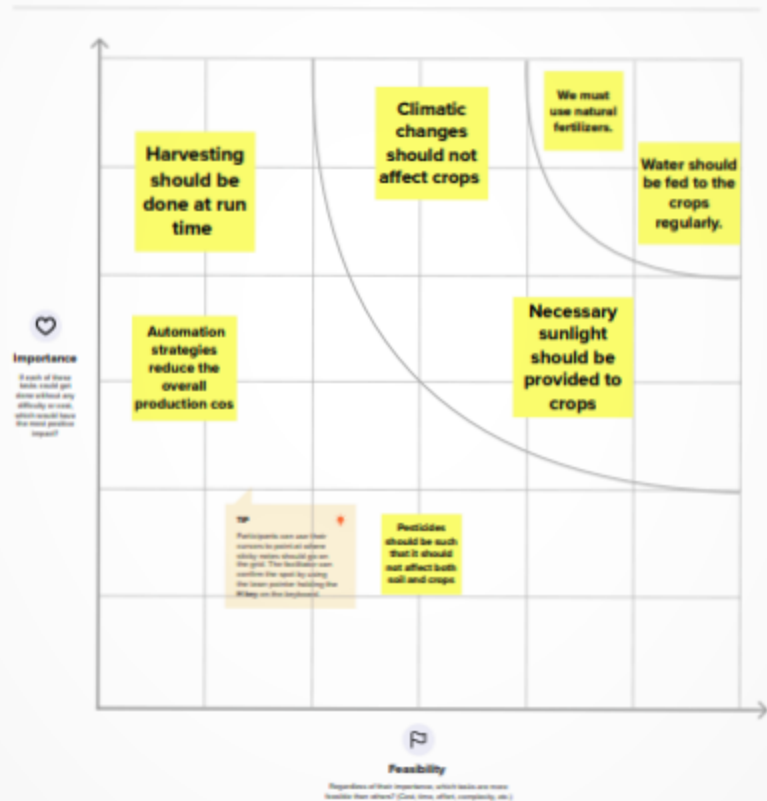
Automation strategies reduce the overall production cos.	Climatic changes should not affect crops	crops should not get damaged in case of natural disasters like.
We must use natural fertilizers.	Organic fertilizers can be used like cow dung.	Organic fertilizers should be used after each collection.
Modern irrigation methods can be used	Irrigation methods should be followed on the basis of type of crops.	Water should be fed to the crops regularly.

4

## Prioritize

Your team should all be on the same page about what's important moving forward. Place your ideas on this grid to determine which ideas are important and which are feasible.

🕒 20 minutes



### **3.3.Proposed Solution:**

Our proposed solution will relate the current situation to a desired result and describe the benefits that will accrue when the desired result is achieved.

#### **Idea / Solution Description:**

- In this project we will be analysing important visualization by creating a dashboard and by going through these we will get most of the insights of Crop production in India.

#### **Novelty/Uniqueness:**

- Creating interactive Chatbots.
- Attractive dashboard with aesthetic visualisation.
- Providing expert suggestions to farmers.

#### **Social Impact/ Customer Satisfaction:**

- This project is helpful in framing government policies related to crop insurance to farmers.
- This helps the farmers to keep track of previous crop yield.

#### **Business Model:**

- Supply chain operation between farmers and Entrepreneurs.
- Helps the companies in project scheduling.

#### **Scalability of solution:**

- Available for large scale farmers.
- Dataset can be updated according to future needs.

### 3.4.Problem Solution Fit:

Problem-Solution Fit means that have found a problem with the customer and that the solution have been realized for it actually solves the customer's problem.

Project Title: Estimate the crop yield using data analytics      Project Design Phase-I - Solution Fit Template      Team ID: PNT2022TMD44196

Define CS, fit into CC	<b>1. CUSTOMER SEGMENT(S)</b> <span>CS</span> Who is your customer? i.e. working parents of 0-5 y.o. kids  <b>This project is mainly for the farmers who are looking for the best and profitable yield.</b>	<b>6. CUSTOMER CONSTRAINTS</b> <span>CC</span> What constraints prevent your customers from taking action or limit their choices of solutions? i.e. spending power, budget, no cash, network connection, available devices.  <b>The estimation may be wrong at rare cases.</b>	<b>5. AVAILABLE SOLUTIONS</b> <span>AS</span> Which solutions are available to the customers when they face the problem or need to get the job done? What have they tried in the past? What pros & cons do these solutions have? i.e. pen and paper is an alternative to digital notetaking  <b>This project will aims at accurate estimation,using natural resources,GIVING appropriate alerts,etc.,</b>	Explore AS, differentiate
	<b>2. JOBS-TO-BE-DONE / PROBLEMS</b> <span>J&amp;P</span> Which jobs-to-be-done (or problems) do you address for your customers? There could be more than one; explore different sides.  <b>As there is no high yield for farmers these days, this project will help them in a great way.</b>	<b>9. PROBLEM ROOT CAUSE</b> <span>RC</span> What is the real reason that this problem exists? What is the back story behind the need to do this job? i.e. customers have to do it because of the change in regulations.  <b>Crop production is the important source of income of farmers and india.</b>	<b>7. BEHAVIOUR</b> <span>BE</span> What does your customer do to address the problem and get the job done? i.e. directly related: find the right solar panel installer, calculate usage and benefits; indirectly associated: customers spend free time on volunteering work (i.e. Greenpeace)  <b>Existing solutions will estimate with low accuracy at sometimes but also helpful for farmers.</b>	
Focus on J&P, tap into BE, understand RC	<b>3. TRIGGERS</b> <span>TR</span> What triggers customers to act? i.e. seeing their neighbour installing solar panels, reading about a more efficient solution in the news. <b>For yielding good crops and high income, correct estimation of crop is needful.</b>	<b>10. YOUR SOLUTION</b> <span>SL</span> If you are working on an existing business, write down your current solution first, fill in the canvas, and check how much it fits reality. If you are working on a new business proposition, then keep it blank until you fill in the canvas and come up with a solution that fits within customer limitations, solves a problem and matches customer behaviour.  <b>Estimating the crop yield in a field for farmers to increase their yield and revenue.</b>	<b>8. CHANNELS of BEHAVIOUR</b> <span>CH</span> <b>8.1 ONLINE</b> What kind of actions do customers take online? Extract online channels from #7  <b>8.2 OFFLINE</b> What kind of actions do customers take offline? Extract offline channels from #7 and use them for customer development.  <b>Farmers are the customers.</b>	Identify strong TR & EM
	<b>4. EMOTIONS: BEFORE / AFTER</b> <span>EM</span> How do customers feel when they face a problem or a job and afterwards? i.e. lost, insecure > confident, in control - use it in your communication strategy & design.  <b>Agriculture with new perspective; agriculture with difference.</b>			

### 4.REQUIREMENT ANALYSIS:

#### 4.1.Functional Requirement:

A functional requirement defines a system or its component and it is defined by the user.The Functional Requirements Definition reports and tracks the basic information expected to effectively portray business and handy necessities. The Functional Requirements Definition report is made in the midst of the Planning Phase of the endeavor. Its objective gathering is the endeavor boss, errand gathering, wander bolster, client/customer, and any accomplice whose information/respect into the necessities definitions system is required.

#### 4.2.Non-Functional Requirements:

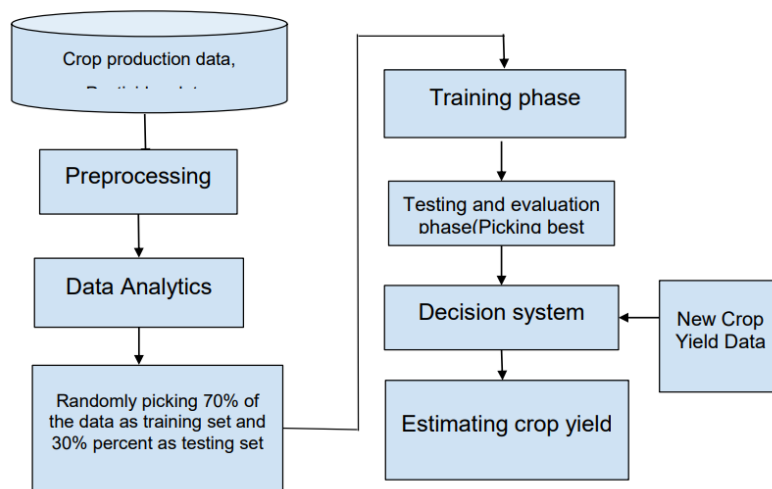
A non-functional requirement defines the quality attribute of a software system.A non-

functional requirement (NFR) is a requirement that specifies criteria that can be used to judge the operation of a system, rather than specific behaviors. Non functional requirements specifies the quality attribute of a software system. The software system can be judged based on reliability, security, maintainability, performance, portability, scalability and flexibility.

## 5.PROJECT DESIGN:

### 5.1.Data Flow Diagrams:

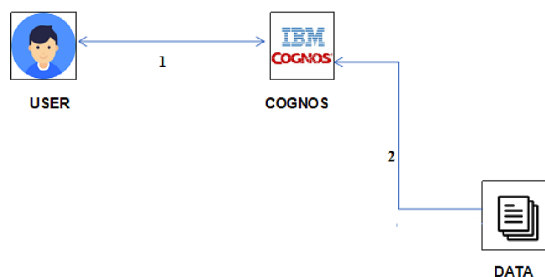
A Data Flow Diagram (DFD) is a traditional visual representation of the information flows within a system. A neat and clear DFD can depict the right amount of the system requirement graphically. It shows how data enters and leaves the system, what changes the information, and where data is stored.



### 5.2.Solution and Technical Architecture:

Solution architects and technical architects develop systems that help organizations focus and simplify their operations.

#### Technical Architecture:



### 5.3.User Stories:

A user story is an informal, general explanation of a software feature written from the perspective of the end user. Its purpose is to articulate how a software feature will provide value to the customer. As our end users are FARMERS, they can get more yield and healthy crops with the help of our project.

## 6.PROJECT PLANNING & SCHEDULING:

### 6.1.Sprint Planning & Estimation:

S.no	Milestone	Activities	Start Date	End Date
1	Solution Requirement	Creating the IBM Cognos for creating dashboard and data visualization charts.	22-Aug-2022	24-Aug-2022
2	Project Objectives	Prepare the project objectives.	22-Aug-2022	24-Aug-2022
3	Project Flow	Prepare the project flow.	22-Aug-2022	24-Aug-2022
4	IBM Cloud Account	Creating IBM cloud account.	22-Aug-2022	24-Aug-2022
5	IBM Cognos Analytics	Creating IBM cognos account.	22-Aug-2022	24-Aug-2022
6	Working with the Dataset	Understanding The Dataset Loading The Dataset.	24-oct-2022	19-Nov-2022
7	Data Visualization Charts	<ul style="list-style-type: none"> <li>Seasons With Average Productions</li> <li>With Years Usage of Area And Production</li> <li>Top 10 States with Most Area</li> <li>State With Crop Production</li> <li>States With the Crop Production Along with Season</li> </ul>	24-oct-2022	19-Nov-2022
8	Creating the Dashboard	Creating The Dashboard	24-oct-2022	19-Nov-2022
9	Export the Analytics	Export The Analytics	24-oct-2022	19-Nov-2022



10	Ideation Phase	<ul style="list-style-type: none"> <li>● Literature Survey On The Selected Project &amp; Information Gathering Prepare</li> <li>● Empathy Map</li> <li>● Ideation</li> </ul>	22-Aug-2022	17-Sept-2022
----	----------------	--	-------------	--------------

S.no	Milestone	Activities	Start Date	End Date
11	Project Design Phase - I	<ul style="list-style-type: none"> <li>● Proposed Solution</li> <li>● Problem Solution Fit</li> <li>● Solution Architecture</li> </ul>	22-Aug-2022	17-Sept-2022
12	Project Design Phase - II	<ul style="list-style-type: none"> <li>● Customer Journey</li> <li>● Functional Requirement</li> <li>● Data Flow Diagrams</li> <li>● Technology Architecture</li> </ul>	22-Aug-2022	01-Oct-2022
13	Project Planning Phase	<ul style="list-style-type: none"> <li>● Prepare Milestone &amp; Activity List</li> <li>● Sprint Delivery Plan</li> </ul>	17-Oct-2022	22-Oct-2022
14	Project Development Phase	<ul style="list-style-type: none"> <li>● Project Development - Delivery of Sprint-1</li> <li>● Project Development - Delivery of Sprint-2</li> <li>● Project Development - Delivery of Sprint-3</li> <li>● Project Development - Delivery of Sprint-4</li> </ul>	24-Aug-2022	19-Nov-2022

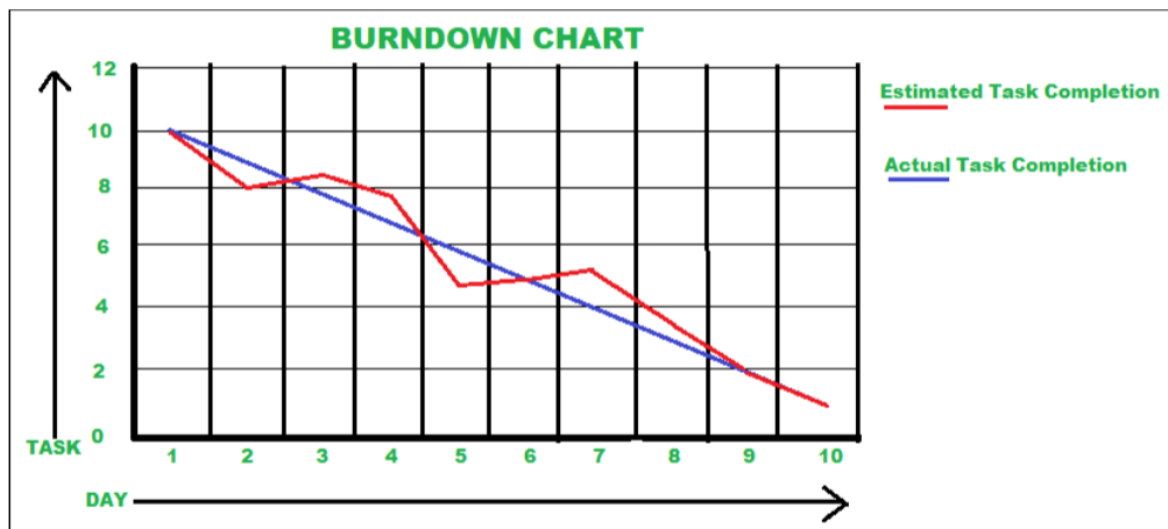
## 6.2.Sprint Delivery Schedule:

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 by entering my id card and request..	2	High	Deevija T Kowsalya A Kavishankari S P Nivethitha M
		USN-2	As a user, I can register for the application through Gmail	2	Medium	Deevija T Kowsalya A
	Login	USN-3	As a user, I can Call and request or Approach for dataset	4	High	Kavishankari S P Nivethitha M
	Working with the Dataset	USN-4	To work on the given dataset, Understand the Dataset.	2	High	Deevija T Kowsalya A Kavishankari S P Nivethitha M
		USN-5	Load the dataset to Cloud platform then Build the required Visualizations.	10	High	Kowsalya A Kavishankari S P
Sprint-2	Data Visualization Chart	USN-6	Using the Crop production in Indian dataset, create various graphs and charts to highlight the insights and visualizations. *Build a Visualization to showcase Average Crop Production by Seasons.	4	Medium	Deevija T Kowsalya A
			*Showcase the Yearly usage of Area in Crop Production.	4	Medium	Deevija T Kowsalya A

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
			Build a visualization to show case top 10 States in Crop Yield Production by Area.	4	Medium	Deevija T Kavishankari S P
			Build the required Visualization to showcase the Crop Production by State.	4	Medium	Deevija T Nivethitha M
			Build Visual analytics to represent the Sates with Seasonal Crop Production using a Text representation.	4	Medium	Kowsalya A Kavishankari S P
Sprint-3	Creating The dashboard	USN-8	Create the Dashboard by using the created visualizations.	20	High	Kowsalya A Deevija T
Sprint-4	Export The Analytics	USN-9	Export the created Dashboard	20	High	Kavishankari S P Nivethitha M

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	20	6 Days	24 Oct 2022	29 Oct 2022	20	29 Oct 2022
Sprint-2	20	6 Days	31 Oct 2022	05 Nov 2022	20	05 Nov 2022
Sprint-3	20	6 Days	07 Nov 2022	12 Nov 2022	20	12 Nov 2022
Sprint-4	20	6 Days	14 Nov 2022	19 Nov 2022	20	19 Nov 2022

### 6.3.Reports from JIRA:

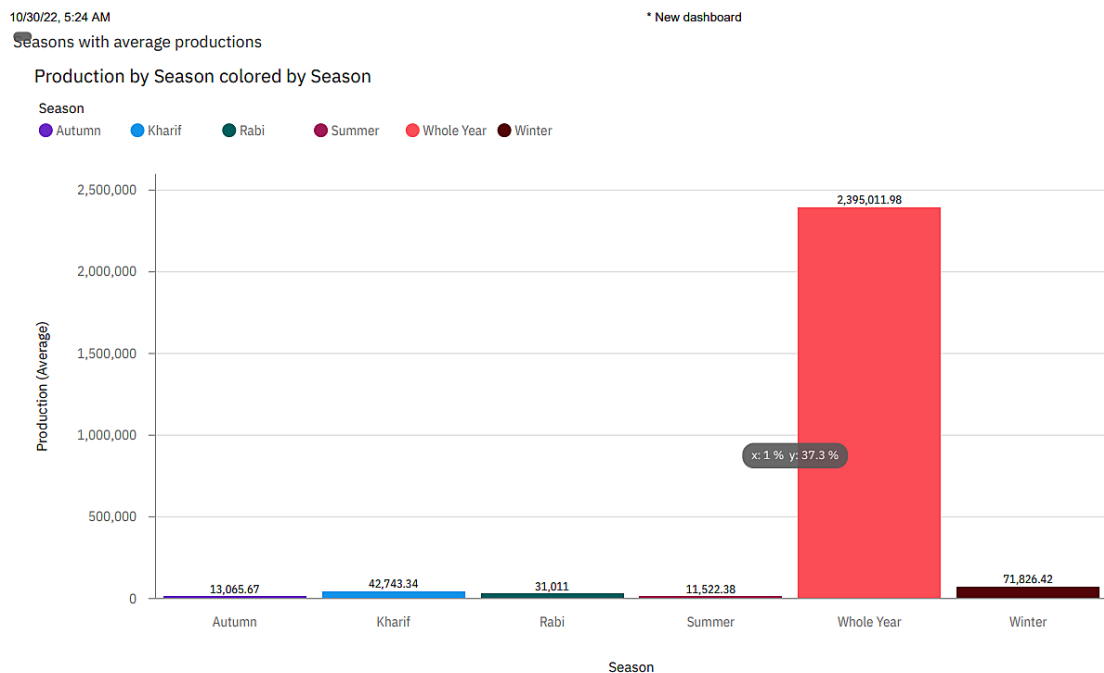


## 7.CODING & SOLUTIONING:

### 7.1.Feature 1:

In this part, we have created five visualizations as follows:

- The first visualization is different seasons with average crop production.



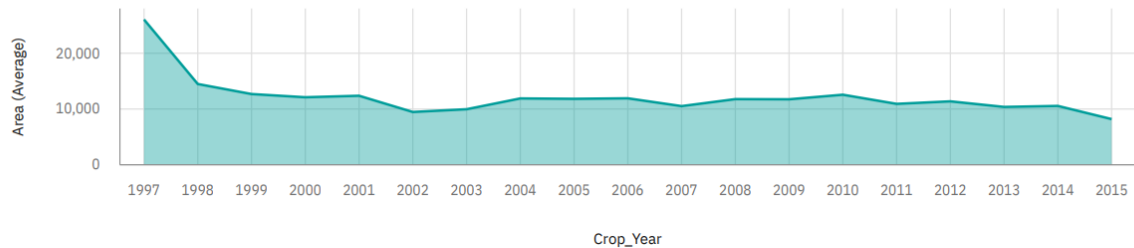
- The second visualization is with the years usage of area and production.

10/30/22, 5:24 AM

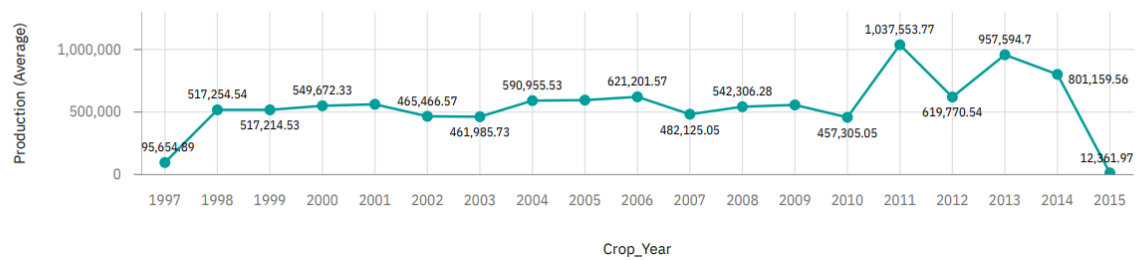
\* New dashboard

With Years Usage Of Area And Production

Area by Crop\_Year



Production by Crop\_Year



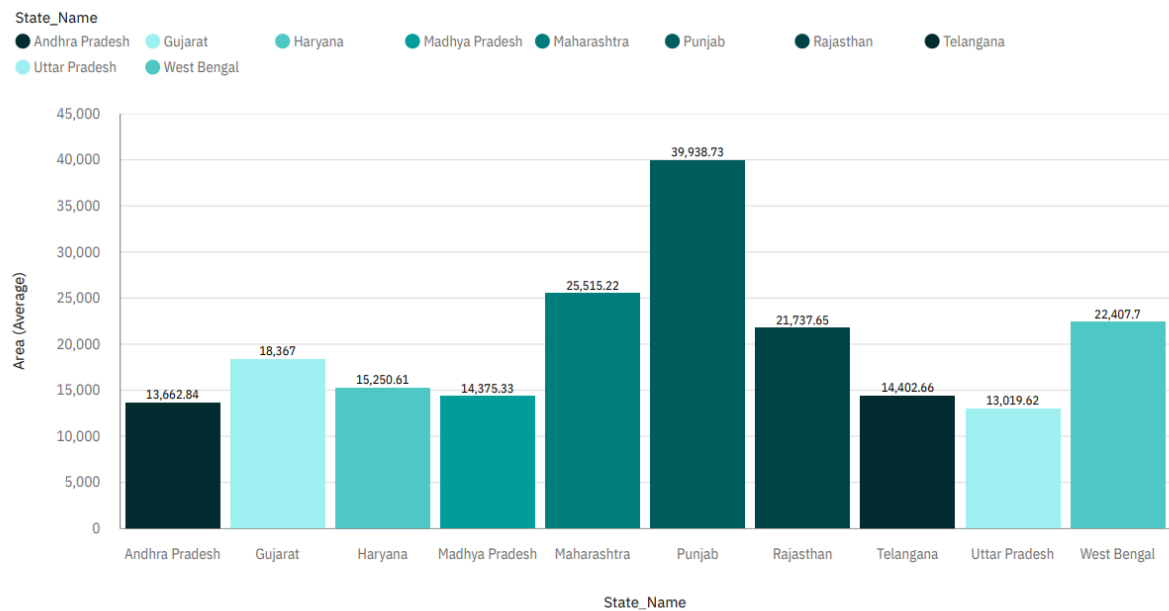
- In our third visualization, we have visualized the top 10 states with most area.

10/30/22, 5:24 AM

\* New dashboard

Top 10 States With Most Area

Area by State\_Name colored by State\_Name



- 10/30/22, 5:24 AM

\* New dashboard

State With Crop Production

State\_Name for State\_Name regions

State\_Name

  - Assam
  - Bihar
  - Madhya Pradesh
  - Andhra Pradesh
  - Rajasthan
  - Gujarat
  - Chhattisgarh
  - Meghalaya
  - Dadra and Nagar Haveli
  - Maharashtra
  - Karnataka
  - Tamil Nadu
  - Goa
  - Telangana
  - Manipur
  - Uttar Pradesh
  - Kerala
  - West Bengal
  - Haryana
  - Andaman and Nicobar Islands
  - Puducherry

- 10/30/22, 5:24 AM

\* New dashboard

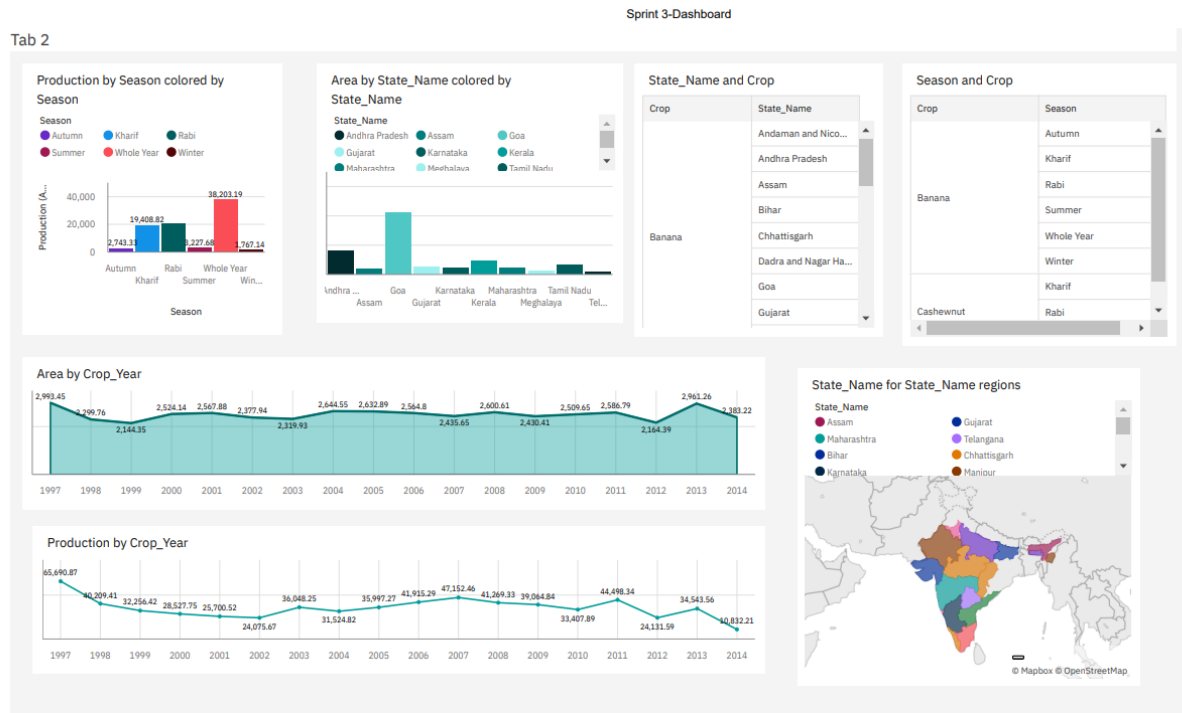
States With The Crop Production Along With Season (Text Table)

State\_Name and Crop

Crop	State_Name
Banana	Andaman and Nicobar Islands
	Andhra Pradesh
	Assam
	Bihar
	Chhattisgarh
	Dadra and Nagar Haveli
	Goa
	Gujarat
	Haryana
	Karnataka
	Kerala
	Madhya Pradesh
	Maharashtra
	Manipur

## 7.2.Feature 2:

Using these five visualizations, we have created a dashboard as shown below.



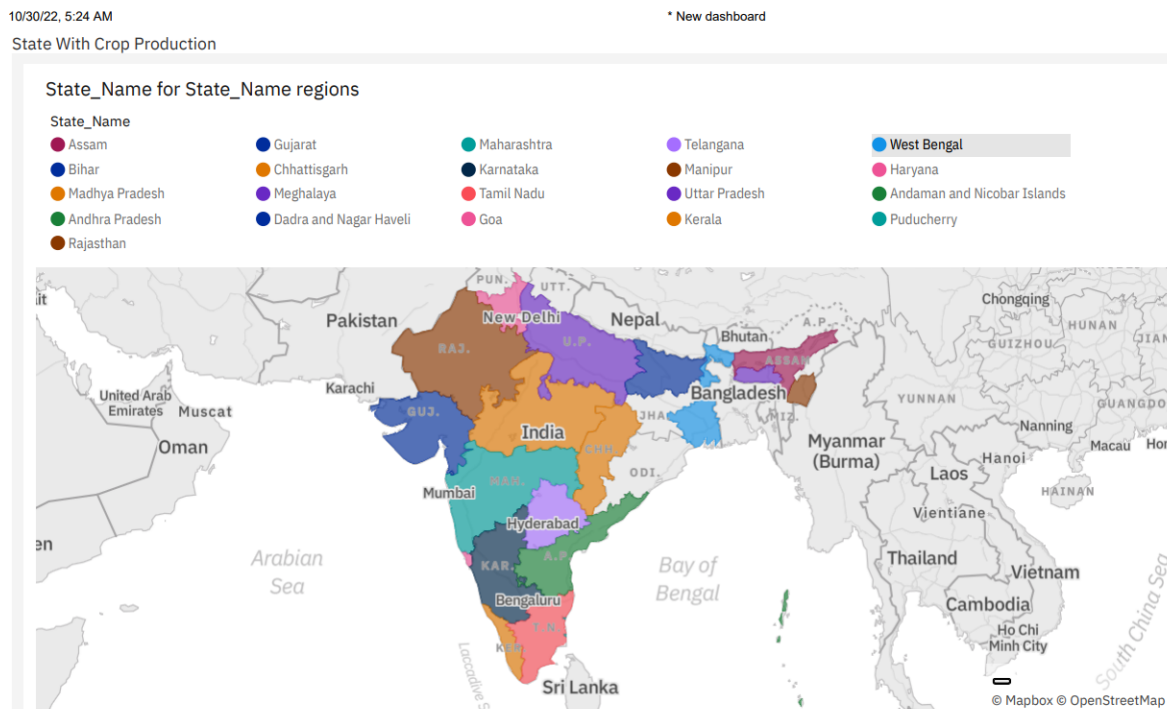
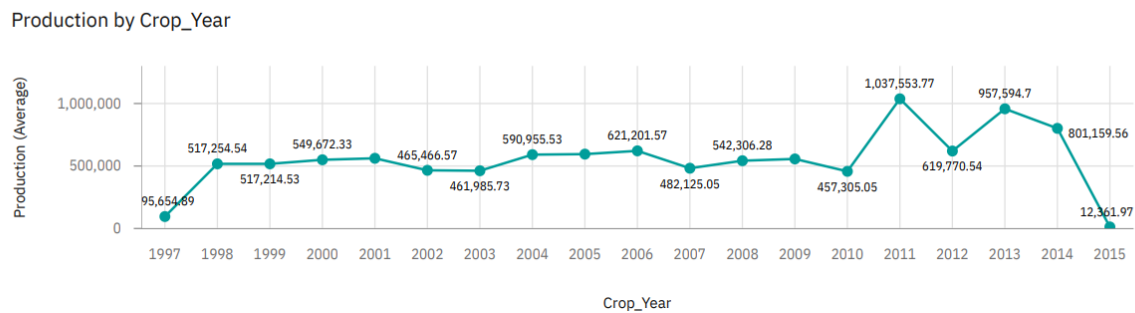
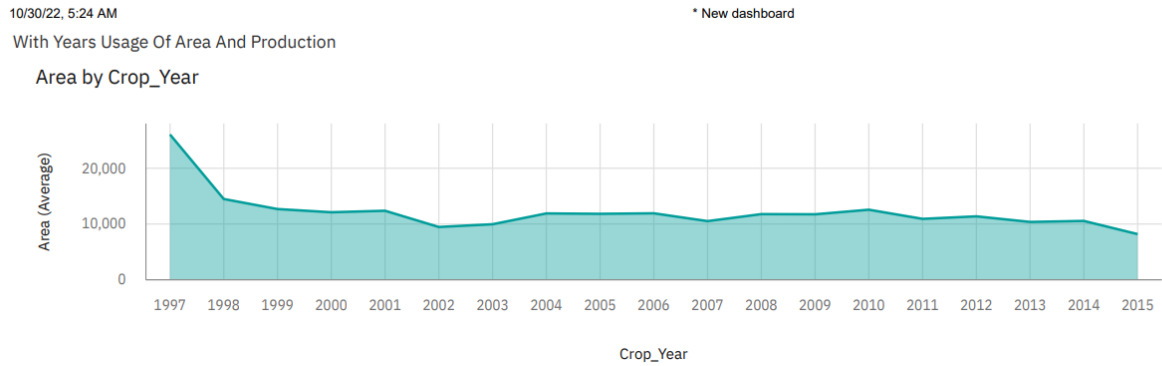
## 7.3.Database Schema:

By using the below dataset, we have visualized the crop yield and created the dashboard.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	State_Name	District_Name	Crop_Year	Season	Crop	Area	Production							
2	Andaman and Nicobar Islands	NICOBARS	2000	Kharif	Areca nut	1254	2000							
3	Andaman and Nicobar Islands	NICOBARS	2000	Kharif	Other Kharif pulses	2	1							
4	Andaman and Nicobar Islands	NICOBARS	2000	Kharif	Rice	102	321							
5	Andaman and Nicobar Islands	NICOBARS	2000	Whole Year	Banana	176	641							
6	Andaman and Nicobar Islands	NICOBARS	2000	Whole Year	Cashewnut	720	165							
7	Andaman and Nicobar Islands	NICOBARS	2000	Whole Year	Coconut	18168	65100000							
8	Andaman and Nicobar Islands	NICOBARS	2000	Whole Year	Dry ginger	36	100							
9	Andaman and Nicobar Islands	NICOBARS	2000	Whole Year	Sugarcane	1	2							
10	Andaman and Nicobar Islands	NICOBARS	2000	Whole Year	Sweet potato	5	15							
11	Andaman and Nicobar Islands	NICOBARS	2000	Whole Year	Tapioca	40	169							
12	Andaman and Nicobar Islands	NICOBARS	2001	Kharif	Areca nut	1254	2061							
13	Andaman and Nicobar Islands	NICOBARS	2001	Kharif	Other Kharif pulses	2	1							
14	Andaman and Nicobar Islands	NICOBARS	2001	Kharif	Rice	83	300							
15	Andaman and Nicobar Islands	NICOBARS	2001	Whole Year	Cashewnut	719	192							
16	Andaman and Nicobar Islands	NICOBARS	2001	Whole Year	Coconut	18190	64430000							
17	Andaman and Nicobar Islands	NICOBARS	2001	Whole Year	Dry ginger	46	100							
18	Andaman and Nicobar Islands	NICOBARS	2001	Whole Year	Sugarcane	1	1							
19	Andaman and Nicobar Islands	NICOBARS	2001	Whole Year	Sweet potato	11	33							
20	Andaman and Nicobar Islands	NICOBARS	2002	Kharif	Rice	189.2	510.84							
21	Andaman and Nicobar Islands	NICOBARS	2002	Whole Year	Areca nut	1258	2083							
22	Andaman and Nicobar Islands	NICOBARS	2002	Whole Year	Banana	213	1278							
23	Andaman and Nicobar Islands	NICOBARS	2002	Whole Year	Black pepper	63	13.5							
24	Andaman and Nicobar Islands	NICOBARS	2002	Whole Year	Cashewnut	719	208							
25	Andaman and Nicobar Islands	NICOBARS	2002	Whole Year	Coconut	18240	67490000							
26	Andaman and Nicobar Islands	NICOBARS	2002	Whole Year	Dry chillies	413	28.8							
27	Andaman and Nicobar Islands	NICOBARS	2002	Whole Year	Dry ginger	47.3	133							
28	Andaman and Nicobar Islands	NICOBARS	2002	Whole Year	Sugarcane	5	40							
29	Andaman and Nicobar Islands	NICOBARS	2003	Kharif	Rice	52	90.17							

### 8.1. Test cases:

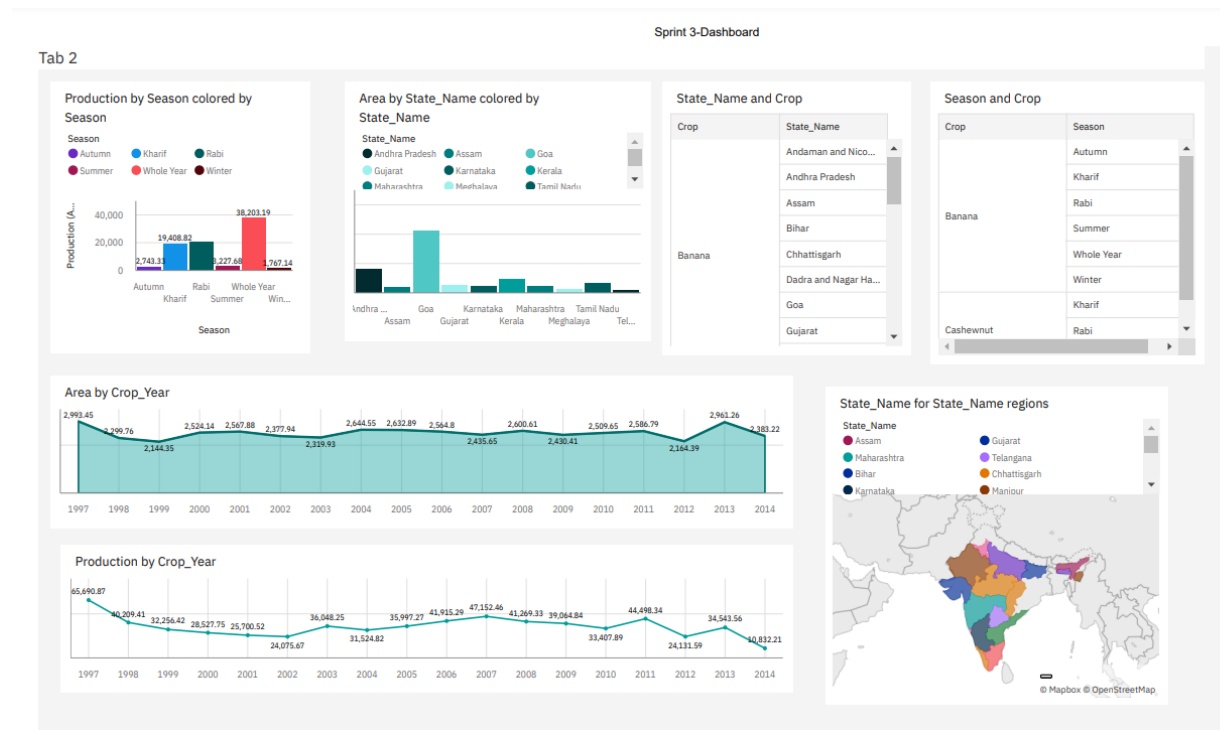
The sample test cases are as follows,





## 8.2.User Acceptance Testing:

User Acceptance Testing (UAT) is a type of testing performed by the end user or the client to verify/accept the software system before moving the software application to the production environment. UAT is done in the final phase of testing after functional, integration and system testing is done.



## 9.RESULTS:

### 9.1.Performance Metrics:

Performance metrics are known as numbers and data representing organizations' abilities, actions, and overall quality. Various forms of performance metrics include profit, sales, customer happiness, return on investment, customer reviews, general quality, personal reviews, along with reputation in marketplaces.



CUSTOMER  
SATISFACTION



INCREASE IN  
PRODUCTION



QUALITY OF  
SERVICE



SOIL

## PERFORMANCE METRICS



ECONOMICAL



RELIABILITY

## **10.ADVANTAGES AND DISADVANTAGES:**

### **Advantages:**

Crop yield prediction is used by farmers to make decisions about when to plant and harvest crops based on

- soil moisture content
- pest infestations
- weather conditions factors and
- fertilizer requirements.

### **Disadvantages:**

- With the changing of climate, agriculture faces increasing problems with extreme weather events leading to considerable yield losses of crops.
- Most often, crop plants are sensitive to stresses since they were mostly selected for high yield, and not for stress tolerance.

## **11.CONCLUSION:**

As a result of penetration of technology into agriculture field, there is a marginal improvement in the productivity. The innovations have led to new concepts like digital agriculture, smart farming, precision agriculture etc. The activities of agriculture field are numerous like weather forecasting, soil quality assessment, seeds selection, crop yield prediction,etc.,Based on the analysis, model will be more accurate if the more datasets are available. So as the data point increases, the system will become more and more accurate.Since displaying the results in the form of graph with actual and predicted in the graphical user interface, it is easy to compare the previous year's data. This model will help farmers to grow the crop which will give more yield so that it will be more profitable.

## **12.FUTURE SCOPE:**

- 1.Predict appropriate crop and maximum yield in the climate change.
- 2.Collection of data, Analysis of it and modification of the algorithm.

3.IOT application in agriculture, automation in production line and man free agriculture which is the future of the world ,this is the first step of it.

4.Find the percentage yield to happen from the match given percentage in terms of % error.

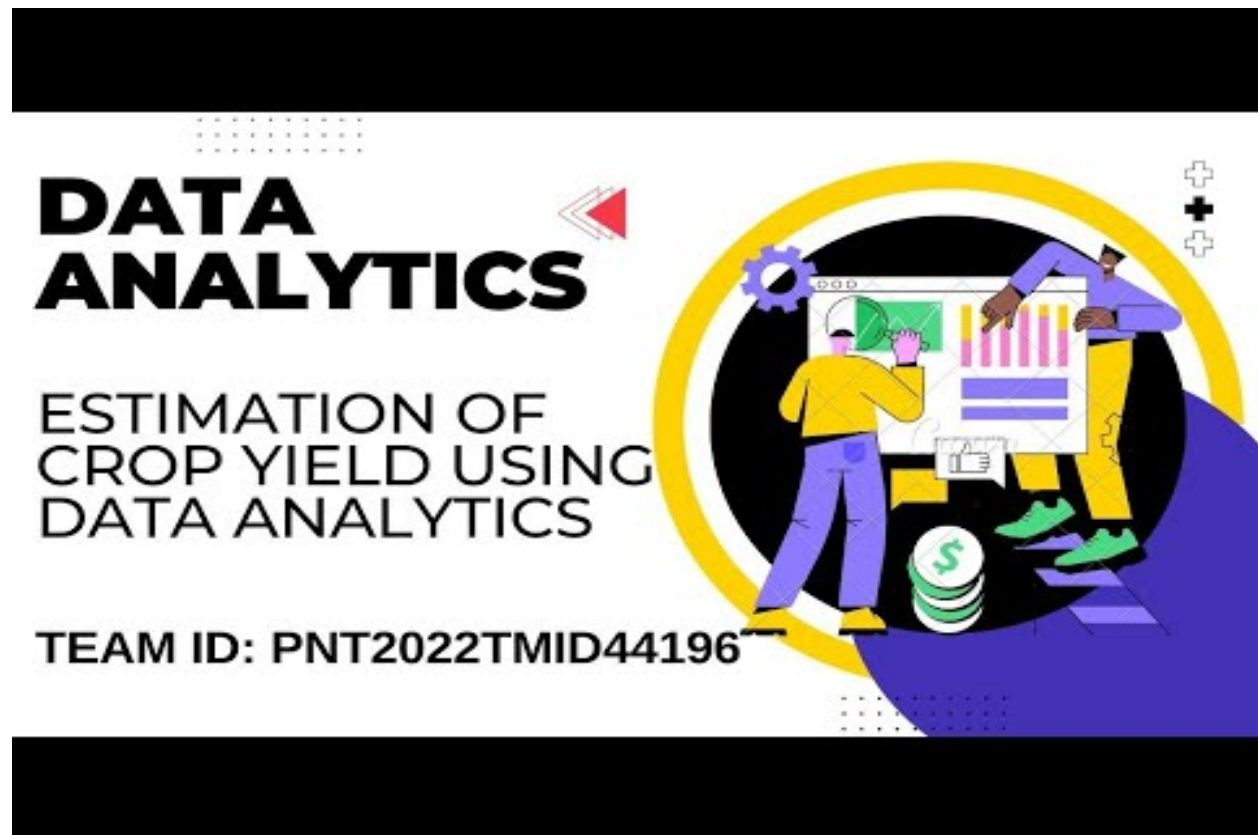
### 13.APPENDIX:

#### GitHub Link:

<https://github.com/IBM-EPBL/IBM-Project-42065-1660648232>

#### Project Demo Link:

#### Youtube link:



#### Drive link:

[https://drive.google.com/file/d/13SRdf6hb3jyijQVGjSOyeSyfTSGPHD5A/view?usp=share\\_link](https://drive.google.com/file/d/13SRdf6hb3jyijQVGjSOyeSyfTSGPHD5A/view?usp=share_link)

