PROJECT REPORT

Estimate The Crop Yield Using Data Analytics

TEAM ID: PNT2022TMID07696

Team Members

S.srinivasan

R.Rithish Mani

S.sathish kumar

S.Pragatheshwari

E.Kanagasabapathi

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

- 7.1 Feature 1
- 7.2 Feature 2
- 7.3 Database Schema (if Applicable)
- 8. TESTING
 - 8.1 Test Cases
 - 8.2 User Acceptance Testing
- 9. RESULTS
 - 9.1 Performance Metrics
- 10. ADVANTAGES & DISADVANTAGES
- 11. CONCLUSION
- 12. FUTURE SCOPE
- 13. APPENDIX

Source Code

GitHub & Project Demo Link

INTRODUCTION:

1.1 Project Overview:

Information on crop area, yield and production plays a vital role in planning and allocating resources for the development of the agricultural sector. Reliable and timely information on crop area, yield and production acts as a fundamental input to the planners and policymakers responsible for formulating efficient agricultural policies, and for making important decisions with respect to procurement, storage, public distribution, import, export and other related issues. The availability of crop area statistics is an essential requirement of the agricultural statistical system of any country, as it is a key variable in estimating crop production and crop yield. For the collection of crop area statistics, both subjective and objective methods are currently used around the world.

The subjective methods, often used in developing countries, include the field reporting system, eye estimation, farmer interview and expert assessment. Thesen methods suffer from certain limitations in terms of the reliability of the data on crop area. Although objective methods of measuring area – such as the polygon method – are expected to provide reliable estimates, they are costly and timeconsuming. Further, under certain unusual and problematic situations (e.g. fields with irregular shapes and boundaries), it becomes difficult to measure area with subjective methods. In these cases, modern technologies such as Global Positioning Systems (GPS) have the potential to provide more accurate estimates of the crop area.

1.2 Purpose:

As noted above, in view of the importance of estimating crop area, yield and production under mixed and continuous cropping, the Global Strategy has awarded the study project entitled "Improving Methods for Estimating Crop Area, Yield and Production under Mixed and Continuous Cropping" to ICARIASRI. The study project has the following objectives:

- 1. Critically review the literature pertaining to crop area and yield under mixed and continuous cropping;
- 2. Identify the gaps relating to the estimation of crop area and yield under mixed and continuous cropping;
- 3. Develop a standard statistical methodology for the estimation of the area and yield rate under mixed and continuous cropping;
- 4. Test the developed methodology in three field-testing countries in AsiaPacific, Africa and the Latin Caribbean region (one country in each region);
 - 5. Identify issues and challenges and provide suitable guidelines for the implementation of the developed methodology in developing countries.

2. LITERATURE SURVEY:

2.1 Existing Problem:

Companies and industries are faced with a huge amount of raw data, which have information and knowledge in their hidden layer. Also, the format, size, variety, and velocity of generated data bring complexity for industries to apply them in an efficient and effective way. So, complexity in data analysis and interpretation incline organizations to deploy advanced tools and techniques to overcome the difficulties of managing raw data. Big data analytics is the advanced method that has the capability for managing data. It deploys machine learning techniques and deep learning methods to benefit from gathered data. In this research, the methods of both ML and DL have been discussed, and an ML/DL deployment model for IOT data has been proposed.

power for corporation. The power of information leads organizations toward being agile and to hit the goals. Big data analytics (BDA) enforces industries to describe, diagnose, predict, prescribe, and cognate the hidden growth opportunities and leads them toward gaining business value. BDA deploys advanced analytical techniques to create knowledge from exponentially increasing amount of data, which will affect the decision-making process in decreasing complexity of the process. BDA needs novel and sophisticated algorithms that process and analyze real-time data and result in high-accuracy analytics. Machine and deep learning allocate their complex algorithms in this process considering the problem approach.

2.2 Reference:

- 1. Bożejko W et al. Parallel tabu search for the cyclic job shop scheduling problem. Computers & Industrial Engineering. 2018;113:512-524
- 2. Kiziloz H, Dokeroglu T. A robust and cooperative parallel tabu search algorithm for the maximum vertex weight clique problem. Computers & Industrial Engineering. 2018;118:54-66
- 3. Acharya U et al. Automated detection of coronary artery disease using different durations of ECG segments with convolutional neural network. Knowledge-Based Systems. 2017;132:62-71

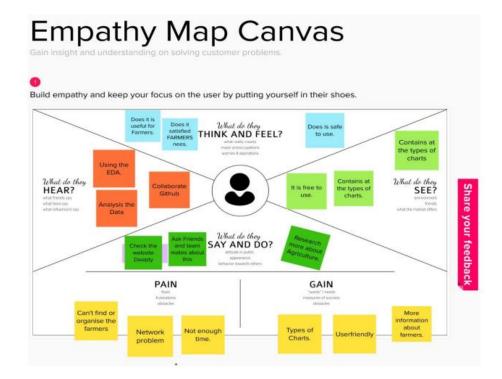
2.3 Problem Statement Definition:

The problem statement draws the reader into your topic by highlighting how important the topic is to social work and overall society. Signposting is an important component of academic writing that helps your reader follow the structure of your argument and literature review.

3. IDEATION & PROPOSED SOLUTION:

3.1 Empathy Map Canvas:

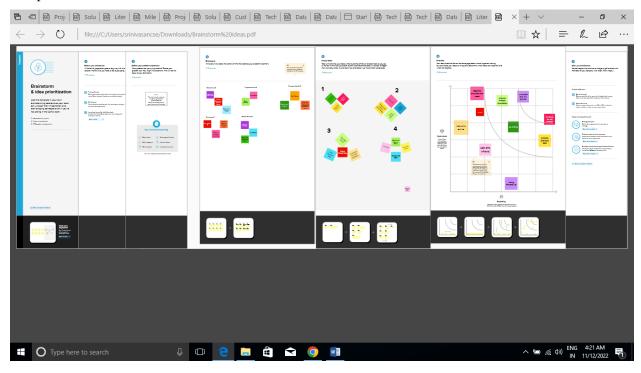
An empathy map canvas helps brands provide a better experience for users by helping teams understand the perspectives and mindset of their customers. Using a template to create an empathy map canvas reduces the preparation time and standardizes the process so you create empathy map canvases of similar quality.



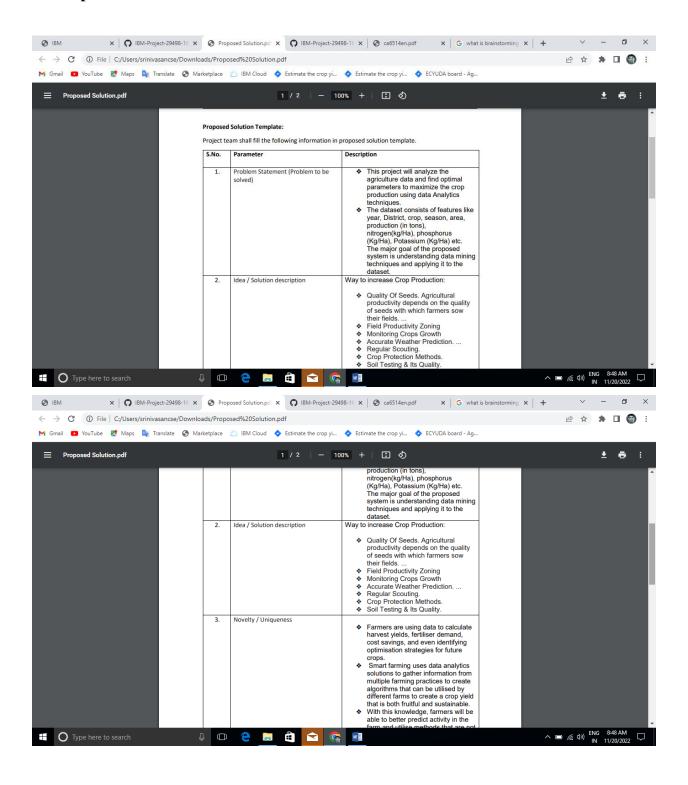
3.2 Ideation & Brainstorming

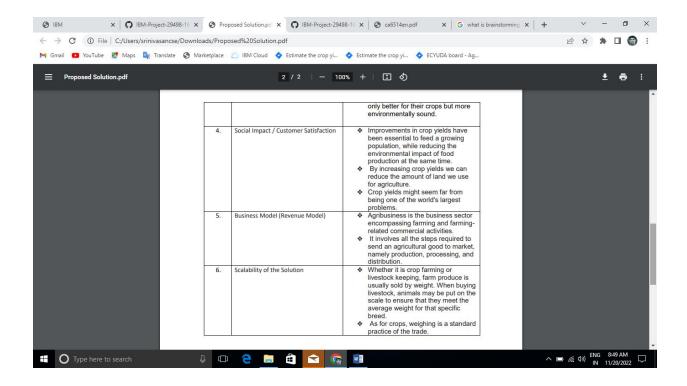
Brainstorming is a group problem-solving method that involves the spontaneous contribution of creative ideas and solutions. This technique requires intensive, freewheeling discussion in which every member of the group is encouraged to think aloud and suggest as many ideas as possible based on their diverse knowledge.

Template View:



3.3 Proposed Solution:

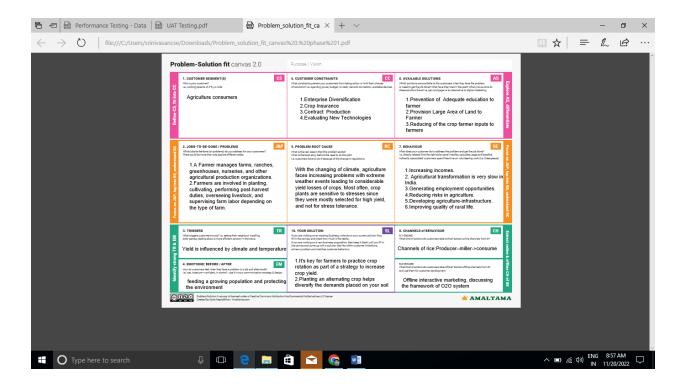




3.4 Problem Solution Fit:

The problem-solution fit canvas is a tool that helps entrepreneurs to assess whether their business idea is likely to be successful. It does this by forcing them to consider the problem that their product or service is solving, the size of the market for that problem, and the competition.

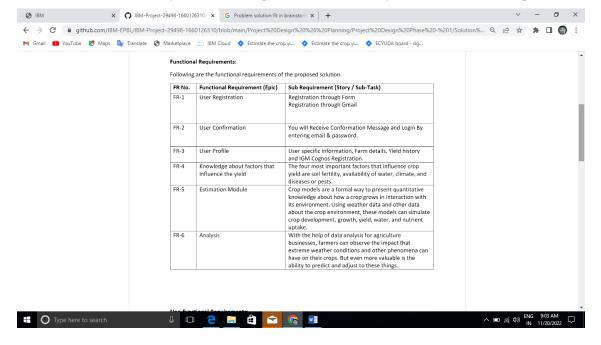
Template View:



4. REQUIREMENT ANALYSIS:

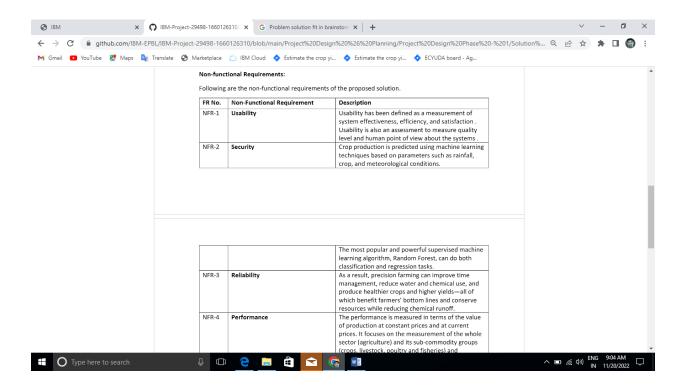
4.1 Functional Requirements:

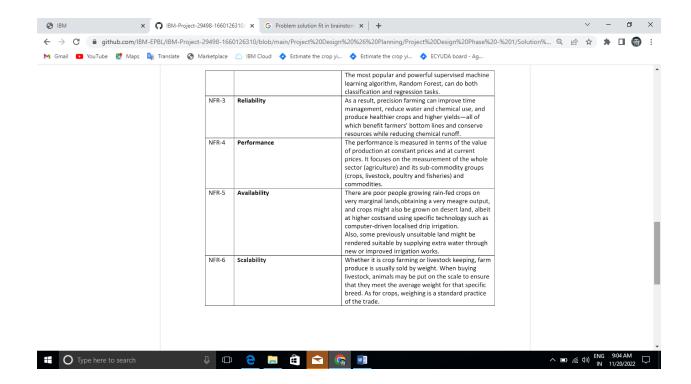
Functional requirements are product features or functions that developers must implement to enable users to accomplish their tasks. So, it's important to make them clear both for the development team and the stakeholders. Generally, functional requirements describe system behavior under specific conditions.



4.2 Non Functional Requirements:

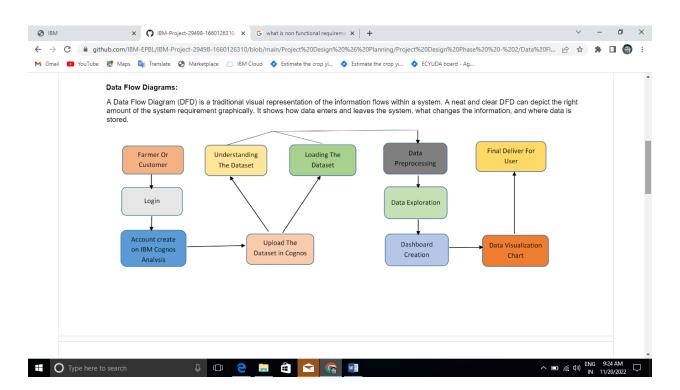
Nonfunctional Requirements (NFRs) define system attributes such as security, reliability, performance, maintainability, scalability, and usability. They serve as constraints or restrictions on the design of the system across the different backlogs





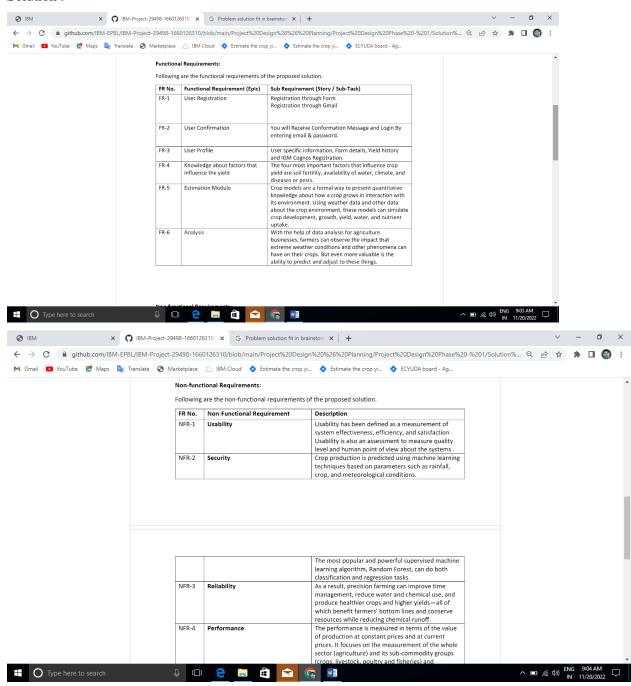
5. PROJECT DESIGN:

5.1 Data flow diagram:

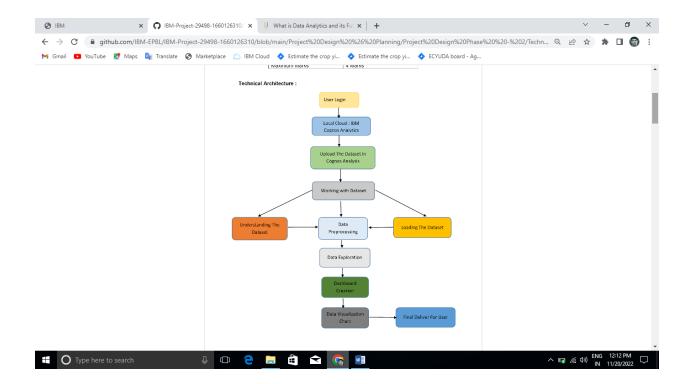


5.2 Solution & Technical Architecture:

Solution:

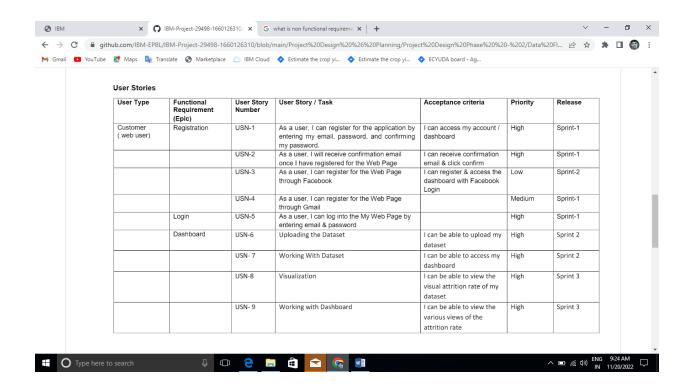


Technical Architecture:



5.3 User Stories:

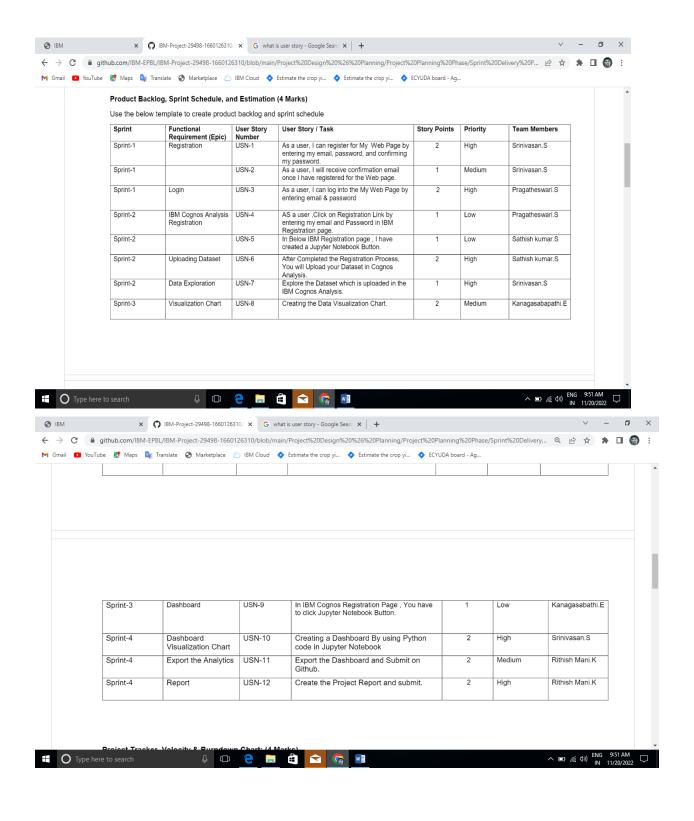
A user story is an informal, general explanation of a software feature written from the perspective of the end user or customer. The purpose of a user story is to articulate how a piece of work will deliver a particular value back to the customer.



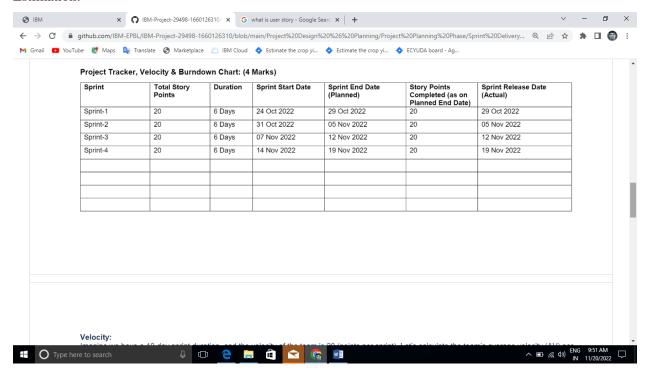
6. PROJECT PLANNING & SCHEDULING:

6.1 Sprint Planning & Estimation:

Planning: 1

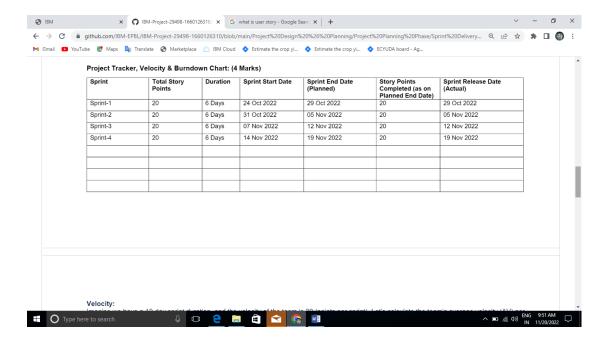


Estimation:

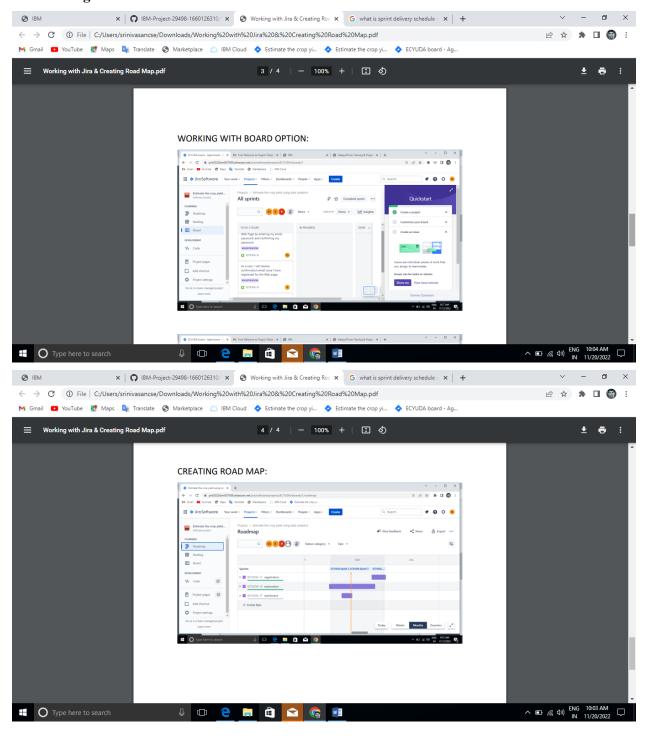


6.2 Sprint Delivery Schedule:

In case you're unfamiliar, a sprint schedule is a document that outlines sprint planning from end to end. It's one of the first steps in the agile sprint planning process—and something that requires adequate research, planning, and communication.

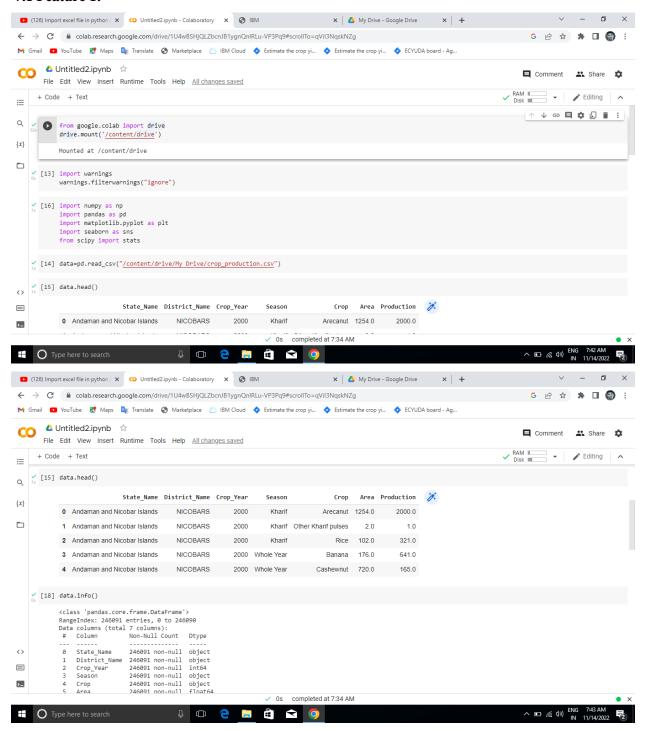


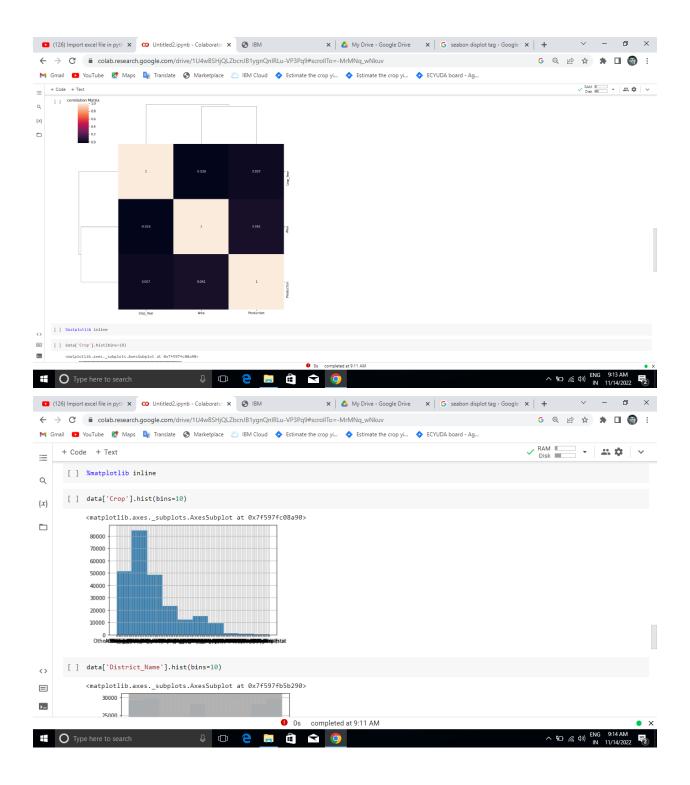
6.3 Working with JIRA:



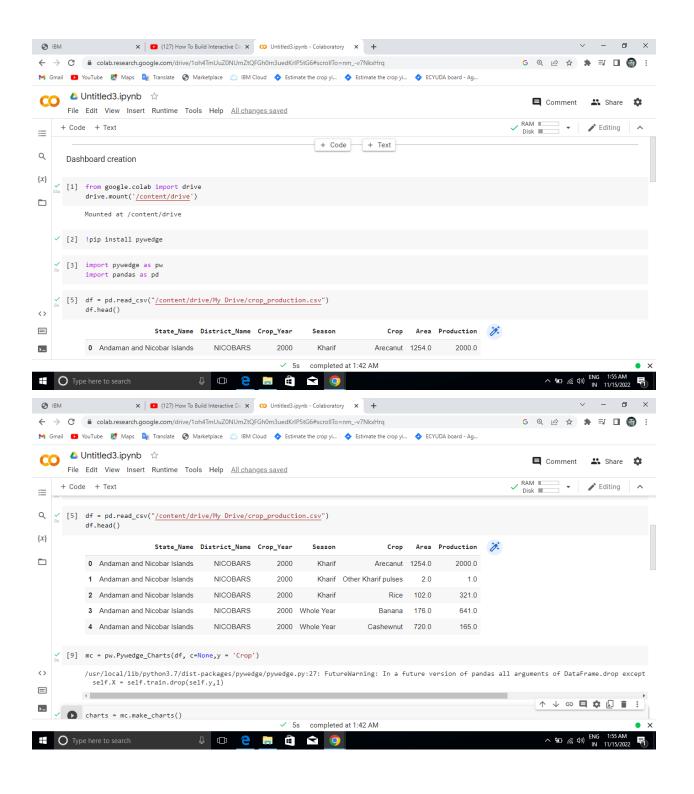
7. CODING & SOLUTION

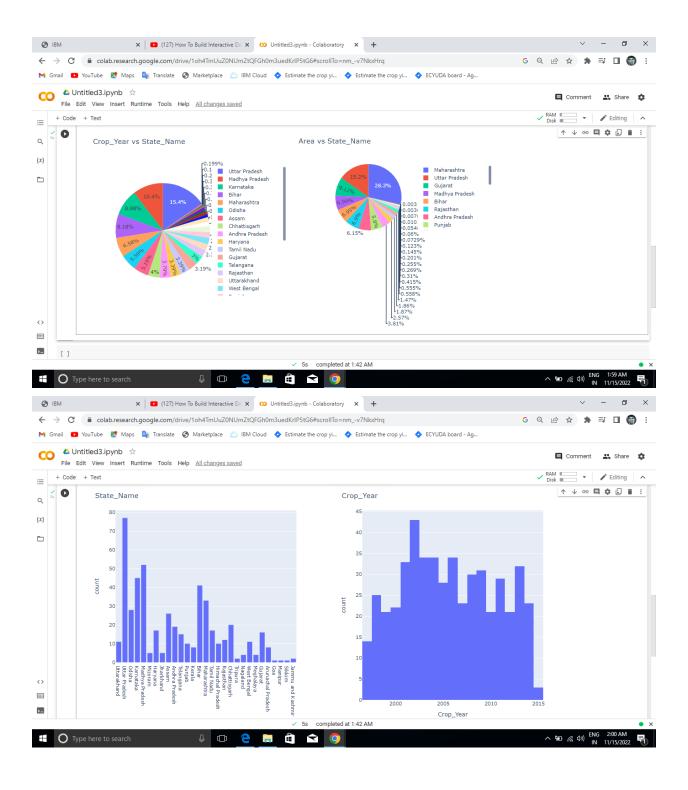
7.1 Feature 1:



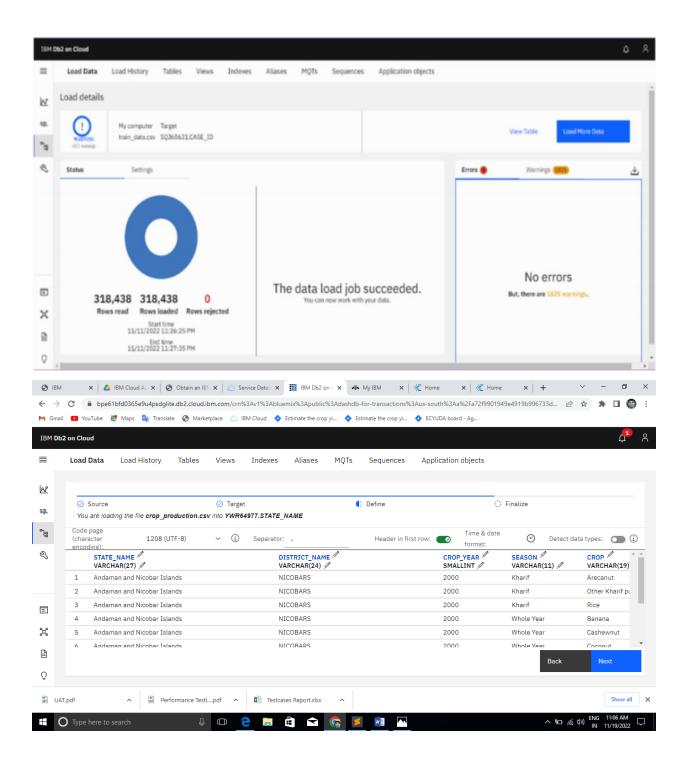


7.2 Feature 2



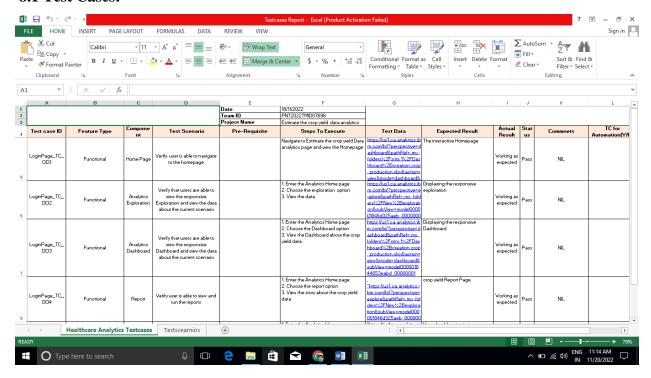


7.3 Database schema:

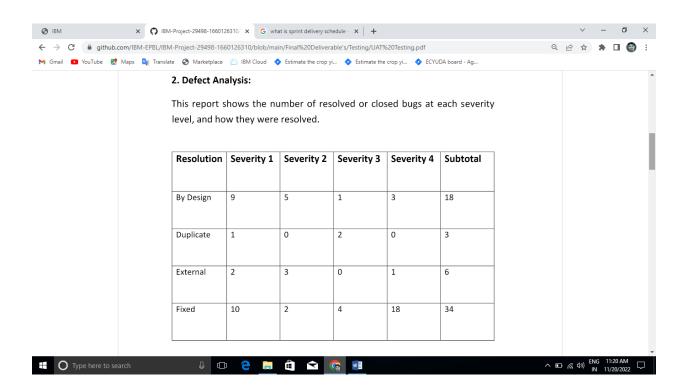


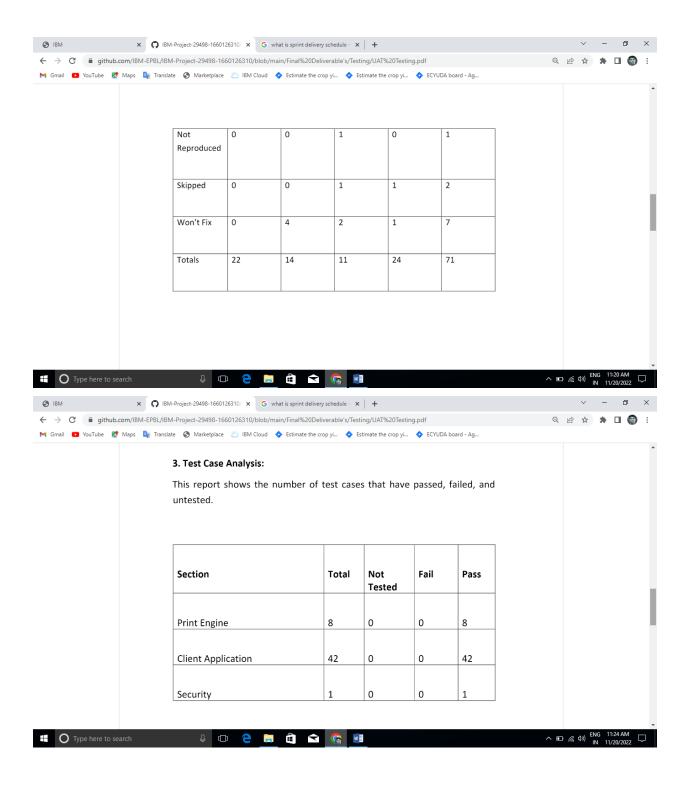
8. TESTING:

8.1 Test Cases:



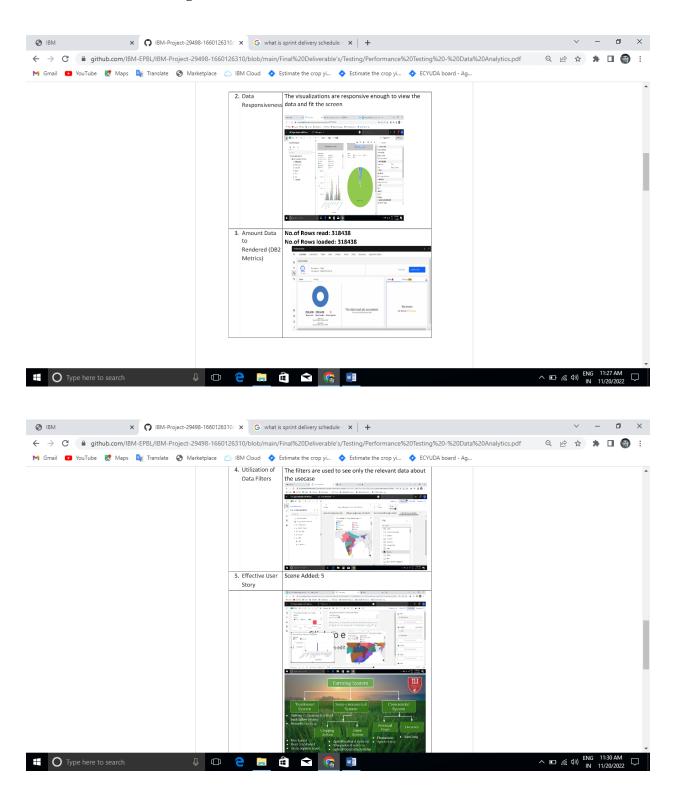
8.2 User Acceptance Testing:





9. RESULTS:

9.1 Performance Testing:



10. ADVANTAGES & DISADVANTAGES:

Advantages

This method is relatively less time-consuming and inexpensive. Farmer assessment does not require the enumerator to visit the individual plots, which is cost-effective particularly if the plots are located far away from the location of the initial interview. Furthermore, farmer assessments of crop area can serve as a baseline for imputation where objective measurements are missing

Disadvantages

This method is highly subjective, as it depends on farmers' knowledge and experience. Furthermore, any nonstandard units of measurement used by farmers may be difficult to standardize. The farmers may also have incentives to misreport crop area for reasons such as taxation. The data analysis conducted within the World Bank's study of four African countries (Carletto et al., 2015) indicate that self-reported land areas systematically differ from GPS land measurements, and that this difference leads to biased estimates of the relationship between land and productivity and consistently low estimates of land inequality. Furthermore, results from methodological experiments carried out by both the World Bank and the Global Strategy indicate that farmers tend to overreport plot area for small plots, and underreport area for very large plots.

11. CONCLUTION:

On the basis of a gap analysis, this Technical Report an attempts to address the problems relating to the estimation of crop area and crop yield in mixed and continuous cropping. Accordingly, an appropriate methodology has been developed to estimate crop area and crop yield in the contexts of mixed cropping and intercropping. To estimate the crop area of component crops in crop mixtures, the domain estimation approach has been proposed. The various crop mixtures are considered as domains. To estimate crop area and crop yield, the sample survey approach is proposed. Various measurement methods – both subjective and objective –to determine crop area and yield are explored, as well as their respective advantages and disadvantages, in light of the results of the field test.

12. FUTURE SCOPE:

Data analytics is used in almost every sector of business, let's discuss a few of them:

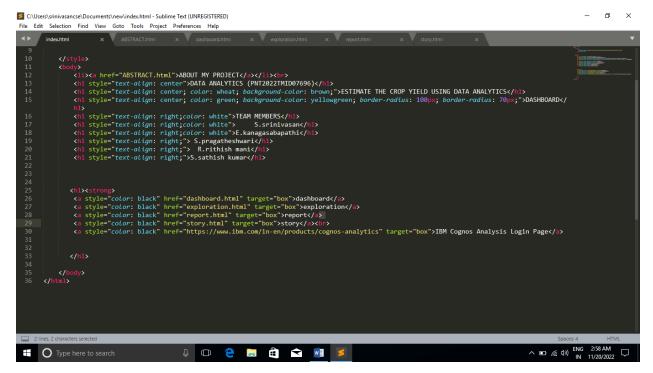
1. Retail: Data analytics helps retailers understand their customer needs and buying habits to predict trends, recommend new products, and boost their business.

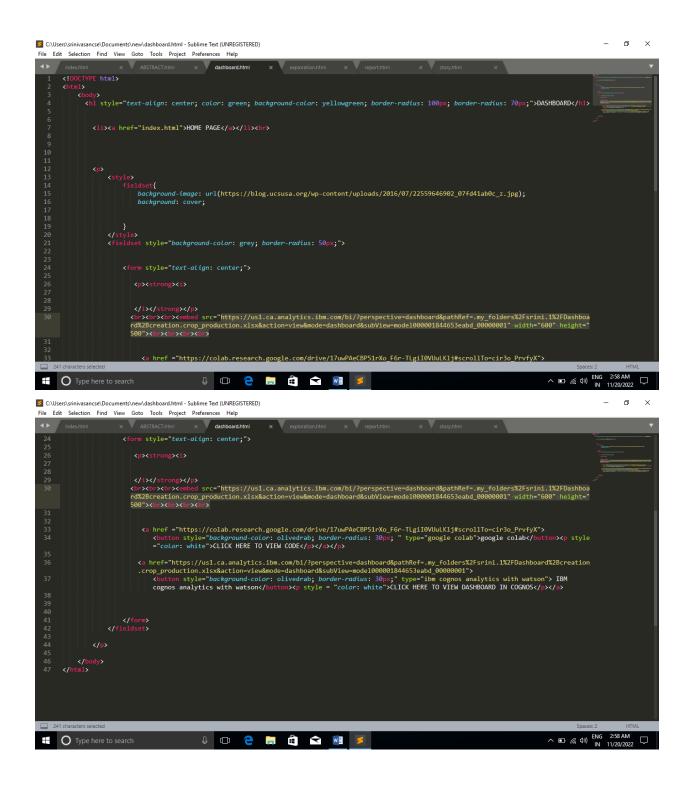
They optimize the supply chain, and retail operations at every step of the customer journey.

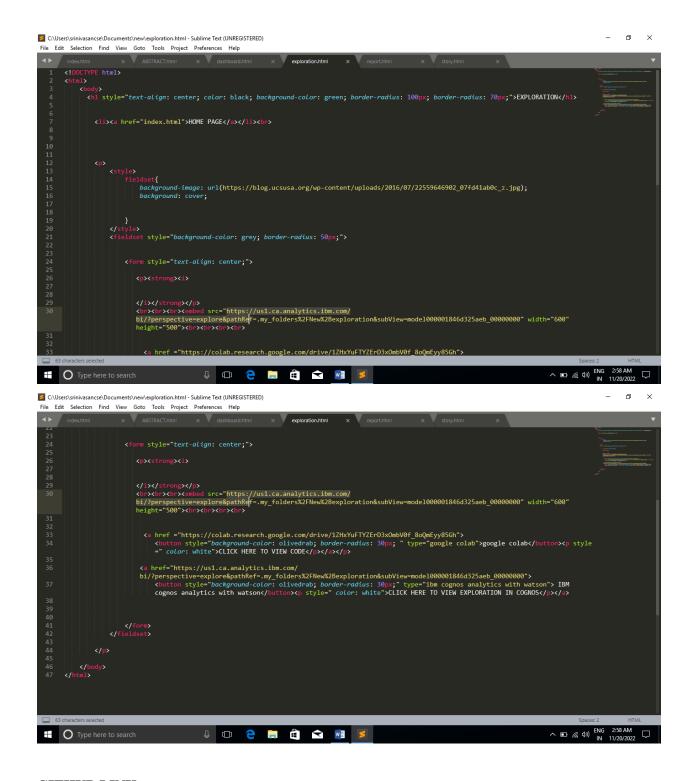
- 2. Healthcare: Healthcare industries analyze patient data to provide lifesaving diagnoses and treatment options. Data analytics help in discovering new drug development methods as well.
- 3. Manufacturing: Using data analytics, manufacturing sectors can discover new cost-saving opportunities. They can solve complex supply chain issues, labor constraints, and equipment breakdowns.
- 4. Banking sector: Banking and financial institutions use analytics to find out probable loan defaulters and customer churn out rate. It also helps in detecting fraudulent transactions immediately.

14. APPENDIX:

Source Code:







GITHUB LINK:

https://github.com/IBM-EPBL/IBM-Project-29498-1660126310

DEMONSTRATION VIDEO LINK:

 $https://drive.google.com/file/d/1RLrVHfgr9LthWE4_SVI5ezVv0NowvG0i/view?usp=share_link$