Project Report Format

Team Id	**PNT2022TMID19454**
Project Name	**Estimate crop yield using data analytics**

1. **INTRODUCTION**
   - 1.1 Project Overview
   - 1.2 Purpose
2. **LITERATURE SURVEY**
   - 2.1 Existing problem
   - 2.2 References
   - 2.3 Problem Statement Definition
3. **IDEATION & PROPOSED SOLUTION**
   - 3.1 Empathy Map Canvas
   - 3.2 Ideation & Brainstorming
   - 3.3 Proposed Solution
   - 3.4 Problem Solution fit
4. **REQUIREMENT ANALYSIS**
   - 4.1 Functional requirement
   - 4.2 Non-Functional requirements
5. **PROJECT DESIGN**
   - 5.1 Data Flow Diagrams
   - 5.2 Solution & Technical Architecture
   - 5.3 User Stories
6. **PROJECT PLANNING & SCHEDULING**
   - 6.1 Sprint Planning & Estimation
   - 6.2 Sprint Delivery Schedule
   - 6.3 Reports from JIRA
7. **DATA MODULATION & EXPLORATION (Explain the features added in the project along with visualization)**
   - 7.1 Data Modulation
   - 7.2 Data exploration
8. **DATA VISUALIZATION CHART CREATION**
   - 8.1 Final Dashboard result
9. **ADVANTAGES & DISADVANTAGES**
10. **CONCLUSION**
11. **FUTURE SCOPE**

12. APPENDIX

Dataset link

Dashboard link

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. Agricultural data is being produced constantly and enourmosly. As a result, agricultural data has come in the era of big data. Smart technologies contribute in data collection using electronic devices. In our project we are going to analyse and mine this agricultural data to get useful results using technologies like data analytics this result will be given to farmers for better crop yield in terms of efficiency and productivity.

1.2 Purpose:

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. The present study gives insights on various data analytics methods applied to crop yield prediction. The work also signifies the important lacunae points' in the proposed area of research. Data analytics provides farmers granular data on rainfall patterns, water cycles, fertilizer requirements, and more. This enables them to make smart decisions, such as what crops to plant for better profitability and when to harvest. Today, there is a need to produce more food to feed the growing population while using less land for it. However, the government is confronted with challenges due to changing climate and environmental issues that have a direct impact on food production. To accomplish this food production target, policymakers and industry leaders are taking support from technological innovations such as Data analytics, Cloud Computing, etc. These technological advancements prove highly beneficial in improving operations and boosting productivity.

2. LITERATURE SURVEY:

2.1 Existing problem:

Agriculture is the backbone of India. It is the major support to the Indian economy. The money flow starts from the farmer's hand as that is where the essential food products make their entry. Almost 70% of the people live in rural basins with more than 50% of the whole Indian population taking agriculture as their main work. India is the second-largest producer of fruits and vegetables in the world. Still, the sector suffers from a number of issues including ignorance and non-recognition. Climate change, unpredictable monsoon, drought, floods, migration of farmers towards cities for better jobs are some of the major distress that agriculture industry goes through. Even the

most acknowledged platform like media fail to cover field agriculture and go after agriculture ministers in the country to cover the issues which lack the farmers part. As institutions failed to provide loans and farmer welfare schemes. Every time a crop is reaped from the land, the soil structure changes. It is hard to find the crop that would next suit the soil type. Some people in the agriculture industry maintain acres of land which makes it difficult to penetrate the potential problems in the other corner of their land piece.

2.2 References:

PAPER 1:
TITLE: Rice Crop Yield Prediction using Data Mining Techniques
AUTHOR NAME: Dakshayini Patil, Dr. M.S.Shidhonkar
PUBLICATION YEAR: 2017
DESCRIPTION:

Data Analytics has emerged as a potential tool for the crop yield prediction, allowing the model to extract features and predict from the datasets. Meanwhile, smart farming technology enables the farmers to achieve maximum crop yield by extracting essential parameters of crop growth.

PAPER 2:
TITLE: A Survey on Crop Yield Prediction based on Agricultural Data
AUTHOR NAME: Dhivya B H, Manjula R, Siva Bharathi S, Madhumathi R
PUBLICATION YEAR: 2017
DESCRIPTION:

The study focuses on the advantages of using data analytics in crop yield prediction . There have been a number of research studies undertaken that focus on the importance of data analytics as a supplementary tool in transforming large volumes of agricultural data into meaningful information.

PAPER 3:
TITLE: Prediction of crop yield using regression analysis
AUTHOR NAME: V Sellamand, E Poovammal
PUBLICATION YEAR: 2016
DESCRIPTION:

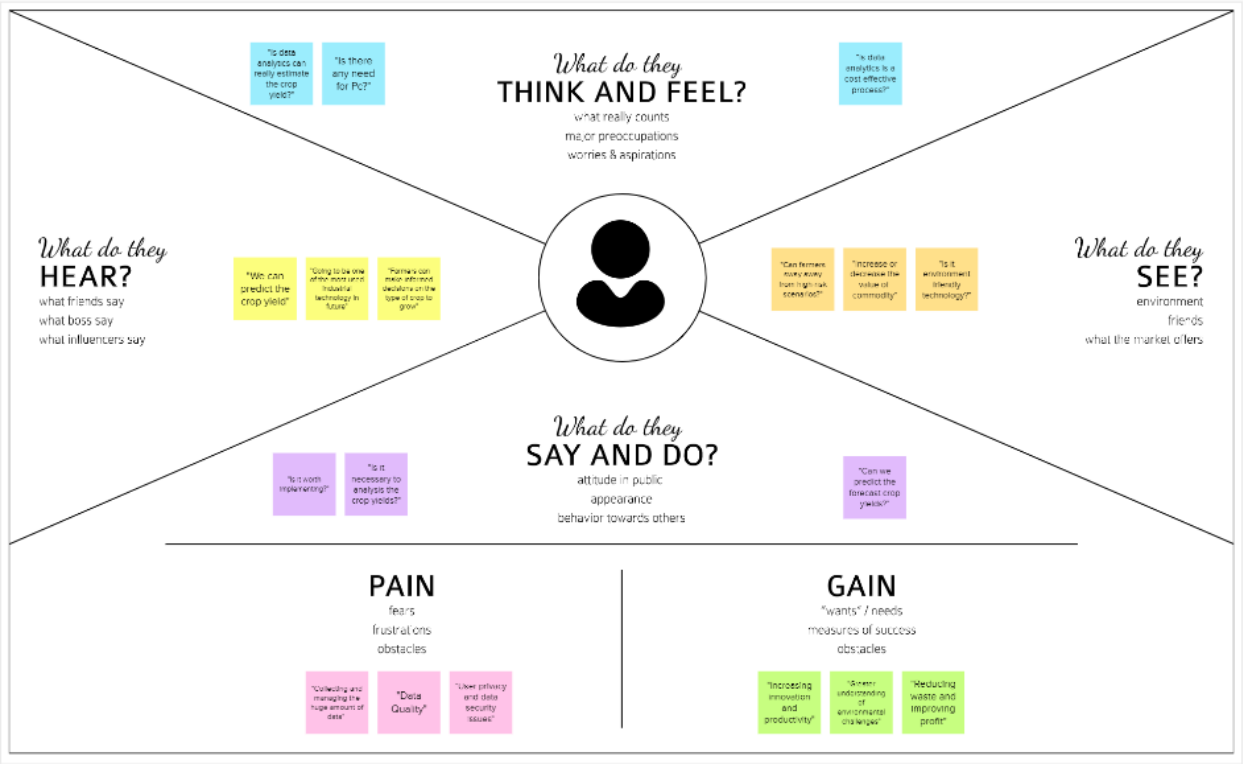
Many researchers have been contributed their previous knowledge towards data analytics in agriculture. There are many simulation model available for crop productivity predictions. As it depends on economical and environmental parameters so we can apply these existing models or methods to any other area.

2.3 Problem Statement Definition:

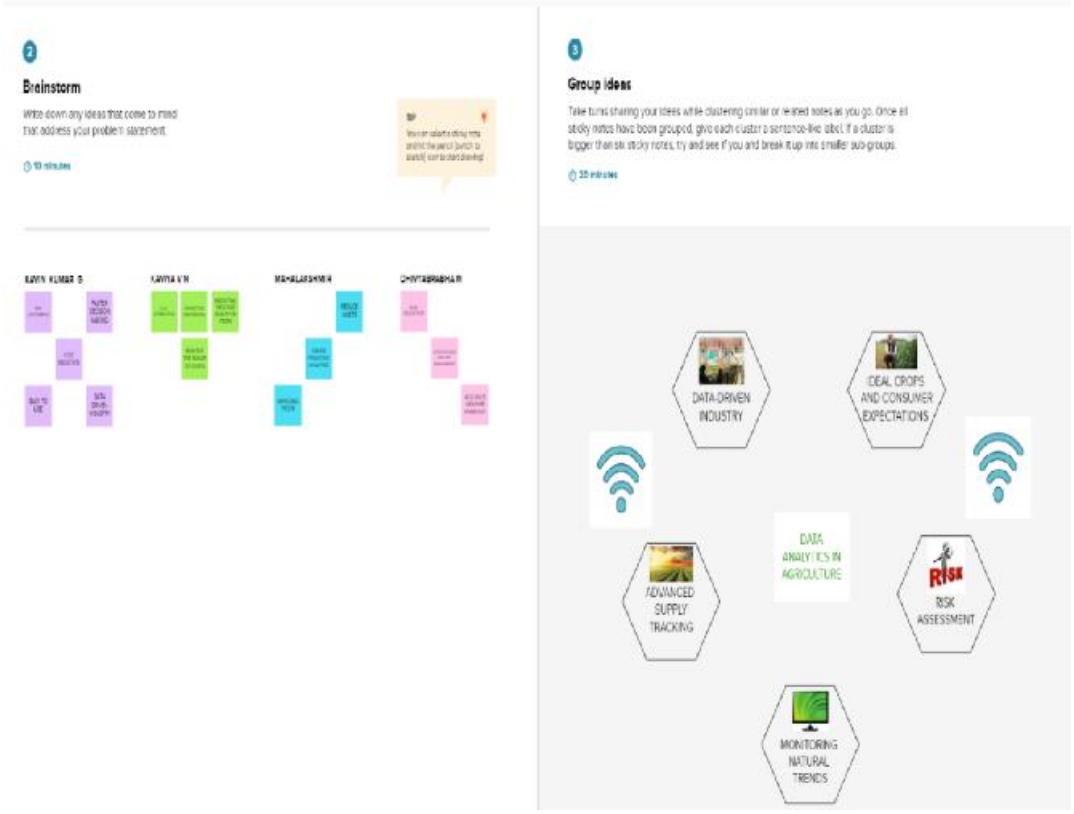
Problem Statement (PS)	I am (Customer)	I'm trying to	But	Because	Which makes me feel
PS-1	A farmer	Cultivate crops	It is a tedious process	Of rising food demand and climate change	Scared
PS-2	A consumer	Purchase groceries	It is a baffling process	Of limited variety Of products and often expensive	Frustrated

3. IDEATION & PROPOSED SOLUTION

3.1 Empathy Map Canvas

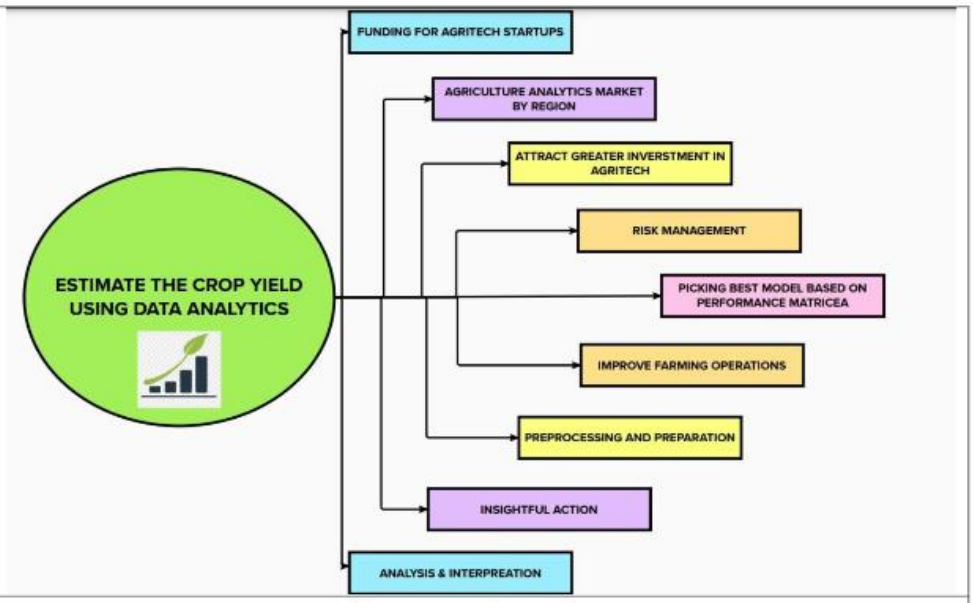


3.2 Ideation & Brainstorming

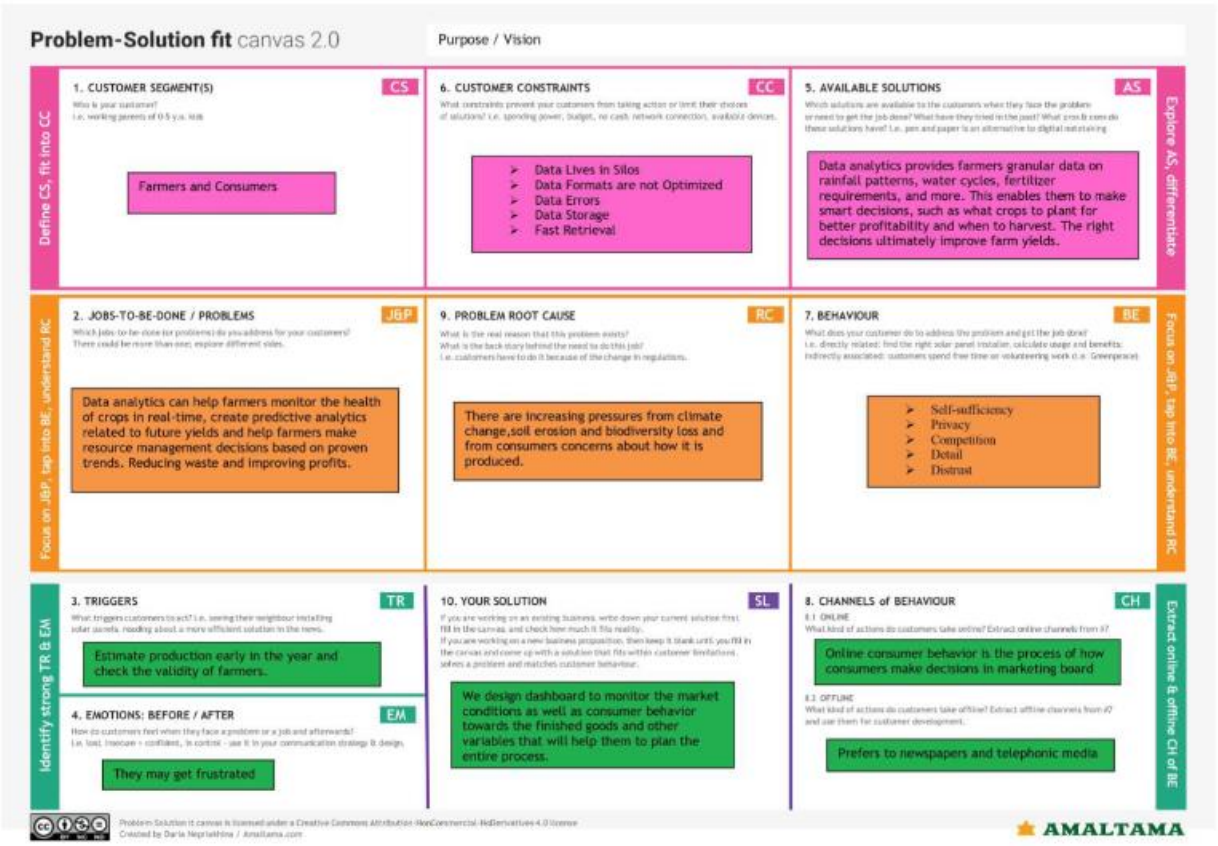


3.3 Proposed Solution

S.No.	Parameter	Description
1.	Problem Statement (Problem to be solved)	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.
2.	Idea / Solution description	One of the most important features of data analytics is data processing. Data processing involves raw data collection and organization to derive inferences. Farmers are now empowered with insights that can help them predict the market conditions, consumer behavior towards the finished goods, factors in inflation, and other variables that will help them plan the entire process.
3.	Novelty / Uniqueness	For products or services, completeness can suggest vital attributes that help customers compare and choose. Accuracy of data represents real-world scenario and confirms with a verifiable source. With data analytics crop yield predictions can be done earlier even before sowing seeds which results in more productivity.
4.	Social Impact / Customer Satisfaction	Meeting customer expectations on the quality of agricultural product, increasing the quality product instruments and reducing the complexity of the current. The quality debate concerns information on product characteristics, on farming methods, and on origin that a farmer wants to communicate and a consumer wants to know.

5.	Business Model (Revenue Model)	
6.	Scalability of the Solution	Automate data science and data engineering tasks. Train, test and deploy models seamlessly across multiple enterprise applications. Extend common data science capabilities across hybrid, multi-cloud environments.

3.4 Problem Solution fit



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	User Profile Update	Updating personal details Updating educational details
FR-4	User Login	Login with username and password
FR-5	User Dashboard	Performing required operations
FR-6	Project Upload	Uploading the project according to the problem statement provided

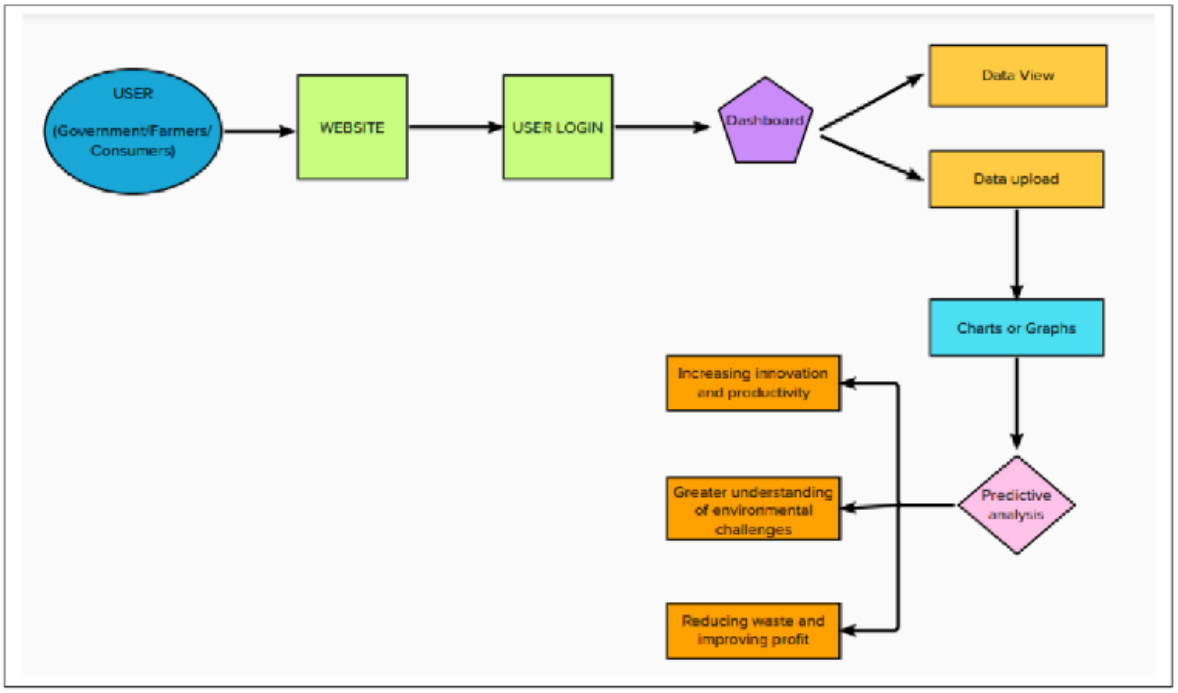
4.2 Non-Functional requirements

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

FR No.	Non-Functional Requirement	Description
NFR-1	**Usability**	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.
NFR-2	**Security**	Data analytics has a positive and significant relationship with a firms ability to manage data security and a positive impact on service supply chain innovation capabilities and service supply chain performance.
NFR-3	**Reliability**	Farmers are now empowered with insights that can help them predict the market conditions, consumer behaviour towards the finished goods, factors in inflation and other variables that helps them.
NFR-4	**Performance**	Data analytics helps in executing the existing algorithms faster with large data sets. One of the most important features of data analytics is data processing. Data processing involves raw data collections and organization to derive inferences.
NFR-5	**Availability**	For products are services, completeness can suggest vital attributes that helps customer compare and choose. With data analytics crop yield predictions can be done earlier even before sowing seeds which results more productivity.
NFR-6	**Scalability**	Automate data science and data engineering tasks. Train, test and deploy models seamlessly across multiple enterprise applications. Extend common data science capabilities across hybrid, multi-cloud environments.

5. PROJECT DESIGN

5.1Data Flow Diagrams



5.2 Solution & Technical Architecture:

Table-1 : Components & Technologies:

S.No	Component	Description	Technology
1.	User Interface	How user interacts with dashboard e.g. Filter, Graph, charts, predictive analysis etc.	IBM Cognos Analytics
2.	Dashboard Logic-1	Logic for a process in the dashboard	IBM Cognos Analytics
3.	Dashboard Logic-2	Logic for a process in the dashboard	IBM Cognos Analytics
4.	Database	Data Type, Configurations etc.	IBM
5.	Cloud Database	Database Service on Cloud	IBM DB2, IBM Cloudant etc.
6.	File Storage	File storage requirements	IBM Block Storage or Other Storage Service or Local Filesystem
7.	Data Analytics Model	Purpose of Data Analytics Model	Predictive Analysis Recognition Model, etc.
8.	Infrastructure (Server / Cloud)	Application Deployment on Local System / Cloud Local Server Configuration: Cloud Server Configuration :	Local, Cloud Foundry, Kubernetes, etc.

Table-2: Application Characteristics:

S.No	Characteristics	Description	Technology
1.	Open-Source Frameworks	List the open-source frameworks used	IBM Cognos Analytics
2.	Security Implementations	Security in IBM Cognos Analytics is optional. Typically, anonymous users have limited read-only access.	Authentication providers, Authorization, Cognos namespace, IBM Cognos Application Firewall
3.	Scalable Architecture	You can enable or disable services run by the dispatcher on individual servers to balance the load for a given computer by request type	XML, SOAP, WSDL
4.	Availability	Web based data modelling, Interactive dashboards and enterprise reports, Data exploration and prediction	XML, SOAP, WSDL

S.No	Characteristics	Description	Technology
5.	Performance	User population grow, processing requests tend to increase in number and complexity and network capacity and other aspects of infrastructure may be modified. These changes can affect IBM Cognos BI performance.	XML, SOAP, WSDL

6. PROJECT PLANNING & SCHEDULING

6.1 Sprint Planning & Estimation

Project Tracker, Velocity & Burndown Chart: (4 Marks)

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

6.2. Sprint Delivery Schedule

Product Backlog, Sprint Schedule, and Estimation (4 Marks)

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 personal details and creating account using username and password.	2	High	Kaviya V N
Sprint-1		USN-2	As a user, After completed the registration I will receive the confirmation message through my registered email id.	1	High	Mahalakshmi R
Sprint-2		USN-3	As a user, I successfully installed the application which is provided.	2	Low	Kavin Kumar G
Sprint-1	Login	USN-4	As a user, I can logged in my account using my username and password.	5	High	Dhivyabrabha M
Sprint-1	Dashboard	USN-5	As a user, I can view and access the information in the dashboard.	1	Medium	kaviya V N
Sprint-3	Building of IBM Cognos analytics	USN-6	Creation of Charts and Graphs	8	High	Kavin Kumar G
Sprint-3	Modelling of Assistant	USN-7	Adding responses to	2	High	Kaviya V N

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-1		USN-8	Account Creation/As a user, I can see a Dashboard which helps to create an account. As a user, I can see a dashboard which helps to solve the queries in crop production.	1	Medium	Mahalakshmi R
Sprint-4		USN-9	As a user, I can see a Dashboard which helps to access charts and graphs.	2	High	Dhivyabrabha M
Sprint-3		USN-10	As a user, I can see a Charts and Graphs which helps to monitor and manage the trending activities in market.	3	Low	Kaviya V N
Sprint-1	Testing & Deployment Phase-I	USN-11	Testing the Dashboard performance with the Administrator.	8	High	Kavin Kumar G
Sprint-2		USN-12	Integration of Flask webpage with the dashboard assistant to provide a framework, As a user, I can see a webpage to access the dashboard.	2	High	Mahalakshmi R
Sprint-4	Deployment Phase-II & Model Improvement	USN-13	Deployment of IBM Cognos analytics based dashboard for crop yield production	1	Low	Kavin Kumar G

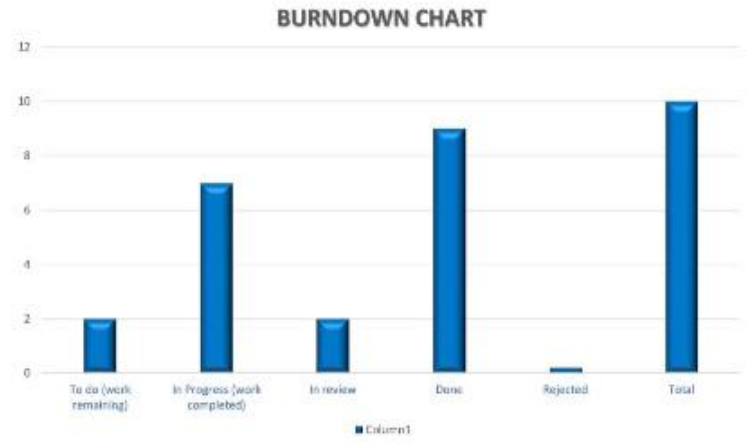
Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
			services/As a user, I can monitor and access the information in the dashboard.			
Sprint-1		USN-14	Improving the model efficiency whenever needed/As a user, I can see new updated dashboard in Future days.	2	Moderate	Dhivyabrabha M
Sprint-2	Verification	USN-15	Administrator can completely verify the submitted application.	5	High	Kavin Kumar G
Sprint-3	Approval	USN-16	After completion, new updated dashboard is provided to the customers.	2	High	Kaviya V N

Velocity:

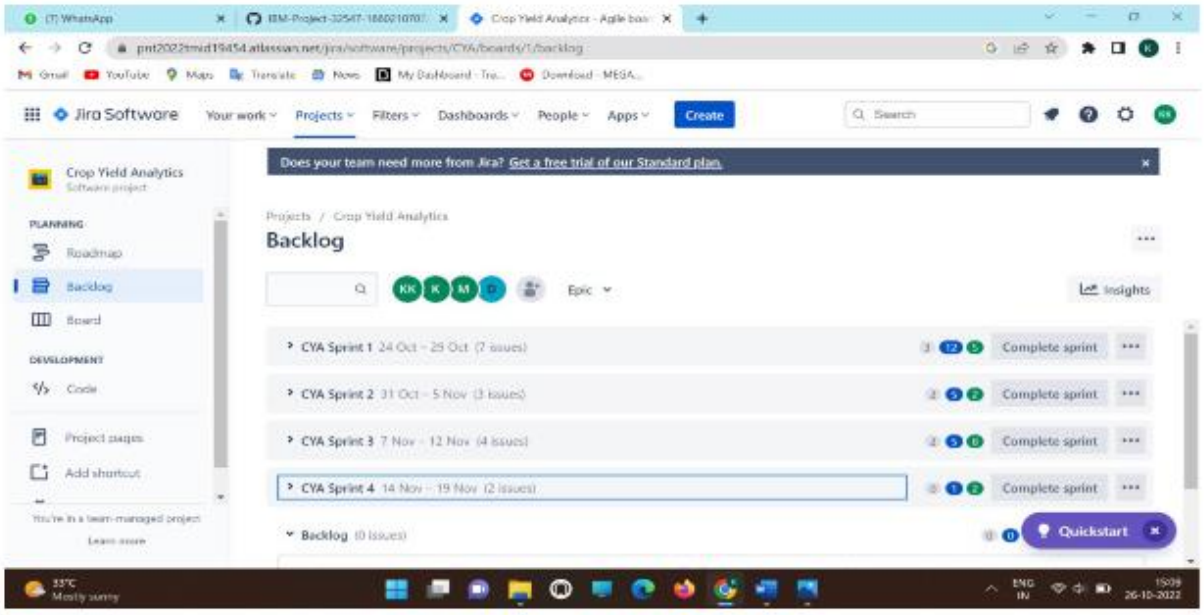
The team's average velocity (AV) per iteration unit (story points per day)

$$AV = \frac{sprint\ duration}{velocity} = 20/6 = 3.34$$

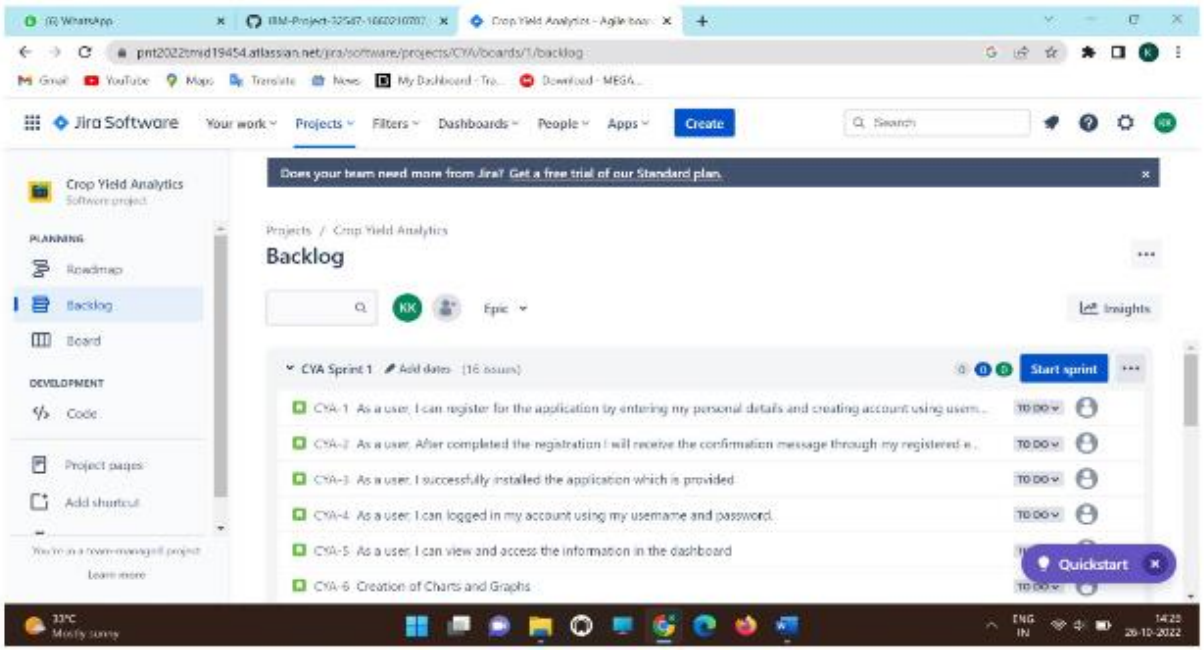
Burndown Chart:

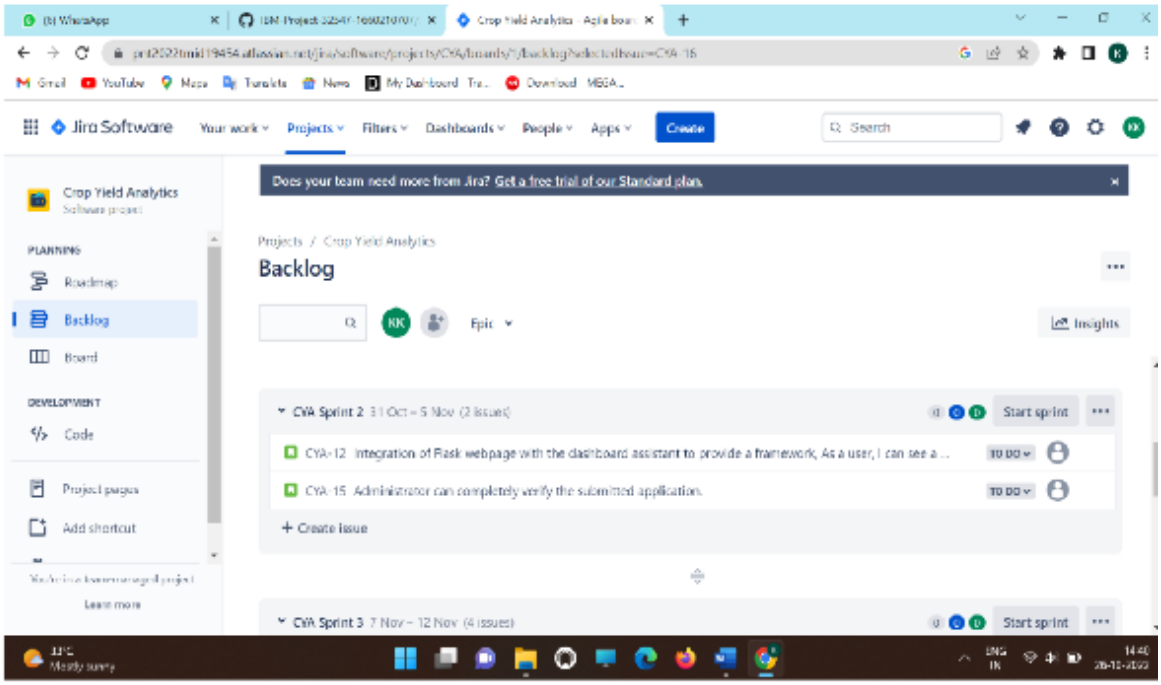


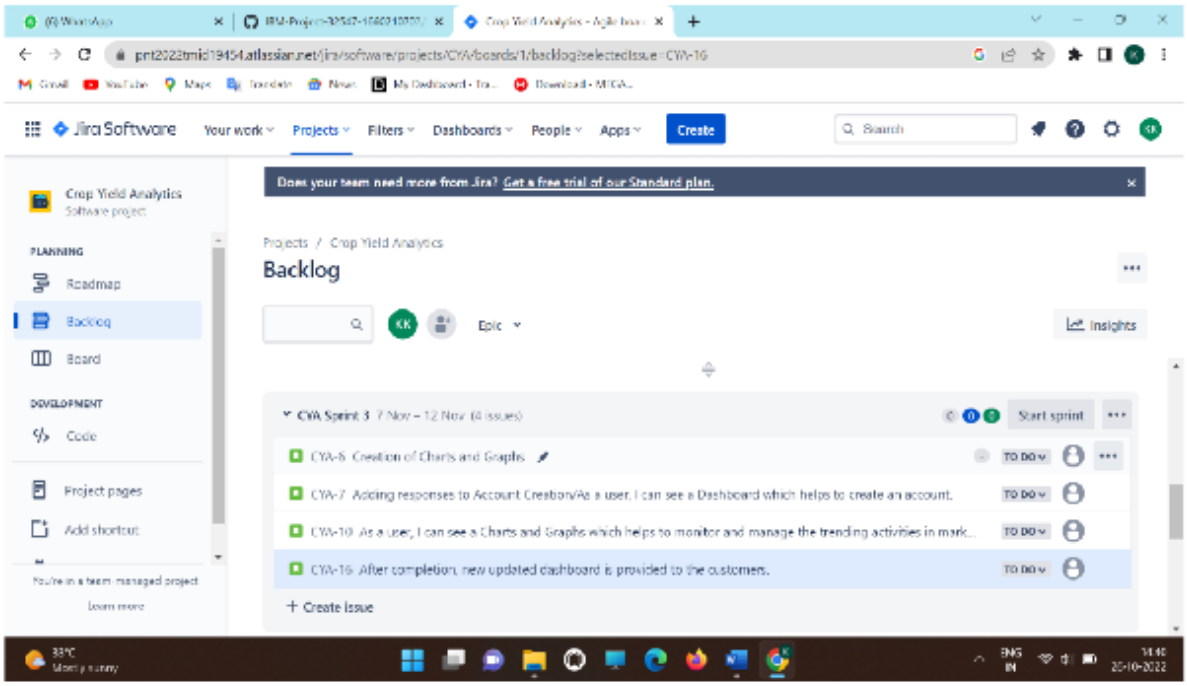
6.3 Reports from JIRA

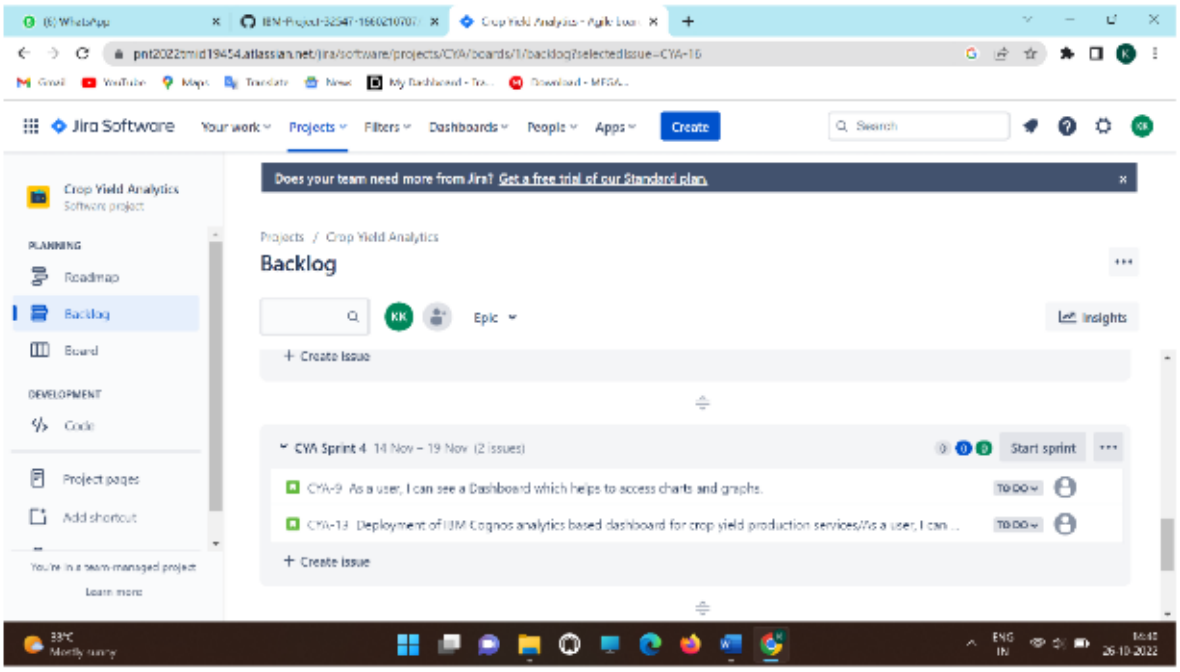


2. **Showing the respective User Stories in their respective Sprints.**

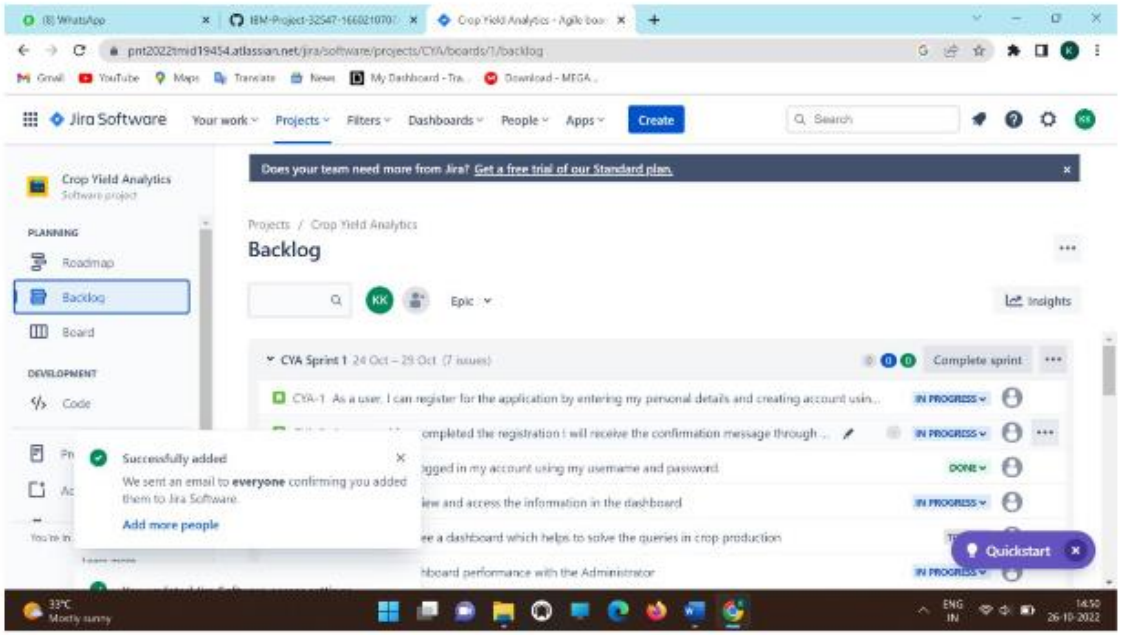



Jira Software
Your work
Projects
Filters
Dashboards
People
Apps
Create
Search
Crop Yield Analytics
Software project
PLANNING
Roadmap
Backlog
Board
DEVELOPMENT
Code
Project pages
Add shortcut
Learn more
Does your team need more from Jira? Get a free trial of our Standard plan.
Projects / Crop Yield Analytics
Backlog
Epic
Insights
CYA Sprint 2 31 Oct – 5 Nov (2 issues)
Start sprint
CYA-12 Integration of Flask webpage with the dashboard assistant to provide a framework, As a user, I can see a ...
CYA-15 Administrator can completely verify the submitted application.
TO DO
+ Create issue
CYA Sprint 3 7 Nov – 12 Nov (4 issues)
Start sprint

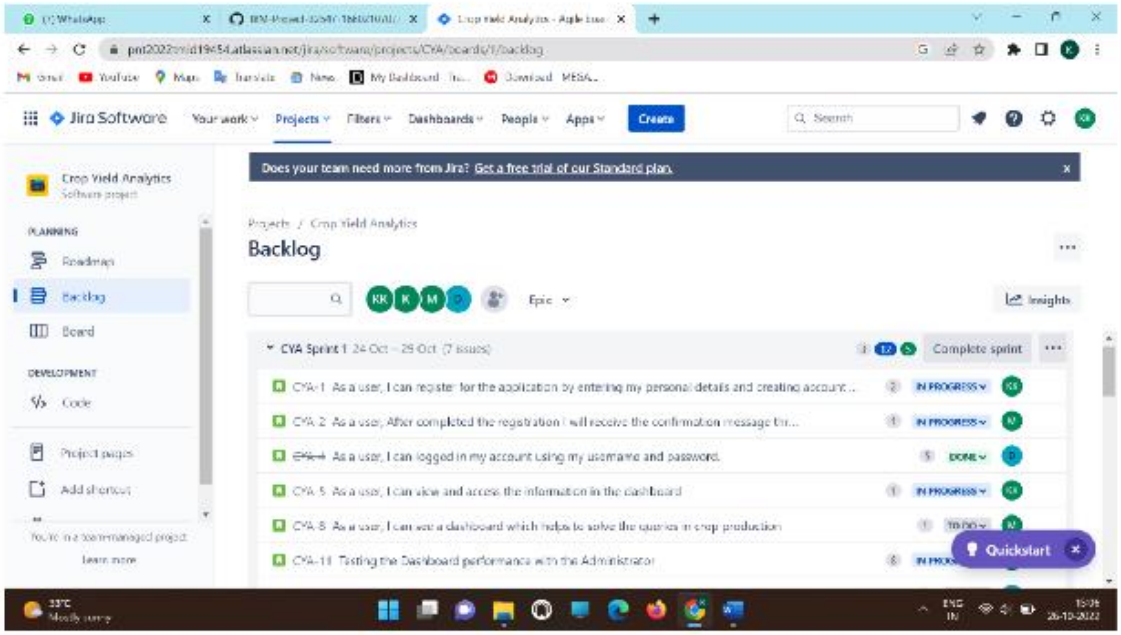

Jira Software
Your work
Projects
Filters
Dashboards
People
Apps
Create
Search
Crop Yield Analytics
Software project
PLANNING
Roadmap
Backlog
Board
DEVELOPMENT
Code
Project pages
Add shortcut
Learn more
Does your team need more from Jira? Get a free trial of our Standard plan.
Projects / Crop Yield Analytics
Backlog
Epic
Insights
CYA Sprint 3 7 Nov – 12 Nov (4 issues)
Start sprint
CYA-6 Creation of Charts and Graphs
CYA-7 Adding responses to Account Creation/As a user, I can see a Dashboard which helps to create an account.
CYA-10 As a user, I can see a Charts and Graphs which helps to monitor and manage the trending activities in mark...
CYA-16 After completion, new updated dashboard is provided to the customers.
TO DO
+ Create issue


Jira Software
Your work
Projects
Filters
Dashboards
People
Apps
Create
Crop Yield Analytics
Software project
PLANNING
Roadmap
Backlog
Board
DEVELOPMENT
Code
Project pages
Add shortcut
Does your team need more from Jira? Get a free trial of our Standard plan.
Projects / Crop Yield Analytics
Backlog
Epic
Insights
+ Create issue
CYA Sprint 4 14 Nov – 19 Nov (2 issues)
Start sprint
CYA-9 As a user, I can see a Dashboard which helps to access charts and graphs.
TO DO
CYA-14 Deployment of IBM Cognos analytics based dashboard for crop yield production services/As a user, I can ...
TO DO
+ Create issue

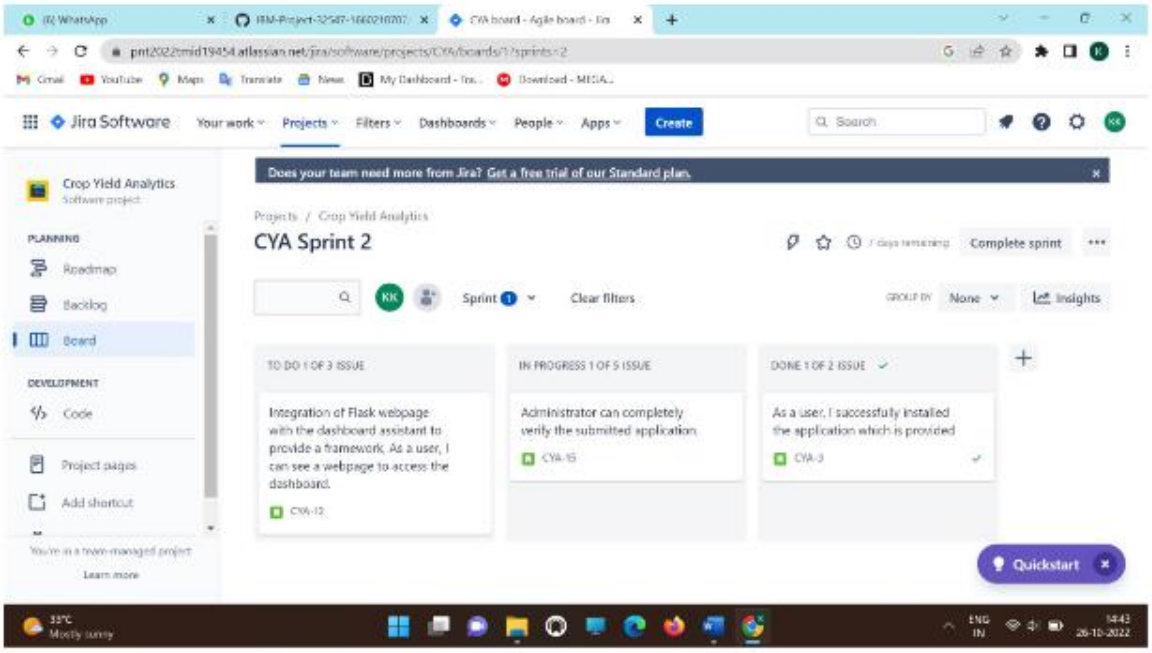
3. Inviting team members in Jira Platform .



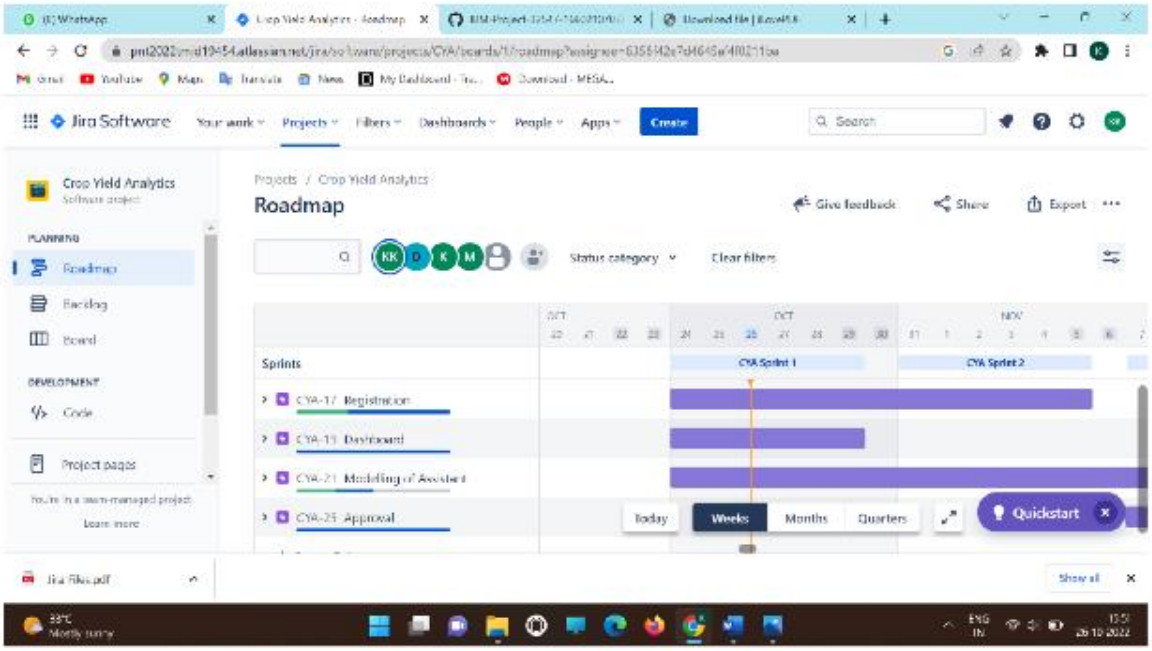
4. Allocating the story points and assigning tasks to team members .



5. **Image showing the working progress of the Sprint.**

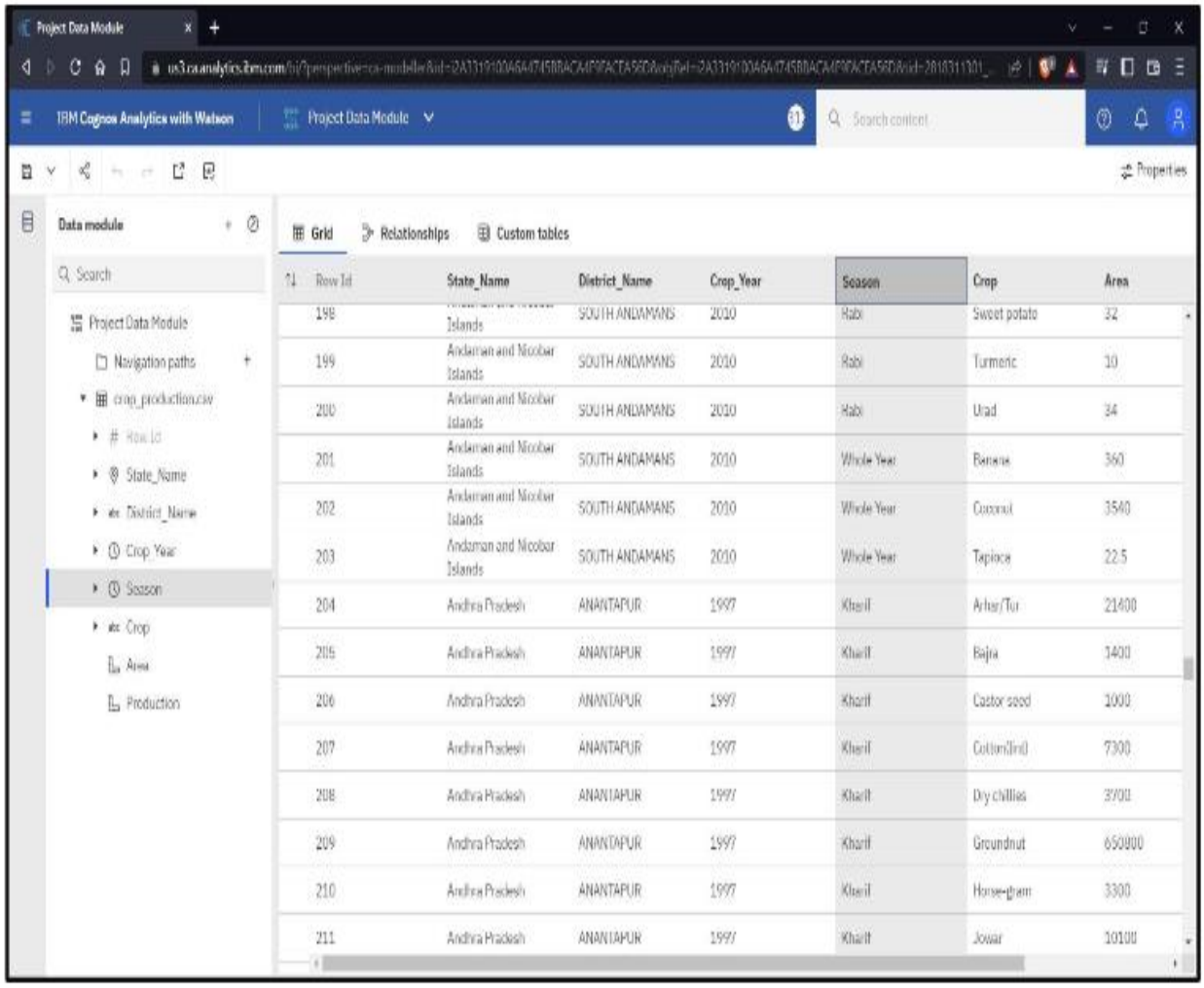


6. **Imaging showing the Roadmap of the Sprints.**

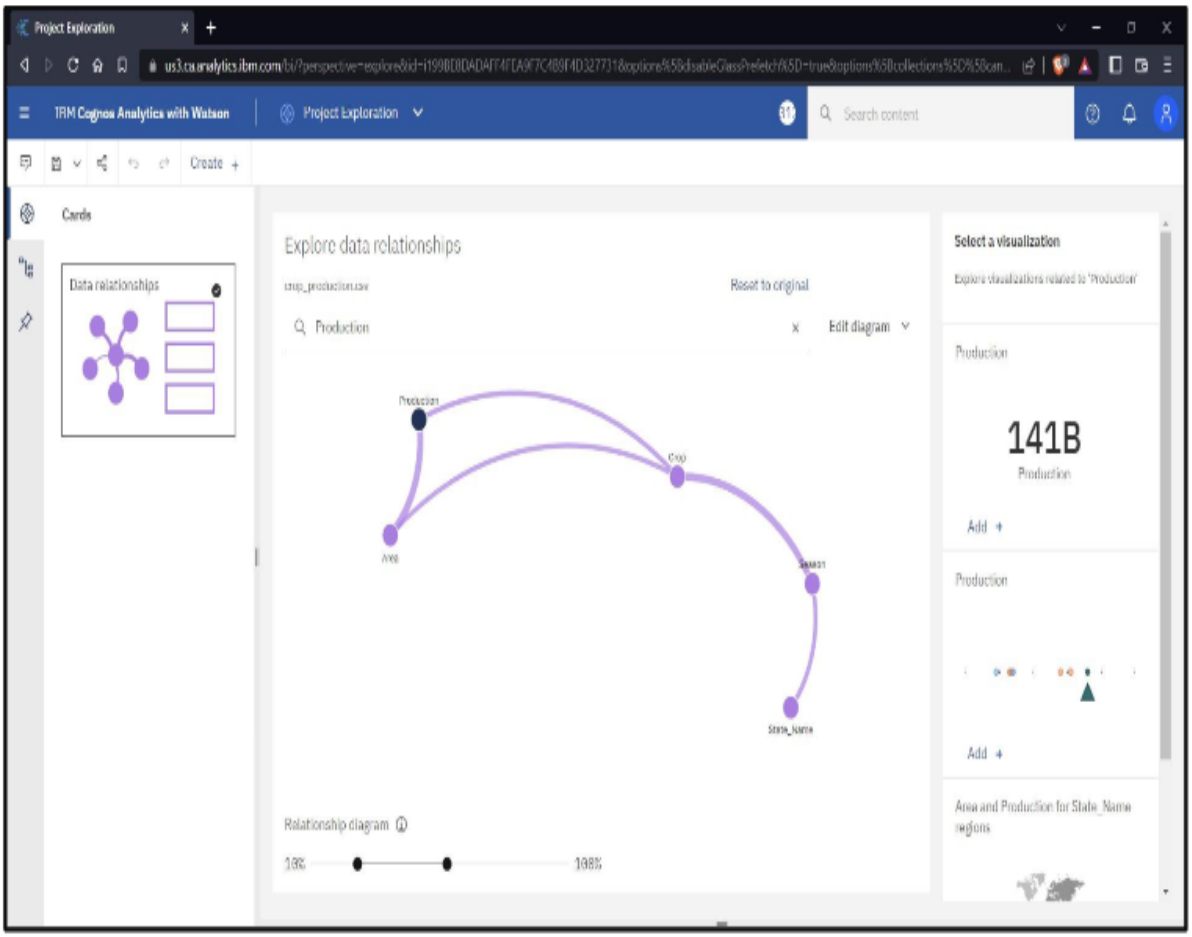


7. DATA MODULE CREATION & EXPLORATION (Explain the features added in the project along with visualization)

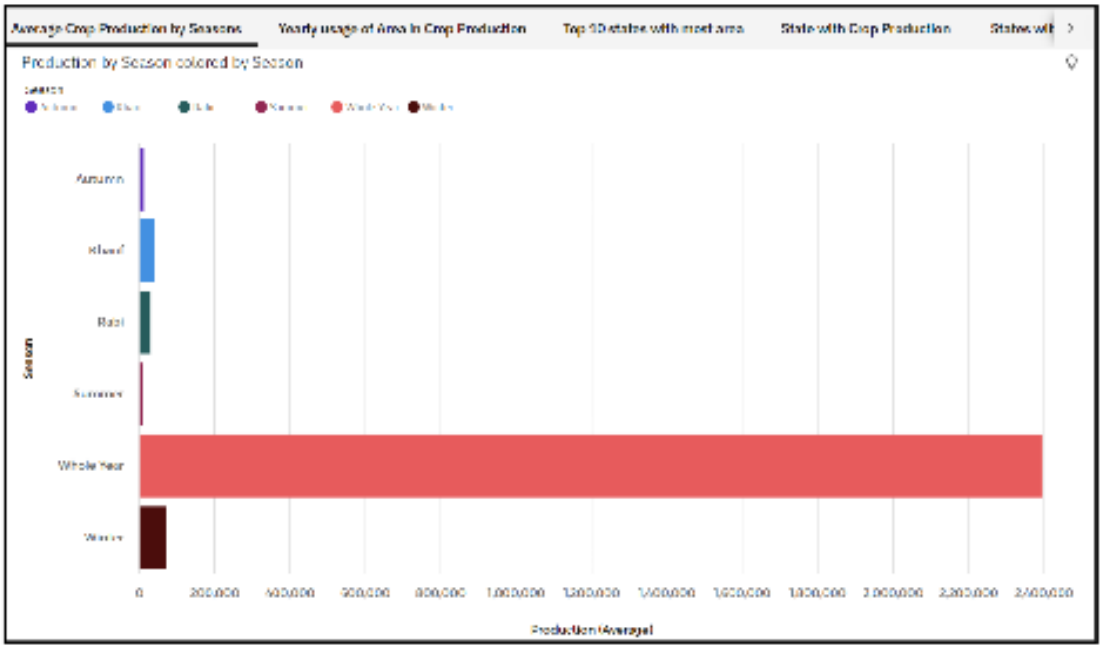
7.1 Data Modulation

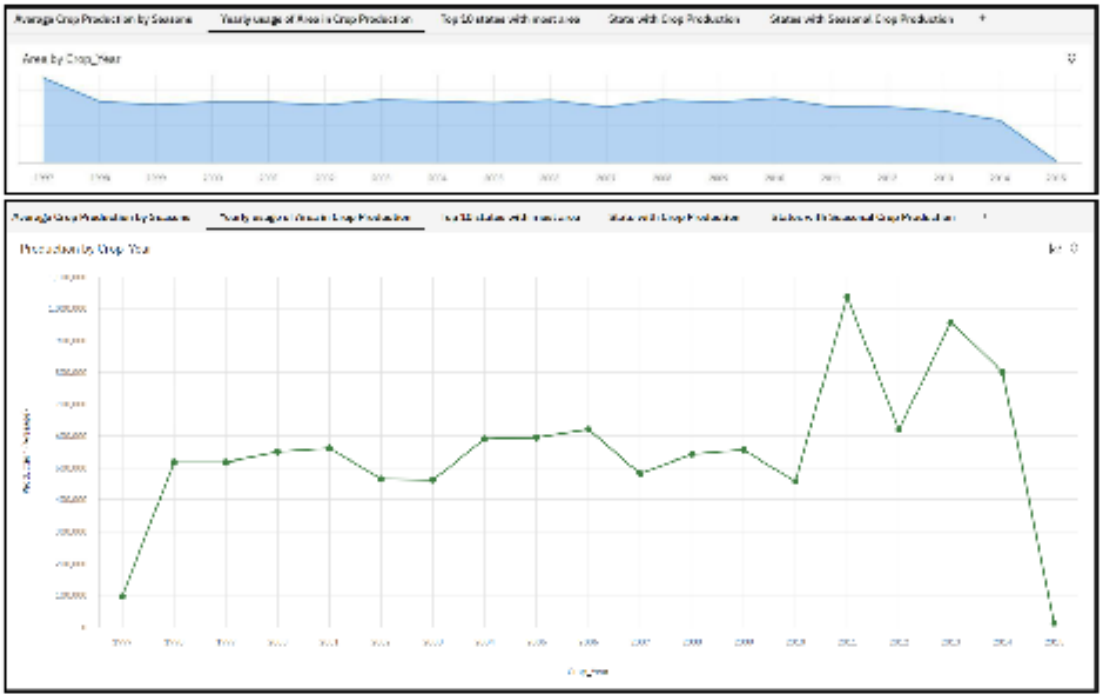


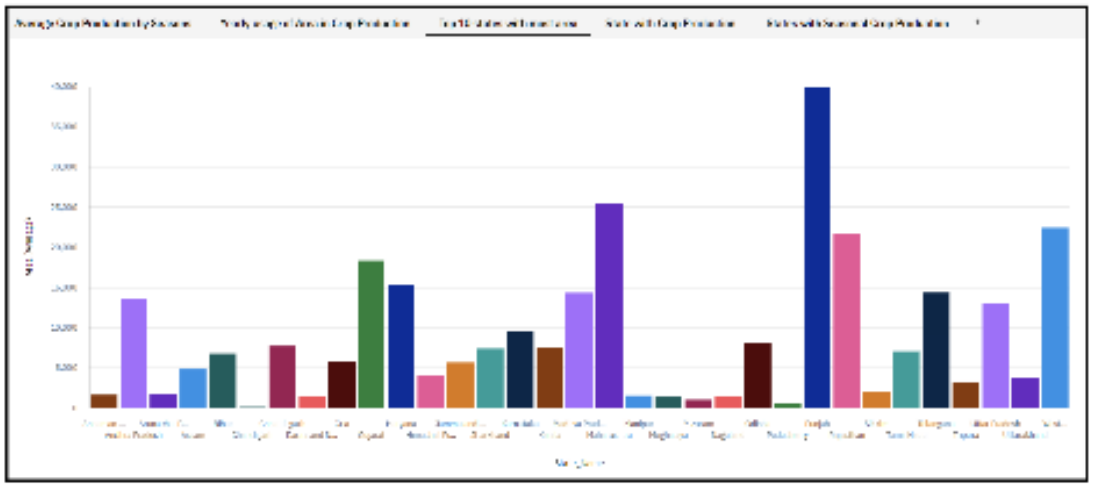
7.2 Data exploration

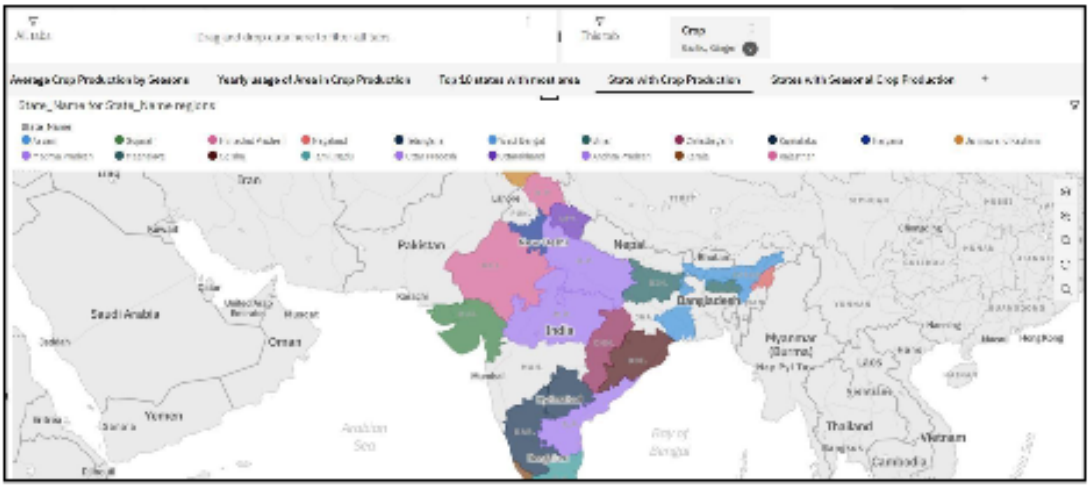


8. DATA VISUALIZATION CHARTS CREATION

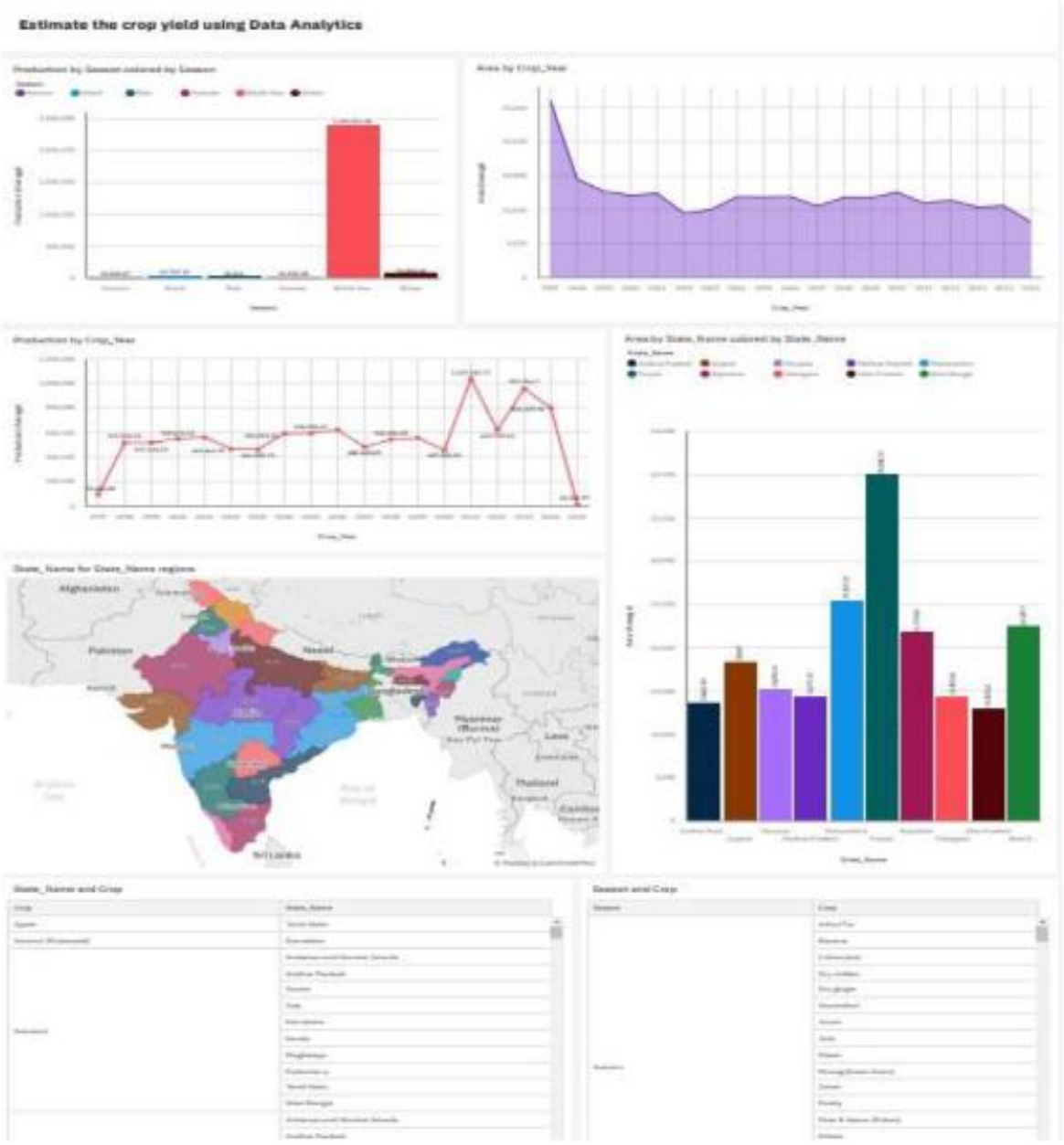







Crop
Average Crop Production by Seasons
Yearly average of Area in Crop Production
Top 10 states with most area
State with Crop Production
States with Seasonal Crop Production
Pakistan
Nepal
Bangladesh
India
Saudi Arabia
Oman
Yemen
Thailand
Vietnam
Myanmar (Burma)
Arabian Sea
Bay of Bengal

8.1 Final Dashboard (Result)



9. ADVANTAGES & DISADVANTAGES

Advantages

- Farmers can easily understand the crop yield production
- The dataset of the crop yield is manipulated and analyzed easily even the dataset is have more data.
- By using Data Analytics , Farmers can predict their crop production using forecast method
- It helps in agriculture technology to manage the crop yield production.

Disadvantages

- Crop yields production may be miscalculated
- Improper fields in the data leads to wrong analysis

10. CONCLUSION:

Our project Estimation of Crop yield using Data Analytics deals with the identifying the how the crop yield production done by a farmer in state wise production. The main purpose of this project is to create a dashboard which helps the farmer to easily analyze the different crops produced in different states in the country.

11. FUTURE SCOPE:

The Dashboard contains the actual value of crop production in the country. In future we can predict the crop yield production for need five years using the actual values with user interface dashboard that helps in betterment of the farmers.

12. APPENDIX:

Dataset link: click here

Dashboard link: click here