

# **ESTIMATE THE CROP YIELD USING DATA ANALYTICS**

**TEAMID: PNT2022TMID15473**

**TEAM LEADER - YOGESH BALAN A**

**TEAM MEMBER 1 - THYAANESHWAR S**

**TEAM MEMBER 2 - VINAY S**

**TEAM MEMBER 3 - SRIVARSHAN S**

# ABSTRACT

Agriculture is important for human survival because it serves the basic need. A well-known fact that the majority of population ( $\geq 55\%$ ) 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.

In this project, we use a dashboard to view, analyze, and extract the majority of the findings.

# TABLE OF CONTENTS

Chapter	Title	Page No.
1.	<b>ABSTRACT</b>	2
2.	<b>INTRODUCTION</b>	4
3.	<b>LITERATURE SURVEY</b>	6
	3.1. Existing System or review of Literature survey.	
	3.2. Problem Statement	
4.	<b>IDEATION PHASE &amp; PROPOSED SOLUTION</b>	9
	4.1. Empathy Map Canvas	
	4.2. Ideation & Brainstorming	
	4.3. Proposed Solution	
	4.4. Problem Solution fit	
5.	<b>REQUIREMENT ANALYSIS</b>	13
	5.1. Functional requirements	
	5.2. Non-Functional requirements	
6.	<b>PROJECT DESIGN</b>	14
	6.1. Data Flow Diagrams	
	6.2. Solution & Technical Architecture	
	6.3. User Stories	
7.	<b>PROJECT PLANNING AND SCHEDULING</b>	17
	7.1. Sprint Planning & Estimation	
	7.2. Sprint Delivery Schedule	
	7.3. Reports from JIRA	
8.	<b>CODING AND SOLUTIONING</b>	17
	8.1. IBM Cognos With Watson	
	8.2. Source code	
9.	<b>TESTING</b>	20
	9.1. Test Cases	
	9.2. User Acceptance Testing	
10.	<b>RESULTS</b>	21

	10.1. Performance Metrics	
11.	<b>ADVANTAGES &amp; DISADVANTAGES</b>	<b>25</b>
12.	<b>CONCLUSION</b>	<b>25</b>
13.	<b>FUTURE SCOPE</b>	<b>25</b>
14.	<b>APPENDIX</b>	<b>25</b>

## **1.INTRODUCTION :-**

Agriculture forms the basis for food security and hence it is important. In India, majority of the population i.e., above 55% is dependent on agriculture as per the recent information. Agriculture is the field that enables the farmers to grow ideal crops in accordance with the environmental balance. In India, wheat and rice are the major grown crops along with sugarcane, potatoes, oil seeds etc. Farmers also grow non-food items like rubber, cotton, jute etc. More than 70% of the household in the rural area depend on agriculture.

This domain provides employment to more than 60% of the total population and has a contribution to GDP also (about 17%). In the farm output, India ranks second considering the world wide scenario. This is the widest economic sector and has an important role regarding the framework of socio-economic fabric of India. 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. These companies make use of

agriculture products as raw materials, animal feed, paper production and so on. 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. Crop yield prediction helps the farmers in various ways by providing the record of previous crop yield. This is helpful to government in framing policies related to crops such as crop insurance policies, supply chain operation policies. Knowing what crops has been grown, and how much area of it had been shown historically, combined with the prices at which it could have been sold at the nearest market-place provides the income-growth profile of the farmer. Agriculture sector is struggling to increase the productivity of crop in India.

Monsoon rainfall is the main source of water for more than 60 percent of the crops. Smart agriculture driven by Information Technology is the emerging trend in the research in this area in recent days. One of the areas being explored is the problem of yield prediction which is a major concern. Data mining techniques are being widely used as a part of solution for crop yield prediction. Various data mining techniques are under evaluation for estimation of crop production of the future years [4]. Data mining is the process in which the hidden patterns are discovered using analysis of large data sets. The data mining and data analytics techniques use artificial intelligence, statistics, machine learning and database system. In data mining, unsupervised and supervised methods are being used.

#### **DATASET :-**

❖ Variable	Description
❖ Crop	Crop name
❖ State Name	Indian state name
❖ District Name	District name list of each state
❖ Cost of Cultivation (₹/Hectare)	Cultivation amount for C2 Scheme
❖ Cost of Production (₹/Quintal)	Production amount for A2+FL Scheme
❖ Yield (Quintal/ Hectare)	Yield of crop
❖ Crop year	Crop year list
❖ District Name	District name for each state
❖ Area	Total area of each place
❖ Rainfall	Water availability of each crop

- ❖ Average humidity                      directly influences the water relations of plant and indirectly affects leaf growth
- ❖ Mean Temperature                      Climate of each crop
- ❖ Cost Production of per yield crop    Cost of crop yield

## 2. LITERATURE SURVEY :-

At present we are at the immense need of another Green revolution to supply the food demand of growing population. With the decrease of available cultivable land globally and the decreased cultivable water resources, it is almost impossible to report higher crop yield. Agricultural based big data analytics is one approach, believed to have a significant role and positive impact on the increase of crop yield by providing the optimum condition for the plant growth and decreasing the yield gaps and the crop damage and wastage. With this aim the present paper reviews about the various advances, design models, software tools and algorithms applied in the prediction assessment and estimation of the crop yield.

India is basically agriculture based country and approximately 70% our country economics is directly or indirectly related to the agricultural crops. The principle crop which occupies the highest (60-70%) percentage of cultivable land in the Indian soil is the paddy culture and it is the major crop especially in central and south parts of the India. Rice crop cultivation plays an imperative part in sustenance security of India, contributing over 40% to general yield generation. The enhanced yield of the rice crop depends largely on the water availability and climatic conditions.

For example, low precipitation or temperature extremes can drastically diminish rice yield. Growing better strategies to foresee yield efficiency in a mixture of climatic conditions can help to understand the role of different principle factors that influence the rice crop yield. Big data analytic methods related to the rice crop yield prediction and estimation will certainly support the farmers to understand the optimum condition of the significant factors for the rice crop yield, hence can achieve higher crop yield.

## **2.1. Review Of Literature Survey:-**

### **1. Rice Crop Yield Prediction using Data Mining Techniques: An Overview**

Dakshayini Patil, Dr. M .S, Shirdhonka

2017

Discussed various data mining techniques utilized for prediction of rice crop yield for the state of Maharashtra, India. WEKA tool was applied in dataset processing

### **2. A Survey on Crop Yield Prediction based on Agricultural Data**

Dhivya B H, Manjula R, Siva Bharathi S, Madhumathi R

2017

Presented a survey on the different algorithms applied in the assessment and prediction of crop yield Discussed about the mechanism of knowledge the discovery in Agricultural data mining

### **3. Big Data for weed control and crop protection**

F K Van Evert, S Fountas, D Jakovetic, V Crnojevic, I Travlos & C Kempenaar

2017

Critically discussed about the challenges faced and the profound opportunities lies in the Big Data analytics in agriculture: Outlined Big Data analytics models with numerical algorithms applied Represent the importance of reforming the mined data in the form of understandable information to the farmers. Discussed about various advances, tools and algorithms applied in transforming the data in to easily understandable information to the framers and thrown a light on success story of Netherlands in achieving the maximum crop yield and their smart forming practices. Also discussed about the control of invasive, parasitic and herbicide resistant weeds to improve the overall crop yield applying Big Data analytics

### **4. Prediction of Crop Yield using Regression Analysis**

V. Sellamand E. Poovammal

2016

Regression analysis was carried out to find the relationship among the parameters i.e Area under Cultivation (AUC), Annual Rainfall (AR) and Food Price Index (FPI) which influences the final crop yield and reported that the crop yield principally depends on the Annual Rainfall (AR)

## **5. A Study on Crop Yield Forecasting Using Classification Techniques**

R.Sujatha, Dr.P.Isakki Devi

2016

Discuss the importance of comparing previous agricultural data with present to identify optimum condition favor enhanced crop yield. Envisaged the importance of best crop selection depending on the season and the climatic factors which supports enhanced crop yield.

### **1.2. Problem Statement:-**

Machine Learning based on prior crop prediction, soil quality analysis to achieve high crop yield throughout technology solution. The main objectives of this project is to predict crop-yield which can be extremely useful to farmers in planning for harvest and sale of grain harvest.



### 3. IDEATION & PROPOSED SOLUTION

#### 3.1 Empathy Map Canvas

#### 3.2 Ideation & Brainstorming

#### 3.3 Proposed Solution

#### 3.4 Problem Solution fit

#### EMPATHY MAP:-



### Group Ideas

Take time among your ideas while clustering similar or related notes as you go. Once all sticky notes have been grouped, give each cluster a tentative title. If a cluster is bigger than six sticky notes, try and see if you can break it up into smaller subgroups.

00 minutes

### 2

#### Data Acquisition

To gather the dataset, you need to identify the source of the data. This could be a public dataset, a company database, or a survey. The data should be cleaned and preprocessed before being used for analysis.

#### Understanding Data

To understand the data, you need to explore the dataset and identify the variables. This could be done by looking at the data distribution, calculating summary statistics, or creating visualizations. The goal is to get a sense of the data and identify any patterns or anomalies.

#### Data Analysis and Prediction

To analyze the data, you need to choose a model and fit it to the data. This could be done using a variety of machine learning algorithms. Once the model is trained, you can use it to make predictions on new data.

#### Making available to customers

To make the data available to customers, you need to create a user interface and a data API. This could be done using a variety of web development frameworks. The goal is to make the data easy to access and use for your customers.

### Prioritize

Your team should all be on the same page about what's important. Making forward. Place your ideas on this grid to determine which ideas are important and which are feasible.

00 minutes

### After you collaborate

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

Quick add-ons

- Share the mural: Share a view link to the mural with stakeholders to keep them in the loop about the outcomes of the session.
- Export the mural: Export a copy of the mural as a PNG or PDF to share in emails, include in slides, or share in your drive.

Keep moving forward

- Storyboard: Define the sequence of a new idea or strategy.
- Customer experience journey map: Understand customer needs, motivations, and expectations for an experience.
- Brainstorm, brainstorm, brainstorm: Brainstorm ideas, opportunities, and threats. Identify strengths, weaknesses, opportunities, and threats (SWOT) in strategy a plan.

00 minutes

### Brainstorm & idea prioritization

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

00 minutes to prepare  
1 hour to brainstorm  
20 minutes to present

00 minutes

### 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.

00 minutes

- Team gathering: Define who should participate in the session and send an invite. Share relevant information or previous work.
- Set the goal: Focus on the problem you're trying to solve in the brainstorming session.
- Learn how to use the facilitation tools: Use the Facilitation Tools to set a happy and productive session.

00 minutes

### Define your problem statement

What problem are you trying to solve? Frame your problem as a how might we statement. This will be the focus of your brainstorm.

00 minutes

How might we...  
To make the data more accessible to our customers.

Key rules of brainstorming

- Stay in topic
- Encourage wild ideas
- Build on others
- Defer judgment
- One idea at a time
- Go for quantity
- Don't criticize
- Be realistic
- Be specific, be clear

### Brainstorm

Write down any ideas that come to mind that address your problem statement.

00 minutes

YOGESH BALAN

- To gather the dataset, you need to identify the source of the data. This could be a public dataset, a company database, or a survey. The data should be cleaned and preprocessed before being used for analysis.
- To understand the data, you need to explore the dataset and identify the variables. This could be done by looking at the data distribution, calculating summary statistics, or creating visualizations. The goal is to get a sense of the data and identify any patterns or anomalies.
- To analyze the data, you need to choose a model and fit it to the data. This could be done using a variety of machine learning algorithms. Once the model is trained, you can use it to make predictions on new data.
- To make the data available to customers, you need to create a user interface and a data API. This could be done using a variety of web development frameworks. The goal is to make the data easy to access and use for your customers.

VINAY

- To gather the dataset, you need to identify the source of the data. This could be a public dataset, a company database, or a survey. The data should be cleaned and preprocessed before being used for analysis.
- To understand the data, you need to explore the dataset and identify the variables. This could be done by looking at the data distribution, calculating summary statistics, or creating visualizations. The goal is to get a sense of the data and identify any patterns or anomalies.
- To analyze the data, you need to choose a model and fit it to the data. This could be done using a variety of machine learning algorithms. Once the model is trained, you can use it to make predictions on new data.
- To make the data available to customers, you need to create a user interface and a data API. This could be done using a variety of web development frameworks. The goal is to make the data easy to access and use for your customers.

THYAAANESHWAR

- To gather the dataset, you need to identify the source of the data. This could be a public dataset, a company database, or a survey. The data should be cleaned and preprocessed before being used for analysis.
- To understand the data, you need to explore the dataset and identify the variables. This could be done by looking at the data distribution, calculating summary statistics, or creating visualizations. The goal is to get a sense of the data and identify any patterns or anomalies.
- To analyze the data, you need to choose a model and fit it to the data. This could be done using a variety of machine learning algorithms. Once the model is trained, you can use it to make predictions on new data.
- To make the data available to customers, you need to create a user interface and a data API. This could be done using a variety of web development frameworks. The goal is to make the data easy to access and use for your customers.

SRIVARSHAN

- To gather the dataset, you need to identify the source of the data. This could be a public dataset, a company database, or a survey. The data should be cleaned and preprocessed before being used for analysis.
- To understand the data, you need to explore the dataset and identify the variables. This could be done by looking at the data distribution, calculating summary statistics, or creating visualizations. The goal is to get a sense of the data and identify any patterns or anomalies.
- To analyze the data, you need to choose a model and fit it to the data. This could be done using a variety of machine learning algorithms. Once the model is trained, you can use it to make predictions on new data.
- To make the data available to customers, you need to create a user interface and a data API. This could be done using a variety of web development frameworks. The goal is to make the data easy to access and use for your customers.

PROPOSED SOLUTIONS:-

### **1. Problem Statement (Problem to be solved)**

In this project using visualization technique we create a dashboard which provides insights of Crop production using the statistical report by data analysis.

### **2. Idea / Solution description**

Crop production in India is one of the most important sources of income and India is one of the top countries to produce crops. In this project we will be analyzing important visualization by creating a dashboard and by going through these we will get most of the insights of Crop production in India. Providing proper knowledge to farmers about the crops to be sown in its respective season.

### **3. Novelty / Uniqueness**

It can be a unique and one-stop solution for a better understanding and to get a clear insight about the previous year's data on crop yield.

### **4. Social Impact / Customer Satisfaction**

Farmers can easily understand about the crop yields in India. Customer can easily understand crop production in graphical view. It is available to everyone who needs help and assistance. This is a simple approach and easy to understand.

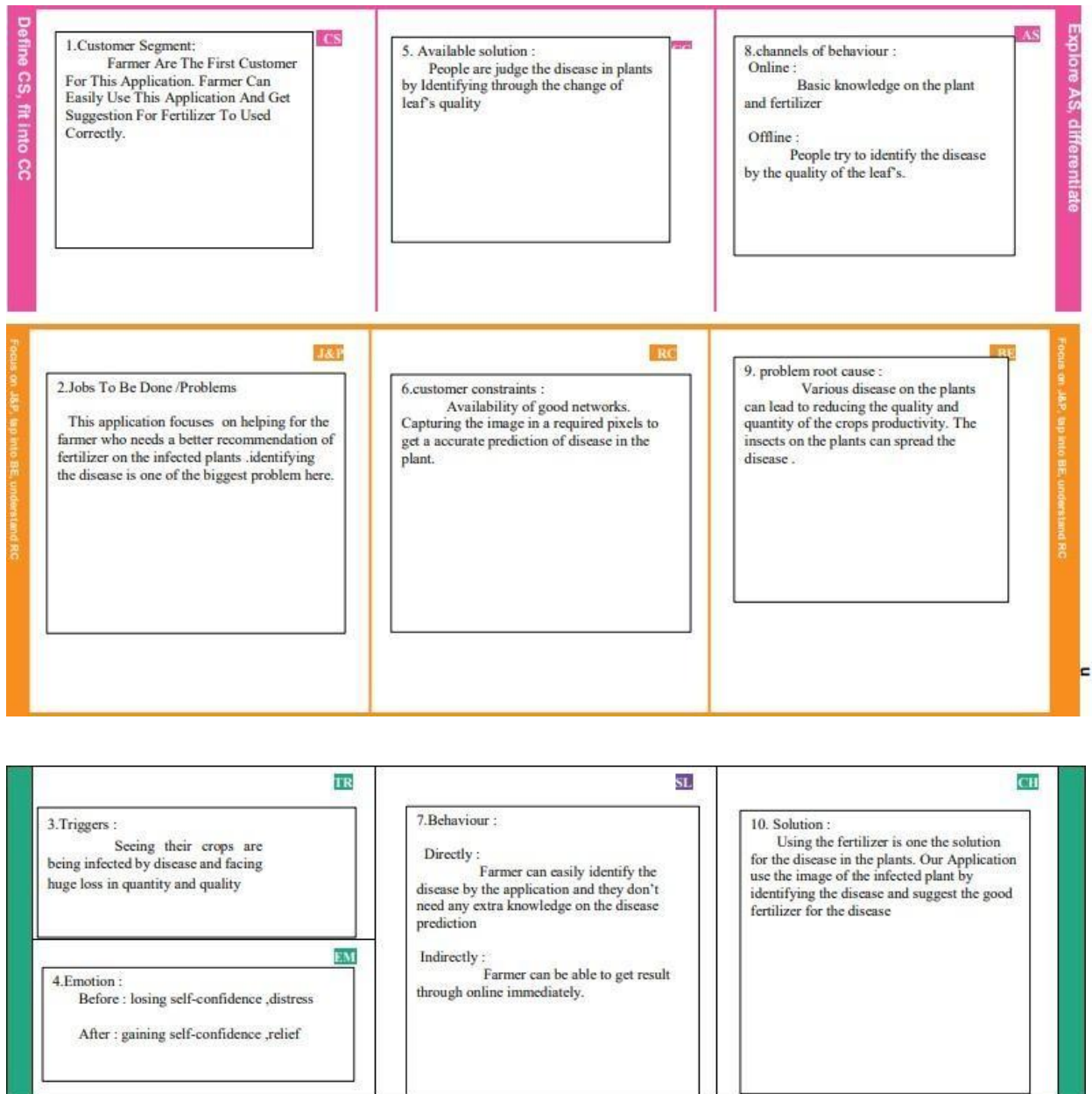
### **5. Business Model (Revenue Model)**

Production resources, agricultural commodities and facilitative services can be improved. We can calculate short term and long term economic potential, viability and future opportunities of chosen commodities

### **6. Scalability of the Solution**

We can increase by including we can include crop insurances, delivery estimates, planning harvest, storage requirements and cash flow budgeting. Dataset can be updated according to future needs.

## PROBLEM SOLUTION FIT:-



## **4. REQUIREMENT ANALYSIS**

### **General:**

Requirements are the basic constraints that are required to develop a system. Requirements are collected while designing the system. The following are the requirements that are to be discussed.

1. Functional requirements
2. Non-Functional requirements
3. Environment requirements

A. Hardware requirements

B. software requirements

### **4.1. Functional requirements:**

The software requirements specification is a technical specification of requirements for the software product. It is the first step in the requirements analysis process. It lists requirements of a particular software system. The following details to follow the special libraries like sk-learn, pandas, numpy, matplotlib.

### **4.2. Non Functional Requirements:-**

IBM Cognos Analytics

A collection of business intelligence tools called IBM Cognos Analytics is offered both on premises and in the cloud. The main emphasis is on descriptive analytics, which uses dashboards, expert reporting, and self-service data exploration to help users understand the information in your data. In this study, we analysed the crop yield data using IBM cognos data analytics.

### **Following are important features of IBM Cognos:**

**1. Get Connected** - Connect your data effortlessly Import data from CSV files and Spread sheets. Connect to cloud or on-premises data sources, including SQL databases, Google Big Query, Amazon, Redshift, and more.

**2. Prepare your data** – Prepare and connect data automatically Save time cleaning your data with AI-assisted data preparation. Clean and prep data from multiple sources, add calculated fields, join data, and create new tables.

**3. Build visualizations** - Create dynamic dashboards easily quickly create compelling, interactive dashboards. Drag and drop data to create auto generated visualizations, drill down for more detail, and share using email or Slack.

**4. Identify Patterns** – Uncover hidden patterns ask the AI assistant a question in plain language, and see the answer in visualization. Use time series modeling to predict seasonal trends.

**5. Generate Personalized Reports** – Create and deliver personalized reports. Keep your stakeholders up-to-date, automatically. Create and share dynamic personalized, multi-page reports in the formats your stakeholders want.

**6. Gain Insights** - Make confident data decisions Get deeper insights without a data science back ground. Validate what you know, identify what you don't with statistically accurate time-series forecasting and pinpoint patterns to consider.

**7. Stay Connected** - Go Mobile Stay connected on the go with the new mobile app. Access data and get alerts right from your phone.

## **5.PROJECT DESIGN**

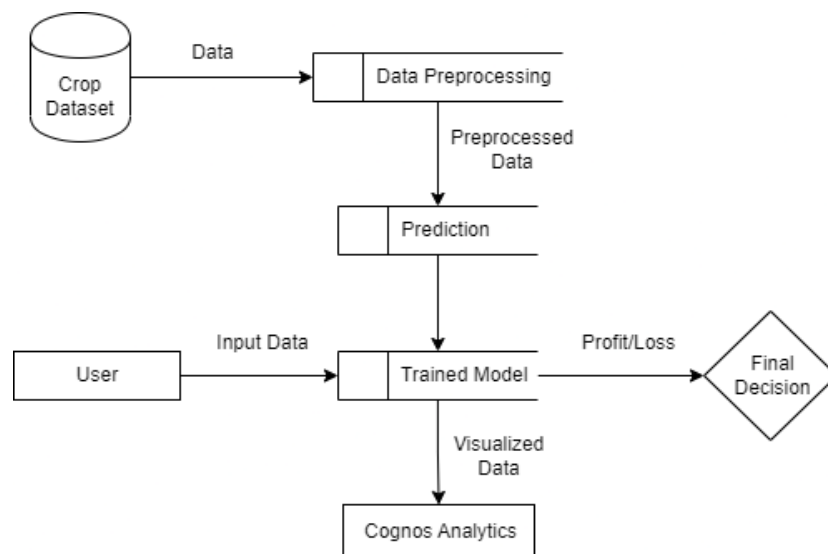
### **Project Flow**

- ☐ Users create multiple analysis graphs/charts.
- ☐ Using the analyzed chart creation of the Dashboard is done.
- ☐ Saving and Visualizing the final dashboard in the IBM Cognos Analytics.

To accomplish this, we have to complete all the activities and tasks listed below

- ☐ IBM Cloud Account
- ☐ Login to Cognos Analytics
- ☐ Working with the Dataset
- ☐ Understand the Dataset
- ☐ Loading the Dataset
- ☐ Data visualization charts
- ☐ Seasons with average productions
- ☐ With years usage of Area and Production
- ☐ Top 10 States with most area
- ☐ State with crop production
- ☐ States with the crop production along with season (Text Table)
- ☐ Dashboard Creation
- ☐ Export the Analytics

### 5.1 Data flow Diagram:-

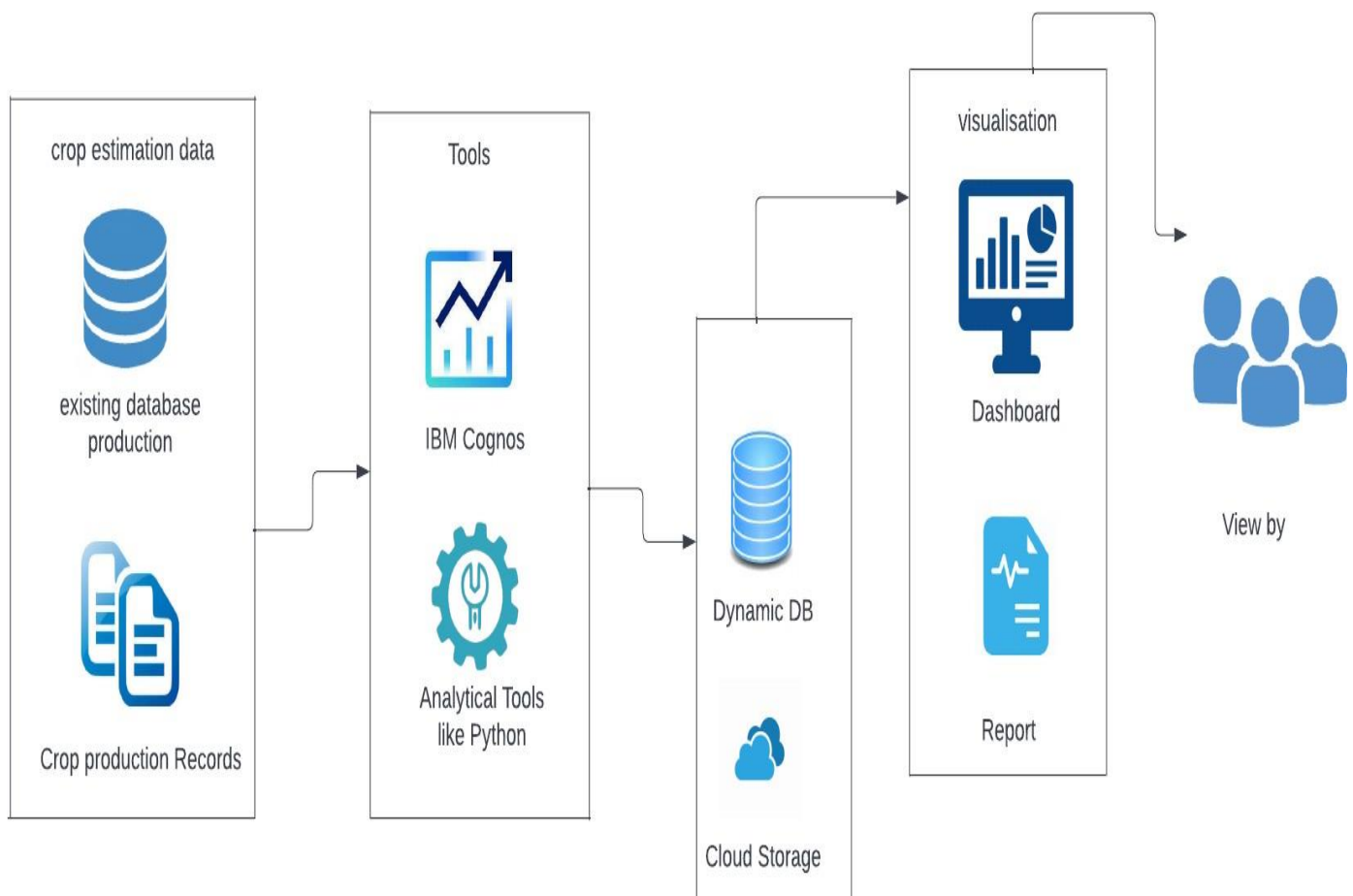


## Solution Architecture:

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

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

Solution Architecture:-





**User Stories:**

User Type	Functional Requirement (Epic)	User Story Number	User Story / Task	Acceptance criteria	Priority	Release
Customer (Web user)	Registration	USN-1	As a user, I can register for the application by entering my email, password, and confirming my password.	I can access my account / dashboard	High	Sprint-1
		USN-2	As a user, I will receive confirmation email once I have registered for the application	I can receive confirmation email & click confirm	High	Sprint-1
		USN-3	As a user, I can register for the application through Facebook	I can register & access the dashboard with Facebook	Low	Sprint-2
		USN-4	As a user, I can register for the application through Gmail		Medium	Sprint-1
	Login	USN-5	As a user, I can log into the application by entering email & password		High	Sprint-1
	Dashboard	USN-6	As a user, I can use the methods used provided in the dashboard		Medium	Sprint-1
		USN-7	As a user, I can view the previous results of predictions done by me		Low	Sprint-3
	Accessing the resources	USN-8	As a user, I can Identify my account with set of unique credentials	The resources can only be accessed by me	High	Sprint-1
Customer Care Executive		USN-9	As customer care executive I will always be available for the interaction with the customer to clarify the queries	An executive will note down the customers complaints and solve their problems	High	Sprint-2
Administrator	Updating data	USN-10	Collecting the data and store it	Checking and updating dataset	High	Sprint-1
Customer Tools	Tools	USN-11	I can perform analysis by tools (cognos and with ML)	Ease of analysis	High	Sprint-2

**7. CODING & SOLUTIONING:-****Model Selection:**

This is the most exciting phase in Applying Machine Learning to any Dataset. It is also known as Algorithm selection for Predicting the best results. Usually Data Scientists use different kinds of Machine Learning algorithms to the large data sets. But, at high level all those different algorithms can be classified in two groups : supervised learning and unsupervised learning. Supervised learning : Supervised learning is a type of system in which both input and desired output data are provided. Input and output data are labeled for classification to provide a

learning basis for future data processing. Supervised learning problems can be further grouped into Regression and Classification problems.

A regression problem is when the output variable is a real or continuous value, such as “salary” or “weight”. A classification problem is when the output variable is a category like filtering emails “spam” or “not spam”

**Unsupervised Learning:** Unsupervised learning is the algorithm using information that is neither classified nor labeled and allowing the algorithm to act on that information without guidance. In our dataset we have the outcome variable or Dependent variable that is why having only two set of values, either M (Malign) or B(Benign). So we will use Classification algorithm of supervised learning.

### **Modules:**

- Data validation and pre-processing technique (Module-01)
- Exploration data analysis of visualization and training a model by given attributes (Module-02)

**Data Validation/ Cleaning/Preparing Process:** Importing the library packages with loading given dataset. To analyzing the variable identification by data shape, data type and evaluating the missing values, duplicate values. A validation dataset is a sample of data held back from training your model that is used to give an estimate of model skill while tuning model's and procedures that you can use to make the best use of validation and test datasets when evaluating your models. Data cleaning /preparing by rename the given dataset and drop the column etc. to analyze the uni-variate, bivariate and multi-variate process. The steps and techniques for data cleaning will vary from dataset to dataset. The primary goal of data cleaning is to detect and remove errors and anomalies to increase the value of data in analytics and decision making.

### **Pre-processing:**

Pre-processing refers to the transformations applied to our data before feeding it to the algorithm. Data Preprocessing is a technique that is used to convert the raw data into a clean data set. In other words, whenever the data is gathered from different sources it is collected in raw format which is not feasible for the analysis.

To achieving better results from the applied model in Machine Learning method of the data has to be in a proper manner. Some specified Machine Learning model needs information in a specified format; for example, Random Forest algorithm does not support null values. Therefore, to execute random forest algorithm null values have to be managed from the original raw data set

### **Exploration data analysis of visualization:**

Data visualization is an important skill in applied statistics and machine learning. Statistics does indeed focus on quantitative descriptions and estimations of data.

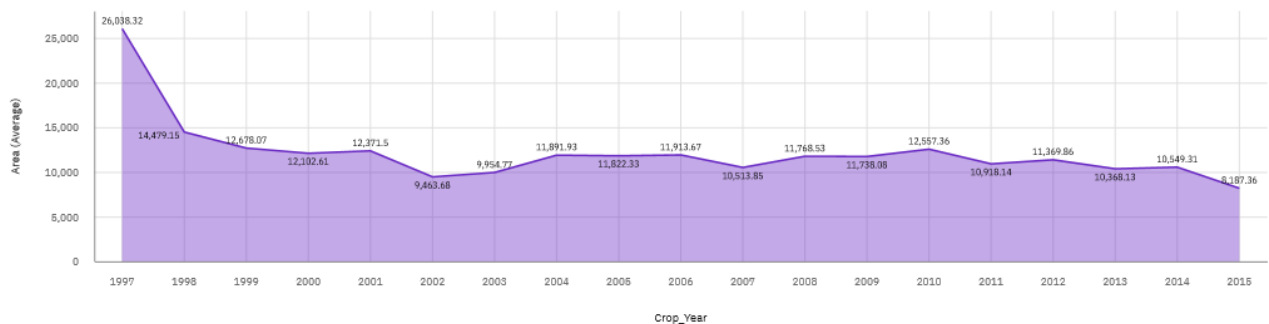
Data visualization provides an important suite of tools for gaining a qualitative understanding. This can be helpful when exploring and getting to know a dataset and can help with identifying patterns, corrupt data, outliers, and much more. With a little domain knowledge, data visualizations can be used to express and demonstrate key relationships in plots and charts that are more visceral and stakeholders than measures of association or significance. Data visualization and exploratory data analysis are whole fields themselves and it will recommend a deeper dive into some the books mentioned at the end.

Sometimes data does not make sense until it can look at in a visual form, such as with charts and plots. Being able to quickly visualize of data samples and others is an important skill both in applied statistics and in applied machine learning. It will discover the many types of plots that you will need to know when visualizing data in Python and how to use them to better understand your own data.

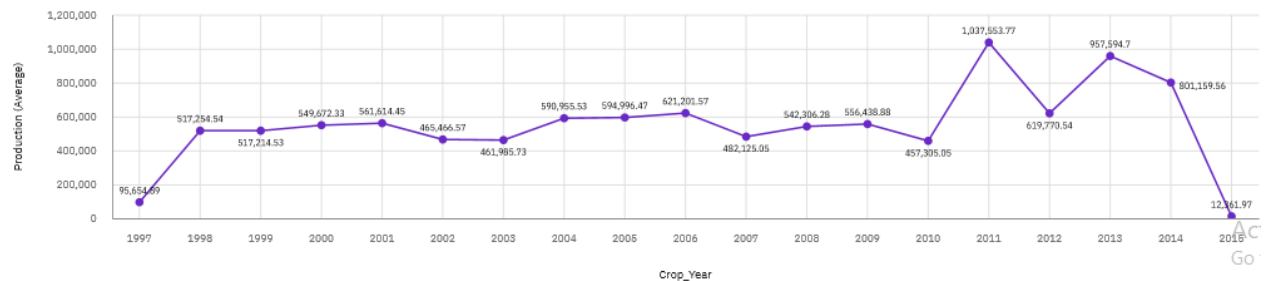
- How to chart time series data with line plots and categorical quantities with bar charts.
- How to summarize data distributions with histograms and box plots.
- How to summarize the relationship between variables with scatter plots.

## With years usage of Area and Productions

Area by Crop\_Year



Production by Crop\_Year



## TESTING:-

### User Acceptance Testing

User Acceptance Testing is a critical phase of any project and requires significant participation by the end user. It also ensures that the system meets the functional requirements.

The purpose of this document is to briefly explain the test coverage and open issues of the [Estimation of crop yield analysis] project at the time of the release to User Acceptance Testing (UAT).

## Test Case Analysis

This report shows the number of test cases that have passed, failed, and untested.

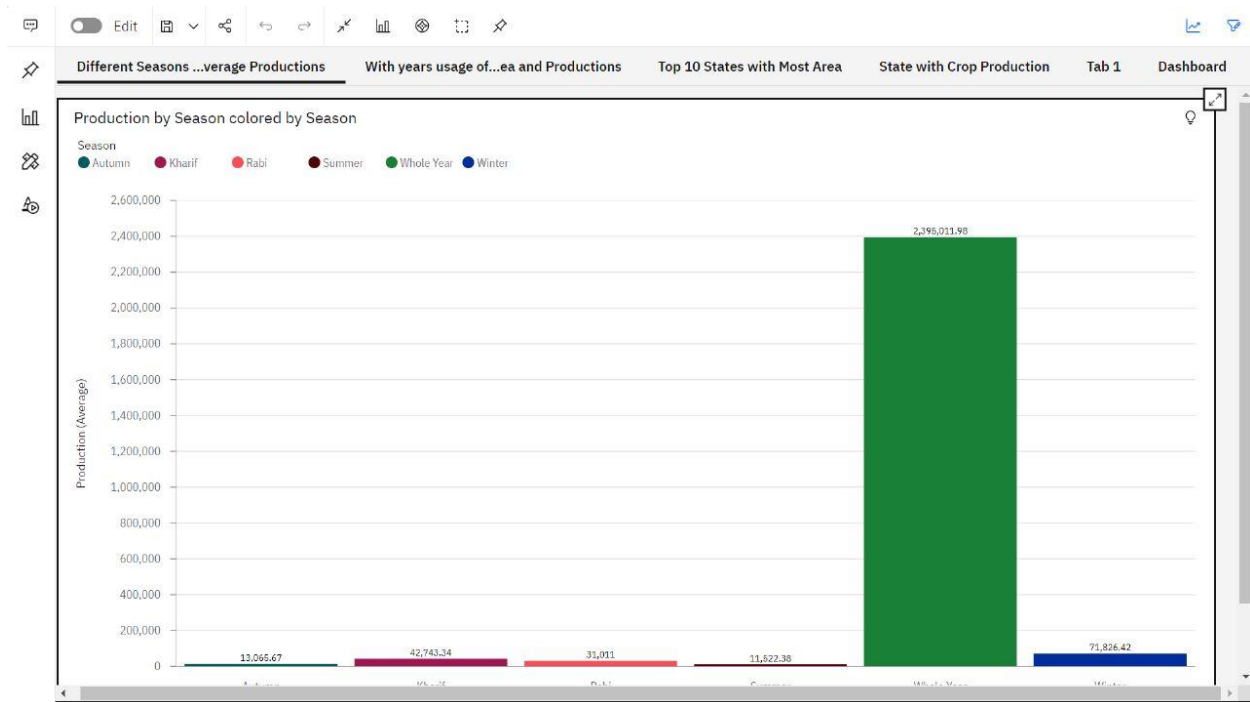
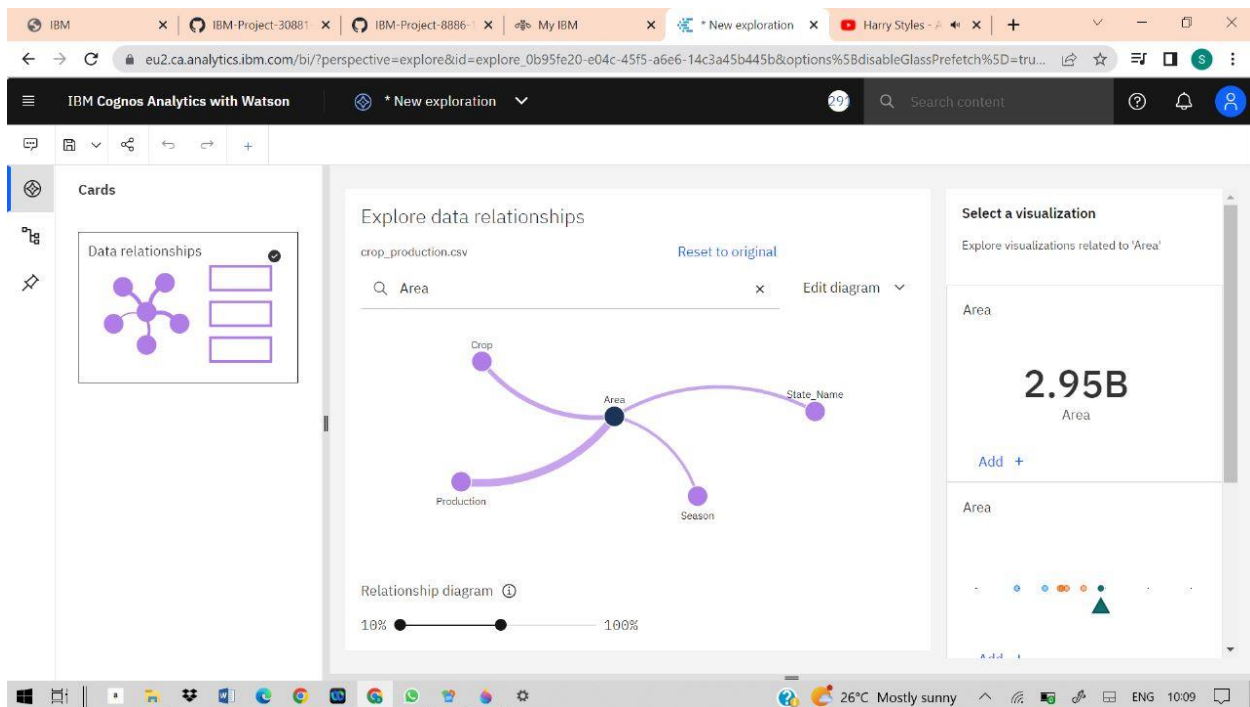
Section	Total Cases	Not Tested	Fail	Pass
User Acceptance	4	0	0	4
Client Application	45	0	2	43
Security	3	0	0	3
Outsource Data	5	0	0	5
Exception Reporting	9	0	0	9
Final Report Output	4	0	0	4
Version Control	2	0	0	2

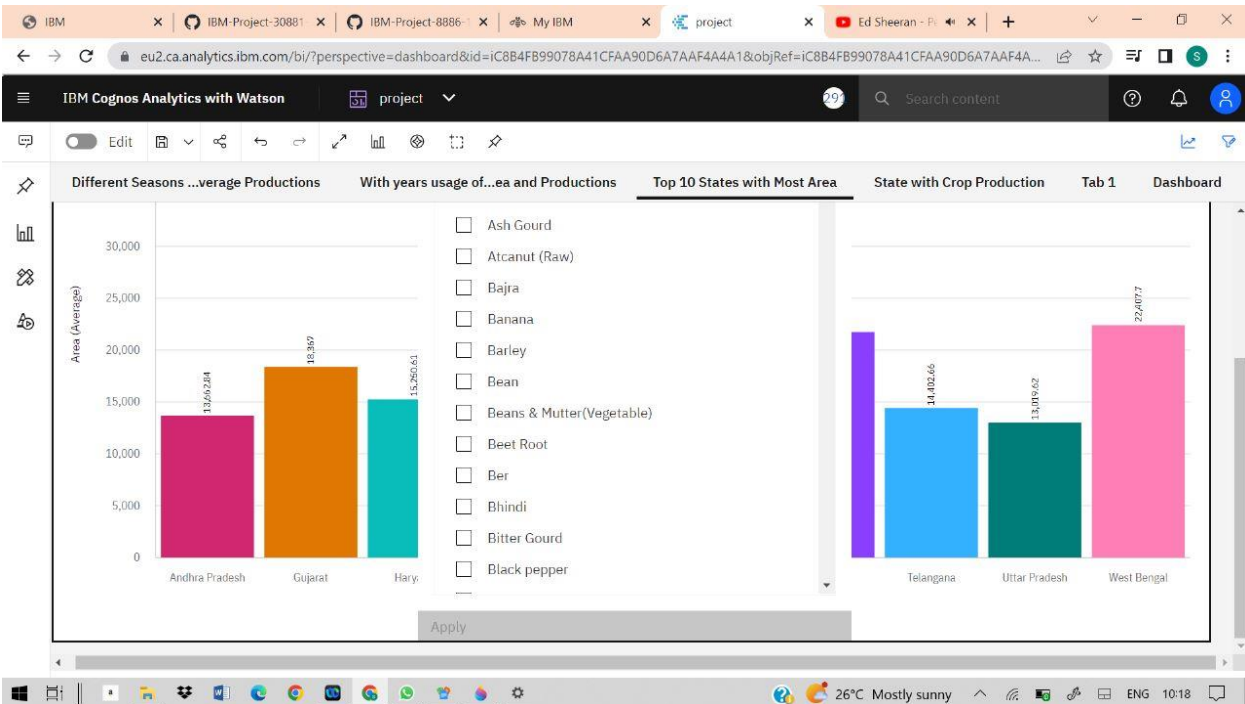
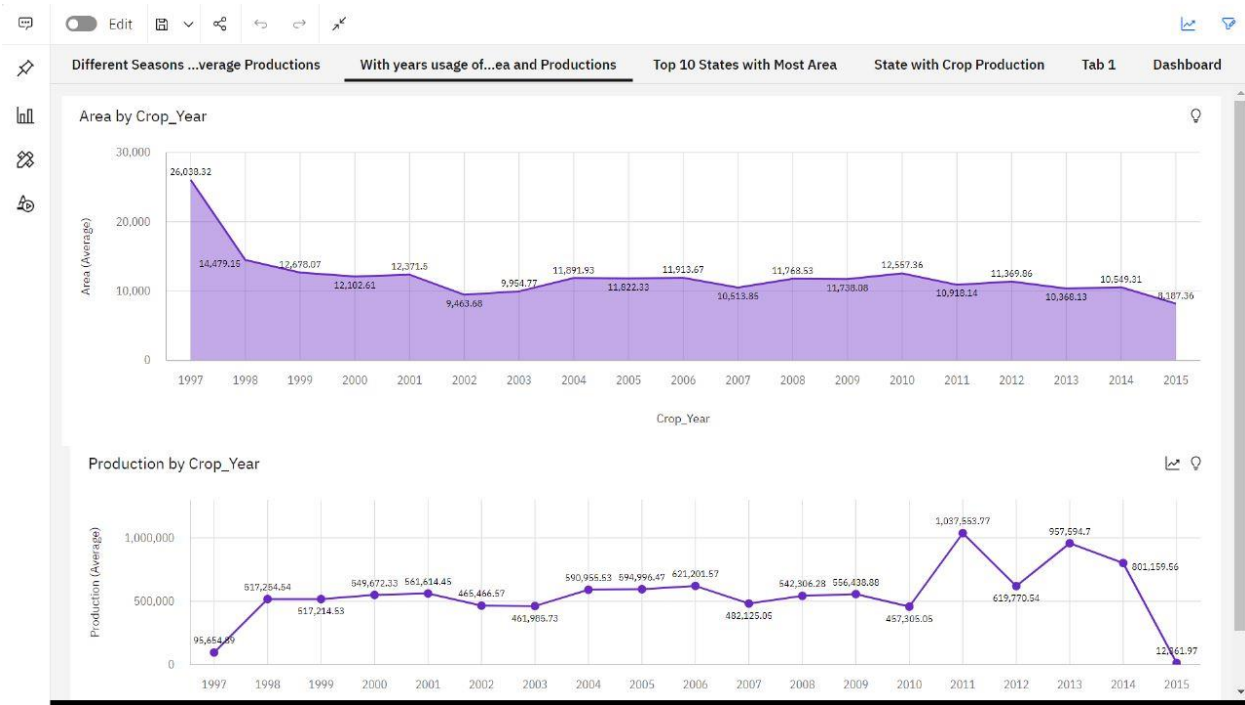
## 9. RESULTS

Using the Crop production in Indian dataset, we plan to create various graphs and charts to highlight the insights and visualizations.

- \* Build a Visualization to showcase Average Crop Production by Seasons.
- \* Showcase the Yearly usage of Area in Crop Production.
- \* Build a visualization to show case top 10 States in Crop Yeild Production by Area.
- \* Build the required Visualization to showcase the Crop Production by State.

\* Build Visual analytics to represent the Sates with Seasonal Crop Production using a Text representation.





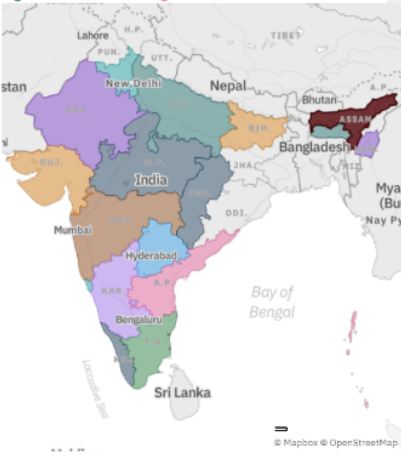
10/25/22, 1:49 PM

project

### State with Crop Production

State\_Name for State\_Name regions

- State\_Name
- Assam
  - Maharashtra
  - Bihar
  - Karnataka
  - Haryana
  - Meghalaya
  - Uttar Pradesh
  - Gujarat
  - Telangana
  - Chhattisgarh
  - Manipur
  - Madhya Pradesh
  - Tamil Nadu
  - Andaman and Nicobar Islands



### Crop

Search

- ☐ Apple
- ☐ Arcanute (Processed)
- ☐ Arcanute
- ☐ Arhar/Tur
- ☐ Ash Gourd
- ☐ Atcanut (Raw)
- ☐ Bajra
- ☒ Banana
- ☐ Barley
- ☐ Bean
- ☐ Beans & Mutter (Vegetable)
- ☐ Beet Root
- ☐ Ber
- ☐ Bhindi
- ☐ Bitter Gourd
- ☐ Black pepper
- ☐ Blackgram
- ☐ Bottle Gourd

Apply

Activate Windows  
Go to Settings to activate Windows.

IBM Cognos Analytics with Watson

project

Search content

Different Seasons ...verage Productions With years usage of...ea and Productions Top 10 States with Most Area State with Crop Production **Tab 1** Dashboard

State\_Name and Crop

Crop	State_Name
Orange	Andhra Pradesh
	Assam
	Madhya Pradesh
	Manipur
	Rajasthan
	Tamil Nadu
	Telangana

Season and Crop

Crop	Season
Orange	Kharif
	Whole Year

26°C Mostly sunny ENG 10:21



## **10.ADVANTAGES & DISADVANTAGES**

### **Advantages:**

- Our goal is push for assisting farmers, government using our predictions. All these publications state they have done better than their competitors but there is no article or public mention of their work being used practically to assist the farmers. If there are some genuine problems in rolling out that work to next stage, then identify those problems and try solving them.
- It is targeted to those farmers who wish to professionally manage their farm by planning, monitoring and analyzing all farming activities.
- Achieving the maximum crop at minimum yield is the ultimate Aim of the project.
- Early detection of problems and management of that problems can help the farmers for better crop yield.

### **Disadvantages:**

- The obtained result for the crop yield prediction using SMO classifier gives less accuracy when compared to naïve Bayes, multilayer perceptron and Bayesian network.
- Previously yield is predicted on the bases of the farmers prior experience but now weather conditions may change drastically so they cannot guess the yield.

## **11. CONCLUSION:**

The analytical process started from data cleaning and processing, missing value, exploratory analysis and finally model building and evaluation. Finally we predict the crop using machine learning algorithm with different results. This brings some of the following insights about crop prediction. As maximum types of crops will be covered under this system, farmer may get to know about the crop which may never have been cultivated and lists out all possible crops, it helps the farmer in decision making of which crop to cultivate. Also, this system takes into

consideration the past production of data which will help the farmer get insight into the demand and the cost of various crops in market.

## 12. FUTURE SCOPE:

- Agricultural department wants to automate the detecting the yield crops from eligibility process (real time).
- To automate this process by show the prediction result in web application or desktop application.

## 13. APPENDIX:-

GitHub Link:-<https://github.com/IBM-EPBL/IBM-Project-30881-1660191752>

Home.html

```
<!DOCTYPE html>
<html lang="en">
<head>
  <meta charset="UTF-8">
  <meta http-equiv="X-UA-Compatible" content="IE=edge">
  <meta name="viewport" content="width=device-width, initial-scale=1.0">
  <title>Home</title>
  <link rel="stylesheet" href="/style/home.css">
  <!-- ===== google font link
===== -->
  <link rel="stylesheet"
href="https://fonts.googleapis.com/css?family=Sofia&effect=3d-float">

<link rel="stylesheet" type="text/css"
href="//fonts.googleapis.com/css?family=Nunito" />

</head>
<body>
  <div class="container">
    <!-- title -->
    <div class="title font-effect-3d-float">Estimate the Crop Yeild using
Data Analytics</div>
    <!--details -->
```

```

<div class="details">
  <p class="teamId">Team ID:PNT2022TMID15473</p>
  <!-- members -->
  <div class="fn">
    <p>YOGESH BALAN A</p>
    <p>THYAANESHWAR S</p>
    <p>VINAY S</p>
    <p>SRIVARSHAN S</p>
  </div>

</div>
<!-- workspace link -->
<div class="workspace-link">
  <a href="workspace.html">Workspace</a>
</div>
</div>
</body>
</html>

```

## Index.html

```

<!DOCTYPE html>
<html lang="en">
<head>
  <meta charset="UTF-8">
  <meta http-equiv="X-UA-Compatible" content="IE=edge">
  <meta name="viewport" content="width=device-width, initial-scale=1.0">
  <link rel="stylesheet" type="text/css" href="./style/index.css">
  <title>Login</title>
  <script type="text/javascript" src="index.js"></script>
</head>
<body>
  <div class="logform">

    <form name="loginForm" >
      <h1>Login</h1>
      <p class="t">User name</p>
      <input type="text" name="uname" id="username" placeholder="Enter User
name">

      <p class="t">Password</p>
      <input type="password" name="upswd" id="password" placeholder="Enter
Password">

```

```

        <div id="ErrorBox"></div><br>

        <button type="button" onclick="login()" class="sub-
btn">Submit</button>
        <input type="reset" class="reset-btn">
        <br>
        <p class="s">Doesn't have a account? New user?<a
href="/register.html">Register</a></p>
    </form>
</div>

</body>
</html>

```

## Workspace.html

```

<!DOCTYPE html>
<html lang="en">
<head>
    <meta charset="UTF-8">
    <meta http-equiv="X-UA-Compatible" content="IE=edge">
    <meta name="viewport" content="width=device-width, initial-scale=1.0">
    <link rel="stylesheet" href="style/workspace.css">
    <title>workspace</title>
</head>
<body>
    <div class="ws-container">
        <div class="ws">
            <iframe
src="https://eu2.ca.analytics.ibm.com/bi/?perspective=dashboard&id=iC8B4FB99078A4
1CFAA90D6A7AAF4A4A1&objRef=iC8B4FB99078A41CFAA90D6A7AAF4A4A1&options%5BdisableGla
ssPrefetch%5D=true&options%5Bcollections%5D%5BcanvasExtension%5D%5Bid%5D=com.ibm.
bi.dashboard.canvasExtension&options%5Bcollections%5D%5BfeatureExtension%5D%5Bid%
5D=com.ibm.bi.dashboard.core-
features&options%5Bcollections%5D%5Bbuttons%5D%5Bid%5D=com.ibm.bi.dashboard.butto
ns&options%5Bcollections%5D%5Bwidget%5D%5Bid%5D=com.ibm.bi.dashboard.widgets&opti
ons%5Bcollections%5D%5BcontentFeatureExtension%5D%5Bid%5D=com.ibm.bi.dashboard.co
ntent-
features&options%5Bcollections%5D%5BsaveServices%5D%5Bid%5D=com.ibm.bi.dashboard.
saveServices&options%5Bcollections%5D%5Btemplates%5D%5Bid%5D=com.ibm.bi.dashboard
.templates&options%5Bcollections%5D%5BvisualizationExtension%5D%5Bid%5D=com.ibm.b
i.dashboard.visualizationExtensionCA&options%5Bcollections%5D%5BboardModel%5D%5Bi

```

```

d%5D=com.ibm.bi.dashboard.boardModelExtension&options%5Bcollections%5D%5BcontentTypes%5D%5Bid%5D=com.ibm.bi.dashboard.contentTypes&options%5Bcollections%5D%5BserviceExtension%5D%5Bid%5D=com.ibm.bi.dashboard.serviceExtension&options%5Bcollections%5D%5BlayoutExtension%5D%5Bid%5D=com.ibm.bi.dashboard.layoutExtension&options%5Bcollections%5D%5BcolorSetExtensions%5D%5Bid%5D=com.ibm.bi.dashboard.colorSetExtensions&options%5Bconfig%5D%5Bproduct%5D=CA&options%5Bconfig%5D%5BeditPropertiesLabel%5D=true&options%5Bconfig%5D%5BenableCustomVisualizations%5D=true&options%5Bconfig%5D%5BassetTags%5D%5B%5D=dashboard&options%5Bconfig%5D%5BfilterDock%5D=true&options%5Bconfig%5D%5BshowMembers%5D=true&options%5Bconfig%5D%5Bupgrades%5D=dashboard-
core%2Fjs%2Fdashboard%2Fupgrades&options%5Bconfig%5D%5BassetType%5D=exploration&options%5Bconfig%5D%5BgeoService%5D=CA&options%5Bconfig%5D%5BsmartTitle%5D=true&options%5Bconfig%5D%5BnavigationGroupAction%5D=true&options%5Bconfig%5D%5BenableDataQuality%5D=false&options%5Bconfig%5D%5BmemberCalculation%5D=false&isAuthoringMode=false&boardId=iC8B4FB99078A41CFAA90D6A7AAF4A4A1" width="100%" height="100%" frameborder="0" gesture="media" allow="encrypted-media" allowfullscreen=""></iframe>
    </div>
    <div class="homepage-link">
        <a href="home.html">Go to Homepage</a>
    </div>
</div>
</body>
</html>

```

## Register.html

```

<!DOCTYPE html>
<html lang="en">
<head>
    <meta charset="UTF-8">
    <meta http-equiv="X-UA-Compatible" content="IE=edge">
    <meta name="viewport" content="width=device-width, initial-scale=1.0">
    <script src="index.js"></script>
    <link rel="stylesheet" href="./style/index.css">
    <title>register</title>
</head>
<body>
    <div class="logform">

        <form name="regForm" >
            <h1>Register</h1>
            <p class="t">User name</p>

```

```
        <input type="text" name="uname1" id="username" placeholder="Enter
user name">

        <p class="t">Password</p>
        <input type="password" name="password1" id="password"
placeholder="Enter Password">
        <p class="p">Confirm Password</p>
        <input type="password" name="repassword1" id="password1"
placeholder="Enter Comfirm password">

        <br>

        <button type="button" onclick="register()"
class="sub-btn">Submit</button>
        <input type="reset" class="reset-btn">

    </form>
    <p class="s">Already user?<a href="index.html">Login</a></p>

</div>

</body>
</html>
```