ESTIMATE THE CROP YIELD USING DATA ANALYTICS

TEAM ID - PNT2022TMID03011

TEAM LEADER – SNEHA R (727719EUIT151)

TEAM MEMBER – SINDHUJA R S (727719EUIT147)

TEAM MEMBER – SREE NANDINI M (727719EUIT158)

TEAM MEMBER – SHRINITHI S (727719EUIT143)

ABSTRACT

Agriculture is important for human survival because it serves the basic need. A well-known fact that the majority of population (≥55%) in India is into agriculture. Due to variations in climatic conditions, there exist bottlenecks for increasing the crop production in India. It has become challenging task to achieve desired targets in Agri based crop yield. Various factors are to be considered which have direct impact on the production, productivity of the crops. Crop yield prediction is one of the important factors in agriculture practices. Farmers need information regarding crop yield before sowing seeds in their fields to achieve enhanced crop yield. 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. The main challenge in using big data in agriculture is identification of effectiveness of big data analytics. Efforts are going on to understand how big data analytics can agriculture productivity. The present study gives insights on various data analytics methods applied to crop yield prediction and also signifies the important lacunae points' in the proposed area of research.

1. INTRODUCTION

1.1 PROJECT OVERVIEW:

In India crop yield is season dependent and majorly influenced by the biological and economic causes of an individual crop. Reporting of progressive agricultural yield in all the seasons is an ample task and an advantageous task for every nation with respect to assesses the overall crop yield prediction and estimation. At present a common issue worldwide is, farmers are stressed in producing higher crop yield due to the influence of unpredictable climatic changes and significant reduction of water resource worldwide. A

study was carried out to collect the data on world climatic changes and the available water resources which can be used to encourage advanced and novel approaches such as big data analytics to retrieve the information of the previous results to the crop yield prediction and estimation. Study imported that the selection and usage of the most desirable crop according to the existing conditions, support to achieve the higher and enhanced crop yield [11]. 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.

1.2 PURPOSE:

Agriculture is the widest economic sector and has an important role regarding the framework of socio-economic fabric of India. 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. The estimation of production of crop helps these companies in planning supply chain decision like production scheduling.The industries such as fertilizers, seed, agrochemicals and agricultural machinery plan production and activities like marketing based on the estimates of crop yield. Farmers experience was the only way for prediction of crop yield in the past days. Technology penetration into agriculture field has led to automation of the activities like yield estimation, crop health monitoring etc.

2. LITERATURE SURVEY:

2.1 EXISTING PROBLEM:

A) P. VINDHYA "CROP YIELD PREDICTION USING BIG DATA ANALYTICS" ANNA UNIVERSITY, TRICHY, TAMIL NADU, INDIA, 5 MAY 2015.

The proposed system suggests the accurate prediction of crop yield certainly benefits the farmers in choosing the right method to reduce the crop damage and get best prices for their crops. The factors involved in this method are Area under Cultivation (AUC) interims of hectors, Annual Rainfall (AR) rates and Food Price Index (FPI) and to develop relationships among these parameters. Regression Analysis (RA) methodology was applied to examine the selected factors and their impact on crop prediction and final yield. RA methodology is a multivariable investigation practice which can categorize the factors into groups such as explanatory and response variables and helps to assess their interaction to obtain a resolution. Crop yield gaps, measured as difference between expected yields based on the potency and actual farm yield received. In order to achieve the higher crop yield, farmers must tackle the influencing factors such as influence of change in climate conditions on the prospects of crop yields, and change in the usage of agricultural land to assess and ultimately reduce the crop yield gaps. Several researchers reported the applications of bio simulation models to estimate the crop yield gaps in the last decade. The critical challenge remaining with these methods is scaling up of these approaches to assess the data collated between different time intervals from the broader geographical regions.

B) M. A. JAYARAM AND NETRA MARAD, "FUZZY INFERENCE SYSTEM FOR CROP PREDICTION", JOURNAL OF INTELLIGENT SYSTEMS, 2012.

The proposed system suggests an attempt to develop fuzzy inference systems for crop yield prediction. Physio morphological features of Sorghum were considered. A huge database (around 1000 records) of physio morphological features such as days of 50 percent showering, dead heart percentage, plant height, panicle length, panicle weight and number of primaries and the corresponding yield were considered for the development of the model. In order to

and out the sensitivity of parameters, one-to-one, two-to-one and three-to-one combinations of input and output were considered. The results have clearly shown that panicle length contributes forth yield as the lone parameter with almost one-to-one matching between predicted yield and actual value while panicle length and panicle weight in combination seemed to play a decisive role in contributing for the yield with the prediction accuracy rejected by very low RMS value. In hybrid plants, the morphological features such as plant height, panicle length, panicle weight, number of primaries and length of the leaves cannot be determined or predicted accurately. Therefore, sometimes this becomes a failure model.

C) A. D. BOSE, ''BIG DATA ANALYTICS IN AGRICULTURE''.

The proposed system suggests how Big Data Analytics combined with various structured and unstructured data helps in providing insight to farmers to make a decision as to which crops to grow and reduce losses due to unexpected or unpredictable disasters. In Section I the paper states that we can collect the data produced by sensors from the official databases that are usually maintained and governed by institutions. Here the author suggests we can collect and analyze the data in different stages in agriculture and see their influence in the big picture. It is dependent on two major factors, the push and pull factor. Visualization of agricultural data is done to simplify the complex, structured, and unstructured data. Interpretation of data can be done using methods like overviews, verifiable models, or in an Ad-Hoc manner graphs. the implementation of analytic techniques in agriculture had been discussed. The first method is an Intelligent crop recommendation system that considers all the factors such as soil conditions, temperature, rainfall and location. This system is further split into two different systems: the crop predictor, whose main task is to help agriculturists by recommending crops and the rainfall prediction system that predicts the occurrence of rainfall for each month across the year. The next method discussed was Precision Agriculture using Map-Reduce used to allow variable rates and inputs which help in the understanding of time and space variability in criterion. Here the data is obtained and pre-processed. Then map-reduce is performed, and 3D visualization is done to visualize the output. Further crop prediction using various machine learning approaches were discussed. A few of them were 1) Grey wolf optimization (GWO) technique 2)

K-means clustering 3) Apriori algorithm 4) Naive Baye. The author states that obstacles faced for agriculture are usually Technical or Organizational problems. The paper further mentions the problems faced in the big data analysis of agriculture data, majorly, availability, accessibility and scalability of data for analysis.

2.2 REFERENCE:

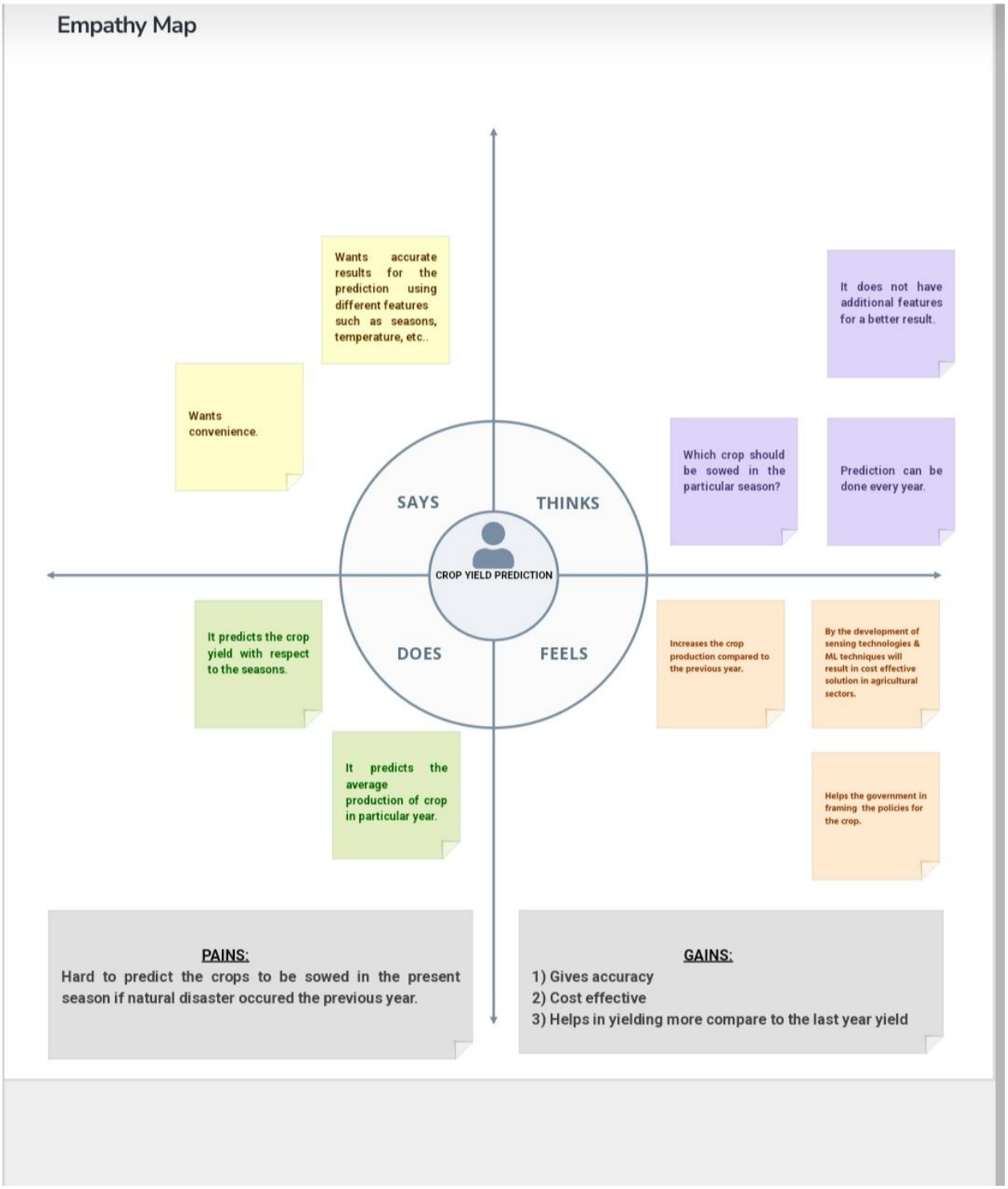
- **https://www.degruyter.com/document/doi/10.1515/jisys-2012- 0016/html**
- **https://ieeexplore.ieee.org/document/8697806**
- **https://www.researchgate.net/publication/339102917_Big_data_analytics_in_Agriculture**

2.3 PROBLEM STATEMENT DEFINITION:

To create a dashboard and perform analysis of crop production in India using IBM Cognos analytic platform. 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.

3. IDEATION & PROPOSED SYSTEM:

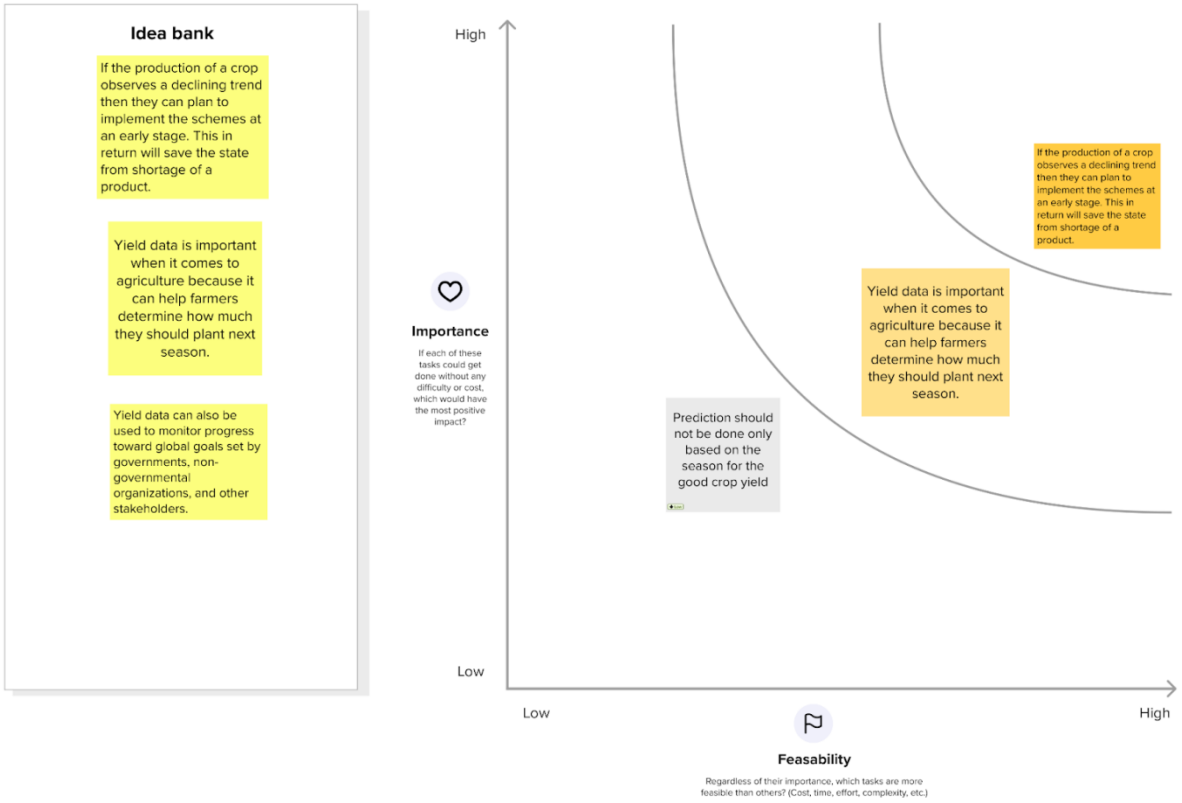
3.1 EMPATHY MAP CANVAS:



3.2 IDEATION & BRAINSTORMING:

A) BRAINSTORMING:

B) IDEA PRIORITIZATION:

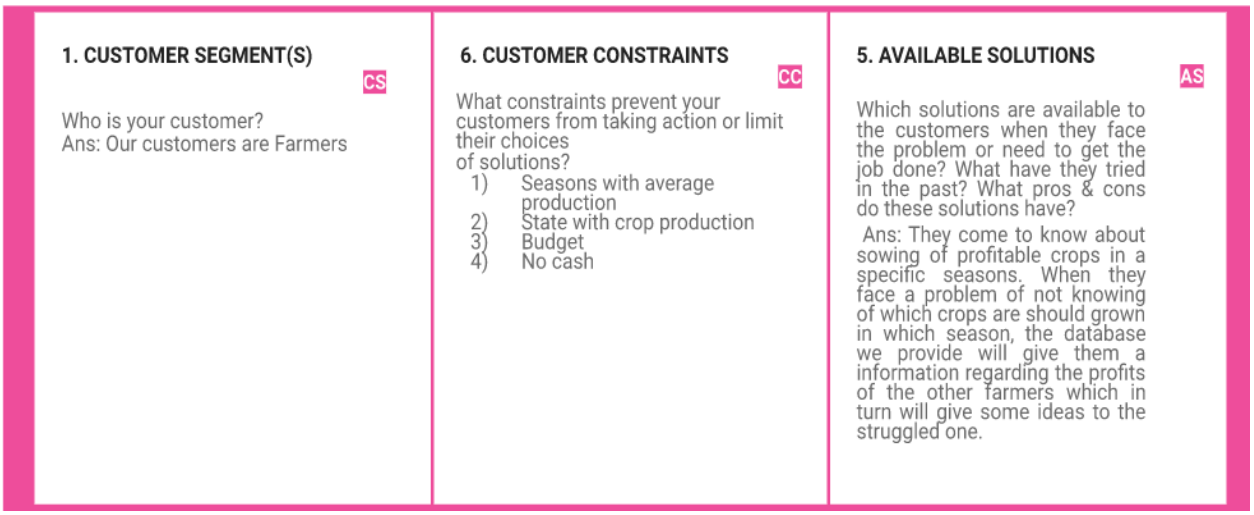


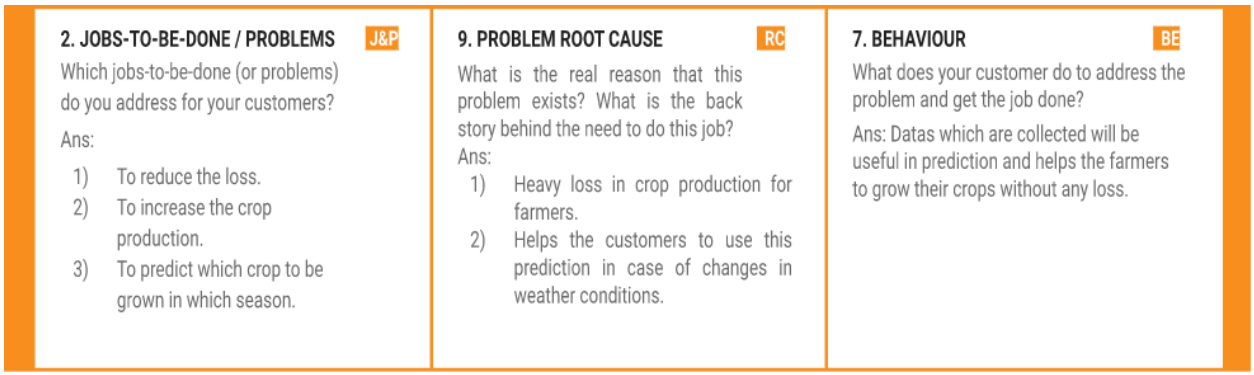
3.3 PROPOSED SOLUTION:

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	The accurate prediction of crop yield certainly benefits the farmers in choosing the right method to reduce the crop damage and get best prices for the crops.
3.	Novelty / Uniqueness	If the production of a crop observes a declining trend then they can plan to implement the **schemes** at an **early stage**. This in return will **save** the state from **shortage of a product**
4.	Social Impact / Customer Satisfaction	It is used to monitor progress towards a global set by governments, non-governmental organizations, and other stakeholders.
5.	Business Model (Revenue Model)	According to the revenue side it will yield more revenue to the farmers as well as to the governments.
6.	Scalability of the Solution	The data which are present in the datasets will be up to date. So it will help the customer to act accordingly

3.4 PROBLEM SOLUTION FIT:

The Problem-Solution Fit simply means that you have found a problem with your customer and that the solution you have realized for it actually solves the customer's problem. It helps entrepreneurs, marketers and corporate innovators identify behavioral patterns and recognize what would work and why.





3. TRIGGERS TR

What triggers customers to act?

Ans: The data which are present in the datasets in an appropriate manner. Due to the diagrammatic representation of the models, the farmers who are not well educated can also be able to work on this and get benefits.

4. EMOTIONS: BEFORE / AFTER EM

How do customers feel when they face a problem or a job and afterwards?

Before: Depressed-> Due to heavy loss in crop production and revenue.

After: Satisfied-> Due to loss reduction in crop production and increment in revenue.

10. YOUR SOLUTION SL

Our projects will enlighten the lives of many farmers by providing suitable information regarding the crop yield in a particular season, which in turn gives many profits to them even from the government side by giving agricultural loans etc,..

8. CHANNELS of BEHAVIOR CH

8.1 ONLINE

What kind of actions do customers take online?

In online mode they are given access to use all the options like crop yield, production, pesticides, agricultural loan, database, etc,..

8.2 OFFLINE

What kind of actions do customers take offline?

In Offline mode they won't be able to access some of the resources, but they can see the databases of the crop production.

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	User Details Farm Details
FR-4	Required Data	The previous year crop yield data set Farm yield methodology User data of the farmer Details of the Seasons and the Regions
FR-5	Analysis	Cleaning and analysis of the past year crop yields Visualizing the datasets using IBM Cognos
FR-6	Estimation	Creating the perfect data module through attractive stories, dashboard and reports to increase the understandability of data.

4.2 NON – FUNCTIONAL REQUIREMENT:

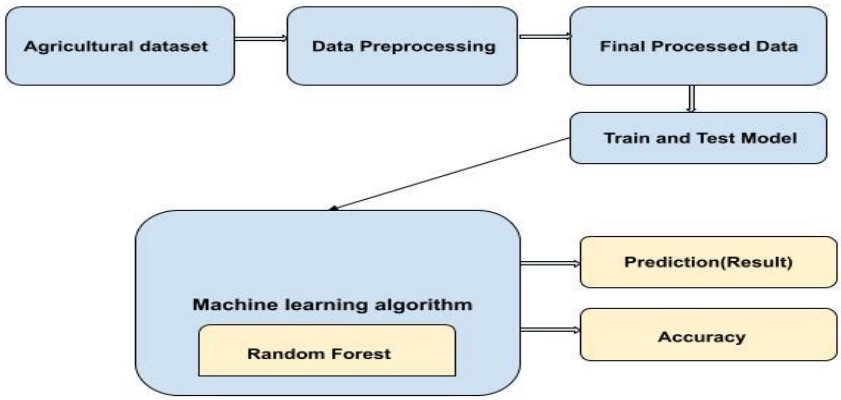
FR No.	Non-Functional Requirement	Description
NFR-1	**Usability**	From the given datasets , analysis is done and a report is created. Accordingly, sowing of crops is recommended.
NFR-2	**Security**	Usage of IBM COGNOS, will provide secure user information(Data Visualization)
NFR-3	**Reliability**	Using the interactive data visual dashboards , we can easily understand the data reports.
NFR-4	**Performance**	Interaction makes better performance between all users and impresses by the data visuals advice.
NFR-5	**Availability**	The dashboard is easily available and accessible in smart phones and PC's.
NFR-6	**Scalability**	Prediction of crops for the forthcoming year can be done .It gives you a variety of crops to choose from our region. Also to know the better profitability of crops.

5. PROJECT DESIGN:

5.1 DATA FLOW DIAGRAM:

Project flow describes a preset sequence of activities required to plan, produce, deliver and maintain project product, along with information, materials, and resources required by the project. Project flow is a convenient way to define and plan projects.

Project flow for estimating the crop yield using data analytics is shown below.



5.2 SOLUTION AND TECHNOLOGY ARCHITECTURE:

The Deliverables hall include the architectural diagram as below and the information as per the table 1 & table 2.

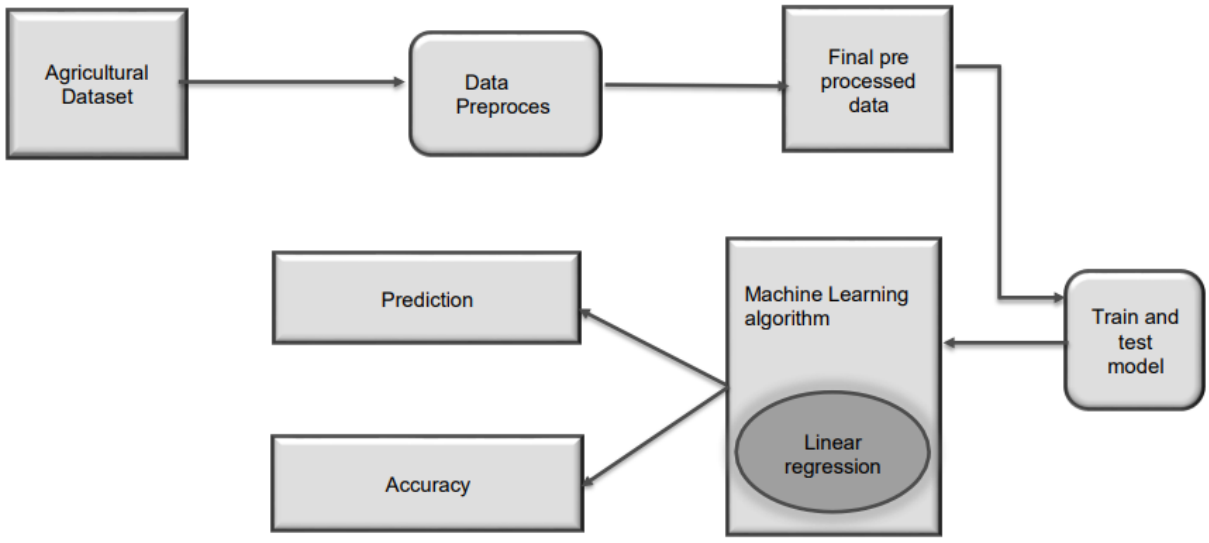


Table-1: Components & Technologies:

S. No	Component	Description	Technology
1.	User Interface	How user interacts with application e.g. Web UI, Mobile App ,Chat bot etc.	HTML, CSS, JavaScript.
2.	Applicationlogic1	Login as a user in the application	Java/Python
3.	Applicationlogic2	Login as admin in the application	IBM Watson STT service
4.	Applicationlogic3	Login as merchants in the application	IBM Watson Assistant
5.	Database	Data related to crop production in previous and also crop data.	MySQL , NoSQL , etc.
6.	Cloud Database	Database Service on Cloud	IBMDB2,IBM Cloudant etc.
7.	File Storage	File storage requirements	IBM Block Storage or Other Storage Service or Local File system
8.	ExternalAPI-1	Weather API are application programming interface that allow you to connect to large databases.	IBM Weather API ,etc.

9.	ExternalAPI-2	Soil testing is a quick and accurate method to determine the relative acidity of the soil and the level of several essential nutrient needed.	Soil API, etc.
10.	Machine Learning Model	It is mostly used for finding out the relationship between variables and forecasting	Linear Regression
11.	Infrastructure(Server/Cloud)	Application Deployment on Local System/Cloud Local Server Configuration CloudServerConfiguration:l1	Local , Cloud Foundry ,Kubernetes ,etc.

Table-2:ApplicationCharacteristics:

S. No	Characteristics	Description	Technology
1.	Open-Source Frameworks	Bootstrap is a free ,open source front-end development frame work	Bootstrap ,React etc.,
2.	Security Implementations	Improves user experience and provides greater security.	Authentication etc.
3.	Scalable Architecture	A3-tier architecture where in application gets data from various sources ,manipulates it, stores the min IBM Cloud and Cognos.	IBM Cloud, IBM Cognos.
4.	Availability	The application is being developed is made available to all users	Cognos Analytics

5.	Performance	Multiple technologies and services that will improve the usability in agriculture activities.	Robots, IOT agriculture sensors.

5.3 USER STORIES:

User Story Number	User Story / Task
USN-1	Understanding the data set .
USN-2	Loading the data set.
USN-3	Convert the data into required format
USN-4	Explore the data's which is uploaded in the IBM cognos
USN-5	Creating the data visualization chart
USN-6	Creating a dashboard
USN-7	Estimation of accuracy using random forest algorithm
USN-8	Export the analytics

6.PROJECT PLANNING AND SCHEDULING:

6.1 SPRINT PLANNING AND ESTIMATION:

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-1	Working with the data set	USN-1	Understanding the data set .	10	Medium	Sneha
Sprint-1	Working with the data set	USN-2	Loading the data set.	10	High	Sindhuja
Sprint-2	Prepare the data	USN-3	Convert the data into required format	10	Medium	Sree Nandini
Sprint-2	Data exploration	USN-4	Explore the data's which is uploaded in the IBM cognos	10	Medium	Shrinithi
Sprint-3	Data visualization	USN-5	Creating the data visualization chart	10	High	Sindhuja
Sprint-3	Dashboard	USN-6	Creating a dashboard	10	High	Sree Nandini
Sprint-3	Visualization	USN-7	Estimation of accuracy using random forest algorithm	10	High	Shrinth
Sprint-4	Export	USN-8	Export the analytics	10	High	Sneha

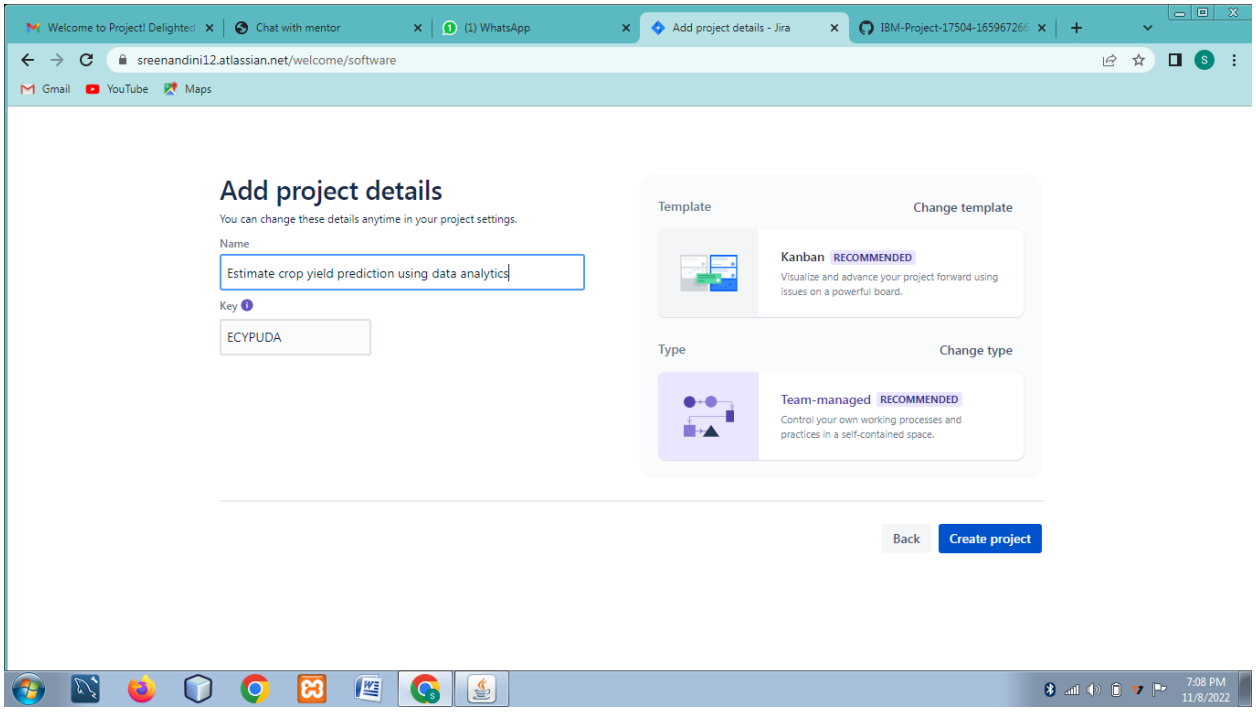
6.2 SPRINT DELIVERY SCHEDULE:

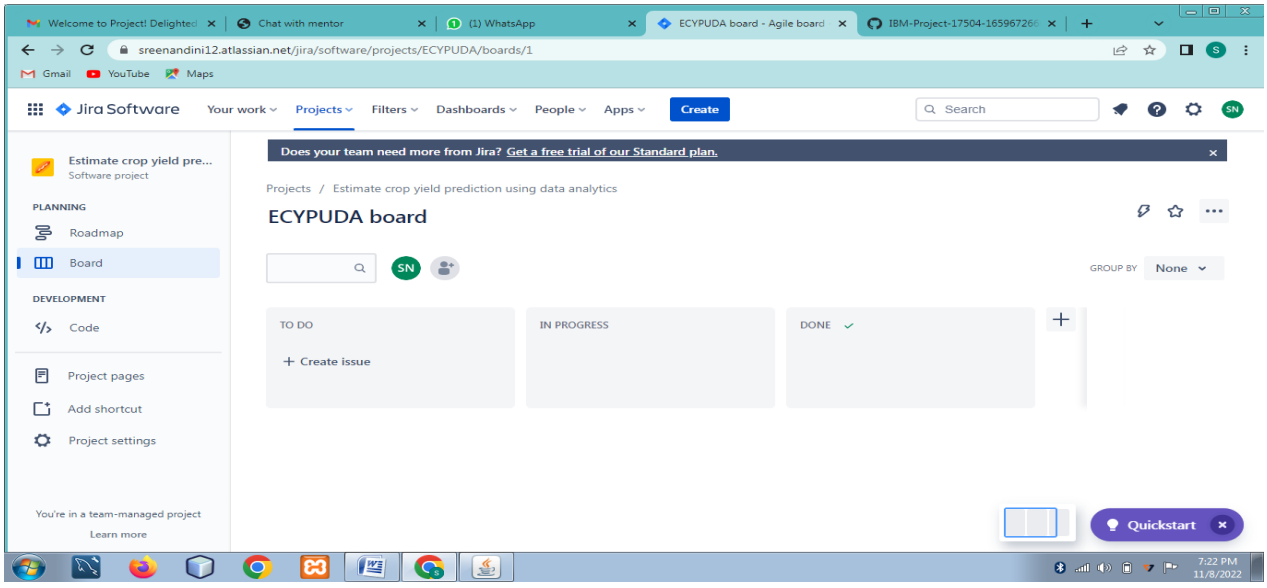
Sprint	Total storyp oints	Duratio n	Sprint Start Date	Sprint End Date (Planned)	Story Points Completed (as on Planned End Date)	Sprint Release Date (Actual)
Sprint-1	20	4 Days	01 Nov 2022	04 Nov 2022	20	08 - 09 Nov 2022
Sprint-2	20	5 Days	05 Nov 2022	10 Nov 2022	20	10 - 13 Nov 2022
Sprint-3	20	4 Days	11 Nov 2022	14 Nov 2022	20	14 - 16 Nov 2022

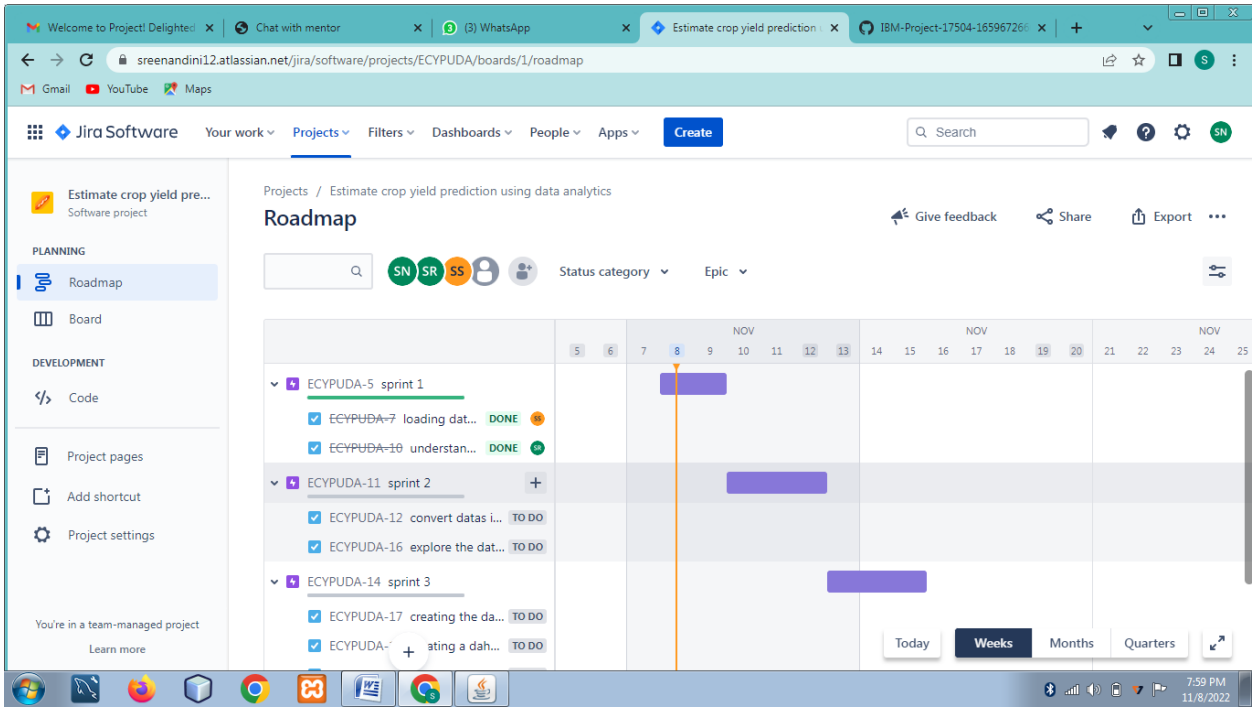
Sprint-4	20	4 Days	15 Nov 2022	19 Nov 2022	20	17 - 19 Nov 2022

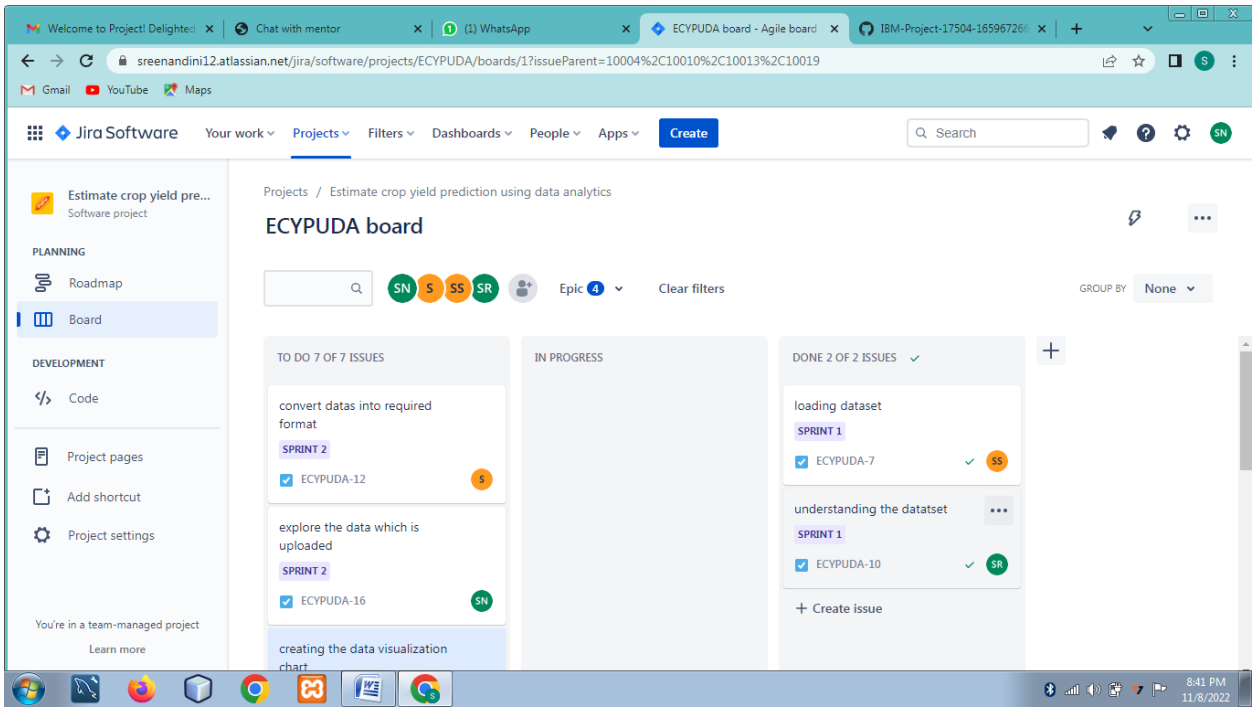
6.3 REPORT FROM JIRA:

A) SPRINT 1:

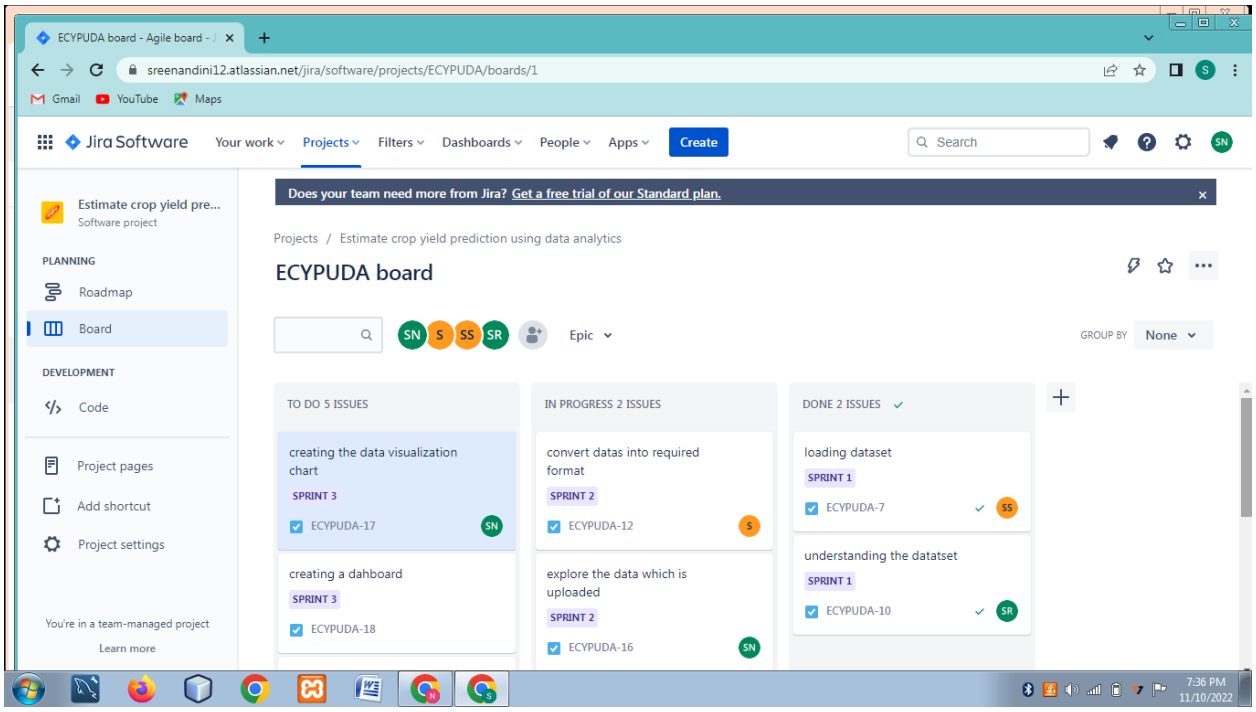


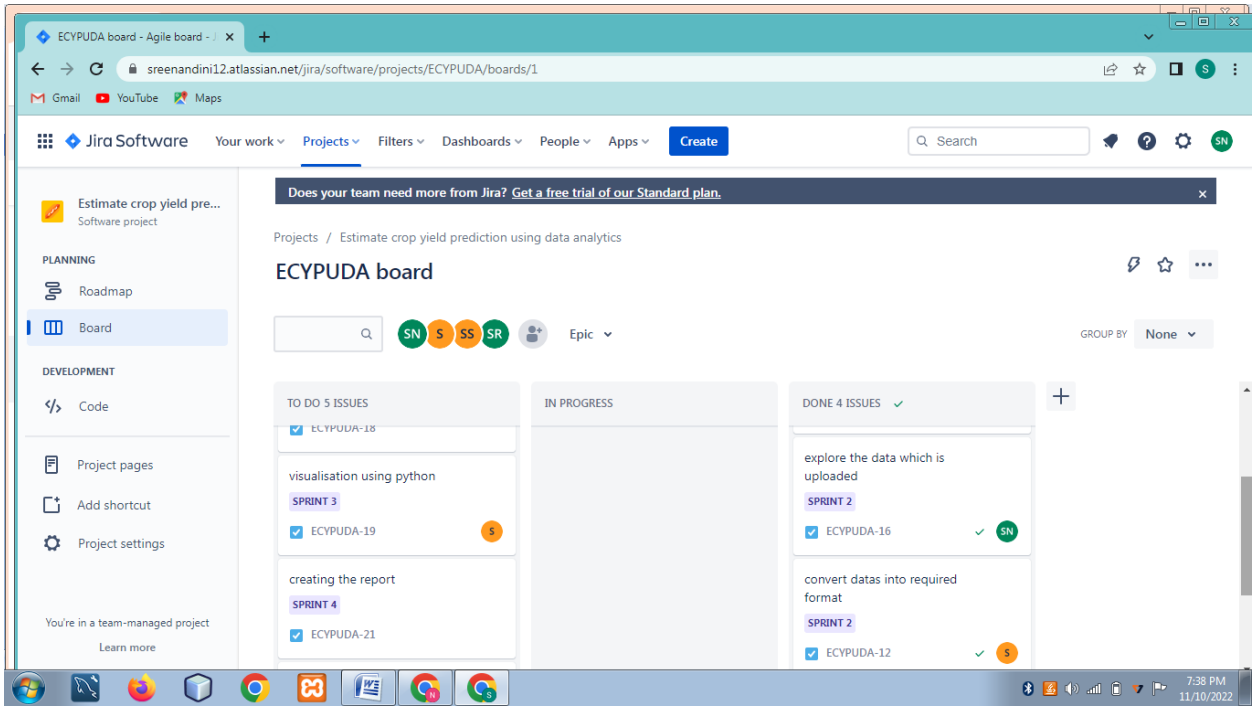


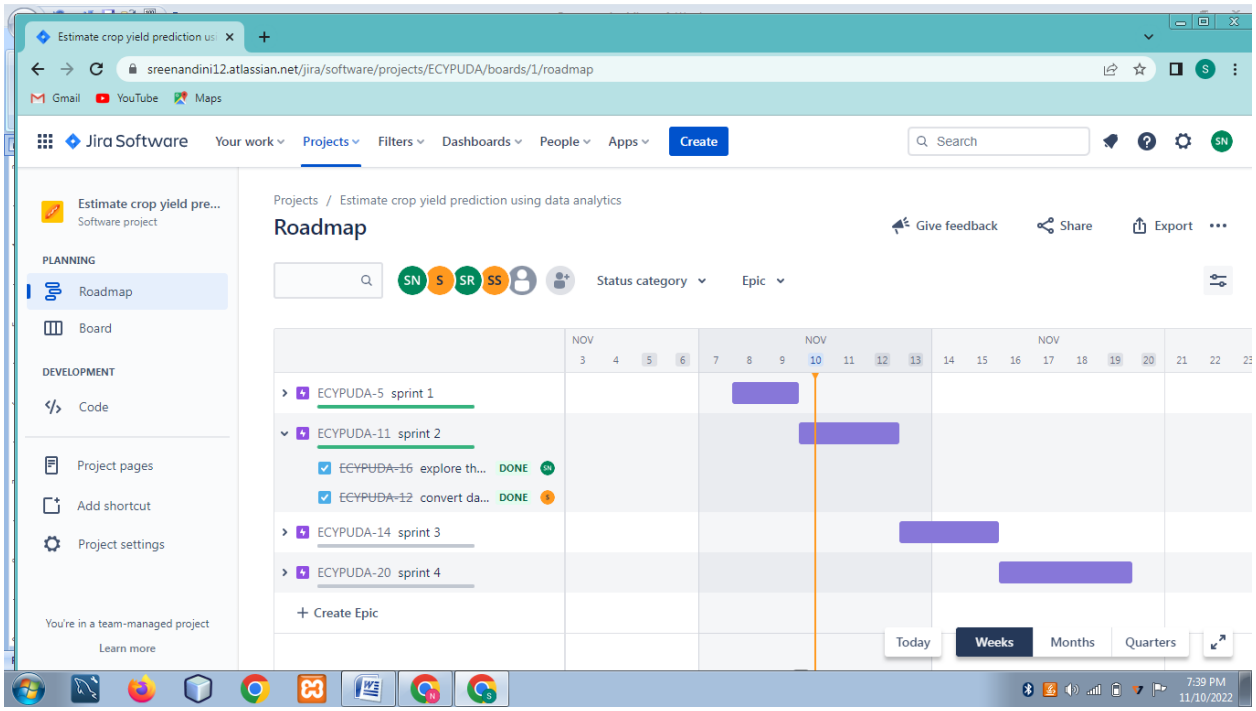

Jira Software
Your work
Projects
Filters
Dashboards
People
Apps
Create
Search
Estimate crop yield pre...
Software project
PLANNING
Roadmap
Board
DEVELOPMENT
Code
Project pages
Add shortcut
Project settings
You're in a team-managed project
Learn more
Projects / Estimate crop yield prediction using data analytics
Roadmap
Give feedback
Share
Export
Status category
Epic
NOV
ECYPUDA-5 sprint 1
ECYPUDA-7 loading dat... DONE
ECYPUDA-10 understan... DONE
ECYPUDA-11 sprint 2
ECYPUDA-12 convert datas i... TO DO
ECYPUDA-16 explore the dat... TO DO
ECYPUDA-14 sprint 3
ECYPUDA-17 creating the da... TO DO
Today
Weeks
Months
Quarters


Jira Software
Your work
Projects
Filters
Dashboards
People
Apps
Create
Search
Estimate crop yield pre...
Software project
PLANNING
Roadmap
Board
DEVELOPMENT
Code
Project pages
Add shortcut
Project settings
You're in a team-managed project
Learn more
Projects / Estimate crop yield prediction using data analytics
ECYPUDA board
Epic 4
Clear filters
GROUP BY
None
TO DO 7 OF 7 ISSUES
IN PROGRESS
DONE 2 OF 2 ISSUES
convert datas into required format
SPRINT 2
ECYPUDA-12
explore the data which is uploaded
SPRINT 2
ECYPUDA-16
creating the data visualization chart
loading dataset
SPRINT 1
ECYPUDA-7
understanding the datatset
SPRINT 1
ECYPUDA-10
Create issue

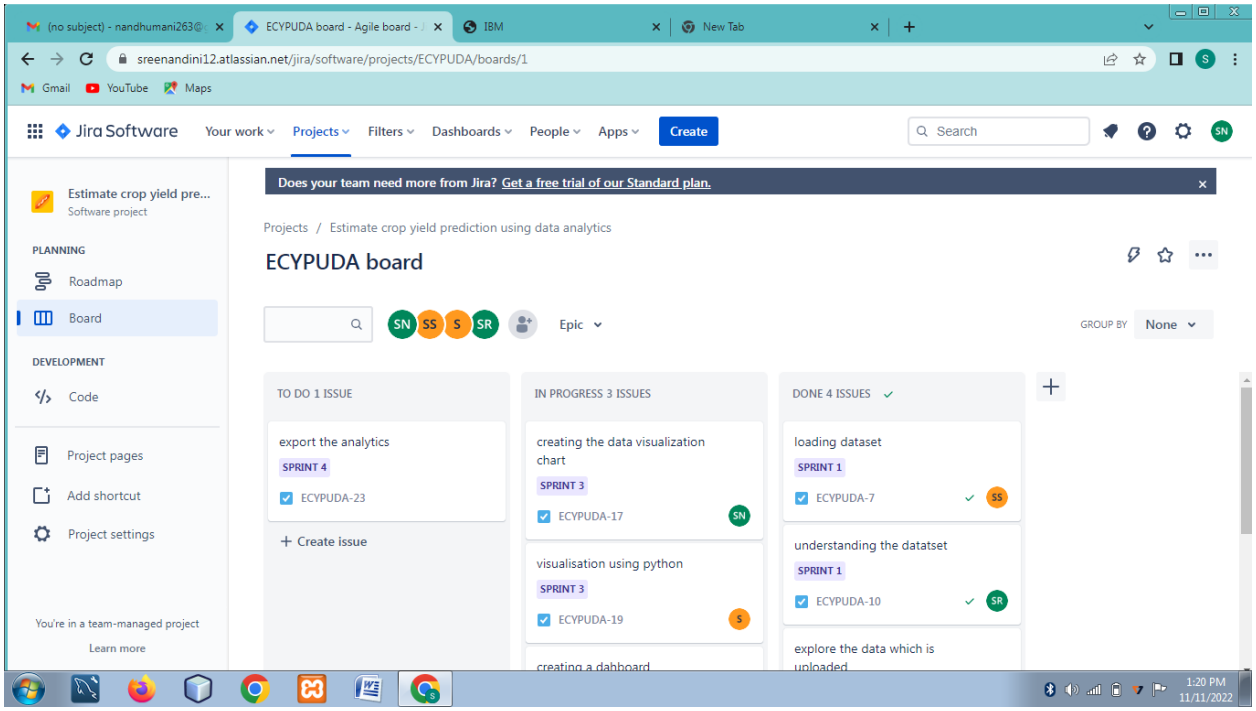
B) SPRINT 2:

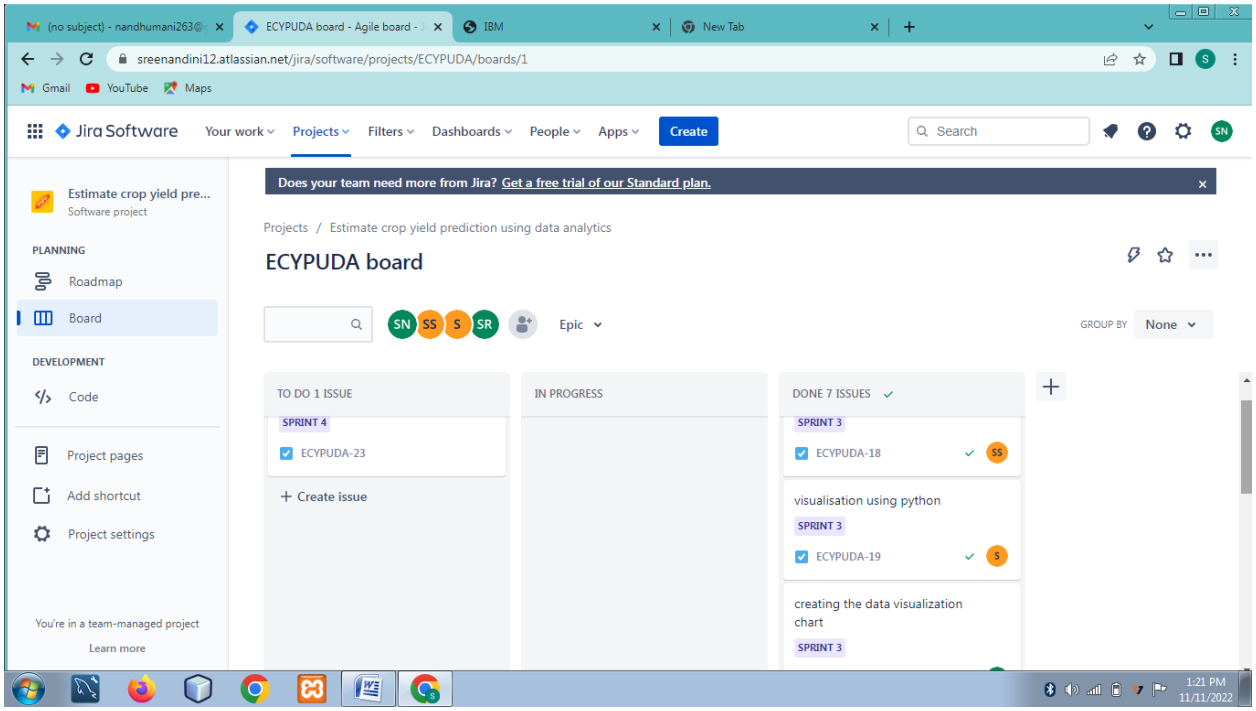

ECYPUDA board - Agile board -
sreenandini12.atlassian.net/jira/software/projects/ECYPUDA/boards/1
Gmail
YouTube
Maps
Jira Software
Your work
Projects
Filters
Dashboards
People
Apps
Create
Search
Estimate crop yield pre...
Software project
PLANNING
Roadmap
Board
DEVELOPMENT
Code
Project pages
Add shortcut
Project settings
You're in a team-managed project
Learn more
Does your team need more from Jira? Get a free trial of our Standard plan.
Projects / Estimate crop yield prediction using data analytics
ECYPUDA board
Epic
GROUP BY
None
TO DO 5 ISSUES
creating the data visualization chart
SPRINT 3
ECYPUDA-17
creating a dahboard
SPRINT 3
ECYPUDA-18
IN PROGRESS 2 ISSUES
convert datas into required format
SPRINT 2
ECYPUDA-12
explore the data which is uploaded
SPRINT 2
ECYPUDA-16
DONE 2 ISSUES
loading dataset
SPRINT 1
ECYPUDA-7
understanding the datatset
SPRINT 1
ECYPUDA-10
7:36 PM
11/10/2022

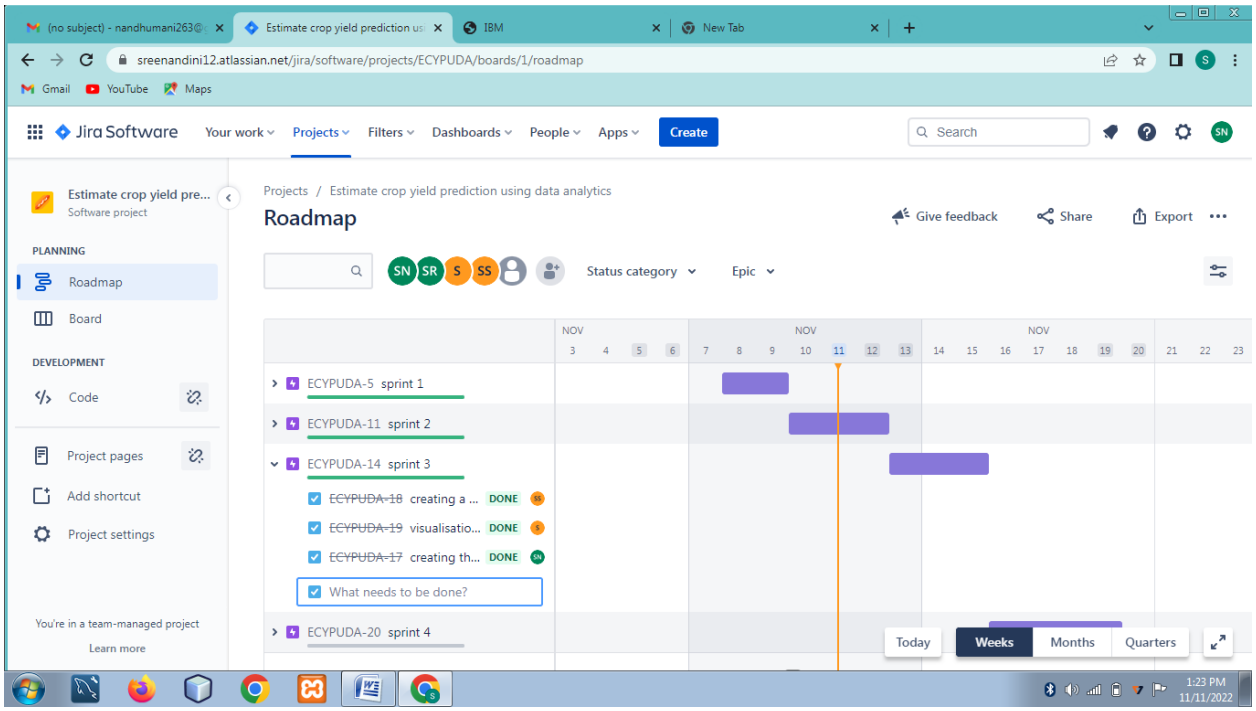

ECYPUDA board - Agile board -
sreenandini12.atlassian.net/jira/software/projects/ECYPUDA/boards/1
Gmail
YouTube
Maps
Jira Software
Your work
Projects
Filters
Dashboards
People
Apps
Create
Search
Estimate crop yield pre...
Software project
PLANNING
Roadmap
Board
DEVELOPMENT
Code
Project pages
Add shortcut
Project settings
You're in a team-managed project
Learn more
Does your team need more from Jira? Get a free trial of our Standard plan.
Projects / Estimate crop yield prediction using data analytics
ECYPUDA board
Epic
GROUP BY
None
TO DO 5 ISSUES
visualisation using python
SPRINT 3
ECYPUDA-19
creating the report
SPRINT 4
ECYPUDA-21
IN PROGRESS
DONE 4 ISSUES
explore the data which is uploaded
SPRINT 2
ECYPUDA-16
convert datas into required format
SPRINT 2
ECYPUDA-12
7:38 PM
11/10/2022


Estimate crop yield prediction us
sreenandini12.atlassian.net/jira/software/projects/ECYPUDA/boards/1/roadmap
Gmail
YouTube
Maps
Jira Software
Your work
Projects
Filters
Dashboards
People
Apps
Create
Search
Estimate crop yield pre...
Software project
PLANNING
Roadmap
Board
DEVELOPMENT
Code
Project pages
Add shortcut
Project settings
You're in a team-managed project
Learn more
Projects / Estimate crop yield prediction using data analytics
Roadmap
Give feedback
Share
Export
Status category
Epic
NOV
ECYPUDA-5 sprint 1
ECYPUDA-11 sprint 2
ECYPUDA-16 explore th... DONE
ECYPUDA-12 convert da... DONE
ECYPUDA-14 sprint 3
ECYPUDA-20 sprint 4
Create Epic
Today
Weeks
Months
Quarters
7:39 PM
11/10/2022

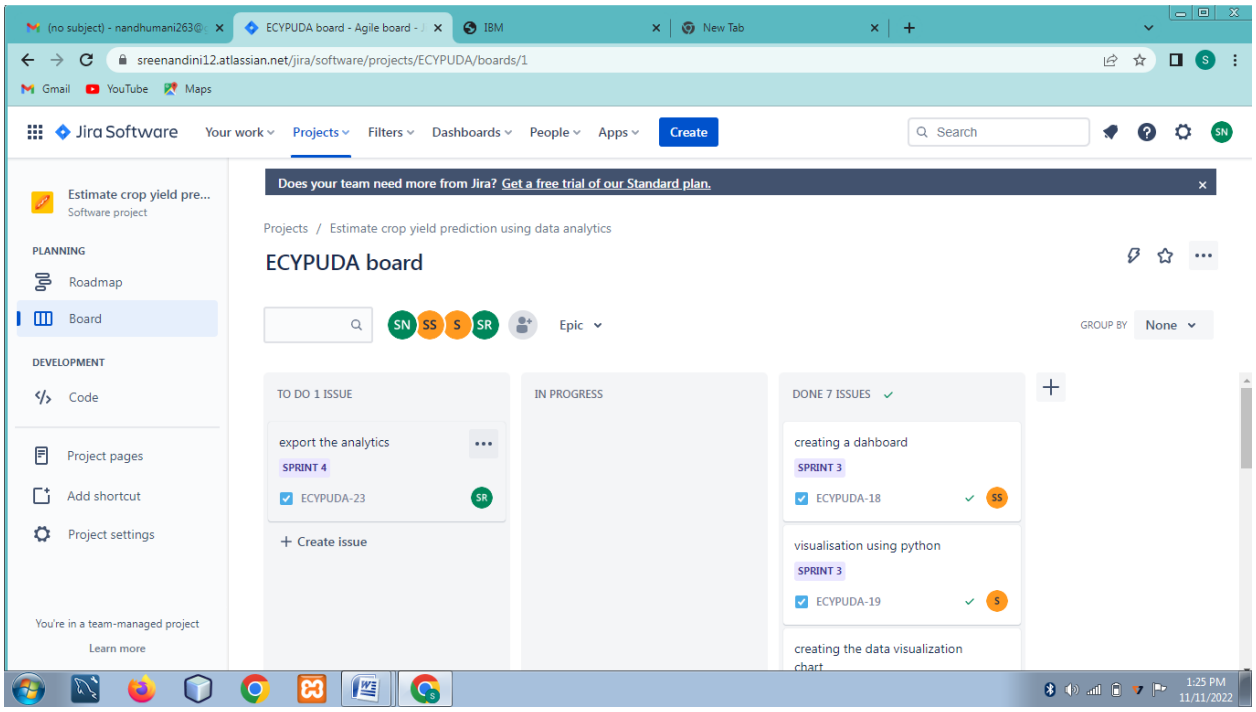
C) SPRINT 3:

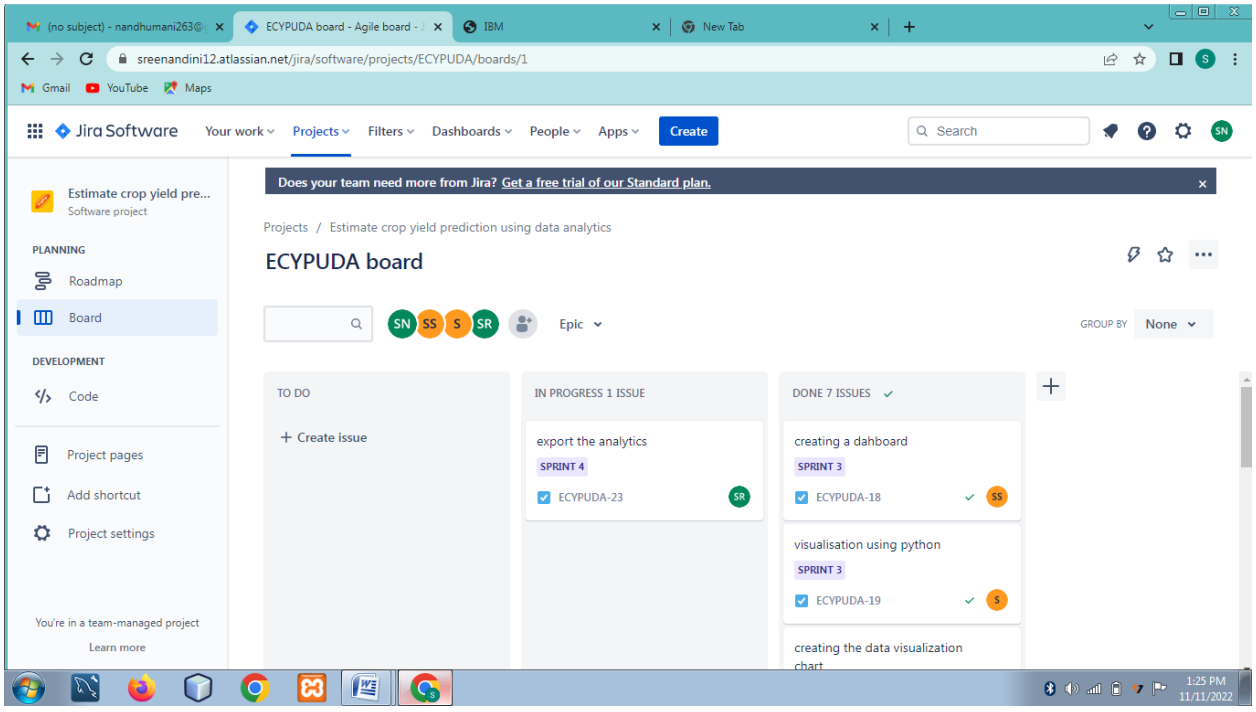

sreenandini12.atlassian.net/jira/software/projects/ECYPUDA/boards/1
Jira Software
Your work
Projects
Filters
Dashboards
People
Apps
Create
Search
Does your team need more from Jira? Get a free trial of our Standard plan.
Estimate crop yield pre...
Software project
PLANNING
Roadmap
Board
DEVELOPMENT
Code
Project pages
Add shortcut
Project settings
You're in a team-managed project
Learn more
Projects / Estimate crop yield prediction using data analytics
ECYPUDA board
Epic
GROUP BY
None
TO DO 1 ISSUE
export the analytics
SPRINT 4
ECYPUDA-23
Create issue
IN PROGRESS 3 ISSUES
creating the data visualization chart
SPRINT 3
ECYPUDA-17
visualisation using python
SPRINT 3
ECYPUDA-19
DONE 4 ISSUES
loading dataset
SPRINT 1
ECYPUDA-7
understanding the datatset
SPRINT 1
ECYPUDA-10
explore the data which is uploaded
1:20 PM
11/11/2022

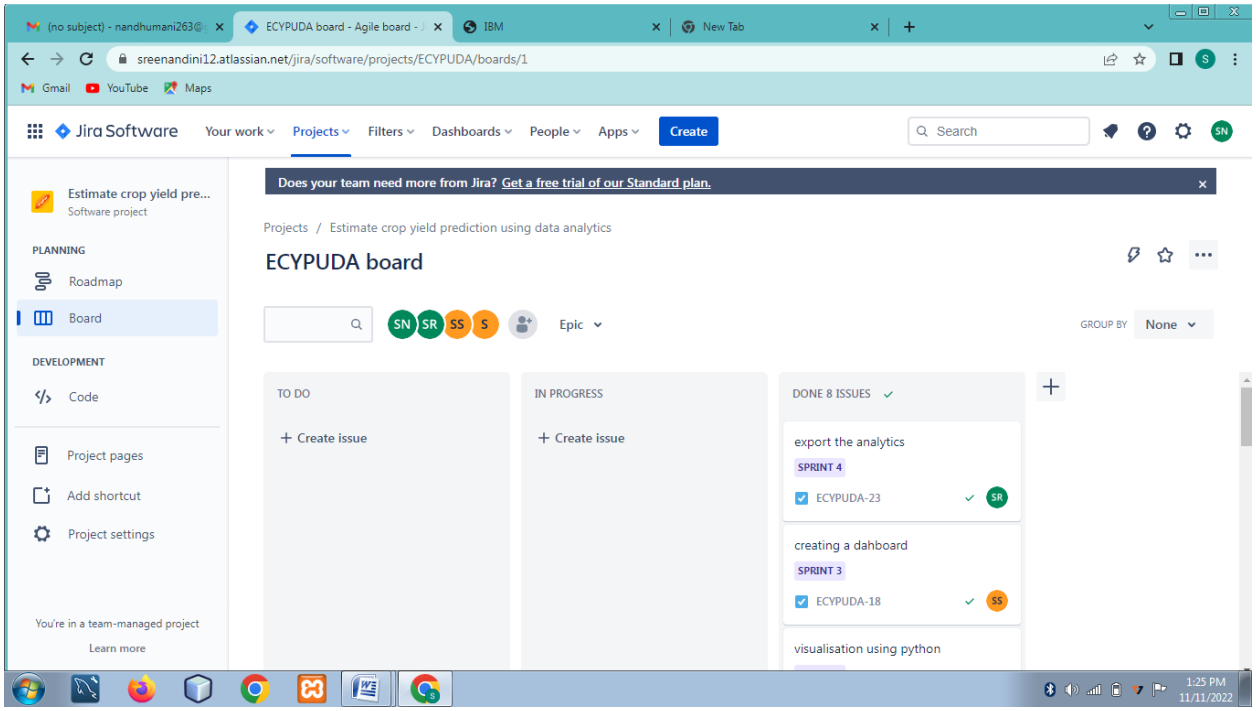

ECYPUDA board - Agile board
sreenandini12.atlassian.net/jira/software/projects/ECYPUDA/boards/1
Jira Software
Your work
Projects
Filters
Dashboards
People
Apps
Create
Search
Estimate crop yield pre...
Software project
PLANNING
Roadmap
Board
DEVELOPMENT
Code
Project pages
Add shortcut
Project settings
You're in a team-managed project
Learn more
Does your team need more from Jira? Get a free trial of our Standard plan.
Projects / Estimate crop yield prediction using data analytics
ECYPUDA board
Epic
GROUP BY
None
TO DO 1 ISSUE
SPRINT 4
ECYPUDA-23
Create issue
IN PROGRESS
DONE 7 ISSUES
SPRINT 3
ECYPUDA-18
visualisation using python
SPRINT 3
ECYPUDA-19
creating the data visualization chart
SPRINT 3
1:21 PM
11/11/2022

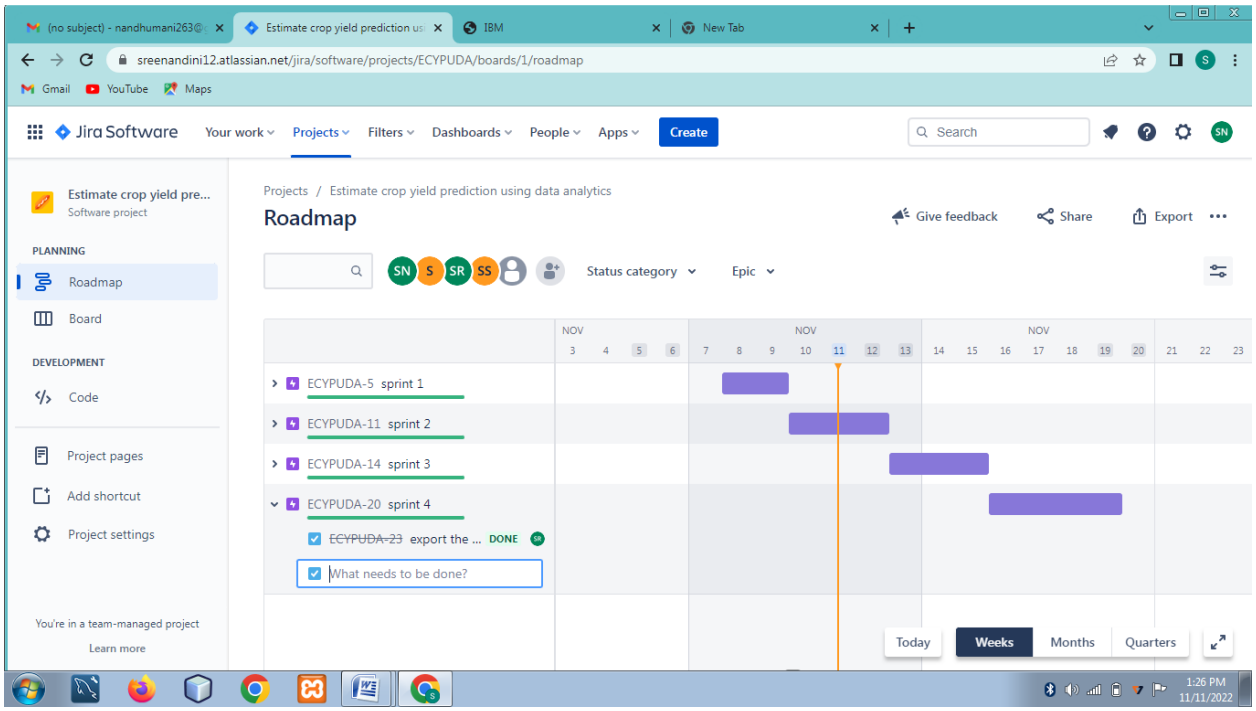

sreenandini12.atlassian.net/jira/software/projects/ECYPUDA/boards/1/roadmap
Jira Software
Your work
Projects
Filters
Dashboards
People
Apps
Create
Search
Estimate crop yield pre...
Software project
PLANNING
Roadmap
Board
DEVELOPMENT
Code
Project pages
Add shortcut
Project settings
You're in a team-managed project
Learn more
Projects / Estimate crop yield prediction using data analytics
Roadmap
Give feedback
Share
Export
Status category
Epic
NOV
ECYPUDA-5 sprint 1
ECYPUDA-11 sprint 2
ECYPUDA-14 sprint 3
ECYPUDA-18 creating a ...
ECYPUDA-19 visualisatio...
ECYPUDA-17 creating th...
DONE
What needs to be done?
ECYPUDA-20 sprint 4
Today
Weeks
Months
Quarters
1:23 PM
11/11/2022

D) SPRINT 4:


sreenandini12.atlassian.net/jira/software/projects/ECYPUDA/boards/1
Jira Software
Your work
Projects
Filters
Dashboards
People
Apps
Create
Search
Estimate crop yield pre...
Software project
PLANNING
Roadmap
Board
DEVELOPMENT
Code
Project pages
Add shortcut
Project settings
You're in a team-managed project
Learn more
Does your team need more from Jira? Get a free trial of our Standard plan.
Projects / Estimate crop yield prediction using data analytics
ECYPUDA board
Epic
GROUP BY
None
TO DO 1 ISSUE
export the analytics
SPRINT 4
ECYPUDA-23
+ Create issue
IN PROGRESS
DONE 7 ISSUES
creating a dahboard
SPRINT 3
ECYPUDA-18
visualisation using python
SPRINT 3
ECYPUDA-19
creating the data visualization
chart
1:25 PM
11/11/2022


sreenandini12.atlassian.net/jira/software/projects/ECYPUDA/boards/1
Jira Software
Your work
Projects
Filters
Dashboards
People
Apps
Create
Search
Estimate crop yield pre...
Software project
PLANNING
Roadmap
Board
DEVELOPMENT
Code
Project pages
Add shortcut
Project settings
You're in a team-managed project
Learn more
Does your team need more from Jira? Get a free trial of our Standard plan.
Projects / Estimate crop yield prediction using data analytics
ECYPUDA board
Epic
GROUP BY
None
TO DO
+ Create issue
IN PROGRESS 1 ISSUE
export the analytics
SPRINT 4
ECYPUDA-23
DONE 7 ISSUES
creating a dahboard
SPRINT 3
ECYPUDA-18
visualisation using python
SPRINT 3
ECYPUDA-19
creating the data visualization
chart
1:25 PM
11/11/2022


(no subject) - nandhumani263@
ECYPUDA board - Agile board -
IBM
New Tab
sreenandini12.atlassian.net/jira/software/projects/ECYPUDA/boards/1
Gmail
YouTube
Maps
Jira Software
Your work
Projects
Filters
Dashboards
People
Apps
Create
Search
Estimate crop yield pre...
Software project
PLANNING
Roadmap
Board
DEVELOPMENT
Code
Project pages
Add shortcut
Project settings
You're in a team-managed project
Learn more
Does your team need more from Jira? Get a free trial of our Standard plan.
Projects / Estimate crop yield prediction using data analytics
ECYPUDA board
Epic
GROUP BY
None
TO DO
+ Create issue
IN PROGRESS
+ Create issue
DONE 8 ISSUES
export the analytics
SPRINT 4
ECYPUDA-23
creating a dahboard
SPRINT 3
ECYPUDA-18
visualisation using python
1:25 PM
11/11/2022

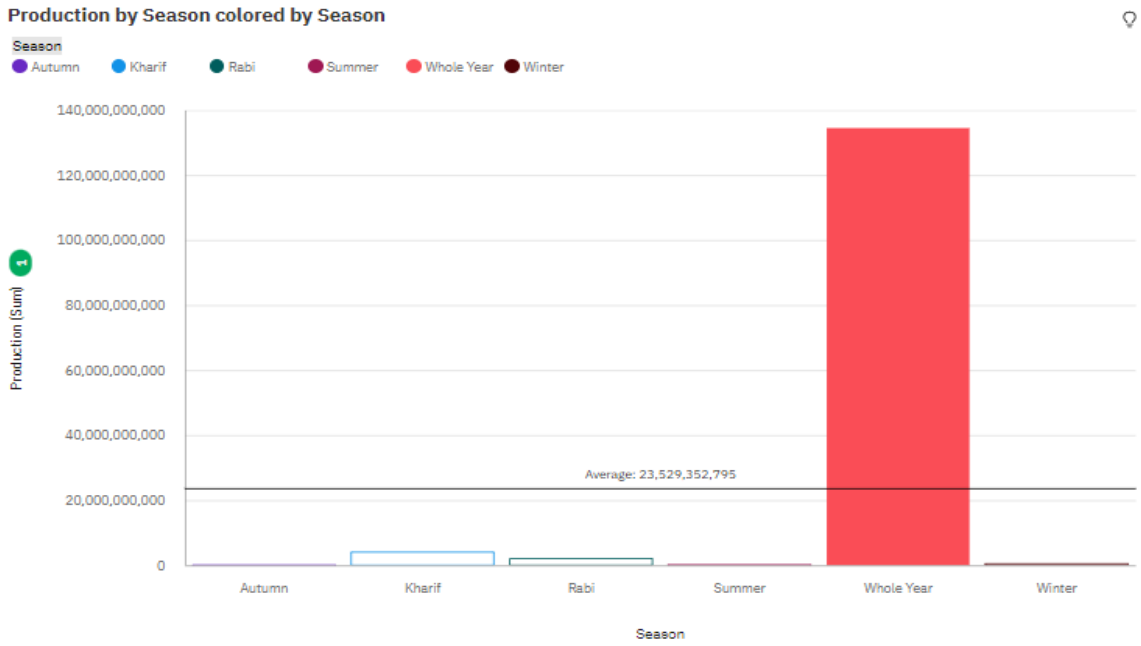

(no subject) - nandhumani263@
Estimate crop yield prediction us
IBM
New Tab
sreenandini12.atlassian.net/jira/software/projects/ECYPUDA/boards/1/roadmap
Gmail
YouTube
Maps
Jira Software
Your work
Projects
Filters
Dashboards
People
Apps
Create
Search
Estimate crop yield pre...
Software project
PLANNING
Roadmap
Board
DEVELOPMENT
Code
Project pages
Add shortcut
Project settings
You're in a team-managed project
Learn more
Projects / Estimate crop yield prediction using data analytics
Roadmap
Give feedback
Share
Export
Status category
Epic
NOV
ECYPUDA-5 sprint 1
ECYPUDA-11 sprint 2
ECYPUDA-14 sprint 3
ECYPUDA-20 sprint 4
ECYPUDA-23 export the ...
DONE
What needs to be done?
Today
Weeks
Months
Quarters
1:26 PM
11/11/2022

7. CODING AND SOLUTIONING:

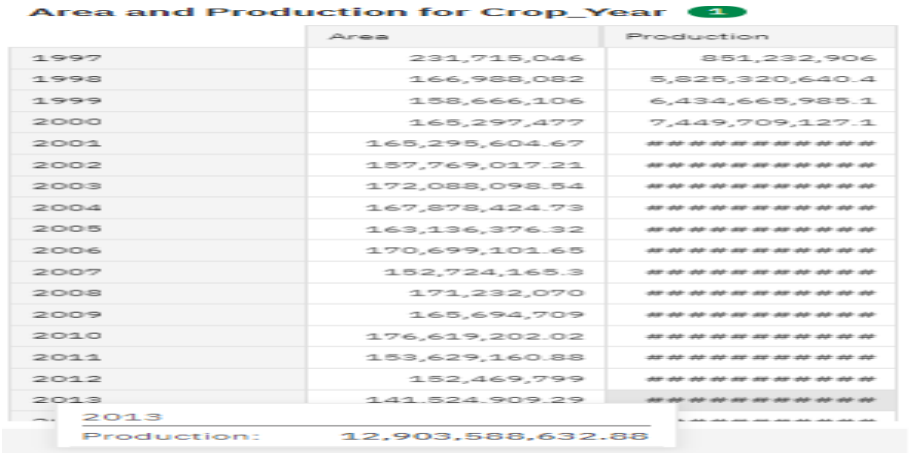
7.1 FEATURE 1:

A) DATA EXPLORATION:

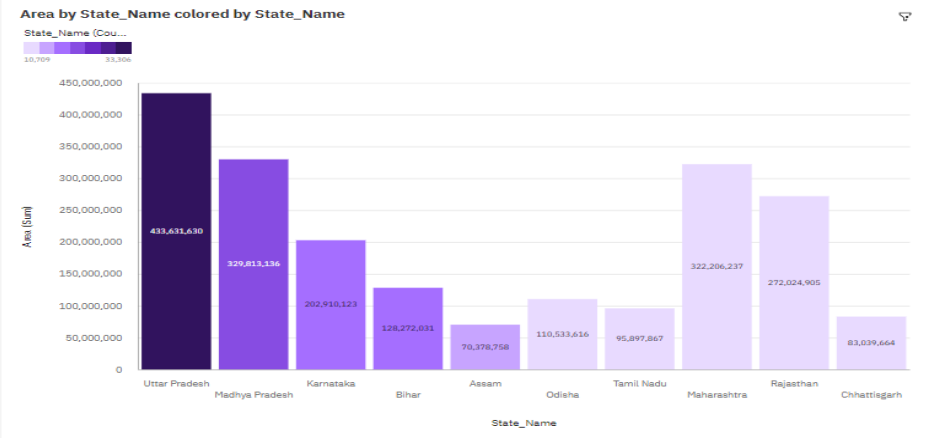
1) Seasons with average production:



2) With years usage of area and production:



3) Top 10 states with the most area.



4) States with crop production with seasons:

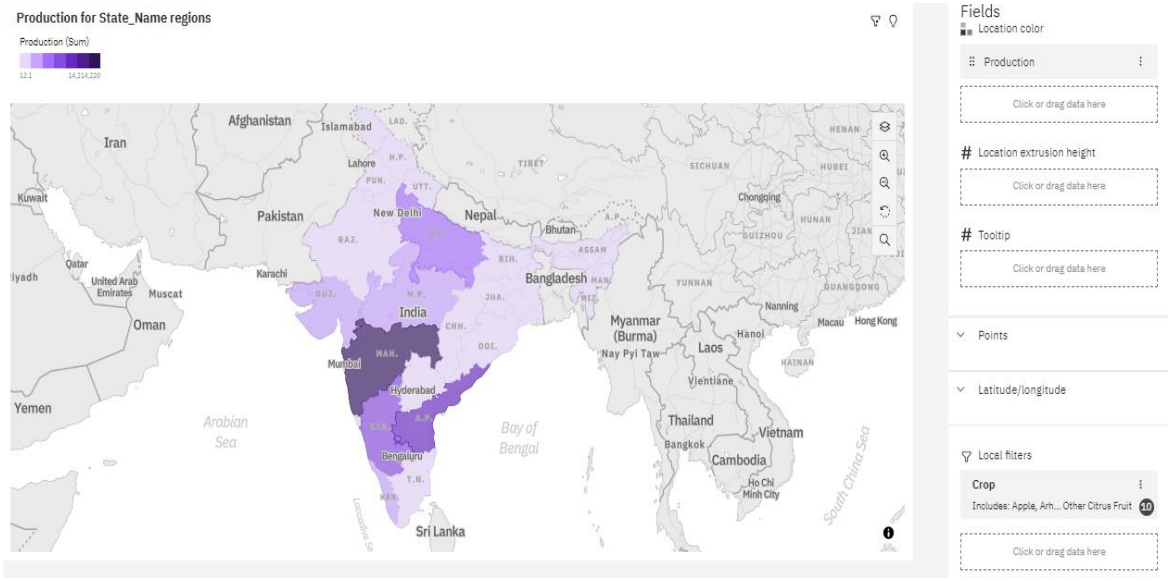
Production for Season, State_Name and Crop

Production		Autumn	Kharif	Rabi	Summer	Whole Year	Winter	Summary
Andaman and Ni...	Arecanut	(no value)	14,500	5,800	(no value)	27,735.81	(no value)	**48,035.81**
	Arhar/Tur	(no value)	(no value)	104	(no value)	(no value)	(no value)	**104**
	Banana	(no value)	(no value)	(no value)	(no value)	97,424.65	(no value)	**97,424.65**
	Black pepper	(no value)	(no value)	120	(no value)	604.5	(no value)	**724.5**
	Cashewnut	(no value)	(no value)	310	(no value)	1,374.79	(no value)	**1,684.79**
	Coconut	(no value)	(no value)	(no value)	(no value)	717,790,000	(no value)	**717,790,000**
	Dry chillies	(no value)	(no value)	575	(no value)	3,443.3	(no value)	**4,018.3**
	Dry ginger	(no value)	(no value)	1,850	(no value)	10,825.6	(no value)	**12,675.6**
	Groundnut	(no value)	(no value)	14.4	(no value)	(no value)	(no value)	**14.4**
	Maize	(no value)	(no value)	367.62	(no value)	(no value)	(no value)	**367.62**
	Moong(Green Gr...	(no value)	(no value)	575.5	(no value)	(no value)	(no value)	**575.5**
	Other Kharif puls...	(no value)	649	(no value)	(no value)	(no value)	(no value)	**649**
	Rice	23,916	199,090.06	(no value)	(no value)	(no value)	(no value)	**223,006.06**
	Sugarcane	1,332.95	(no value)	(no value)	(no value)	29,305.72	(no value)	**30,638.67**
	Sunflower	(no value)	(no value)	2.4	(no value)	(no value)	(no value)	**2.4**
	Sweet potato	(no value)	(no value)	923	(no value)	2,142.35	(no value)	**3,065.35**

Production for Season, State_Name and Crop 1

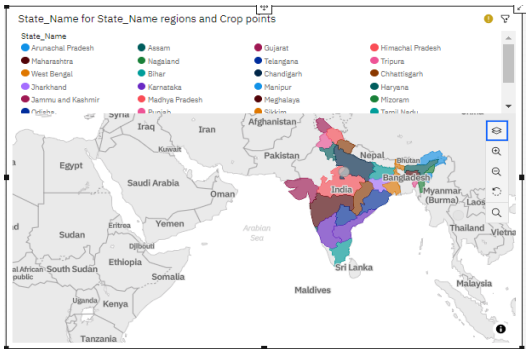
Production		Autumn	Kharif	Rabi	Summer	Whole Year	Winter	Summary
Haryana	Peas & beans (P...	(no value)	(no value)	19,624	(no value)	(no value)	(no value)	**19,624**
	Potato	(no value)	(no value)	(no value)	(no value)	3,621,500	(no value)	**3,621,500**
	Rapeseed &Must...	(no value)	(no value)	10,803,800	(no value)	(no value)	(no value)	**10,803,800**
	Rice	(no value)	49,318,300	(no value)	(no value)	(no value)	(no value)	**49,318,300**
	Sannhamp	(no value)	29	(no value)	(no value)	1,800	(no value)	**1,829**
	Sesamum	(no value)	18,379	(no value)	(no value)	(no value)	(no value)	**18,379**
	Soyabean	(no value)	(no value)	(no value)	(no value)	(no value)	(no value)	**(no value)**
	Sugarcane	(no value)	(no value)	(no value)	(no value)	112,680,900	(no value)	**112,680,900**
	Sunflower	(no value)	18,900	146,500	(no value)	(no value)	(no value)	**165,400**
	Sweet potato	(no value)	(no value)	(no value)	(no value)	16,900	(no value)	**16,900**
	Tobacco	(no value)	(no value)	(no value)	(no value)	(no value)	(no value)	**(no value)**
	Turmeric	(no value)	(no value)	(no value)	(no value)	965	(no value)	**965**
	Urad	(no value)	11,318	(no value)	(no value)	(no value)	(no value)	**11,318**
	Wheat	(no value)	(no value)	158,647,000	(no value)	(no value)	(no value)	**158,647,000**
	other oilseeds	(no value)	(no value)	(no value)	(no value)	(no value)	(no value)	**(no value)**
	Summary	**(no value)**	**88,593,481**	**173,272,098**	**(no value)**	**119,408,311**	**(no value)**	**381,273,890**
	Arhar/Tur	(no value)	591	(no value)	(no value)	(no value)	(no value)	**591**

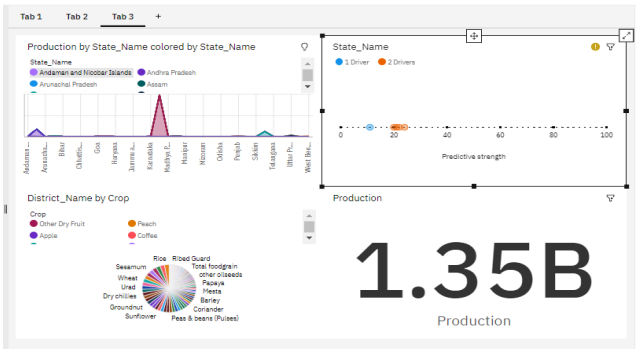
5) States with crop production:

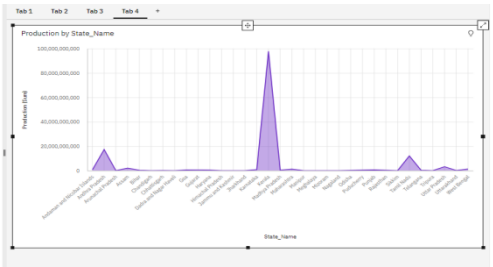


7.2 FEATURE 2:

B) CREATING THE DASHBOARD:







7.3 CODING:

- https://colab.research.google.com/drive/14py8wcogYJRqZNE1vGeyyIa2XAUI0H8H

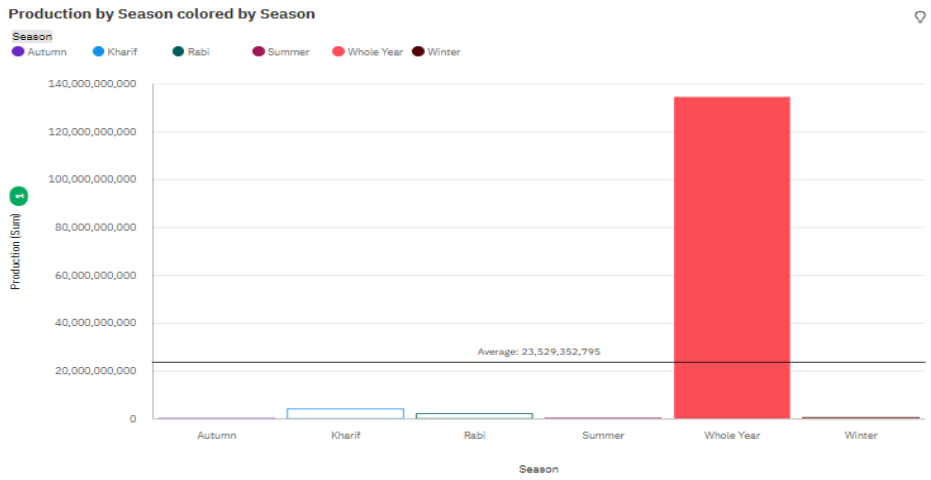
8.TESTING:

S.No.	Parameter	Screenshot / Values
1.	Dashboard design	No of Visulizations / Graphs - 6
2.	Data Responsiveness	With minimal delay our model is responding to the large sets of inputs from the user.
3.	Amount Data to Rendered (DB2 Metrics)	We rendered 70% of data from the dataset.
4.	Utilization of Data Filters	We have chosen the particular state (Maharashtra) and year (>=2004) for easy analysis.
5.	Effective User Story	No of Scene Added - 9
6.	Descriptive Reports	No of Visulizations / Graphs - 6

9. RESULTS:

A) DATA EXPLORATION:

1) Seasons with average production:



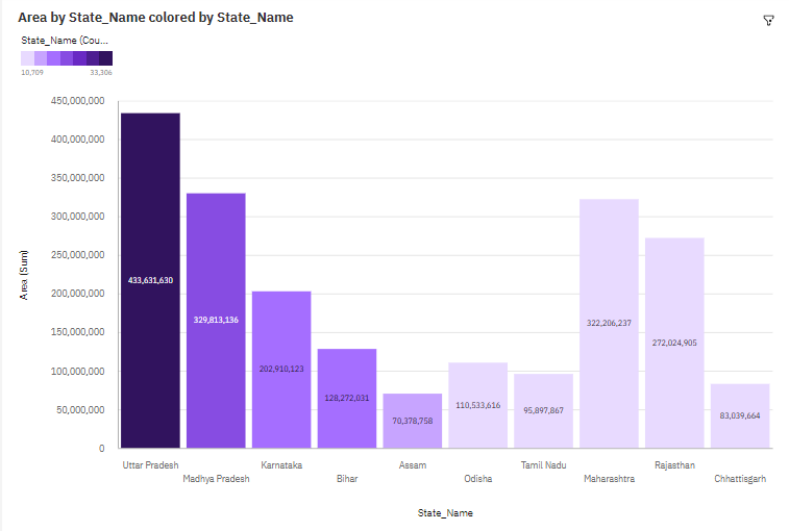
2) With years usage of area and production:

Area and Production for Crop_Year 1

	Area	Production
1997	231,715,046	851,232,906
1998	166,988,082	5,825,320,640.4
1999	158,666,106	6,434,665,985.1
2000	165,297,477	7,449,709,127.1
2001	165,295,604.67	##########
2002	157,769,017.21	##########
2003	172,088,098.54	##########
2004	167,878,424.73	##########
2005	163,136,376.32	##########
2006	170,699,101.65	##########
2007	152,724,165.3	##########
2008	171,232,070	##########
2009	165,694,709	##########
2010	176,619,202.02	##########
2011	153,629,160.88	##########
2012	152,469,799	##########
2013	141,524,909.29	##########

2013
Production: 12,903,588,632.88

3) Top 10 states with the most area.

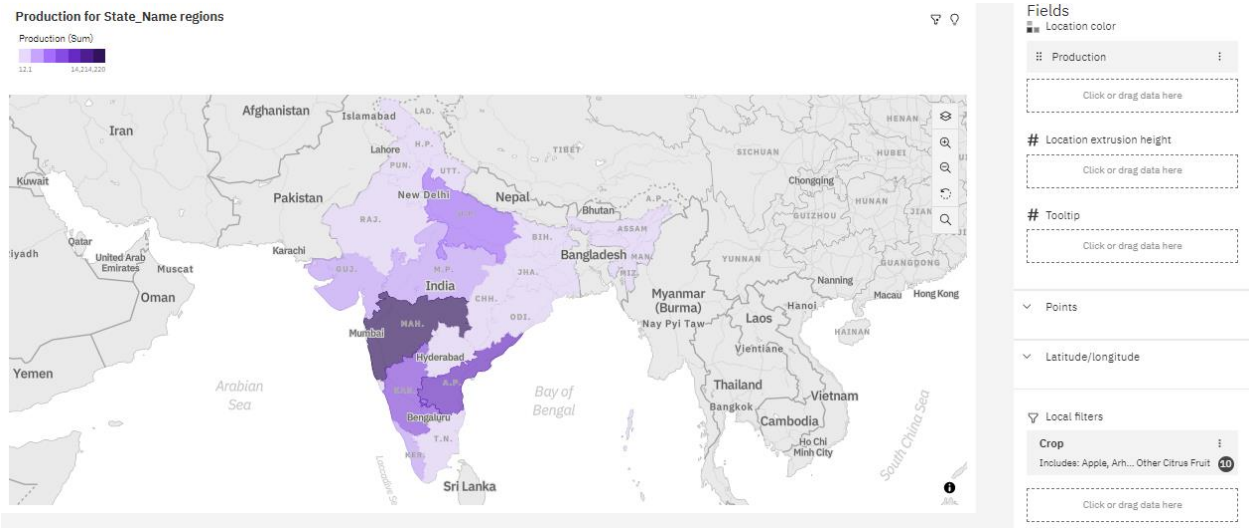


4) States with crop production with seasons:

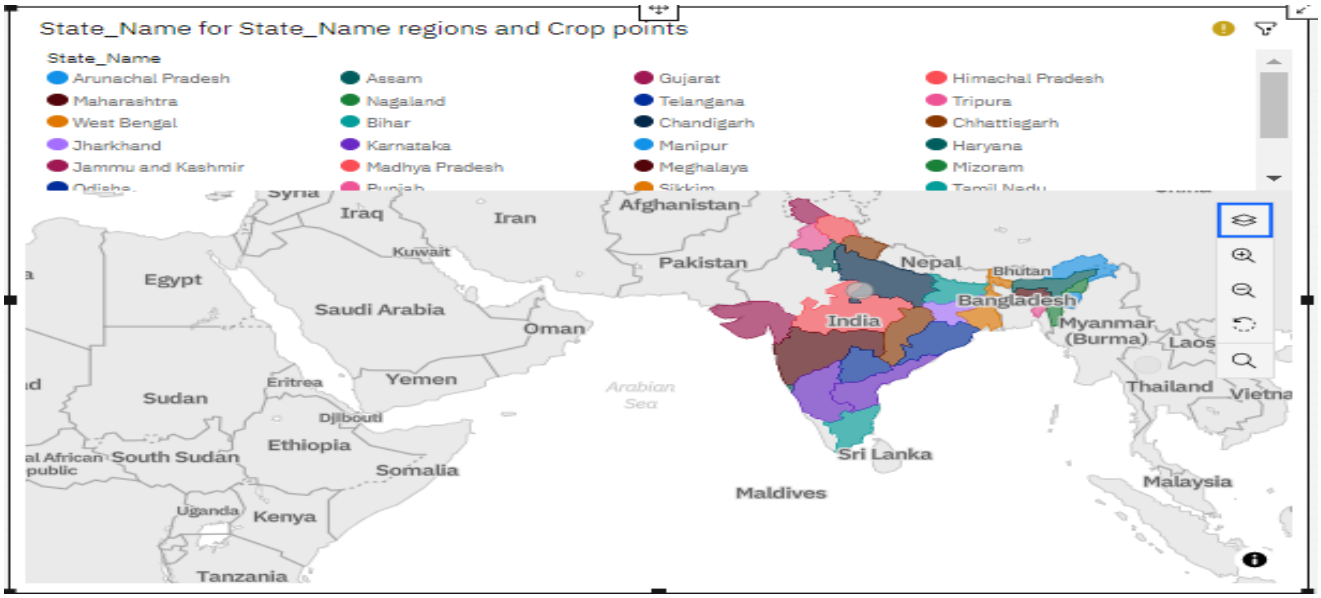
Production for Season, State_Name and Crop

Production		Autumn	Kharif	Rabi	Summer	Whole Year	Winter	Summary
Andaman and Ni...	Arecanut	(no value)	14,500	5,800	(no value)	27,735.81	(no value)	**48,035.81**
	Arhar/Tur	(no value)	(no value)	104	(no value)	(no value)	(no value)	**104**
	Banana	(no value)	(no value)	(no value)	(no value)	97,424.65	(no value)	**97,424.65**
	Black pepper	(no value)	(no value)	120	(no value)	604.5	(no value)	**724.5**
	Cashewnut	(no value)	(no value)	310	(no value)	1,374.79	(no value)	**1,684.79**
	Coconut	(no value)	(no value)	(no value)	(no value)	717,790,000	(no value)	**717,790,000**
	Dry chillies	(no value)	(no value)	575	(no value)	3,443.3	(no value)	**4,018.3**
	Dry ginger	(no value)	(no value)	1,850	(no value)	10,825.6	(no value)	**12,675.6**
	Groundnut	(no value)	(no value)	14.4	(no value)	(no value)	(no value)	**14.4**
	Maize	(no value)	(no value)	367.62	(no value)	(no value)	(no value)	**367.62**
	Moong(Green Gr...	(no value)	(no value)	575.5	(no value)	(no value)	(no value)	**575.5**
	Other Kharif puls...	(no value)	649	(no value)	(no value)	(no value)	(no value)	**649**
	Rice	23,916	199,090.06	(no value)	(no value)	(no value)	(no value)	**223,006.06**
	Sugarcane	1,332.95	(no value)	(no value)	(no value)	29,305.72	(no value)	**30,638.67**
	Sunflower	(no value)	(no value)	2.4	(no value)	(no value)	(no value)	**2.4**
	Sweet potato	(no value)	(no value)	923	(no value)	2,142.35	(no value)	**3,065.35**

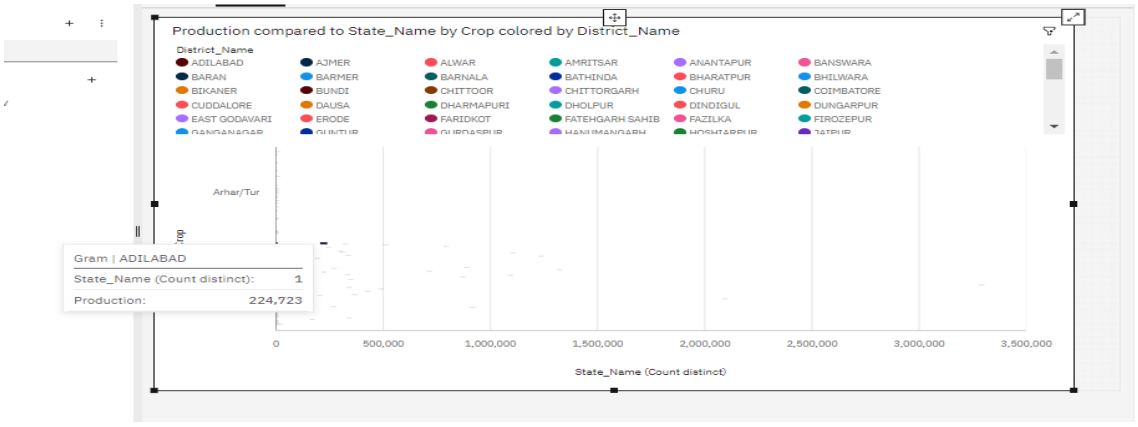
5) States with crop production:

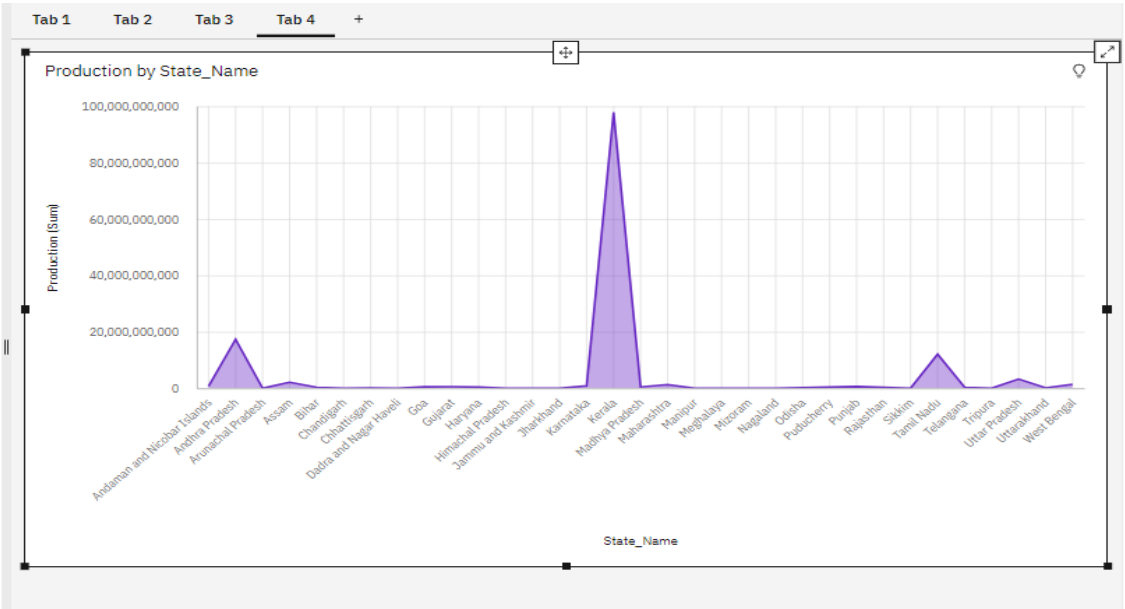


B)DASHBOARD:




Tab 1
Tab 2
Tab 3
Production by State_Name colored by State_Name
State_Name
Andaman and Nicobar Islands
Andhra Pradesh
Arunachal Pradesh
Assam
State_Name
1 Driver
2 Drivers
0
20
40
60
80
100
Predictive strength
District_Name by Crop
Crop
Other Dry Fruit
Peach
Apple
Coffee
Rice
Ribed Guard
Total foodgrain
other oilseeds
Papaya
Mesta
Barley
Coriander
Peas & beans (Pulses)
Sunflower
Groundnut
Dry chillies
Urad
Wheat
Sesamum
Production
1.35B
Production


Production compared to State_Name by Crop colored by District_Name
District_Name
ADILABAD
AJMER
ALWAR
AMRITSAR
ANANTAPUR
BANSWARA
BARAN
BARMER
BARNALA
BATHINDA
BHARATPUR
BHILWARA
BIKANER
BUNDI
CHITTOOR
CHITTORGARH
CHURU
COIMBATORE
CUDDALORE
DAUSA
DHARMAPURI
DHOLPUR
DINDIGUL
DUNGARPUR
EAST GODAVARI
ERODE
FARIDKOT
FATEHGARH SAHIB
FAZILKA
FIROZEPUR
Arhar/Tur
Gram | ADILABAD
State_Name (Count distinct): 1
Production: 224,723
0
500,000
1,000,000
1,500,000
2,000,000
2,500,000
3,000,000
3,500,000
State_Name (Count distinct)


Tab 1
Tab 2
Tab 3
Tab 4
Production by State_Name
100,000,000,000
80,000,000,000
60,000,000,000
40,000,000,000
20,000,000,000
0
Production (Sum)
Andaman and Nicobar Islands
Andhra Pradesh
Arunachal Pradesh
Assam
Bihar
Chandigarh
Chhattisgarh
Dadra and Nagar Haveli
Goa
Gujarat
Haryana
Himachal Pradesh
Jammu and Kashmir
Jharkhand
Karnataka
Kerala
Madhya Pradesh
Maharashtra
Manipur
Meghalaya
Mizoram
Nagaland
Odisha
Puducherry
Punjab
Rajasthan
Sikkim
Tamil Nadu
Telangana
Tripura
Uttar Pradesh
Uttarakhand
West Bengal
State_Name

10.ADVANTAGES:

- Predicting productivity of crop in various climatic conditions can **help farmer and other partners in essential basic leadership as far as agronomy and product decision**.
- This model can be used to select the most excellent crops for the region and also its yield thereby improving the values and gain of farming also.
- This will help the policy makers of the state to determine the budget.
- If the production of a crop observes a declining trend then, they can plan to implement the schemes at an early stage. This in return will save the state from shortage of the product.
- Monitors the growth of healthy crops.
- Helps the government to frame the government policies.
- Yield data helps the farmer to determine how much they should plant next year.
- Helps the farmer in Seed Selection, Pest Management, Irrigation Scheduling, etc,…

CHALLENGES:

Challenges are the major basis which imminent the negative impacts on current project. Some of the challenges faced during crop yield prediction are:

- Choosing appropriate dataset, after choosing dataset tuning of the parameters which makes project more efficient to get the desired results.
- Model must be trained by taking consideration of less computational efficiency and power.
- Increase of error rate due to dynamically changing the environment.

11. CONCLUSION:

Our project will make policy maker of the state to determine the budget. If the production of a crop observes a declining trend then, they can plan to implement the schemes at an early stage. This in return will save the state from shortage of the product. Monitors the growth of healthy crops. Helps the government to frame the government policies. The productivity of agriculture has slightly increased as a result of technology's introduction. New ideas like digital agriculture, smart farming, precision agriculture, etc. have been made possible by the innovations. The analysis of agricultural productivity and the uncovering of hidden patterns utilizing data sets related to seasons and crop yields have been noted in the literature. Using IBM Cognos, we have observed and conducted analysis regarding various crops grown, areas, and productions in various states and districts. "**The scope of the project is to determine the crop yield of an area by considering dataset with some features which are important or related to crop production such as temperature, moisture, rainfall, and production of the crop in previous years. To predict a continuous value, regression models are used."**

12. FUTURE SCOPE:

Our future scope is to add many more geographical features and predict using those features.

13. APPENDIX:

- **GITHUB LINK - https://github.com/IBM-EPBL/IBM-Project-24177-1659939177**
- **DEMO LINK - https://drive.google.com/drive/folders/1d7yX8tiQE0y-LGqGVfxGhYdR5Q6QgDW-**