Project Report

1. INTRODUCTION

1.1 Project Overview

Crop production in India is one of the most important sources of income and India is one of the top countries to produce crops. As per 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.

1.2 Purpose :

To reduce the farmer loss , control them . Increase the Crop yield growth by past data sets . The chief aim of data analytics is to apply statistical analysis and technologies on data to find trends and solve problems. Data analytics has become increasingly important in the enterprise as a means for analyzing and shaping business processes and improving decision-making and business results.

2. LITERATURE SURVEY

2.1 Existing problem :

Loss in Crop Yield even after many new ideologies .

2.2 References

Camps-Valls G, Gomez-Chova L, Calpe-Maravilla J, Soria-Olivas E, Martin-Guerrero J D, Moreno J, "Support Vector Machines for Crop Classification using Hyper Spectral Data", Lect Notes Comp Sci 2652, 2003, pages : 134-141.

Araby, A. A., Abd Elhameed, M. M., Magdy, N. M., Abdelaal, N., Abd Allah, Y. T., Darweesh, M. S., ... & Mostafa, H. (2019, May). Smart IoT Monitoring System for Agriculture with Predictive Analysis. In 2019 8th International Conference on Modern Circuits and Systems Technologies (MOCAST)

S.Veenadhari, Dr Bharat Misra, Dr CD Singh.2019."Machine learning approach for forecasting crop yield based on climatic parameters.". International Conference on Computer Communication and Informatics

Brownlee, J. (2016). K Nearest Neighbors for Machine Learning. Retrieved March 23, 2020, from https://machinelearningmastery.com. Chauhan, N. S. (2019). Introduction to Artificial Neural Networks (ANN).

Retrieved March 26, 2020, from https://towardsdatascience.com. Donges, N. (2019). A Complete Guide to the Random Forest Algorithm. Retrieved March 30, 2020, from https://builtin.com/data-science/random-forest_algorithm.html. Gandhi, R. (2018).

Naïve Bayes Classifier. Retrieved March 25, 2020, from https://towardsdatascience.com. Hardesty. L. (2017). Explained: Neural Networks. Retrieved March 26, 2020, from https://news.mit.edu/2017/explained_neural-networks-deep-learning-0414. Harrison, O. (2018).

Machine Learning Basics with the K_Nearest Neighbors Algorithm. Retrieved March 23, 2020, from https://towardsdatascience.com.

Jayalakshmi, R. & Devi, M. S. (2019). Relevance of Machine Learning Algorithms on Soil Fertility Prediction using R. International Journal of Computational Intelligence and Informatics, 8(4), 193-199. Martin, K. Logistic Regression Models for Multinomial and Ordinal Variables. Retrieved March 27, 2020, from https://www.theanalysisfactor.com/logistic-regression_models-for - multinomial-and-ordinal-variables.html

Nagesh, S. (2019). Real world implementation of Logistic Regression. Retrieved March 27, 2020, from http://towardsdatascience.com. Priya, P., Muthaiah, U. & Balamurugan, M. (2018). Predicting Yield of the Crop Using Machine Learning Algorithm. International Journal of Engine

Athmaja S., Hanumanthappa M, "Applications of Mobile Cloud Computing and Big data Analytics in Agriculture Sector: A survey", October 2016
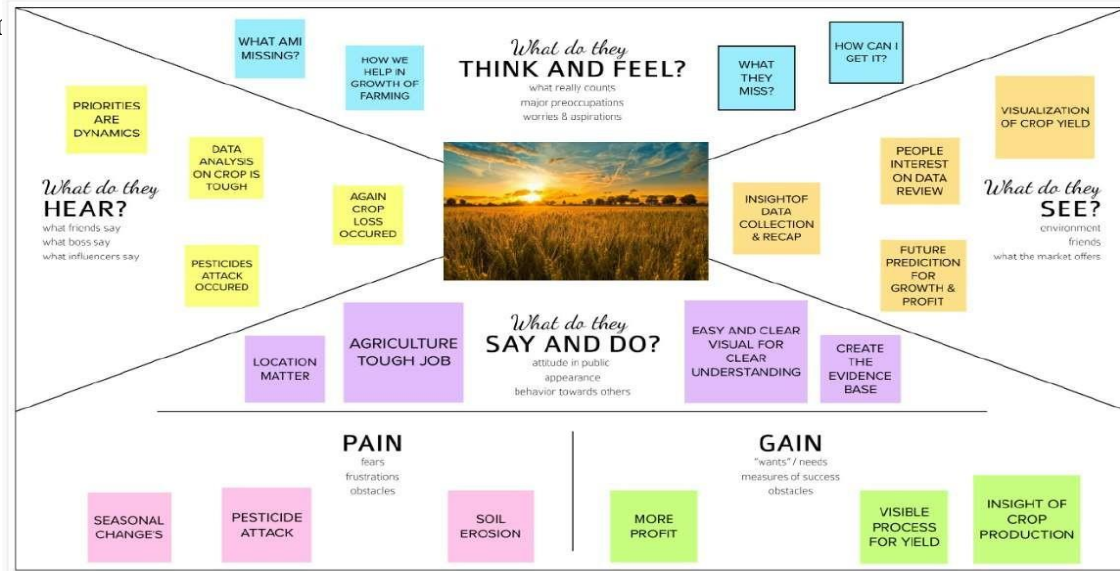
2.3 Problem Statement Definition :

Empathy Map Canvas

Gain insight and understanding on solving customer problems.

3. **IDEA**

Build empathy and keep your focus on the user by putting yourself in their shoes.

3.1 Em

WHAT AM I MISSING?

HOW WE HELP IN GROWTH OF FARMING

What do they
THINK AND FEEL?
what really counts
major preoccupations
worries & aspirations

WHAT THEY MISS?

HOW CAN I GET IT?

PRIORITIES ARE DYNAMICS

VISUALIZATION OF CROP YIELD

DATA ANALYSIS ON CROP IS TOUGH

PEOPLE INTEREST ON DATA REVIEW

What do they
HEAR?
what friends say
what boss say
what influencers say

AGAIN CROP LOSS OCCURED

INSIGHT OF DATA COLLECTION & RECAP

What do they
SEE?
environment
friends
what the market offers

FUTURE PREDICITION FOR GROWTH & PROFIT

PESTICIDES ATTACK OCCURED

LOCATION MATTER

AGRICULTURE TOUGH JOB

What do they
SAY AND DO?
attitude in public
appearance
behavior towards others

EASY AND CLEAR VISUAL FOR CLEAR UNDERSTANDING

CREATE THE EVIDENCE BASE

PAIN
fears
frustrations
obstacles

GAIN
"wants" / needs
measures of success
obstacles

SEASONAL CHANGE'S

PESTICIDE ATTACK

SOIL EROSION

MORE PROFIT

VISIBLE PROCESS FOR YIELD

INSIGHT OF CROP PRODUCTION

3.2 Ideation & Brainstorming

Step-1: Team Gathering, Collaboration and Select the Problem Statement:

Brainstorm & idea prioritization

Use this template in your own brainstorming sessions so your team can unleash their imagination and start shaping concepts even if you're not sitting in the same room.

- 🕐 **10 minutes** to prepare
- ⏳ **1 hour** to collaborate
- 👤 **2-8 people** recommended

💬 Share template feedback

Before you collaborate

A little bit of preparation goes a long way with this session. Here's what you need to do to get going.

🕐 **10 minutes**

A **Team gathering**
Define who should participate in the session and send an invite. Share relevant information or pre-work ahead.

B **Set the goal**
Think about the problem you'll be focusing on solving in the brainstorming session.

C **Learn how to use the facilitation tools**
Use the Facilitation Superpowers to run a happy and productive session.

Open article →

---

1 Define your problem statement

The Indian farmers are the very effective and efficient farmers in these world. But they don't have data report of their crop yield and whole Indian farmer's yield too. According to climate, soil, crop growth & Accurate time to provide the pesticide for crop and these are things which make loses in crop yield. which we are going to provide by our data analytics report (By Visuals).

PROBLEM
How might we [your problem statement]?

Key rules of brainstorming
To run an smooth and productive session

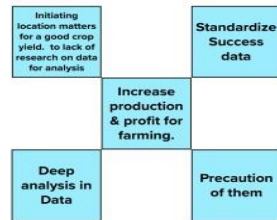|	
Stay in topic.	Encourage wild ideas.
Defer judgment.	Listen to others.
Go for volume.	If possible, be visual.

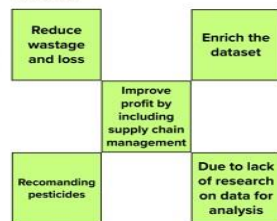# Step-2: Brainstorm, Idea Listing and Grouping:

2

Brainstorm

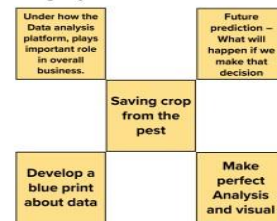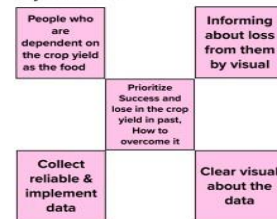**Following ideas are came to our mind that address to our problem statement :**

Gunal K

Initiating location matters for a good crop yield. to lack of research on data for analysis	Standardize Success data
	Increase production & profit for farming.
Deep analysis in Data	Precaution of them

Rangaraj P

Under how the Data analysis platform, plays important role in overall business.	Future prediction – What will happen if we make that decision
	Saving crop from the pest
Develop a blue print about data	Make perfect Analysis and visual

Harish C

Reduce wastage and loss	Enrich the dataset
	Improve profit by including supply chain management
Recomanding pesticides	Due to lack of research on data for analysis

JayaPrasana T

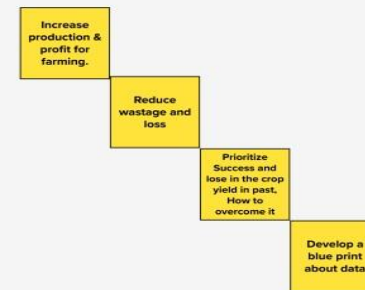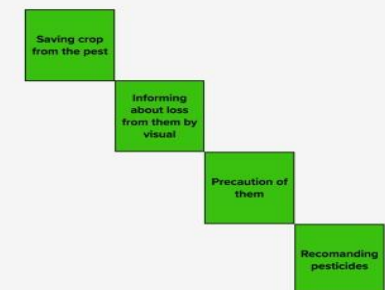|	
People who are dependent on the crop yield as the food	Informing about loss from them by visual
	Prioritize Success and lose in the crop yield in past, How to overcome it
Collect reliable & implement data	Clear visual about the data

3

Group ideas

Take turns sharing your ideas while clustering similar or related notes as you go.
In the last 10 minutes, give each cluster a sentence-like label. If a cluster is bigger
than six sticky notes, try and see if you and break it up into smaller sub-groups.
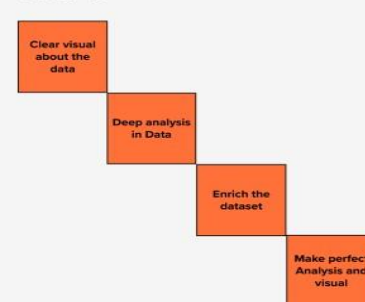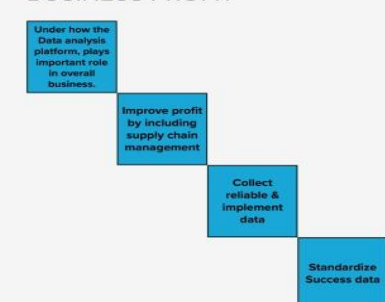
🕐 20 minutes

CROP YIELD DEVELOPMENT

- Increase production & profit for farming.
- Reduce wastage and loss
- Prioritize Success and lose in the crop yield in past, How to overcome it
- Develop a blue print about data

PESTICIDES

- Saving crop from the pest
- Informing about loss from them by visual
- Precaution of them
- Recomanding pesticides

VISUAL

- Clear visual about the data
- Deep analysis in Data
- Enrich the dataset
- Make perfect Analysis and visual

BUSINESS PROFIT

- Under how the Data analysis platform, plays important role in overall business.
- Improve profit by including supply chain management
- Collect reliable & implement data
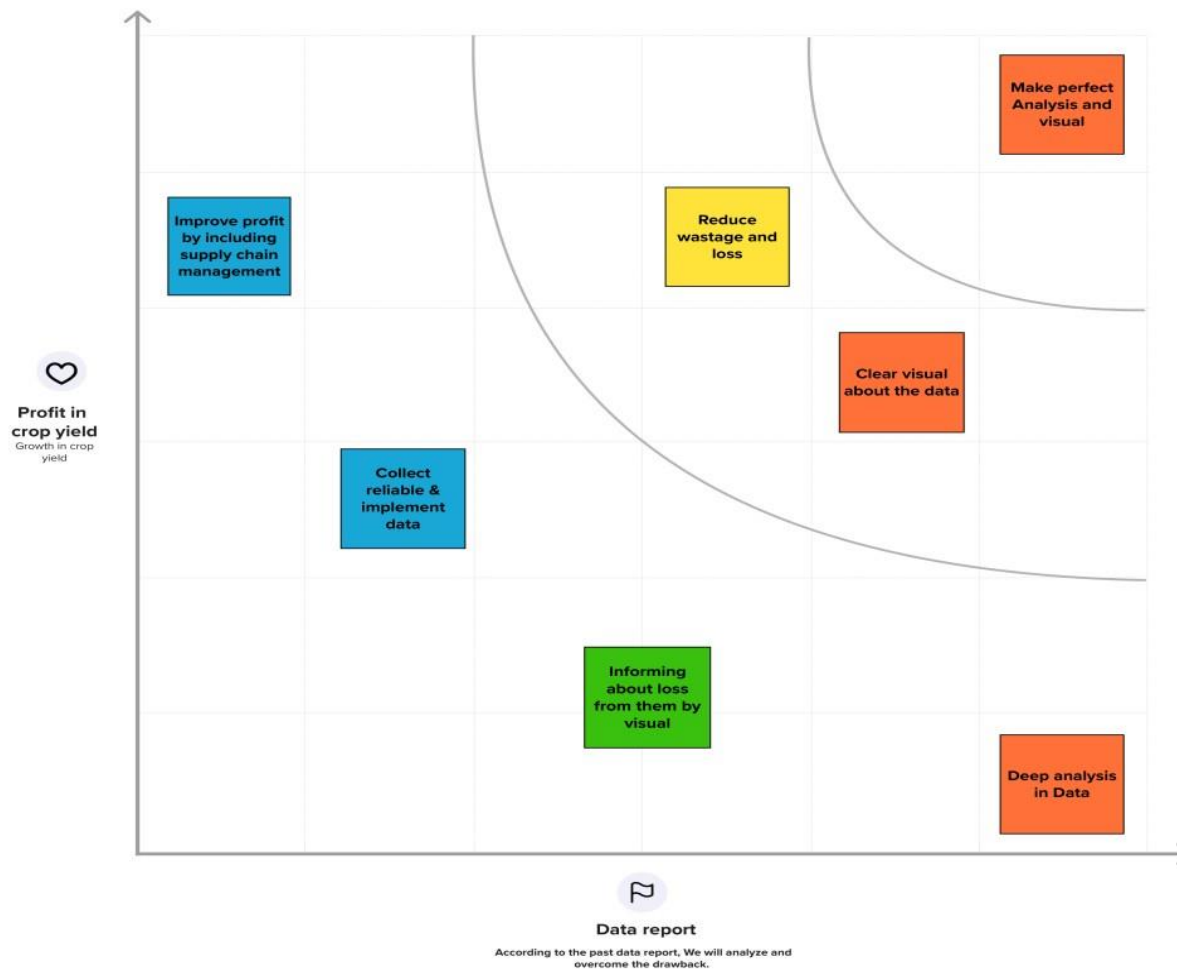- Standardize Success data

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**

Profit in crop yield
Growth in crop yield

Make perfect Analysis and visual

Improve profit by including supply chain management

Reduce wastage and loss

Clear visual about the data

Collect reliable & implement data

Informing about loss from them by visual

Deep analysis in Data

Data report
According to the past data report, We will analyze and overcome the drawback.

After you collaborate

You can export the mural as an image or pdf to share with members of your company who might find it helpful.

3.3 Proposed Solution:

S. No.	Parameter	Description
1.	Problem Statement (Problem to be solved)	Loss in crop yield, Loss of appetite , Increase of Poverty , Decrease depression and Suicide farmer .
2.	Idea / Solution description	Provide perfect data report after deep analyse of past data. Helping them out to overcome loss in farming and business.
3.	Novelty / Uniqueness	According state, district , Climate (Season) and Area and others .We will analytics report and give tips to plant which is comfortable and profitable for them to make profit in the crop yield

4.	Social Impact / Customer Satisfaction	It can solve almost all the problems statement which we analyse. By providing the perfect datavisuals it can create a large impact in crop yield and profit of the farmers.
5.	Business Model (Revenue Model)	We can create large number crop production and other raw materials too. We can sale these data report to the industry which need raw materials continuously for their factories. If the outcome is profited According to the Farmer wealth we cost for the model. we can approachgovernment too buy and share the farmer to get wealthy crop yield .
6.	Scalability of the Solution	With the data visual reports we can cultivate crop according to the Crop, State, district , climate , soil can change the estimation of Cropyield

3.4 Problem Solution fit:

Project Title: Estimate the Crop yield using data analytics **Project Design Phase-I – Solution Fit Template** **Team ID: PNT2022TMID43579**

Define CS, fit into CC | **Explore AS, differentiate**

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

The firmer are our customer. Who work in fields and yield crop

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.

Less knowledge and development towards the current environmental changes and technologies, they follow ancient methods, which is also worthy but, the climatic changes and new kind pesticides attack make more losses

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 is an alternative to digital notetaking

They have the dataset. But they don't have the perfect data report , which can help them overcome there problem

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

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.

Data report should to be created to reduce the loss of the crop and earn more profit in agriculture fields

Loss in agriculture, less amount of crop yield

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.

The problem exist because of the climatic changes, soil condition because of continues cropping, unknown form pest attacks pre-caution them.

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)

Some farmers do get advice from nearby Agriculture welfare office, many of them don't get it properly.

According they cultivate there crops.

Many of the farmers don't have much knowledge to take help from online resources.

They try to contact offline agriculture office itself

Identify strong TR & EM | **Identify strong TR & EM**

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.

Many farmers in India have commited to suicide because of the loss they faced in agriculture

The solution for the problem, creating data report using past datasets, in a understandable way using IBM cognos dashboard (visuals) could make them understand easily

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 Whatsapp Registration through Agri-Consultancy
FR-2	User Confirmation	Confirmation via Email Confirmation via OTP Confirmation via Letter
FR-3	User Profile	User Details Farm Details
FR-4	Required Data	The past crop yield data the user(Farmer) data to analyse

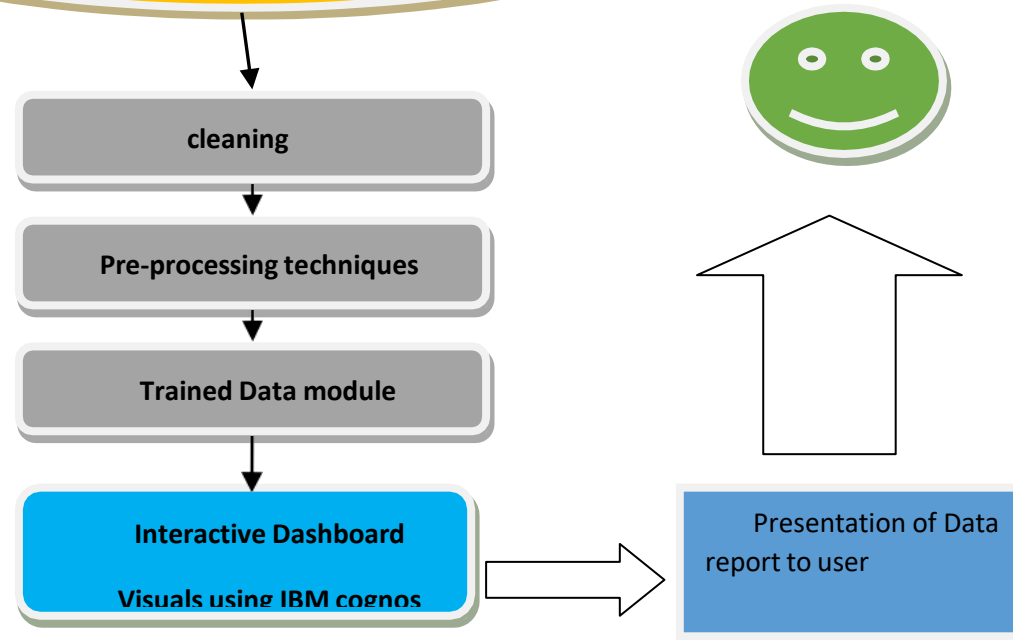| FR-5	Analysis	Clean and analyse the data according to the set of past data of the multiple users(Farmer)
FR-6	Estimation	Creating the perfect data module, visuals using IBM Cognos to increase the estimation of the crop yield

4.2 Non-Functional requirements:

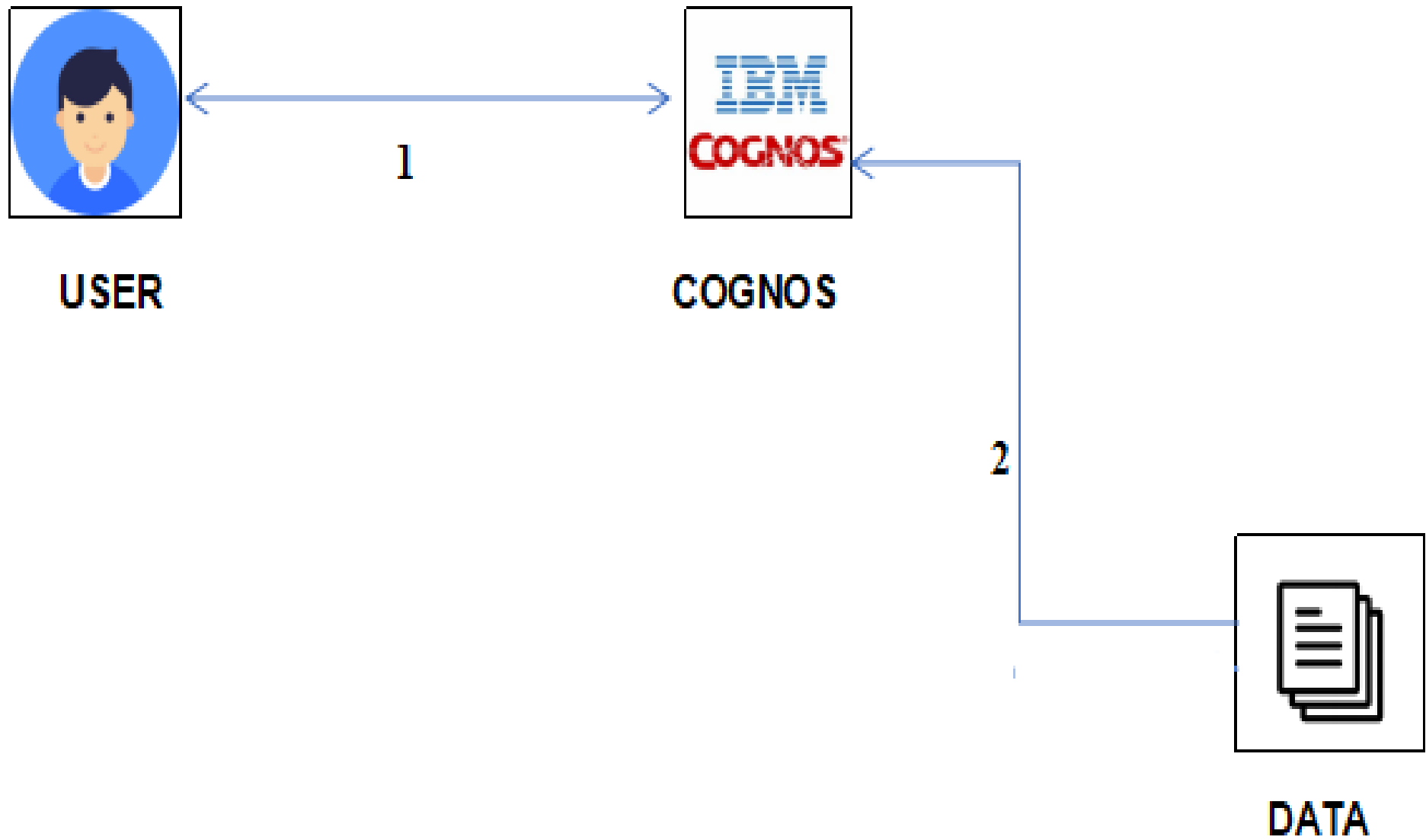| FR No.	Non-Functional Requirement	Description
NFR-1	**Usability**	According to the past data itself, data report is created.By these recommendation the sowing of crops will be advised or consulted
NFR-2	**Security**	IBM Cognos have a secure user information(Data Visuals)
NFR-3	**Reliability**	The interactive data visuals dashboard can make easily understandable of the data report

NFR-4	**Performance**	Interaction makes better performance between all user and impressing by the visual advise
NFR-5	**Availability**	The dashboard could be easily viewed and ...lable ...ps, ems e...
NFR-6	**Scalability**	The flexibility of the methodology to implement the proposed solution is very easy that can make increase in the estimation of crop yield in the differ farms for different user

Farmer(user1)

Farmer(user2)

Farmer(user3)

Farmer(user4)

User stories , Query and past

Data of estimated crop yield

5. PROJECT DESI...

5.1 Data Flow Diagrams:

cleaning

Pre-processing techniques

Trained Data module

Interactive Dashboard

Visuals using IBM cognos

Presentation of Data report to user

5.2 Solution & Technical Architecture

USER

COGNOS

DATA

1

2

5.3 User Stories:

User Type	Functional Requirement (Epic)	User Story Number	User Story / Task	Acceptance criteria	Priority	Release
Customer (Farmer)	Registration	USN-1	As a user, I can register for the application by entering my email, password, and confirmingmy password.	I can access my account / dashboard	High	Sprint-1
		USN-2	As a user, I will receive confirmation email onceI 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 applicationthrough Whatsapp , Facebook	I can register & access the dashboard with Whatsapp , Facebook Login	Low	Sprint-1
	Required Data	-	Cropping history , profit and loss in theirfarming	Past dataset of cropping and field estimation of cropyield	High	Sprint-2
	Analysis		Clean and analyse to data according to the setpast data		HIgh	Sprint-3

			As a user, i can provide support systems for companies that often communicate with the customers	I can maintain strong relationships with customer and client ,so I can ease their queries and increase productivity	medium	Sprint-4
Customer Care Executive	Customer Care Executive (Commun ication)					
Estimator	Estimation		y toestimate this session	I have a feel for the size ofthe various items in the product based	Medium	Sprint-4

6. PROJECT PLANNING & SCHEDULING

6.1 Sprint Planning & Estimation:

	Functional Requireme nt(Epic)	User Story Numb er	User Story / Task	Stor y Poin ts	Prior ity	Team Members
-1	Registration	USN-1	As a user, I can register for by entering my Agri - id card and request..	2	High	Rangaraj P
		USN-3	As a user, I can register for the application through Gmail	2	Medi um	Jayaprasana T

Sprint							
	Logi n	USN-4	As a user, I can Call and request or Approach for dataset	2	High	Guanl K
	Working with the Dataset	USN-5	To work on the given dataset, Understand the Dataset.	2	High	Gunal K Harish C
		USN-6	Load the dataset to Cloud platform then Build the required Visualizations.	10	High	Gunal K Harish C
Sprint -2	Data Visualization Chart	USN-7	Using the Crop production in Indian dataset, create various graphsand charts to highlight the insights and visualizations. *Build a Visualization to showcase Average Crop Production by Seasons.	4	Medi um	Gunal K Harish C
			*Showcase the Yearly usage of Area in Crop Production.	4	Medi um	Rangaraj P Jayaprasana T

6.2 Sprint Delivery Schedule:

| Sprint	Total Story Points	Durati on	Sprint Start Date	Sprint End Date (Planned)	Story Points Completed (as on Planned End	Sprint Release Date (Actual)

|						Date)
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

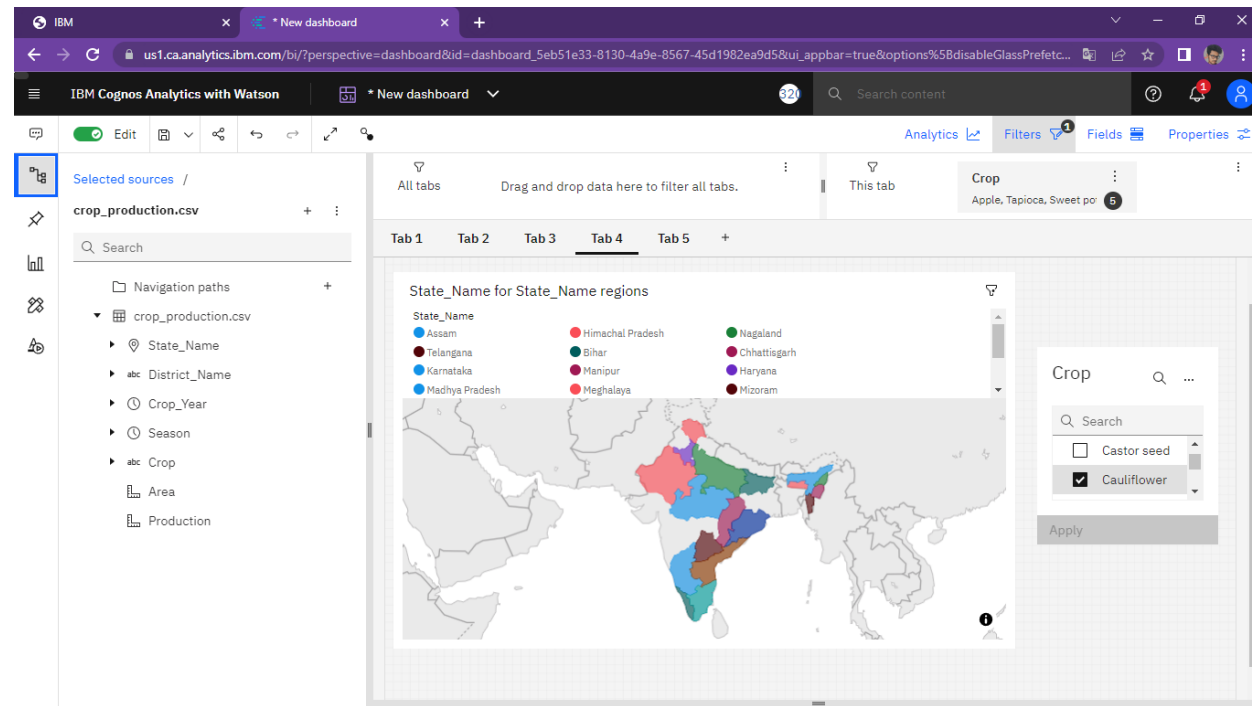7. **CODING & SOLUTIONING (Explain the features added in the project along with code)**

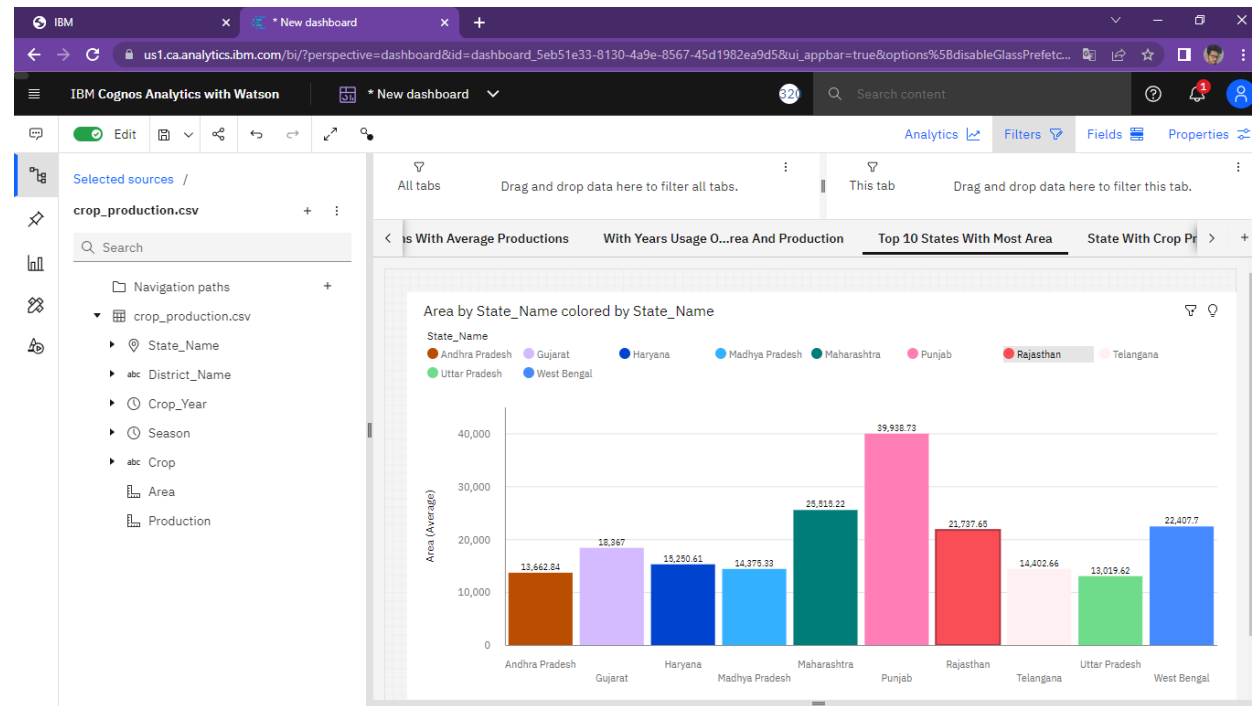https://drive.google.com/file/d/1b68LXWkdxQBIZGy1n2tg0JjP_HqS0_W4/view?usp=sharing
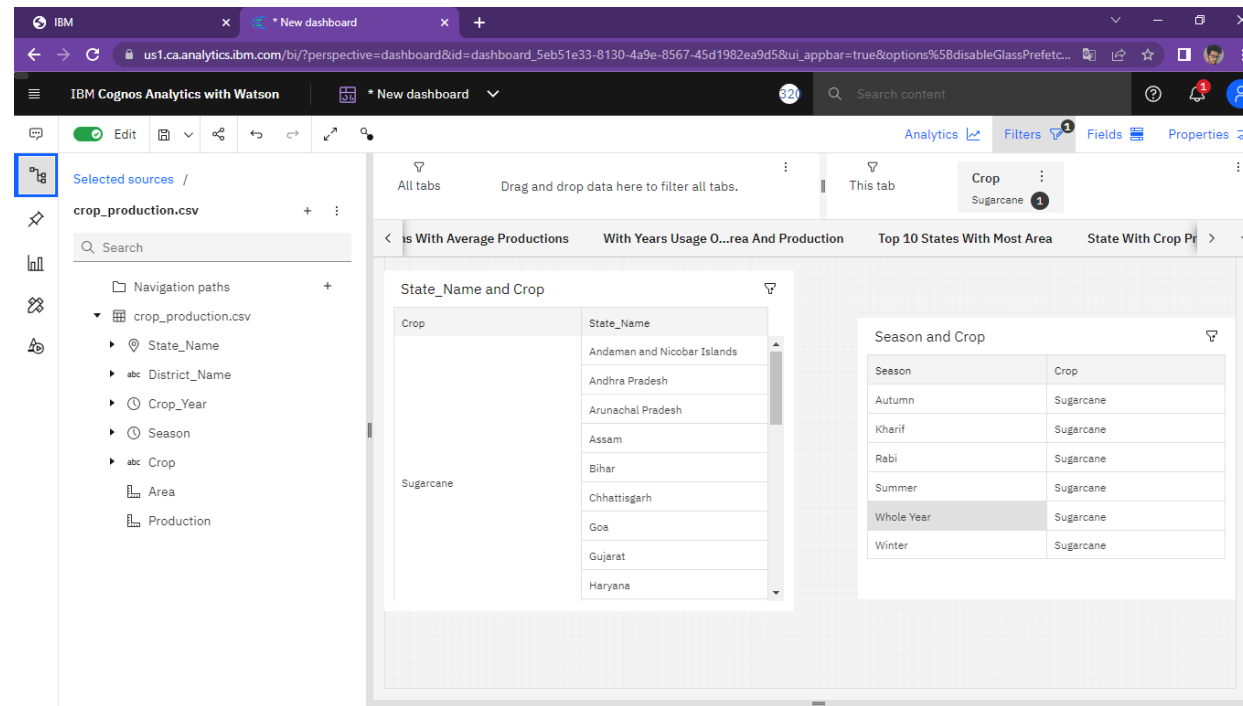
8. **TESTING**

8.1 User Acceptance Testing
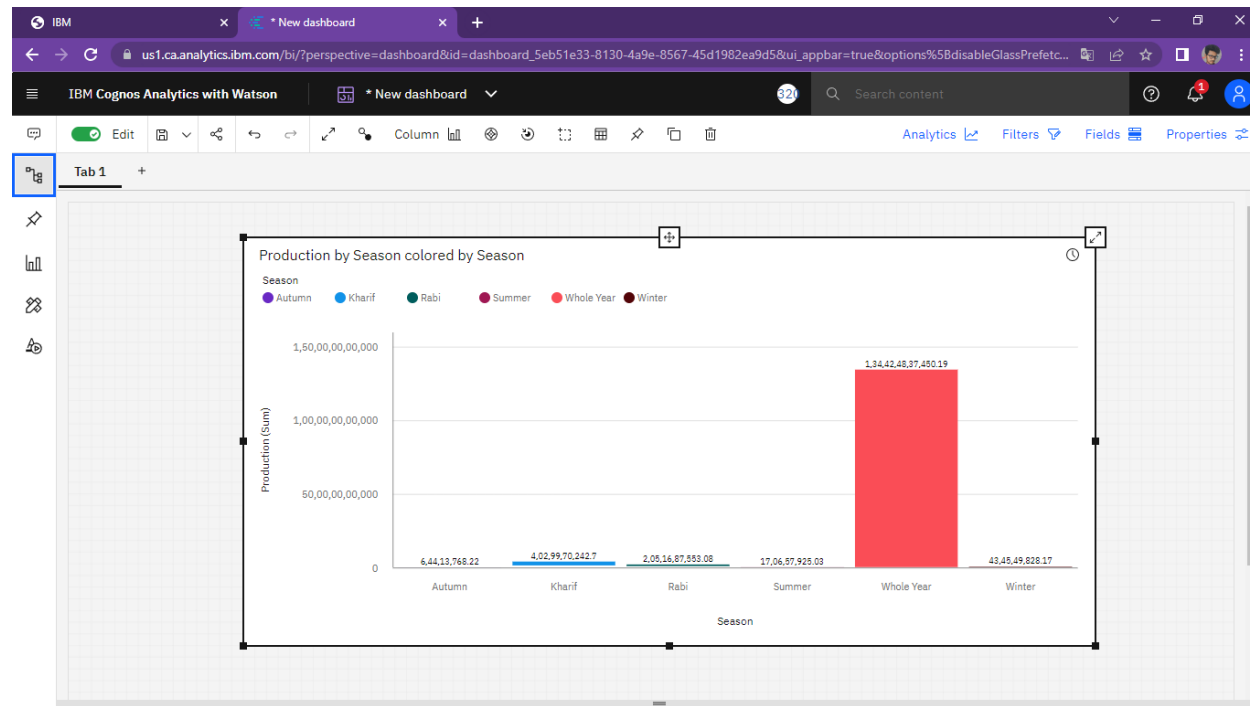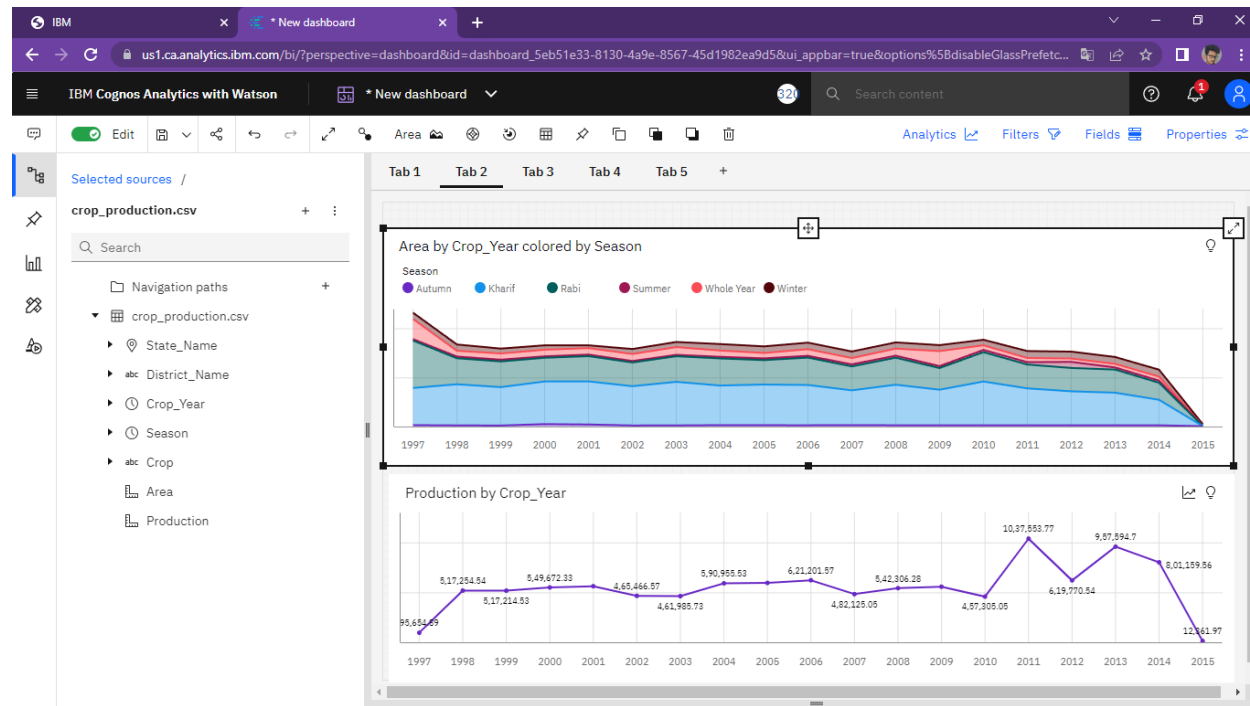
9. RESULTS

As per our prediction choosing what content to create, developing products and more. It gives you a 360-degree view of your customers, which means you understand them more fully, enabling you to better meet their needs.

10. ADVANTAGES :

The agricultural industry is poised for a data driven revolution. Technological advancements and the rise of big data have transformed the way we farm and produce food.

Agriculture is becoming increasingly reliant on technology and innovation in order to maintain production capacity. With depleting natural capital and increasing food demands, farmers will be turning to smart farming internet of things (IoT) solutions to lower cost, mitigate risk and maintain their production capacity.

DISADVANTAGES:

Perfect dataset is required to complete the prediction properly.

False dataset lead to loss

11.CONCLUSION

The proposed "ESTIMATE THE CROP YIELD USING DATA ANALYTICS" is used to predict the crop yield using the attributes such as State_Name ,District_ Name,Crop_ Year,Season, Crop,Area and Production. The proposed model is build with IBM Cognos. As a result of penetration of technology into agricultural field, there is a marginal improvement in the productivity. The innovation have led to new concepts like digital agriculture, smart farming, precision agriculture etc. It has been observed that analysis has been done on crop,

hidden pattern discovery using dataset related to season, area, production data. There

exists a lot of research scope in this research area

12. FUTURE SCOPE

The dashboard creation, visualization have taken lots of procedures and steps.The aim of the future work is to analyze the target attribute by reducing the number of procedures and steps. To improve the accuracy of the analysis algorithm selection procedure need to be optimized.

13. APPENDIX

Git Hub : https://github.com/IBM-EPBL/IBM-Project-37350-1660305426:

Project   Demo   Link   :   https://drive.google.com/drive/folders/1U-1BPKacmM1O9eVlfpXLbdSEX_9AeZ6b?usp=share_link