PROJECT REPORT

ESTIMATION OF CROP PRODUCTION USING DATA ANALYTICS

TEAM MEMBERS :
- **SANJEEV M (TEAM LEAD)**
- **MANIKANDAN P**
- **MOHAMED NOWFAL S**
- **SREEJITH G**

1. Introduction:
1.1 Overview:
The demand for food is found increasing with increase in population. With a rapid demand for cultivation of food, it is highly essential to analyse the growth of their production with respect to seasons, production, area, demands across cities etc. Hence we bring in an exemplary analytical dashboard for the farmers to understand all possible enhancements that has to be done to upscale their production. We have given all possible insights that will help the users to get quick overview.

1.2 Purpose:
Our Analytical Dashboard is completely user friendly and it is designed in a way that grabs the famer's attention as the representation is made appetizing and interesting. . The main problem to be solved is predicting the crop yield using Data Analytics which helps farmers to analyse the Crop Production. To assist him in planning and harvesting.The ultimate purpose for analytics is to build a strong marketing strategy which can be achieved through our project.

2. LITERATURE SURVEY:
2.1 Existing Problem:
Agriculture is the field that enables the farmers to grow ideal crops in accordance with the environmental balance. In India, wheat and rice are the major grown crops along with sugar-cane, potatoes, oil seeds etc. Farmers also grow non-food items like rubber,

cotton, jute etc. More than 70% of the household in the rural area depend on agriculture. This domain provides employment to more than 60% of the total population and has a contribution to GDP also (about 17%). In the farm output, India ranks second considering the world wide scenario. This is the widest economic sector and has an important role regarding the framework of socio-economic fabric of India.

2.2 References:

Aggarwal et al discuss about various Data Mining tools such as Dashboards, TextMining tools. They provide an overview about these tools and the various scenarios in which they can be deployed [2].
The proposed architecture mainly focuses on open source tools for the development of the application. The user can select location from map for which the details are available at one click [3].
K. Sabarina and N. Priya, [2] has presented an efficient strategy for crops Big data for the benefit of precision agricultural lowering data dimensionality. Predictive analytics can be used to make the smartest decision in farming by collection real-time data analysis with streaming data.
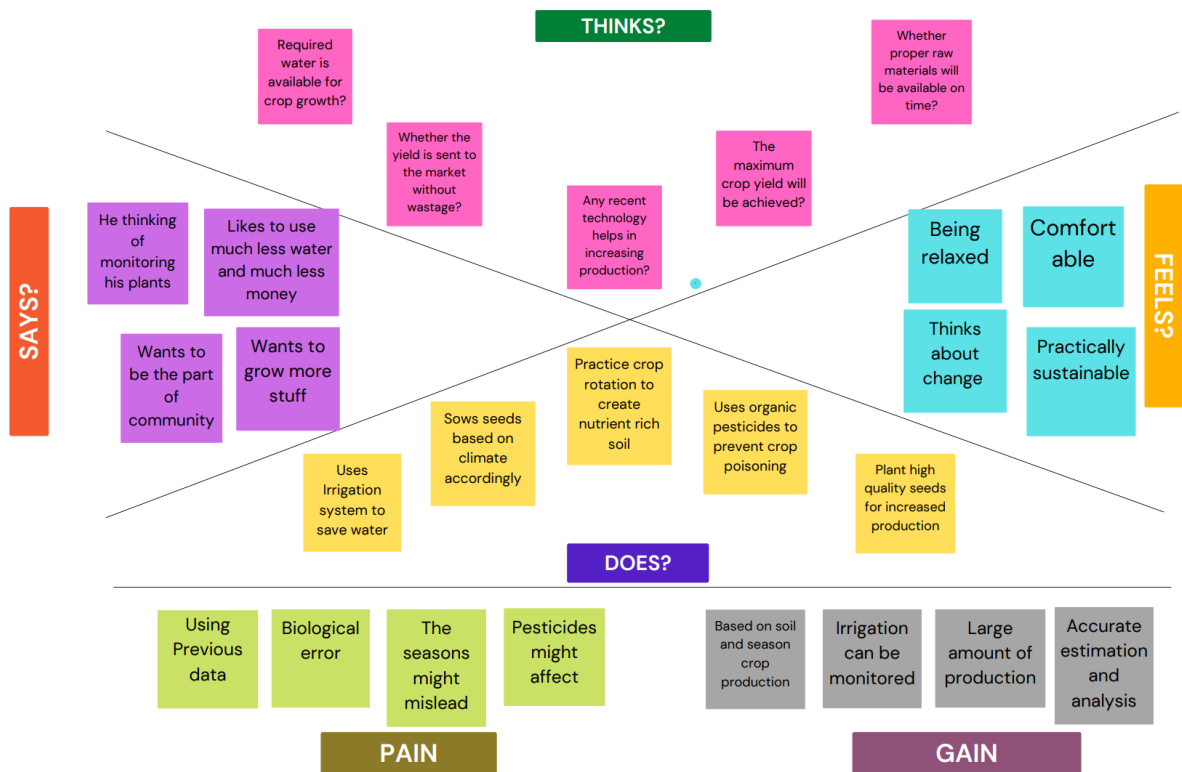
2.3 Problem Statement Definition:

This work is to construct a model for testing the soil fertility. It also suggests the crop which has to be planted depending upon the value obtained from the sensor. It also provides the regional wise information about the crop in the form of graph. We have farmer chat where the farmers can share and get idea from the expert by registering in this application. It also suggests the fertilizer which has to be added to the soil in order to increase the crop productivity. It helps the farmer to analyse the fertility of their yard and plant the better crop to increase their productivity and profit. It also provides the information about the fertilizer to be added in the soil and also provide the information about the nearby fertilizer shop.It's high time we take a call and analyse the farmer needs and increase or decrease the production based on seasons. Hence with the help of Cognos Analytics, we have built a dashboard where there are 5 tabs, each representing a unique relationship that needs to be solved. We have used Data Player to highlight the top values which would be more appetizing for the users. This tool predicts the crop yield based on the parameters like Rainfall, temperature, pesticides, etc. This helps in predicting the yield and increasing production.

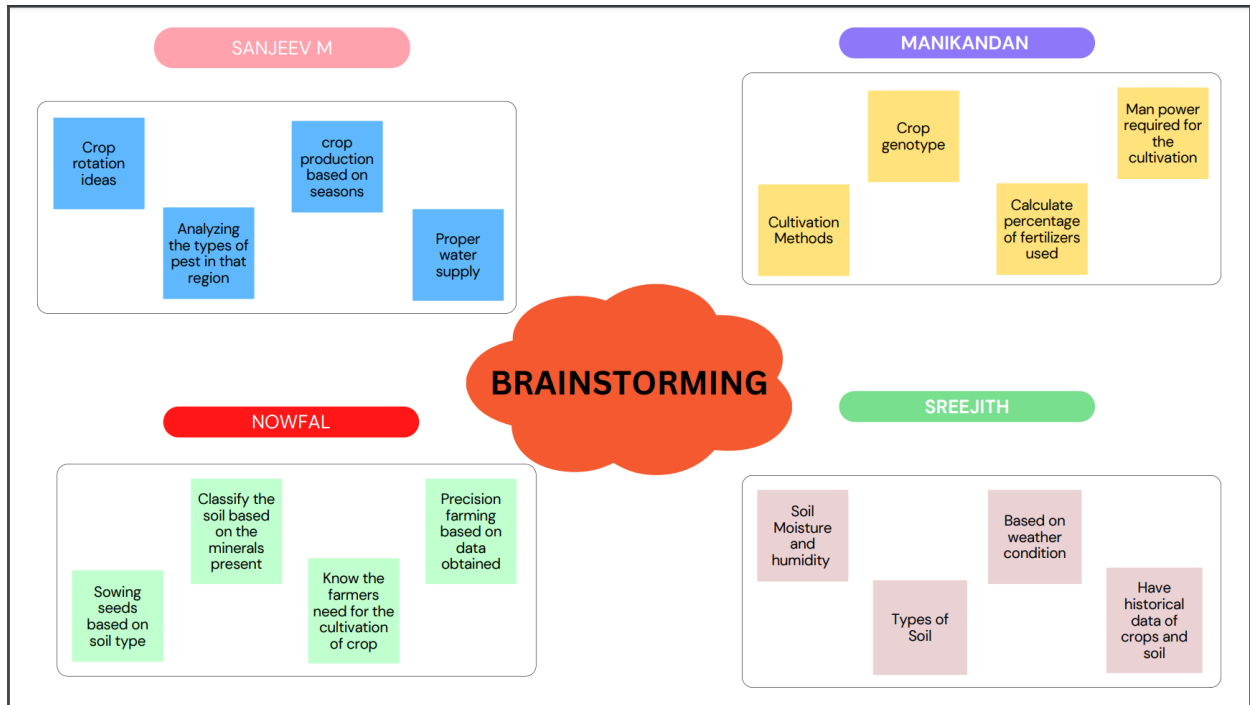## 3.IDEATION AND PROPOSED SOLUTION:
3.1 Empathy Map Canvas:

Estimation crop yield

Empathy Map

THINKS?

- Required water is available for crop growth?
- Whether the yield is sent to the market without wastage?
- Any recent technology helps in increasing production?
- The maximum crop yield will be achieved?
- Whether proper raw materials will be available on time?

SAYS?

- He thinking of monitoring his plants
- Likes to use much less water and much less money
- Wants to be the part of community
- Wants to grow more stuff

DOES?

- Uses Irrigation system to save water
- Sows seeds based on climate accordingly
- Practice crop rotation to create nutrient rich soil
- Uses organic pesticides to prevent crop poisoning
- Plant high quality seeds for increased production

FEELS?

- Being relaxed
- Comfort able
- Thinks about change
- Practically sustainable

PAIN

- Using Previous data
- Biological error
- The seasons might mislead
- Pesticides might affect

GAIN

- Based on soil and season crop production
- Irrigation can be monitored
- Large amount of production
- Accurate estimation and analysis

3.2 Ideation and Brainstorming:

Ideation

Team Members	Idea-1	Idea-2	Idea-3	Idea-4
Sanjeev M	Increase in the overall production	Controlled water consumption	Analysing seeds based on the season	Soil type determination
Mohamed Nowfal S	Collect data	Reduce the wastage	Clean up data for faster accessing	Increased accuracy
Manikandan P	Analyse the gathered data	Understand the needs and came up with a project map	Enrich data	User friendly dashboard
Sreejith G	Using pictorial representation for accurate and easy understanding	Came up with new ideas	Using AI based assistant for easy guide	Increasing the productivity

SANJEEV M

- Crop rotation ideas
- Analyzing the types of pest in that region
- crop production based on seasons
- Proper water supply

MANIKANDAN

- Crop genotype
- Man power required for the cultivation
- Cultivation Methods
- Calculate percentage of fertilizers used

BRAINSTORMING

NOWFAL

- Classify the soil based on the minerals present
- Precision farming based on data obtained
- Sowing seeds based on soil type
- Know the farmers need for the cultivation of crop

SREEJITH

- Soil Moisture and humidity
- Based on weather condition
- Types of Soil
- Have historical data of crops and soil

3.3 PROPOSED SOLUTION:

S.No.	Parameter	Description
1.	Problem Statement (Problem to be solved)	The main problem to be solved using Data analytics is to predict the crop yield primarily based on certain factors which is very useful for farmers.
2.	Idea / Solution description	Analysing past agricultural data related to the crop yields and provide perfect data report of analyse. Helping them out to overcome loss.
3.	Novelty / Uniqueness	This analysed data can be very much helpful for the farmers. Visualize the data and gives farmers to chose which crop to cultivate at certain season/period. This increases the profit and helps us understand why certain crop costs more in one year than the pervious years.
4.	Social Impact / Customer Satisfaction	Analysed data being visualised clearly shows anaytics done on agricluture soil,fertilizers,etc., thus helps in increased yield of crops and gain more profits to the farmers.
5.	Business Model (Revenue Model)	It heps in monitoring of crops in real time, create large number of crop production and other raw materials, predictions analysis of future yields. Increase in natural and organic foods from this benefits farmers in a large scale.
6.	Scalability of the Solution	Based on the data analytics and reports, we can cultivate crop according to the favourable factors like type of soil,climate,different states,etc., thus providing accurate crop yield.

3.4 PROBLEM SOLUTION FIT:

Define CS, fit into CC

Focus on J&P, tap into BE, understand RC

1. CUSTOMER SEGMENT(S) `CS`
Who is your customer?
i.e. working parents of 0-5 y.o. kids

The farmers, who wants to yield the crops on the field .

6. CUSTOMER CONSTRAINTS `CC`
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.

Production constraints have been identified that contribute to explaining the yield gap, i.e. limited water availability, limited nutrient availability, inadequate crop protection

5. AVAILABLE SOLUTIONS `AS`
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

Customers want a great experience, and part of that experience means not losing time, money, and patience. Brands similarly value their time and money, but they must always have the patience to deliver their customers a great experience. Even with the best efforts, however, difficult customer situations are sure to arise, and how they handle these situations can mean the difference between customer churn and long-term loyalty.

Explore AS, differentiate

2. JOBS-TO-BE-DONE / PROBLEMS `J&P`
Which jobs-to-be-done (or problems) do you address for your customers? There could be more than one; explore different sides.

Soil Samples.

As-Applied (planting, nutrient and crop protection) Data.

Harvest Files.

Input Cost Data.

9. PROBLEM ROOT CAUSE `RC`
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.

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.

7. BEHAVIOUR `BE`
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)

Providing the correct data and input with algorithms for a solution

Focus on J&P, tap into BE, understand RC

Identify strong TR & EM

3. TRIGGERS `TR`
What triggers customers to act? i.e. seeing their neighbour installing solar panels, reading about a more efficient solution in the news.

Using high yielding variety seeds. Use of modern irrigation methods so as to obtain more amount of water. Crop rotation so as to increase the fertility of the soil.

4. EMOTIONS: BEFORE / AFTER `EM`
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.

before
    the famer has stress based on production
after
    after the good crop production farmer will be happy

10. YOUR SOLUTION `SL`
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.

monitoring the health of crops in real-time, **create predictive analytics related to future yields** and help farmers make resource management decisions based on proven trends. Reducing waste and improving profits.

8. CHANNELS of BEHAVIOUR `CH`
8.1 ONLINE
What kind of actions do customers take online? Extract online channels from #7

8.2 OFFLINE
What kind of actions do customers take offline? Extract offline channels from #7 and use them for customer development.

online
monitoring the analysis of crop yield
offline
gathering data for outcomes

4. REQUIREMENT ANALYSIS:
4.1 Functional Requirement:

Functional Requirements:

Following are the functional requirements of the proposed solution.

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	User profile	User need to log in Access the profile
FR-4	Take in the required data	The user will give a required date take that data
FR-5	Estimation	Analyse the yield of crop from the data given by the user
FR-6	Analysis	An analysis is done on the given data to gain useful insights

4.2 Non Functional Requirement:

Non-functional Requirements:

Following are the non-functional requirements of the proposed solution.

| FR No.	Non-Functional Requirement	Description
NFR-1	Usability	Crop recommendations are created and saved, the these recommended crops are sown by farmers for increased crop yield.
NFR-2	Security	This software will secure the information .
NFR-3	Reliability	The interactive dashboards which is easy to understand and useful for the users
NFR-4	Performance	The software provide us good performance and user friendly.
NFR-5	Availability	This software is available for access at any time from anywhere
NFR-6	Scalability	The software should be flexible and other developers must be able to improve its capabilities

5. PROJECT DESIGN:

5.1 Data Flow Diagrams:

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.

Example: (Simplified)

Example: DFD Level 0 (Industry Standard)
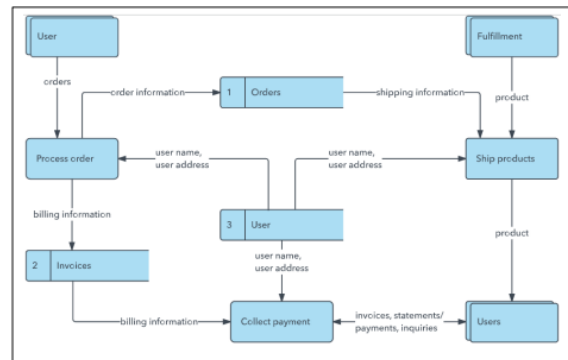
Flow

1. User configures credentials for the Watson Natural Language Understanding service and starts the app.
2. User selects data file to process and load.
3. Apache Tika extracts text from the data file.
4. Extracted text is passed to Watson NLU for enrichment.
5. Enriched data is visualized in the UI using the D3.js library.

5.2 Solution and Technical Architecture:

Solution Architecture:

Solution architecture is a complex process – with many sub-processes – that bridges the gap between business problems and technology solutions. Its goals are to:

- Find the best tech solution to solve existing business problems.
- Describe the structure, characteristics, behavior, and other aspects of the software to project stakeholders.
- Define features, development phases, and solution requirements.
- Provide specifications according to which the solution is defined, managed, and delivered.

Solution Architecture Diagram:
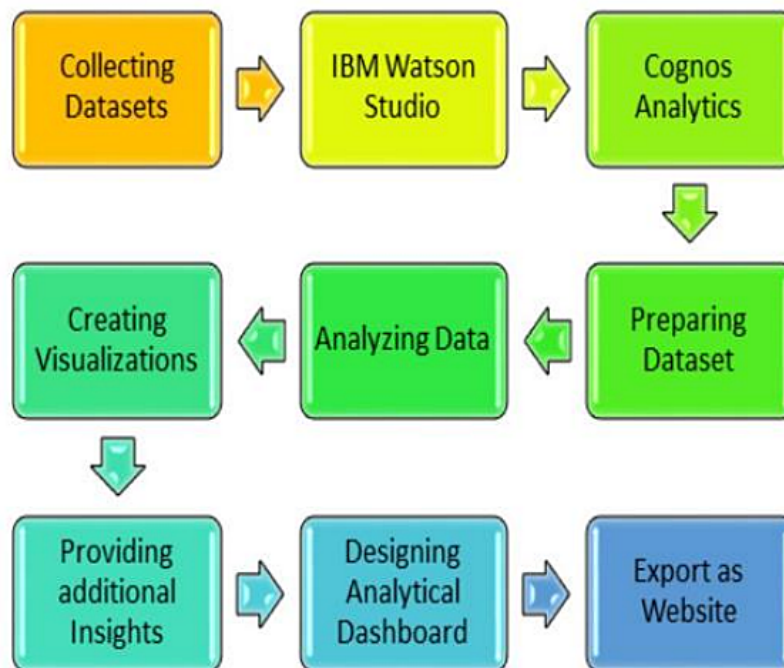
Technical architecture
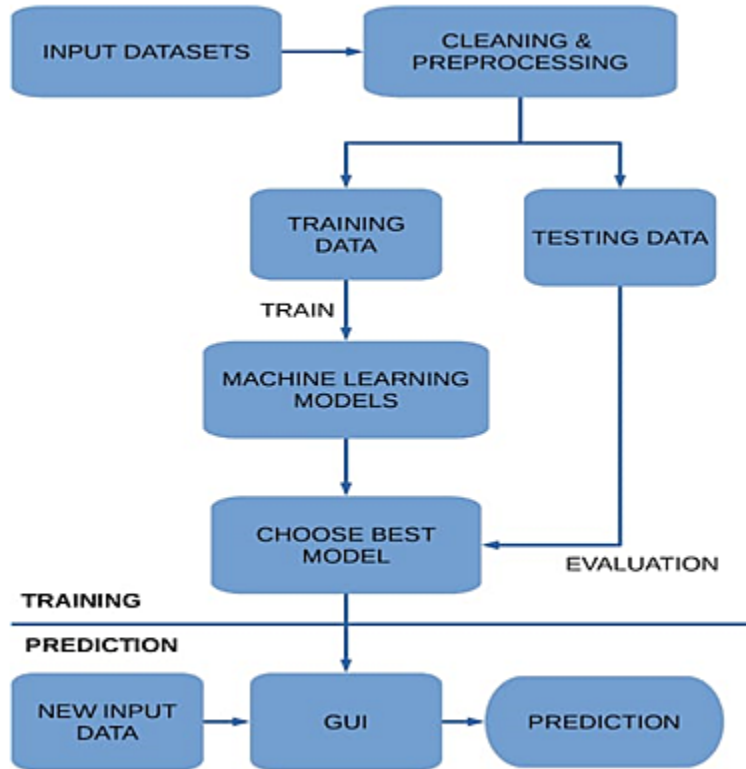
5.3 USER STORIES:

User Stories

Use the below template to list all the user stories for the product.

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	I can able to go through the cart.	Medium	Sprint-1

User Type	Functional Requirement (Epic)	User Story Number	User Story / Task	Acceptance criteria	Priority	Release
	Login	USN-5	As a user, I can log into the application by entering email & password	I can reset my password .if I forgot my password	High	Sprint-1
	My account	USN-6	User can view my personal information and data.	I can edit my profile photo and email and I can logout from the application.	High	Sprint -1
Customer (Web user)	Registration	USN-7	As a user can register the application by using email, password, and confiming my password	In account I can upload profile photo and add my name.	Medium	Sprint - 1
Customer Care Executive	Communication	USN-8	As a user, I can communicate with the customers and provide support system for companies.	I can handle a good communication between customer and client. If customer have queries we will sort it out.	High	Sprint -1
Administrator	Chief Executive	USN -9	As an administrator ,I can modify the list of products .so I can adjust our offerings over time	Modify the product. Add or remove products. select a category for the products .modify category taxonomy	High	Sprint-1
Estimator	Estimation	USN-10	As a user ,i can see all the items we will try to estimate this session	I have a feel for the size of the various items in the product based	Medium	Sprint-1
Moderator	Mediator	USN-10	As a user, i can invite estimators by giving them a url where they can access the product	I can answer the question about the current product	High	Sprint -1

6. Project Planning:

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.	5	High	Manikandan P Mohamed Nowfal S
		USN-2	As a user, I will receive confirmation email once I have registered for the application	4	Medium	Sreejith G
		USN -3	Creating a website for the login application	4	Medium	Sanjeev M
		USN – 4	Upload the dataset into the cognos analytics	2	Low	Sreejith G
		USN – 5	Create a new dashboard	5	High	Manikandan P
Sprint-2	Data Visualization Chart	USN-6	Build a Visualization to showcase Average Crop Production by Seasons.	5	Medium	Mohamed Nowfal S Sreejith G
			Showcase the Yearly usage of Area in Crop Production	4	Medium	Sanjeev M Manikandan P
			Top 10 States in Crop Yield Production by Area.	4	Medium	Sreejith G
			Crop Production by State.	4	Medium	Mohamed Nowfal S

| Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
			States with Seasonal Crop Production using a Text representation.	3	Medium	Manikandan P Sreejith G
Sprint-3	Creating The dashboard	USN-7	Create the Dashboard by using the created visualizations.	20	High	Sanjeev M Manikandan P Mohamed Nowfal S Sreejith G
Sprint-4	Export The Analytics	USN-8	Finally export the created dashboard	20	High	Mohamed Nowfal S Sanjeev M

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	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

9.RESULTS :

The final dashboard created with 5 tabs are attached below:
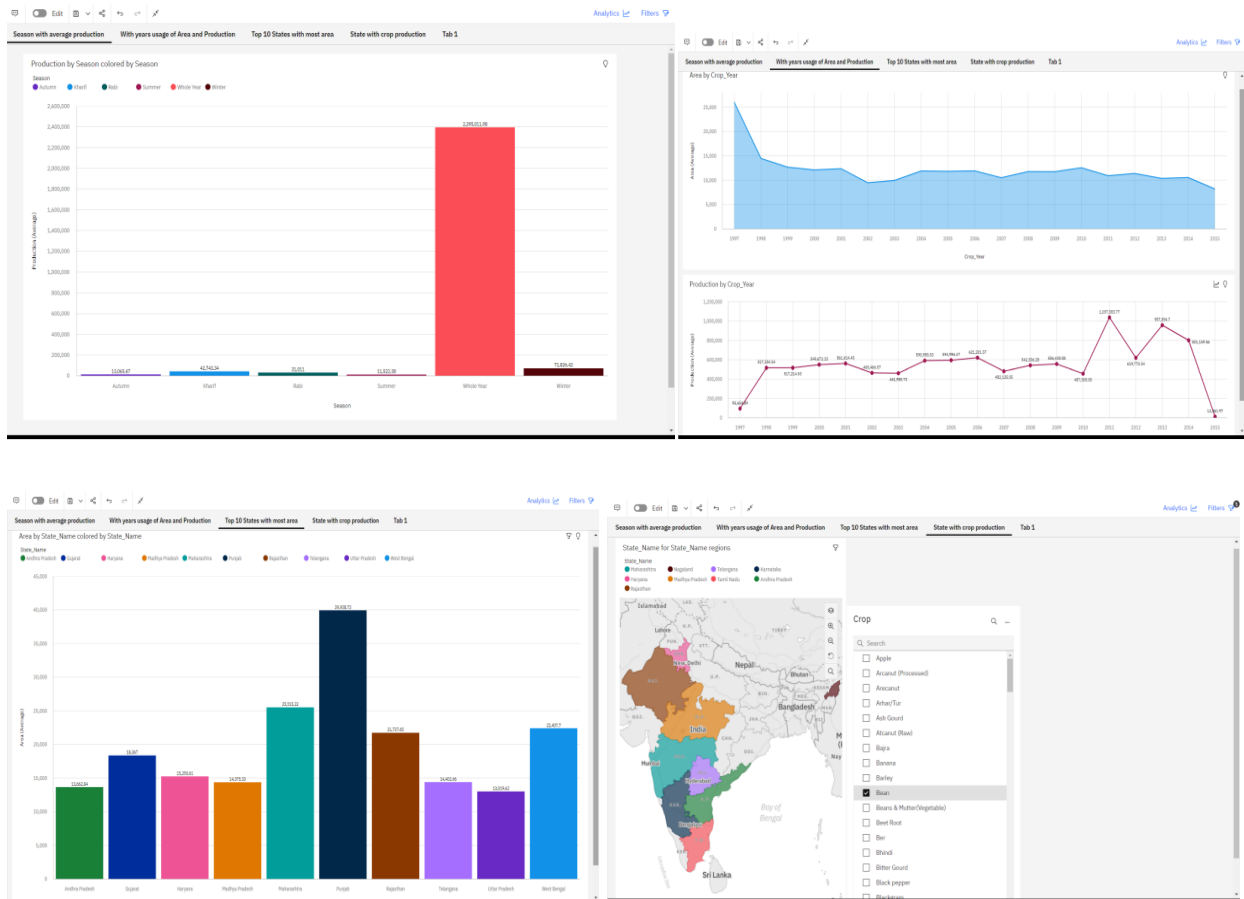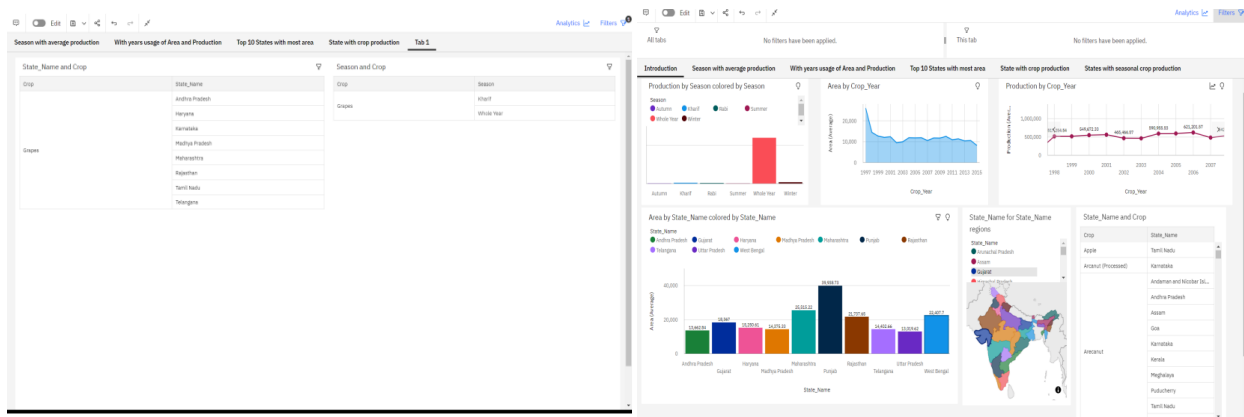
10. ADVANTAGES AND DISADVANTAGES:

ADVANTAGES:
- ➤ Data analytics is a critical part of improving business operations in every industry.
- ➤ An organization can utilize data analytics to improve decision-making, analyze customer trends, track customer satisfaction and identify opportunities for new products and services to meet growing market needs.
- ➤ Data analytics can help farmers monitor the health of crops in real-time, create predictive analytics related to future yields and help farmers make resource management decisions based on proven trends.

DISADVANTAGES:
- ➤ There might be data redundancy and the results might not be accurate.
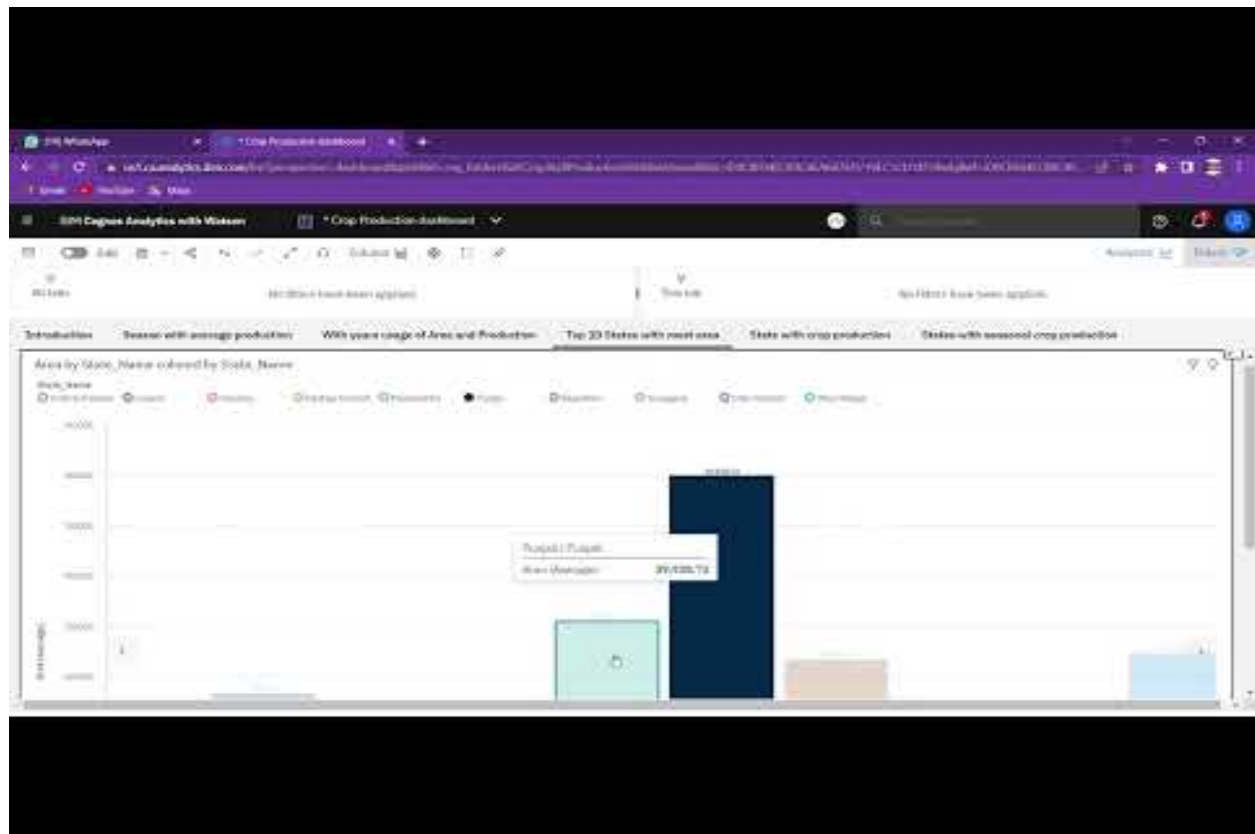- ➤ The seasons might mislead and the prediction goes wrong.

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.In the literature,it has been observed that analysis has been done on agriculture soils,hidden patterns discovery using data set related to climatic conditions and crop yields data.The activities of agriculture field are numerous like weather forecasting, soil quality assessment, seeds selection, crop yield prediction etc. In this survey, the specific activity, crop yield prediction has been surveyed and the major trends have been identified. It can be concluded that the research in the field of agriculture with reference to using IT trends like data analytics is in its infancy. As the food is the basic need of humans, the requirement of getting the maximum yields using optimal resource will become the necessity in near future as a result of growing population. The survey outcomes indicate the need for improved techniques in crop yield analytics.There exists a lot of research scope in this research area.

12.FUTURE SCOPE:

Crop production in India is one of the most important sources of income and India is one of the top countries to produce crops.Our project mainly focuses on getting dominant insights from the data that helps the farmers to choose which crop to use wisely. The operating sector such as Collecting datasets, Analyzing the data, and Providing relationships and Insights. This process makes to he farmers to predict in which season what crop to use and helps to forecast future sales strategically. 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. In coming decades, two most significant and important factors found to influence crop yield is, increase in the global population and economy, which greatly demands the higher and sustainable agricultural based crop yields. The capacities of food production at global level is going to be very limited due to the less availability of cultivable land, water resources, difficulties in maintaining the sustainable crop production levels, effects of changes in the global climatic conditions and also by various biophysical parameters which influence the crop yield. The farmers need to be educated on the application of scientifically proven methods to quantify the crop yield capacities and same need to be informed to higher authorities to maintain transparency in sharing the actual information, intern helps in making the policy based, research oriented, development and investment related decisions that aim to influence future crop yield.

DEMO LINK

Git repo:

https://github.com/IBM-EPBL/IBM-Project-50320-1660903095