LOW LEVEL DESIGN (LLD) DOCUMENT
CROP PRODUCTION DATA ANALYSIS

Revision Number: 1.0

Last date of Revision: 25/09/2021

Submitted by:

Dattatreya Thunuguntla

Document Version Control:

Date Issued	Version	Description	Author
25th September,2021	1.0	First Version of Complete LLD	Dattatreya Thunuguntla

Contents:

S. No	Topic	Page No
I	Document Version Control	2
1	Introduction	4
1.1	What is Low Level Design Document	4
1.2	Scope	4
2	Architecture	5
2.1	Tableau Architecture	5
3	Architecture Description	6
3.1	Data Description	6
3.2	Web Scraping	6
3.3	Data Transformation	6
3.4	Data Exploration & Exploratory Data Analysis:	9
3.5	Creating relationships between Parameters	9
3.5	Deployment	15
4	Unit Test Cases	17

1. Introduction

1.1 What is Low Level Design Document

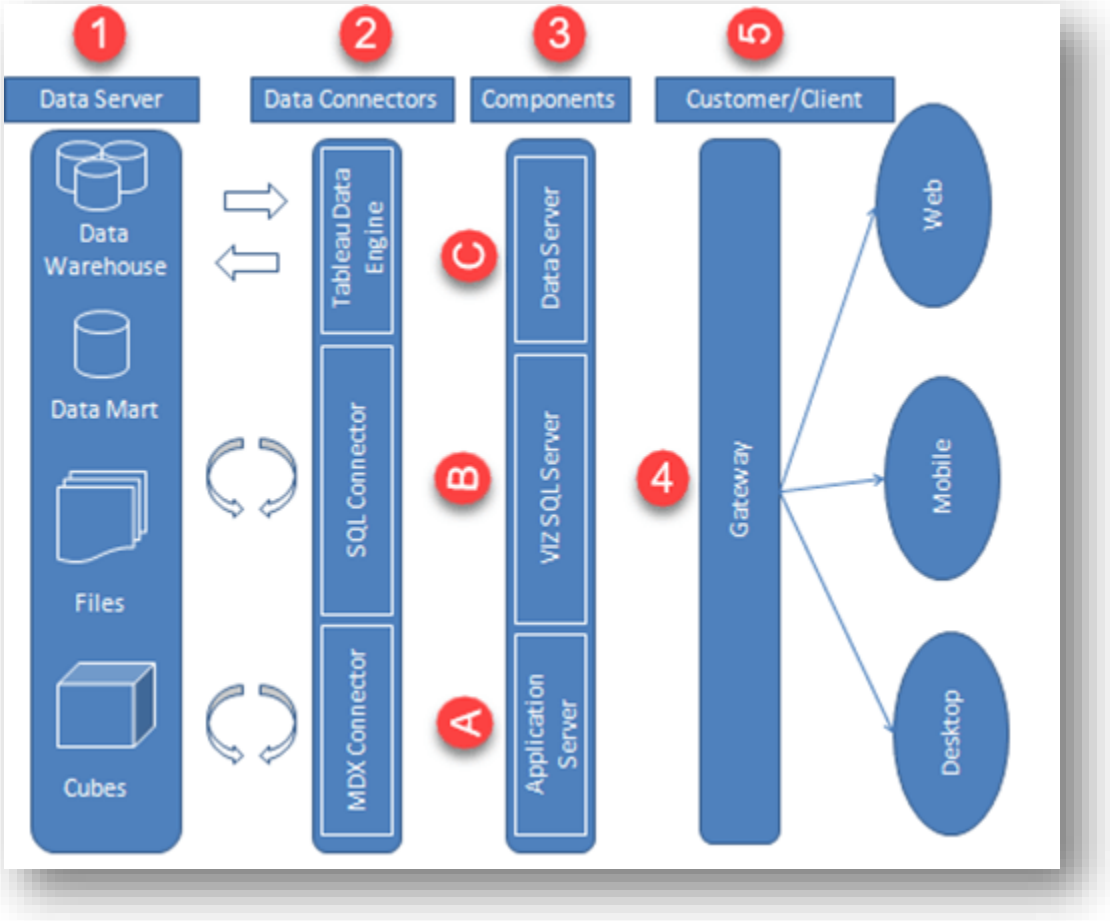
The goal of the LDD or Low-level design document (LLDD) is to give the internal logic design of the actual program code for the House Price Prediction dashboard. LDD describes the class diagrams with the methods and relations between classes and programs specs. It describes the modules so that the programmer can directly code the program from the document.

1.2 Scope

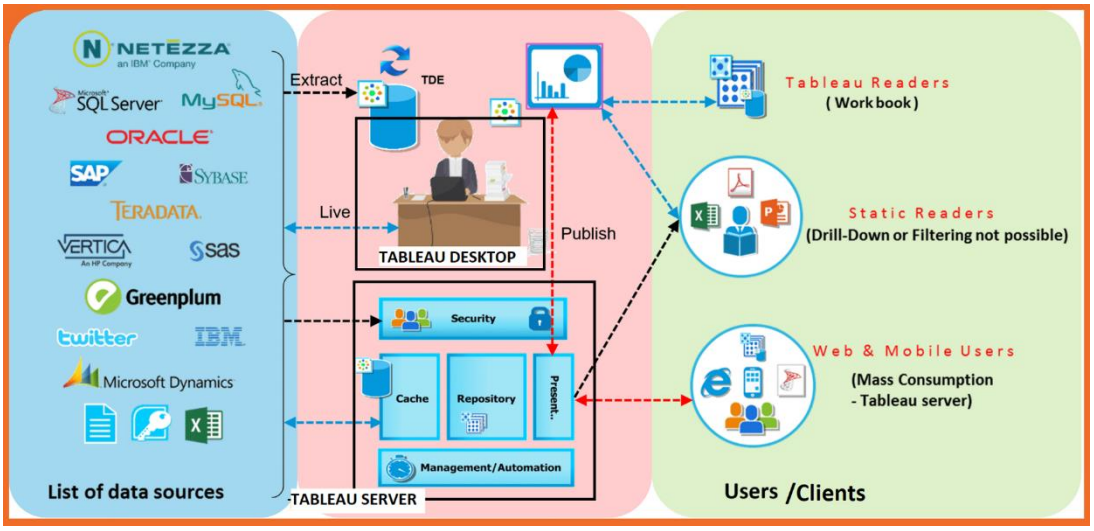
Low-level design (LLD) is a component-level design process that follows a step-by-step refinement process. The process can be used for designing data structures, required software architecture, source code and ultimately, performance algorithms. Overall, the data organization may be defined during requirement analysis and then refined during data design work.

2. Architecture:

2.1 TABLEAU Architecture:



Functional Architecture of Business Intelligence



3. Architecture Description:

3.1 Data Description:

The Crop Production Data consists of following columns:

- **State Name: Name of the State**
- **District Name: Name of the District**
- **Crop Year: What's the year**
- **Season: Crop Season**
- **Crop: Type of Crop grown in the field**
- **Area: In how much area does the crop has grown (in Hectares)**
- **Production: What's the total crop production (in Tonnes)**

3.2 Web Scraping:

Web scraping is a technique to automatically extract content and data from websites using bots. It is also known as web data extraction or web harvesting. Web scrapping is made simple now days, many tools are used for web scrapping. Some of python libraries used for web scrapping are Beautiful Soup, Scrapy, Selenium, etc.

3.3 Data Transformation:

In the Transformation Process, we will convert our original datasets with other necessary attributes format. For the given Data Set names of the Columns have been changed and Null Values, Error Values have been removed from the Data Set.

Data Transformation and manipulation using Pandas library and visualized in Matplotlib library. Firstly, we import the libraries and do the data manipulation & EDA.

Importing Libraries

In [1]:
```
import pandas as pd
import matplotlib.pyplot as plt
import streamlit as st
```

In [2]:
```
import matplotlib
matplotlib.rcParams['font.size'] = 14
matplotlib.rcParams['figure.figsize'] = (9, 5)
matplotlib.rcParams['figure.facecolor'] = '#00000000'
```

In [3]:
```
%matplotlib inline
```

Given Data Set

In [12]:
```
df=pd.read_csv("E:\Downloads_E\crop_production.csv")
```

In [13]:
```
df
```

Out[13]:

	State_Name	District_Name	Crop_Year	Season	Crop	Area	Production
0	Andaman and Nicobar Islands	NICOBARS	2000	Kharif	Arecanut	1254.0	2000.0
1	Andaman and Nicobar Islands	NICOBARS	2000	Kharif	Other Kharif pulses	2.0	1.0
2	Andaman and Nicobar Islands	NICOBARS	2000	Kharif	Rice	102.0	321.0
3	Andaman and Nicobar Islands	NICOBARS	2000	Whole Year	Banana	176.0	641.0
4	Andaman and Nicobar Islands	NICOBARS	2000	Whole Year	Cashewnut	720.0	165.0
...	...	...	...	...	...	...	...
246086	West Bengal	PURULIA	2014	Summer	Rice	306.0	801.0
246087	West Bengal	PURULIA	2014	Summer	Sesamum	627.0	463.0
246088	West Bengal	PURULIA	2014	Whole Year	Sugarcane	324.0	16250.0
246089	West Bengal	PURULIA	2014	Winter	Rice	279151.0	597899.0
246090	West Bengal	PURULIA	2014	Winter	Sesamum	175.0	88.0

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 246091 entries, 0 to 246090
Data columns (total 7 columns):
 #   Column         Non-Null Count   Dtype
---  ------         --------------   -----
 0   State_Name     246091 non-null  object
 1   District_Name  246091 non-null  object
 2   Crop_Year      246091 non-null  int64
 3   Season         246091 non-null  object
 4   Crop           246091 non-null  object
 5   Area           246091 non-null  float64
 6   Production     242361 non-null  float64
dtypes: float64(2), int64(1), object(4)
memory usage: 13.1+ MB
```

```
df.describe()
```

	Crop_Year	Area	Production
count	246091.000000	2.460910e+05	2.423610e+05
mean	2005.643018	1.200282e+04	5.825034e+05
std	4.952164	5.052340e+04	1.706581e+07
min	1997.000000	4.000000e-02	0.000000e+00
25%	2002.000000	8.000000e+01	8.800000e+01
50%	2006.000000	5.820000e+02	7.290000e+02
75%	2010.000000	4.392000e+03	7.023000e+03
max	2015.000000	8.580100e+06	1.250800e+09

There are missing values lets do some data manipulation techniques using Transform & Fillna()

```
df.loc[df.Production.isnull(),'Production']=df.groupby(['Crop_Year','Crop']).Production.transform('mean')
```

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 246091 entries, 0 to 246090
Data columns (total 7 columns):
 #   Column         Non-Null Count   Dtype
---  ------         --------------   -----
 0   State_Name     246091 non-null  object
 1   District_Name  246091 non-null  object
 2   Crop_Year      246091 non-null  int64
 3   Season         246091 non-null  object
 4   Crop           246091 non-null  object
 5   Area           246091 non-null  float64
 6   Production     246035 non-null  float64
dtypes: float64(2), int64(1), object(4)
memory usage: 13.1+ MB
```

```
df.describe()
```

	Crop_Year	Area	Production
count	246091.000000	2.460910e+05	2.460350e+05
mean	2005.643018	1.200282e+04	5.843780e+05
std	4.952164	5.052340e+04	1.696752e+07
min	1997.000000	4.000000e-02	0.000000e+00
25%	2002.000000	8.000000e+01	9.100000e+01
50%	2006.000000	5.820000e+02	7.670000e+02
75%	2010.000000	4.392000e+03	7.099500e+03

```
df['Production']=df['Production'].fillna(df['Production'].median())
```

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 246091 entries, 0 to 246090
Data columns (total 7 columns):
 #   Column         Non-Null Count   Dtype
---  ------         --------------   -----
 0   State_Name     246091 non-null  object
 1   District_Name  246091 non-null  object
 2   Crop_Year      246091 non-null  int64
 3   Season         246091 non-null  object
 4   Crop           246091 non-null  object
 5   Area           246091 non-null  float64
 6   Production     246091 non-null  float64
dtypes: float64(2), int64(1), object(4)
memory usage: 13.1+ MB
```

```
df.describe()
```

	Crop_Year	Area	Production
count	246091.000000	2.460910e+05	2.460910e+05
mean	2005.643018	1.200282e+04	5.842452e+05
std	4.952164	5.052340e+04	1.696559e+07
min	1997.000000	4.000000e-02	0.000000e+00
25%	2002.000000	8.000000e+01	9.100000e+01
50%	2006.000000	5.820000e+02	7.670000e+02
75%	2010.000000	4.392000e+03	7.095500e+03
max	2015.000000	8.580100e+06	1.250800e+09

```
df.to_csv('datasource.csv')
```

3.4 Data Exploration & Exploratory Data Analysis:

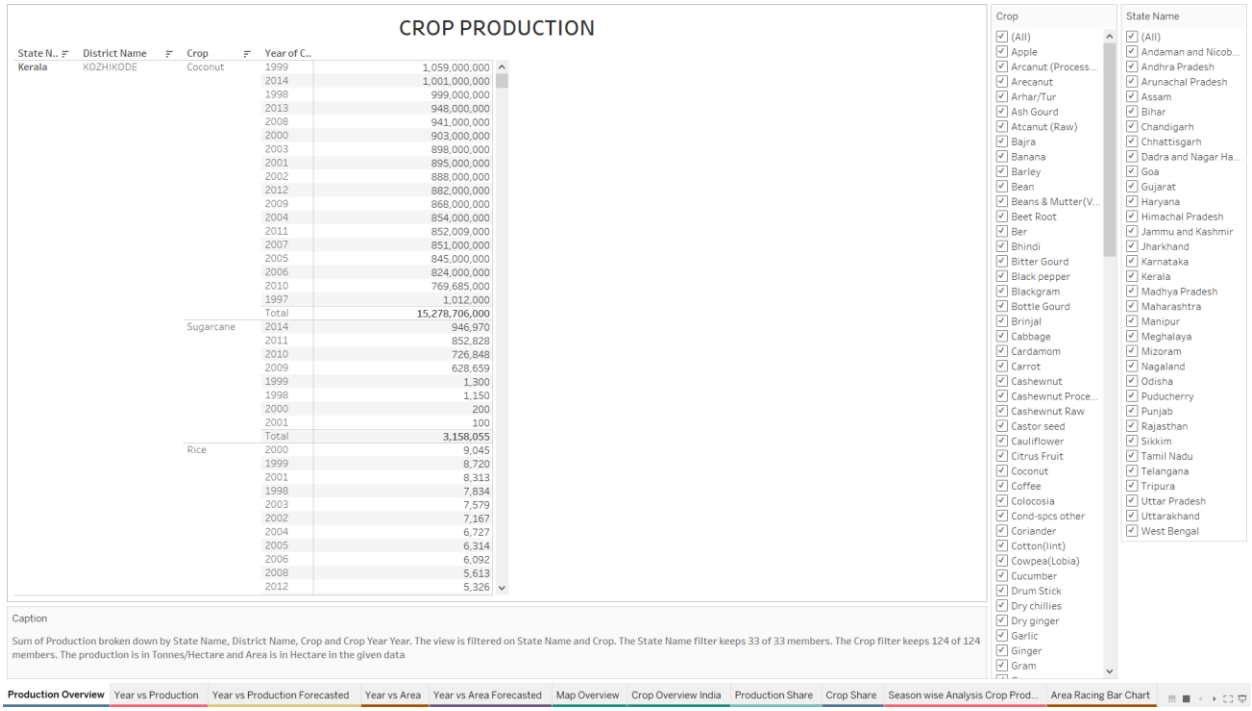
The file which has been saved as Datasource.csv has been analyzed and you can see the EDA operations in the following link: https://github.com/dattu-98/crops_india/blob/master/Crop%20Production%20EDA.ipynb

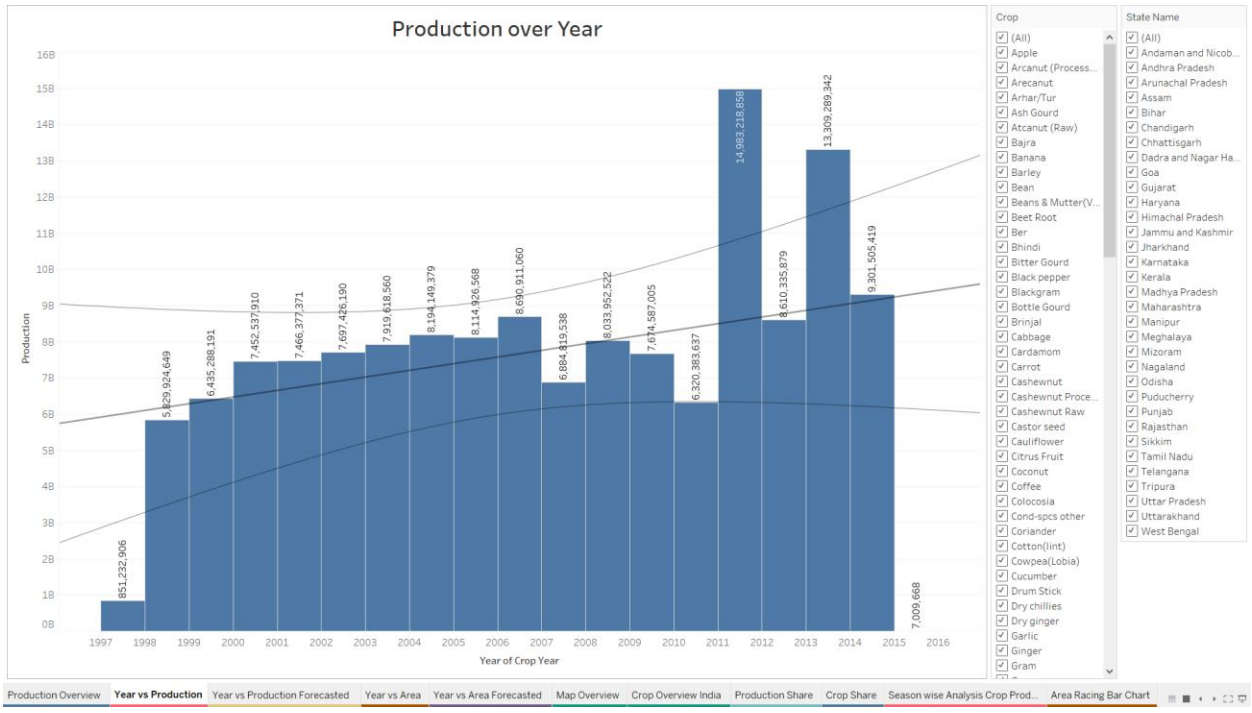
3.5 Creating relations between Parameters.

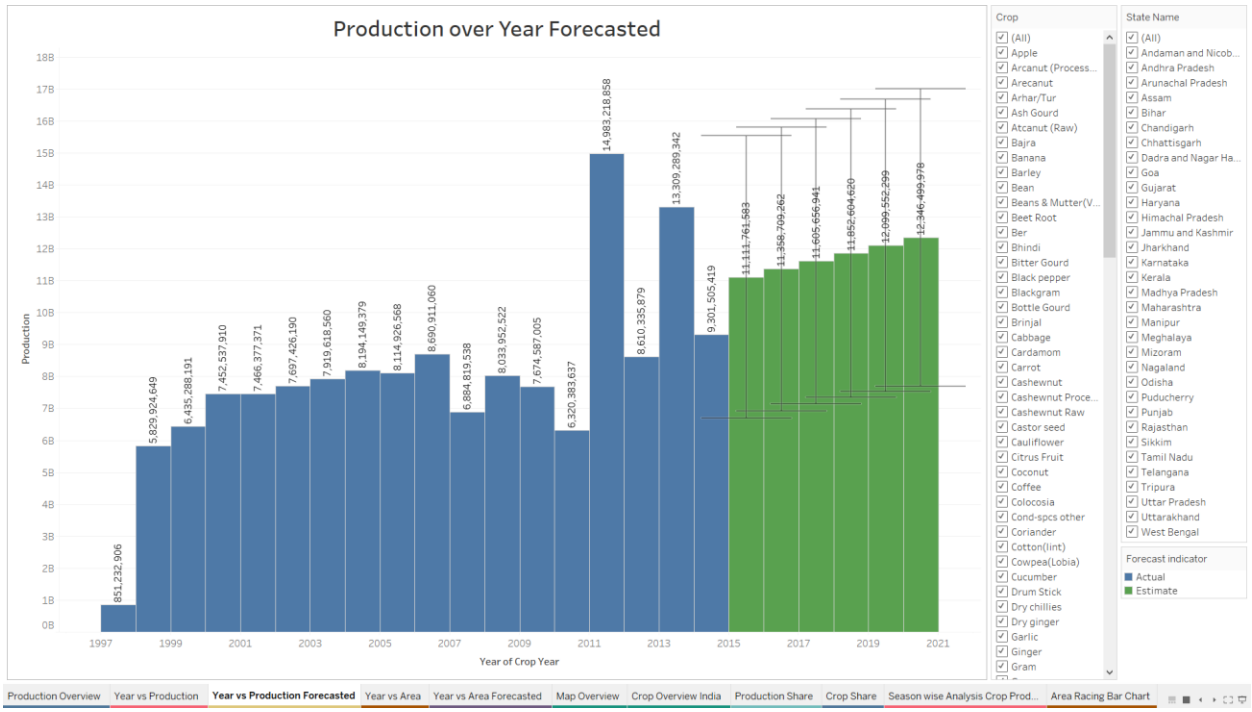
In this Project we had used many types of Visualizations like

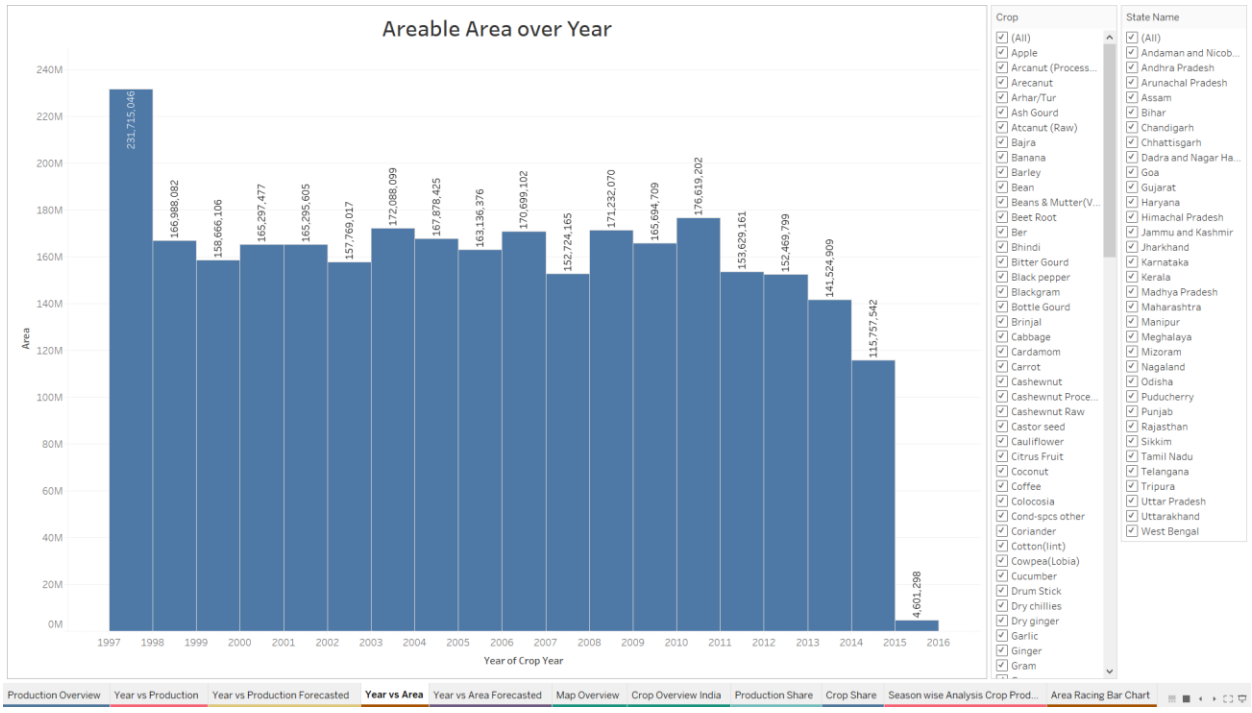
1. Text Tables
2. Symbol Maps
3. Maps
4. Pie Charts
5. Horizontal Bars
6. Treemap
7. Forecasting Models

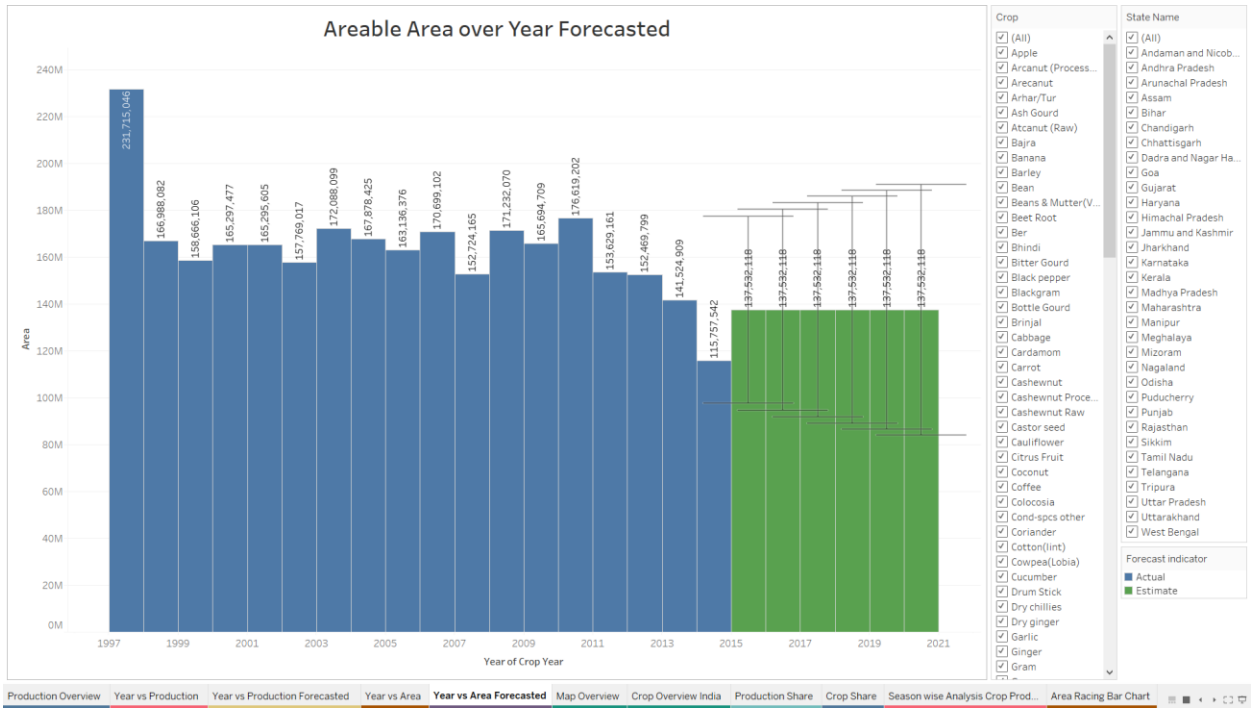
With the use of all the available parameters we had plotted visualizations.

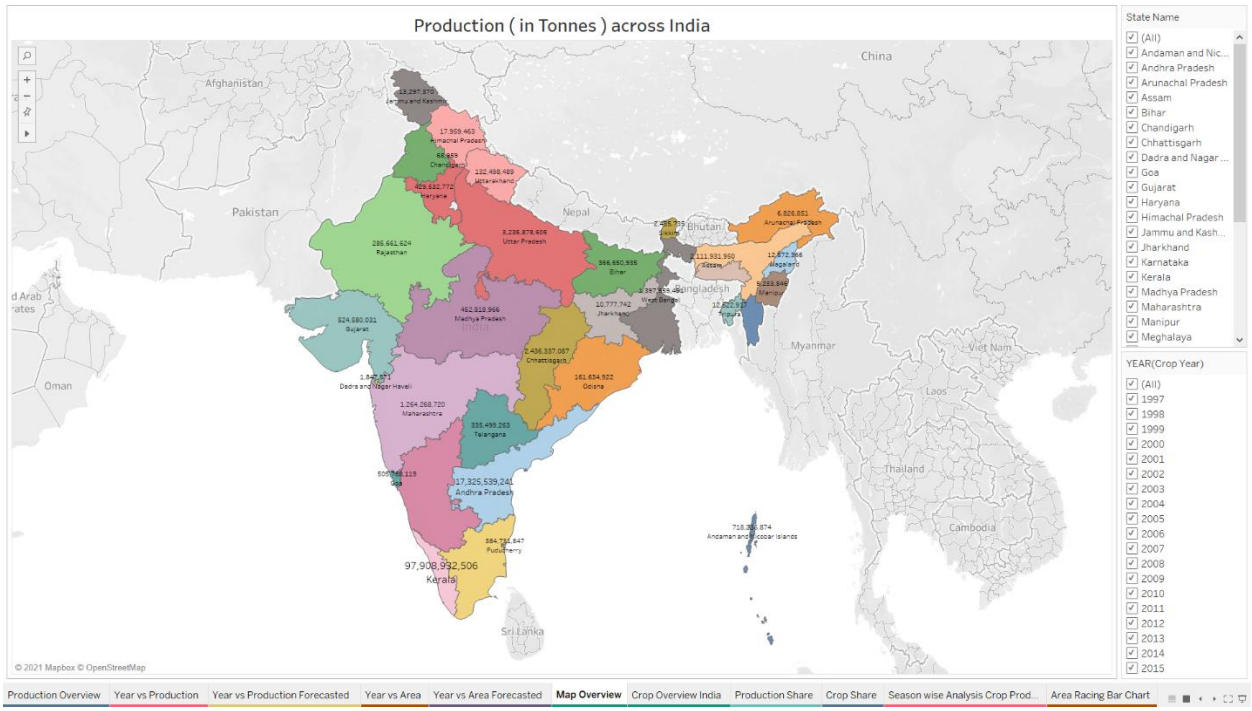


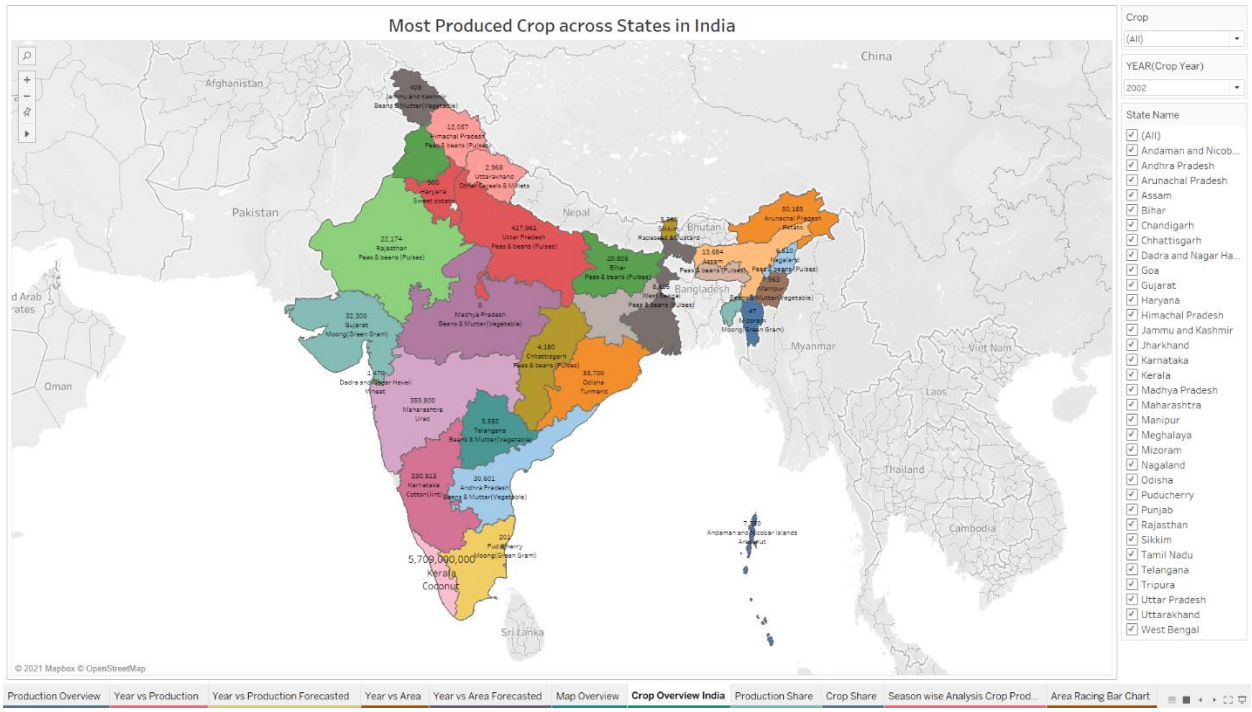

Production over Year
Production
Year of Crop Year
851,232,906
5,829,924,649
6,435,288,191
7,452,537,910
7,466,377,371
7,697,426,190
7,919,618,560
8,194,149,379
8,114,926,568
8,690,911,060
6,884,819,538
8,033,952,522
7,674,587,005
6,320,383,637
14,983,218,858
8,610,335,879
13,309,289,342
9,301,505,419
7,009,668
Crop
(All)
Apple
Arcanut (Process...
Arecanut
Arhar/Tur
Ash Gourd
Atcanut (Raw)
Bajra
Banana
Barley
Bean
Beans & Mutter(V...
Beet Root
Ber
Bhindi
Bitter Gourd
Black pepper
Blackgram
Bottle Gourd
Brinjal
Cabbage
Cardamom
Carrot
Cashewnut
Cashewnut Proce...
Cashewnut Raw
Castor seed
Cauliflower
Citrus Fruit
Coconut
Coffee
Colocosia
Cond-spcs other
Coriander
Cotton(lint)
Cowpea(Lobia)
Cucumber
Drum Stick
Dry chillies
Dry ginger
Garlic
Ginger
Gram
State Name
(All)
Andaman and Nicob...
Andhra Pradesh
Arunachal Pradesh
Assam
Bihar
Chandigarh
Chhattisgarh
Dadra and Nagar Ha...
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
Production Overview
Year vs Production
Year vs Production Forecasted
Year vs Area
Year vs Area Forecasted
Map Overview
Crop Overview India
Production Share
Crop Share
Season wise Analysis Crop Prod...
Area Racing Bar Chart

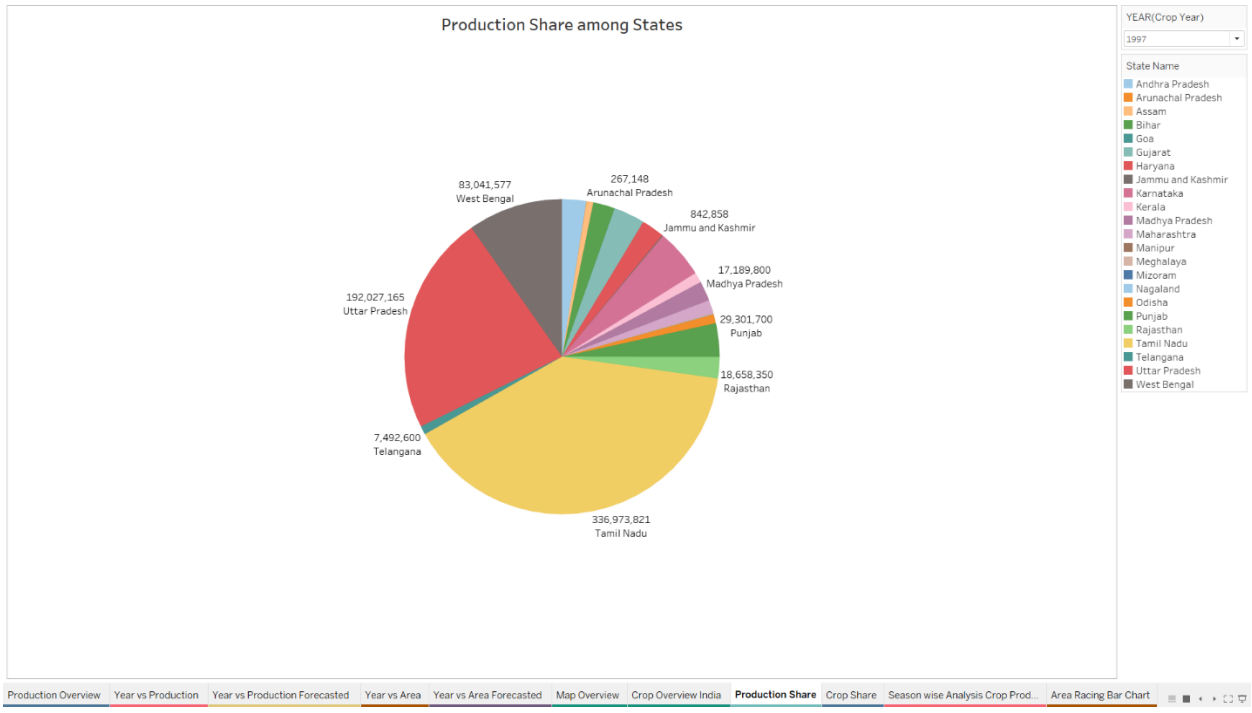

Production over Year Forecasted
Production
Year of Crop Year
851,232,906
5,829,924,649
6,435,288,191
7,452,537,910
7,466,377,371
7,697,426,190
7,919,618,560
8,194,149,379
8,114,926,568
8,690,911,060
6,884,819,538
8,033,952,522
7,674,587,005
6,320,383,637
14,983,218,858
8,610,335,879
13,309,289,342
9,301,505,419
11,111,761,583
11,398,709,262
11,695,656,941
11,852,604,620
12,099,552,299
12,346,499,978
Crop
(All)
Apple
Arcanut (Process...
Arecanut
Arhar/Tur
Ash Gourd
Atcanut (Raw)
Bajra
Banana
Barley
Bean
Beans & Mutter(V...
Beet Root
Ber
Bhindi
Bitter Gourd
Black pepper
Blackgram
Bottle Gourd
Brinjal
Cabbage
Cardamom
Carrot
Cashewnut
Cashewnut Proce...
Cashewnut Raw
Castor seed
Cauliflower
Citrus Fruit
Coconut
Coffee
Colocosia
Cond-spcs other
Coriander
Cotton(lint)
Cowpea(Lobia)
Cucumber
Drum Stick
Dry chillies
Dry ginger
Garlic
Ginger
Gram
State Name
(All)
Andaman and Nicob...
Andhra Pradesh
Arunachal Pradesh
Assam
Bihar
Chandigarh
Chhattisgarh
Dadra and Nagar Ha...
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
Forecast indicator
Actual
Estimate
Production Overview
Year vs Production
Year vs Production Forecasted
Year vs Area
Year vs Area Forecasted
Map Overview
Crop Overview India
Production Share
Crop Share
Season wise Analysis Crop Prod...
Area Racing Bar Chart

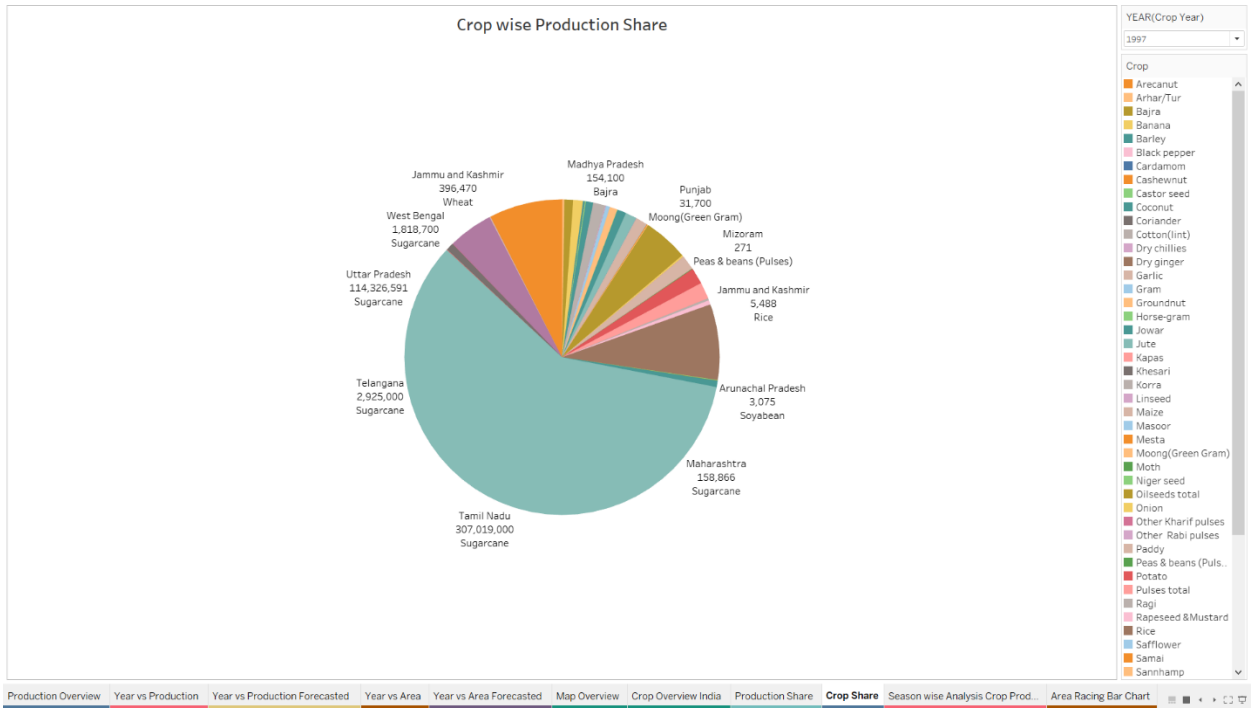

Areable Area over Year
Area
Year of Crop Year
231,715,046
166,988,082
158,666,106
165,297,477
165,295,605
157,769,017
172,088,099
167,878,425
163,136,376
170,699,102
152,724,165
171,232,070
165,694,709
176,619,202
153,629,161
152,469,799
141,524,909
115,757,542
4,601,298
Crop
State Name
Production Overview
Year vs Production
Year vs Production Forecasted
Year vs Area
Year vs Area Forecasted
Map Overview
Crop Overview India
Production Share
Crop Share
Season wise Analysis Crop Prod...
Area Racing Bar Chart

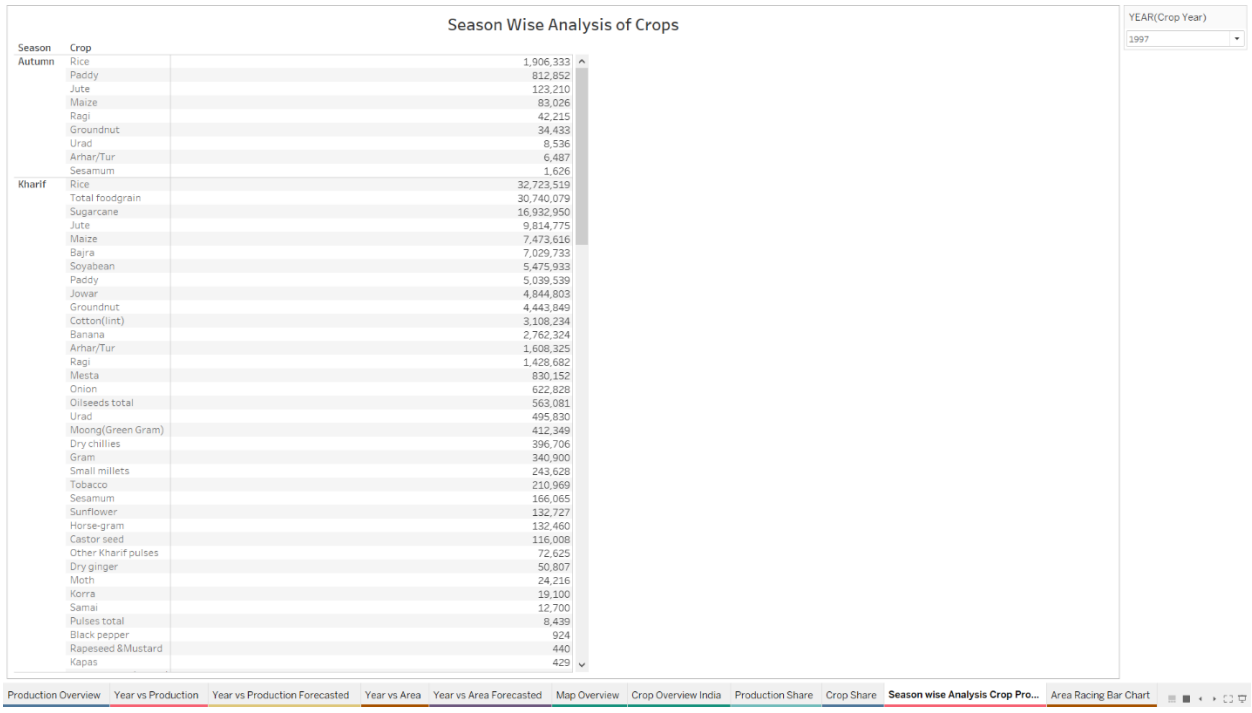

Areable Area over Year Forecasted
Area
Year of Crop Year
231,715,046
166,988,082
158,666,106
165,297,477
165,295,605
157,769,017
172,088,099
167,878,425
163,136,376
170,699,102
152,724,165
171,232,070
165,694,709
176,619,202
153,629,161
152,469,799
141,524,909
115,757,542
137,532,118
Crop
State Name
Forecast indicator
Actual
Estimate
Production Overview
Year vs Production
Year vs Production Forecasted
Year vs Area
Year vs Area Forecasted
Map Overview
Crop Overview India
Production Share
Crop Share
Season wise Analysis Crop Prod...
Area Racing Bar Chart

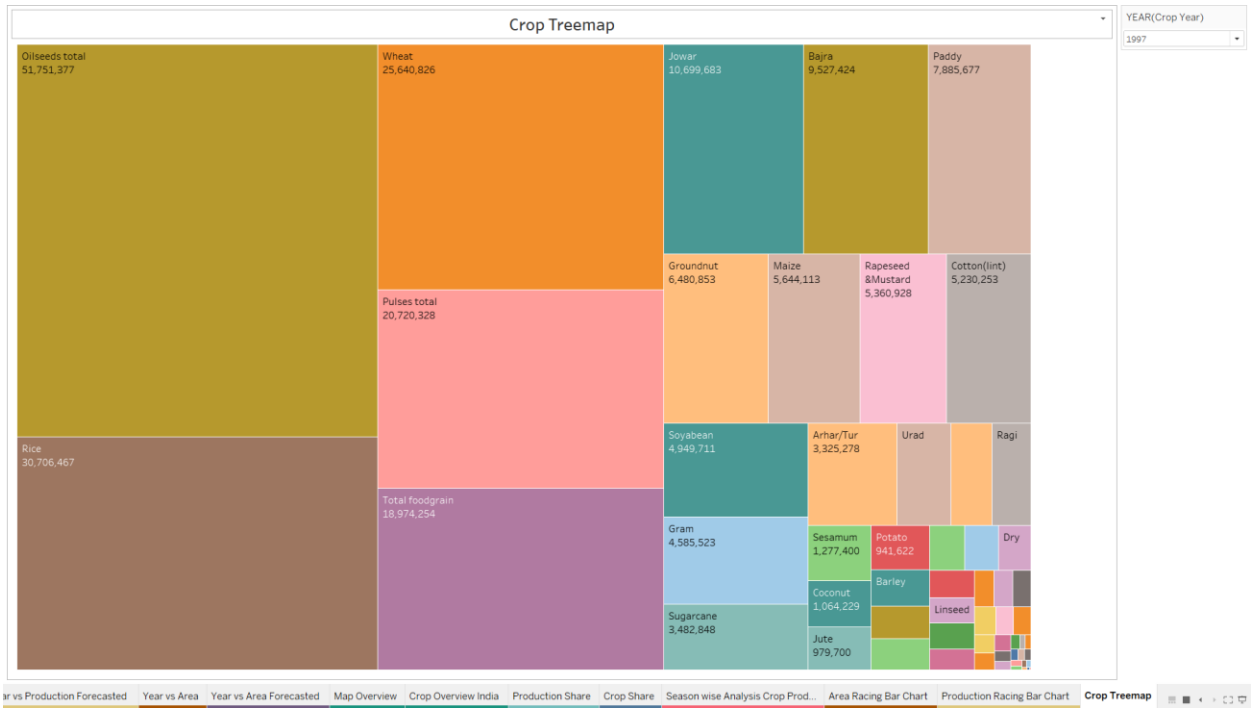

Production (in Tonnes) across India
China
Afghanistan
Pakistan
Nepal
Bhutan
Bangladesh
Myanmar
Viet Nam
Laos
Thailand
Cambodia
Oman
Sri Lanka
Arabian Sea
17,325,539,241
Andhra Pradesh
97,908,932,506
Kerala
State Name
(All)
Andaman and Nic...
Andhra Pradesh
Arunachal Pradesh
Assam
Bihar
Chandigarh
Chhattisgarh
Dadra and Nagar ...
Goa
Gujarat
Haryana
Himachal Pradesh
Jammu and Kash...
Jharkhand
Karnataka
Kerala
Madhya Pradesh
Maharashtra
Manipur
Meghalaya
YEAR(Crop Year)
(All)
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
© 2021 Mapbox © OpenStreetMap
Production Overview
Year vs Production
Year vs Production Forecasted
Year vs Area
Year vs Area Forecasted
Map Overview
Crop Overview India
Production Share
Crop Share
Season wise Analysis Crop Prod...
Area Racing Bar Chart


Most Produced Crop across States in India
China
Afghanistan
Pakistan
Nepal
Bhutan
Bangladesh
Myanmar
Viet Nam
Laos
Thailand
Cambodia
Oman
Sri Lanka
Arabian Sea
5,709,090,000
Kerala
Coconut
Crop
(All)
YEAR(Crop Year)
2002
State Name
(All)
Andaman and Nicob...
Andhra Pradesh
Arunachal Pradesh
Assam
Bihar
Chandigarh
Chhattisgarh
Dadra and Nagar Ha...
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
© 2021 Mapbox © OpenStreetMap
Production Overview
Year vs Production
Year vs Production Forecasted
Year vs Area
Year vs Area Forecasted
Map Overview
Crop Overview India
Production Share
Crop Share
Season wise Analysis Crop Prod...
Area Racing Bar Chart


Production Share among States
YEAR(Crop Year)
1997
State Name
Andhra Pradesh
Arunachal Pradesh
Assam
Bihar
Goa
Gujarat
Haryana
Jammu and Kashmir
Karnataka
Kerala
Madhya Pradesh
Maharashtra
Manipur
Meghalaya
Mizoram
Nagaland
Odisha
Punjab
Rajasthan
Tamil Nadu
Telangana
Uttar Pradesh
West Bengal
83,041,577
West Bengal
267,148
Arunachal Pradesh
842,858
Jammu and Kashmir
17,189,800
Madhya Pradesh
29,301,700
Punjab
18,658,350
Rajasthan
336,973,821
Tamil Nadu
7,492,600
Telangana
192,027,165
Uttar Pradesh
Production Overview
Year vs Production
Year vs Production Forecasted
Year vs Area
Year vs Area Forecasted
Map Overview
Crop Overview India
Production Share
Crop Share
Season wise Analysis Crop Prod...
Area Racing Bar Chart


Crop wise Production Share
YEAR(Crop Year)
1997
Crop
Arecanut
Arhar/Tur
Bajra
Banana
Barley
Black pepper
Cardamom
Cashewnut
Castor seed
Coconut
Coriander
Cotton(lint)
Dry chillies
Dry ginger
Garlic
Gram
Groundnut
Horse-gram
Jowar
Jute
Kapas
Khesari
Korra
Linseed
Maize
Masoor
Mesta
Moong(Green Gram)
Moth
Niger seed
Oilseeds total
Onion
Other Kharif pulses
Other Rabi pulses
Paddy
Peas & beans (Puls..
Potato
Pulses total
Ragi
Rapeseed &Mustard
Rice
Safflower
Samai
Sannhamp
Madhya Pradesh
154,100
Bajra
Jammu and Kashmir
396,470
Wheat
West Bengal
1,818,700
Sugarcane
Uttar Pradesh
114,326,591
Sugarcane
Telangana
2,925,000
Sugarcane
Tamil Nadu
307,019,000
Sugarcane
Maharashtra
158,866
Sugarcane
Arunachal Pradesh
3,075
Soyabean
Jammu and Kashmir
5,488
Rice
Peas & beans (Pulses)
Mizoram
271
Moong(Green Gram)
Punjab
31,700
Production Overview
Year vs Production
Year vs Production Forecasted
Year vs Area
Year vs Area Forecasted
Map Overview
Crop Overview India
Production Share
Crop Share
Season wise Analysis Crop Prod...
Area Racing Bar Chart


Season Wise Analysis of Crops
YEAR(Crop Year)
1997
Season Crop
Autumn Rice 1,906,333
Paddy 812,852
Jute 123,210
Maize 83,026
Ragi 42,215
Groundnut 34,433
Urad 8,536
Arhar/Tur 6,487
Sesamum 1,626
Kharif Rice 32,723,519
Total foodgrain 30,740,079
Sugarcane 16,932,950
Jute 9,814,775
Maize 7,473,616
Bajra 7,029,733
Soyabean 5,475,933
Paddy 5,039,539
Jowar 4,844,803
Groundnut 4,443,849
Cotton(lint) 3,108,234
Banana 2,762,324
Arhar/Tur 1,608,325
Ragi 1,428,682
Mesta 830,152
Onion 622,828
Oilseeds total 563,081
Urad 495,830
Moong(Green Gram) 412,349
Dry chillies 396,706
Gram 340,900
Small millets 243,628
Tobacco 210,969
Sesamum 166,065
Sunflower 132,727
Horse-gram 132,460
Castor seed 116,008
Other Kharif pulses 72,625
Dry ginger 50,807
Moth 24,216
Korra 19,100
Samai 12,700
Pulses total 8,439
Black pepper 924
Rapeseed &Mustard 440
Kapas 429
Production Overview
Year vs Production
Year vs Production Forecasted
Year vs Area
Year vs Area Forecasted
Map Overview
Crop Overview India
Production Share
Crop Share
Season wise Analysis Crop Pro...
Area Racing Bar Chart


Crop Treemap
YEAR(Crop Year)
1997
Oilseeds total
51,751,377
Wheat
25,640,826
Jowar
10,699,683
Bajra
9,527,424
Paddy
7,885,677
Pulses total
20,720,328
Groundnut
6,480,853
Maize
5,644,113
Rapeseed
&Mustard
5,360,928
Cotton(lint)
5,230,253
Rice
30,706,467
Soyabean
4,949,711
Arhar/Tur
3,325,278
Urad
Ragi
Total foodgrain
18,974,254
Gram
4,585,523
Sesamum
1,277,400
Potato
941,622
Dry
Barley
Coconut
1,064,229
Linseed
Sugarcane
3,482,848
Jute
979,700
Year vs Area
Year vs Area Forecasted
Map Overview
Crop Overview India
Production Share
Crop Share
Season wise Analysis Crop Prod...
Area Racing Bar Chart
Production Racing Bar Chart
Crop Treemap

3.5 Deployment:

We can deploy the Tableau report using Tableau Online and Server. But we need a Enterprise version to deploy it in the portal. There are different ways of tableau deployment like Tableau Server - On Premises, Tableau Server - Public Cloud (IaaS), Tableau Online (SaaS).

In this project we had used Heroku to deploy our webapp.

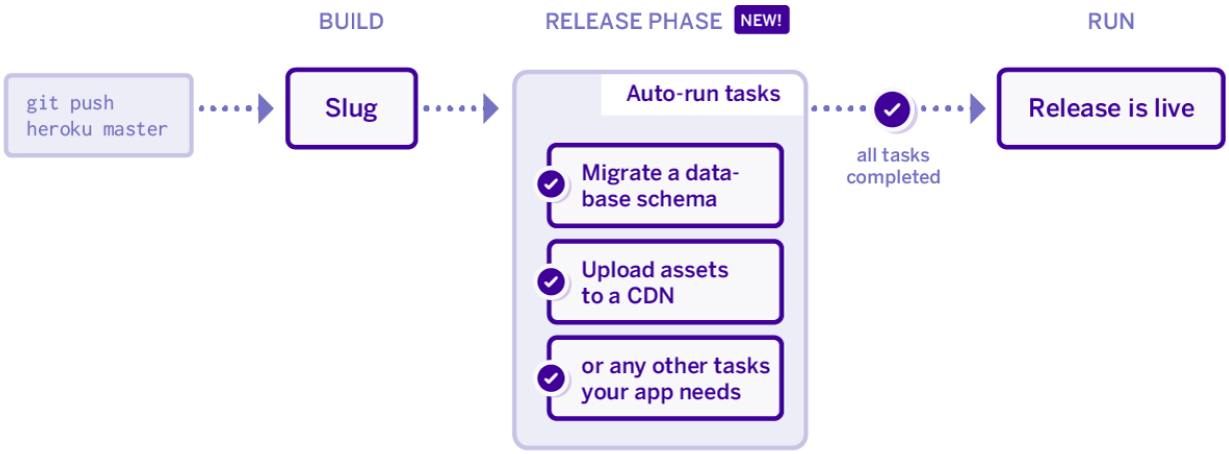
The deployment in the Heroku consists of 3 types like

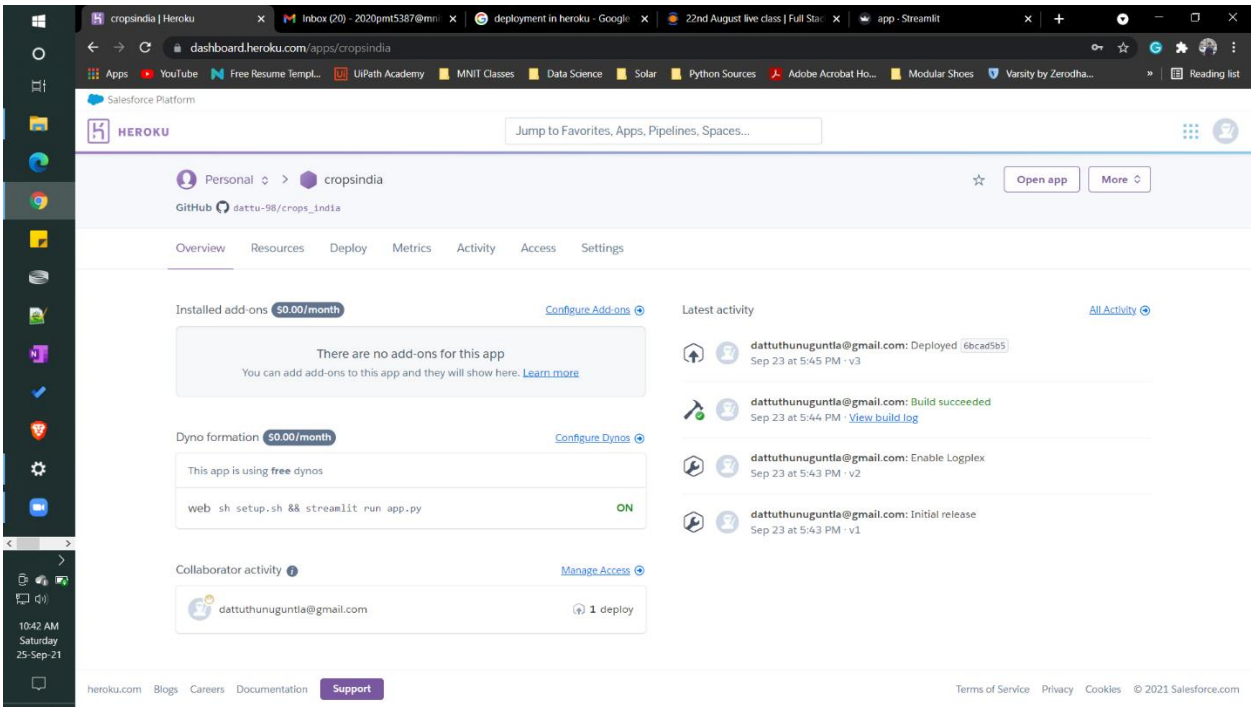
1. Heroku Git
2. Personal GitHub
3. Container Registry

In this project we had used Personal GitHub Deployment method. You can check the source files in the personal GitHub repository in the following link: https://github.com/dattu-98/crops_india

- The GitHub repository consists of Procfile, app.py, requirements.txt, runtime.txt, setup.sh

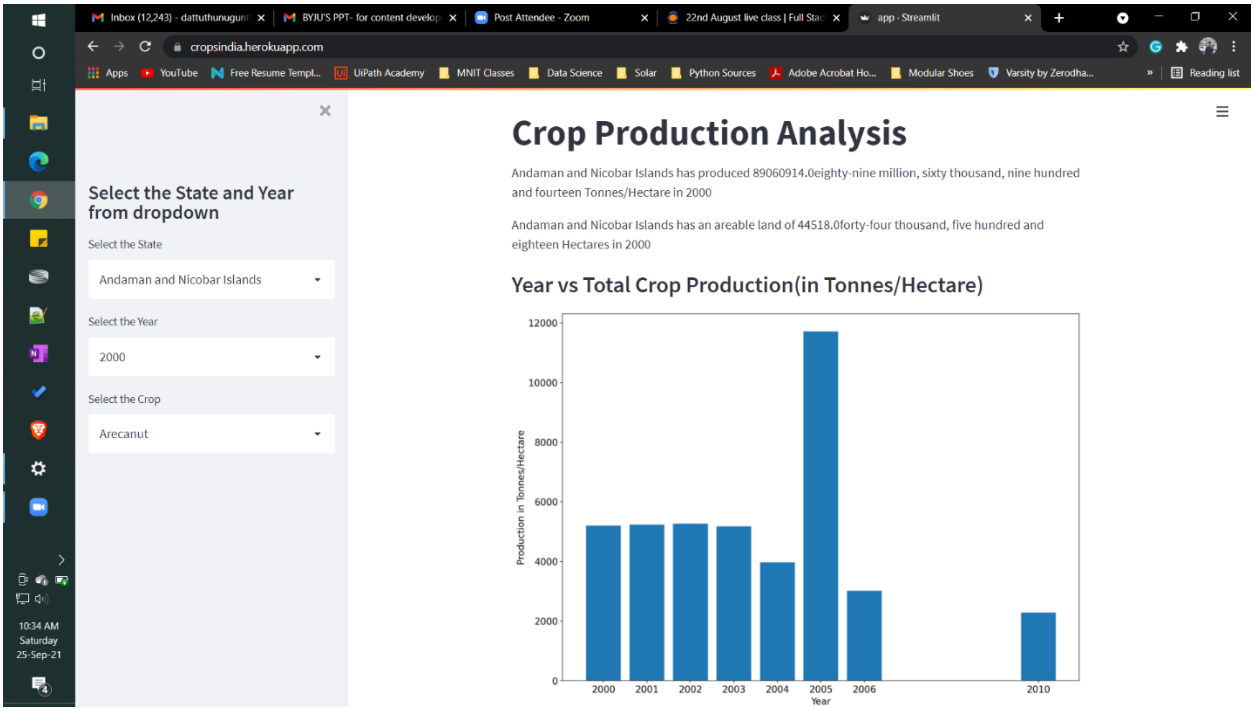
HEROKU DEPLOYMENT PHASES





After deployment in Heroku the Webapp looks in the following fashion:

Heroku Webapp Link: https://cropsindia.herokuapp.com/



4. Unit Test Cases:

S. No	Test Case Description	Expected Results
1	Select Andhra Pradesh as State, year as 2002, crop as Rice	The Stacked bar charts are plotted in the Streamlit Web App
2	Selecting the State Name and Crop in the Production Overview Worksheet in the Tableau report	Results will be appeared according to the State and Crop Selected
3	Playing the Area racing bar chart	The bar chart changes dynamically over years showing values of Arable Area
4	Selecting the State Name and year from the Map Overview Worksheet of the Tableau File	The overall Production appears in the forms of Symbol Maps with Production values
5	Selecting the State Name and year from the Crop Overview Worksheet of the Tableau File	The Highest Crop produced in the state appears in the State Label with its production
6	Selecting the year from Season wise analysis of Crop Production	Text Tables is appeared as soon as we enter the year according to the Season and Crop