Project Title : Agriculture Data Analytics in Crop Yield Estimation using IBM Cognos

Project on Agriculture Data Analytics in Crop Yield Estimation using IBM Cognos

Prof. Chandrika M, Department of MCA, Dayananda Sagar College of Engineering, Bangalore.

1. Introduction

1.1 Overview

India is a country in which huge population is dependant on agriculture for their livelyhood. Analysis of Agricultural data plays a vital role in understanding the food supply chain as well as improvising the agricultural practices. Early estimation of crop yield will also help farmers plan ahead for many related tasks such as storage requirements, transport as well as budgeting.

1.2 Purpose

The major objectives of the project includes:

- Spatial Analysis of crop production
- Seasonal Analysis of crop production
- Annual analysis of crop production
- Statewise crop production analysis

2. Literature Survey

2.1 Existing Problem

As data is the collection of agricultural activities from various states across India, Understanding and identifying patterens, trends and other insights regarding the data requires rigorous analysis and visualization.

2.2 Proposed Solution

To understand the data clearly, IBM Cognos Cloud tool is used to plot the data in various types of charts to get better insights regarding the crop yield.

3. Theoritical Analysis

3.1 Block Diagram

The content is taken from data.world.website
https://data.world/thatzprem/agriculture-india

3.2 Hardware/Software Designing

Data set Information:

```
<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
```

Here is some sample data from the dataset:

	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	NICOBA	2000	Kharif	Rice	102.0	321.0

	Nicobar Islands	RS					
3	Andaman and Nicobar Islands	NICOBA RS	2000	Whole Year	Banana	176.0	641.0
4	Andaman and Nicobar Islands	NICOBA RS	2000	Whole Year	Cashewnut	720.0	165.0

4. Experimental Investigations

Here are some of the statistical information of the data:

	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

The size of the data set is given by:

df.shape

(246091, 7)

Means, 246091 rows and 7 columns.

Analysing the dataset for the presence of null values:

df.isnull().any()

State_Name False
District_Name False
Crop_Year False
Season False
Crop False
Area False
Production True
dtype: bool

df.isnull().any().sum()

1

As the Production column contains one missing value, the value of that cell is replaced by the mean of the Production column:

df.isnull().any()

State_Name False
District_Name False
Crop_Year False
Season False
Crop False
Area False
Production True
dtype: bool

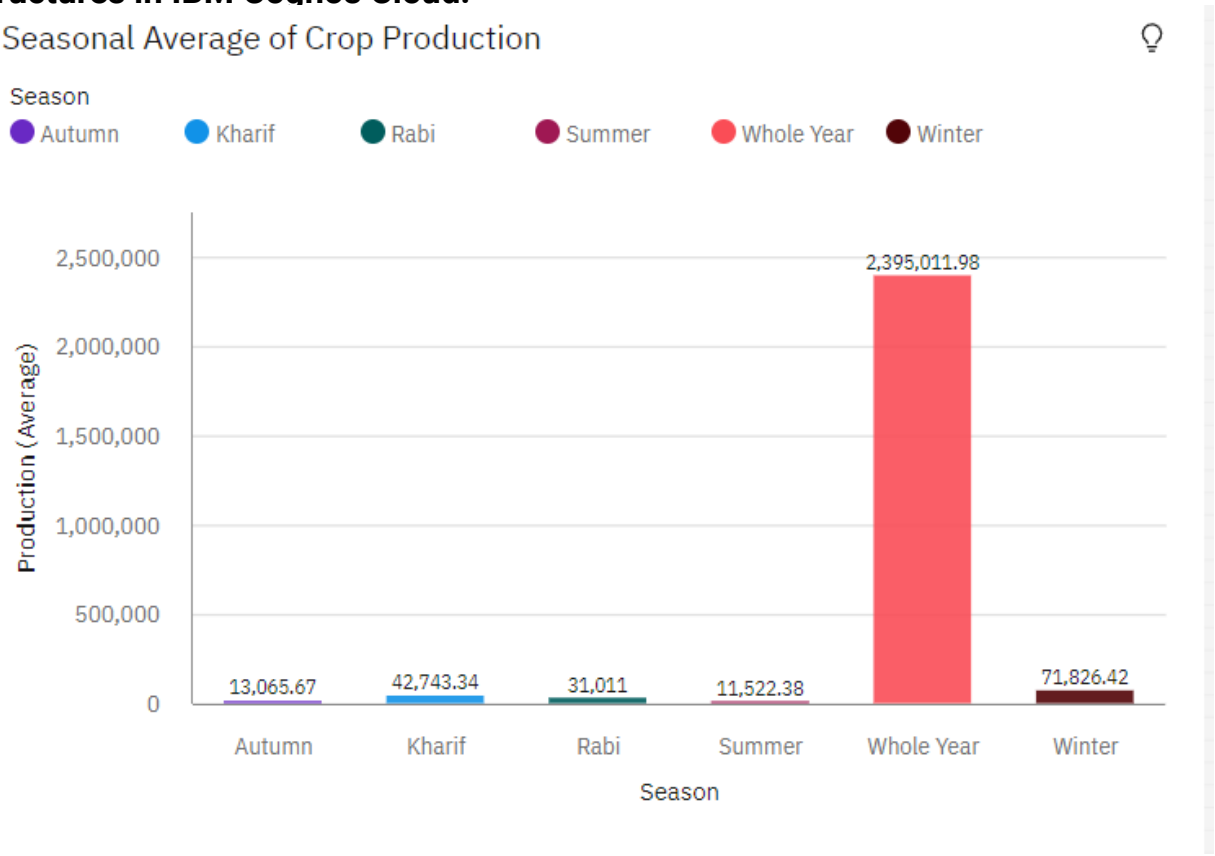
df.isnull().any().sum()

1

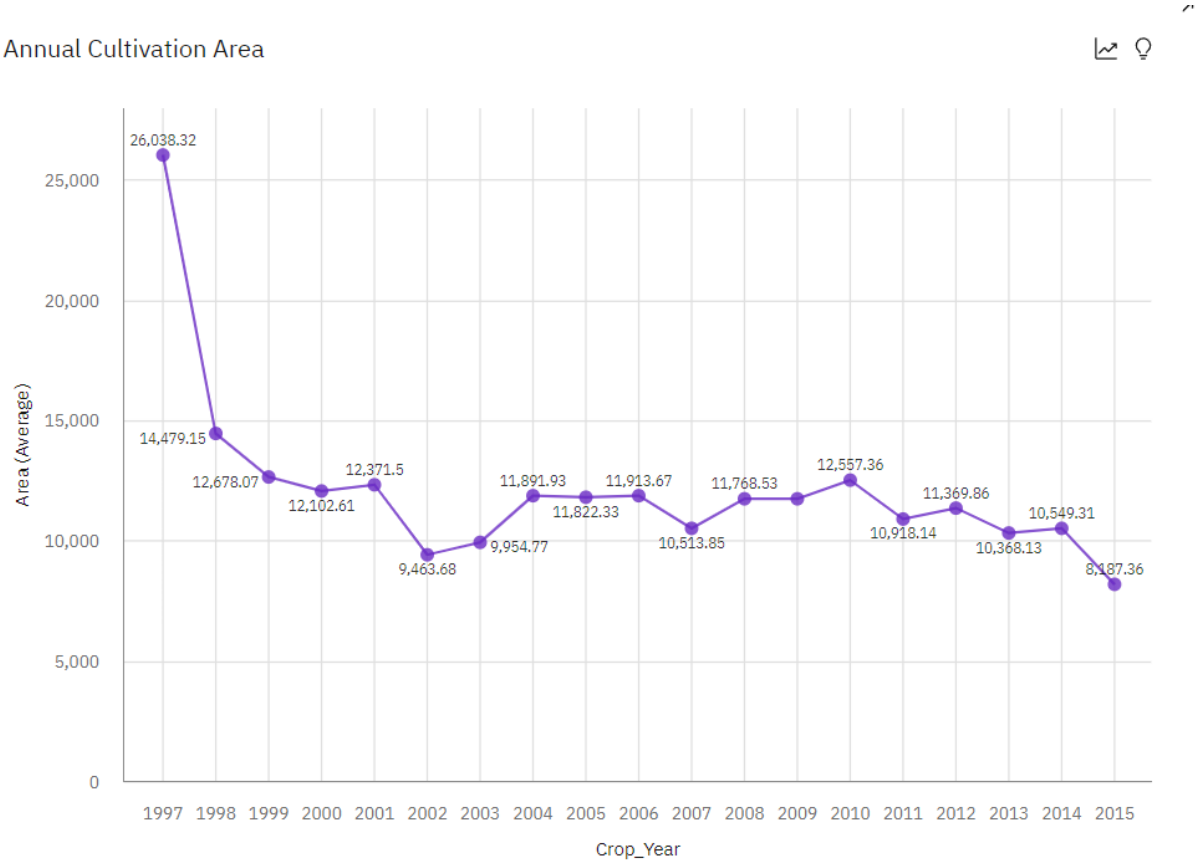
5. Flowchart

Project Title : Agriculture Data Analytics in Crop Yield Estimation using IBM Cognos

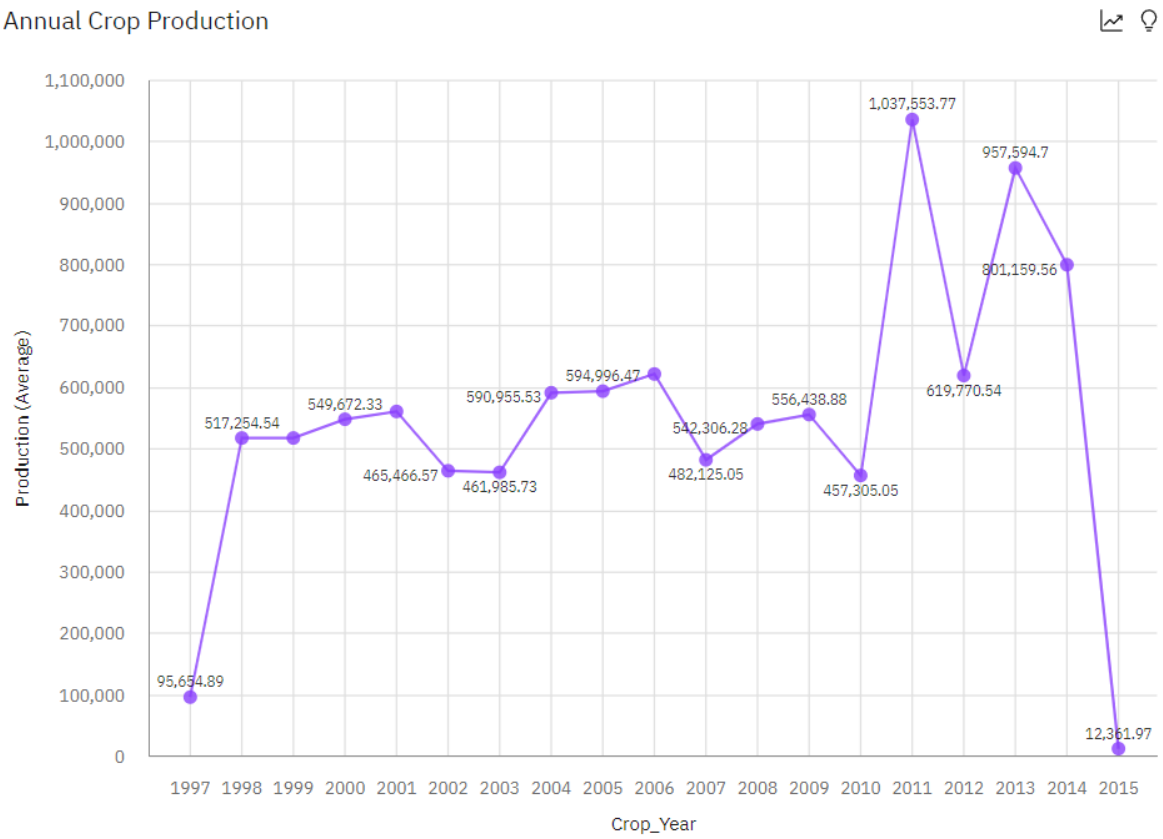
Here are some of the visualization done on the data to understand the structures in IBM Cognos Cloud:



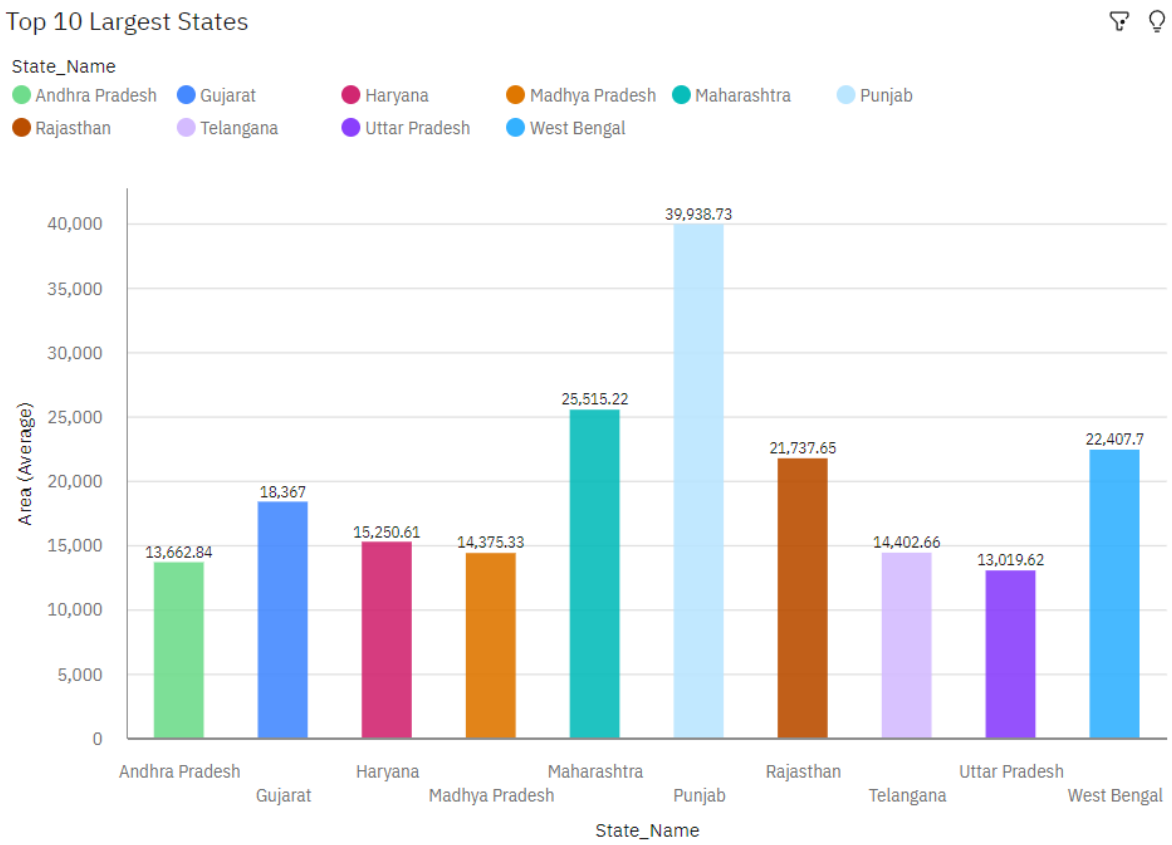
Project Title : Agriculture Data Analytics in Crop Yield Estimation using IBM Cognos



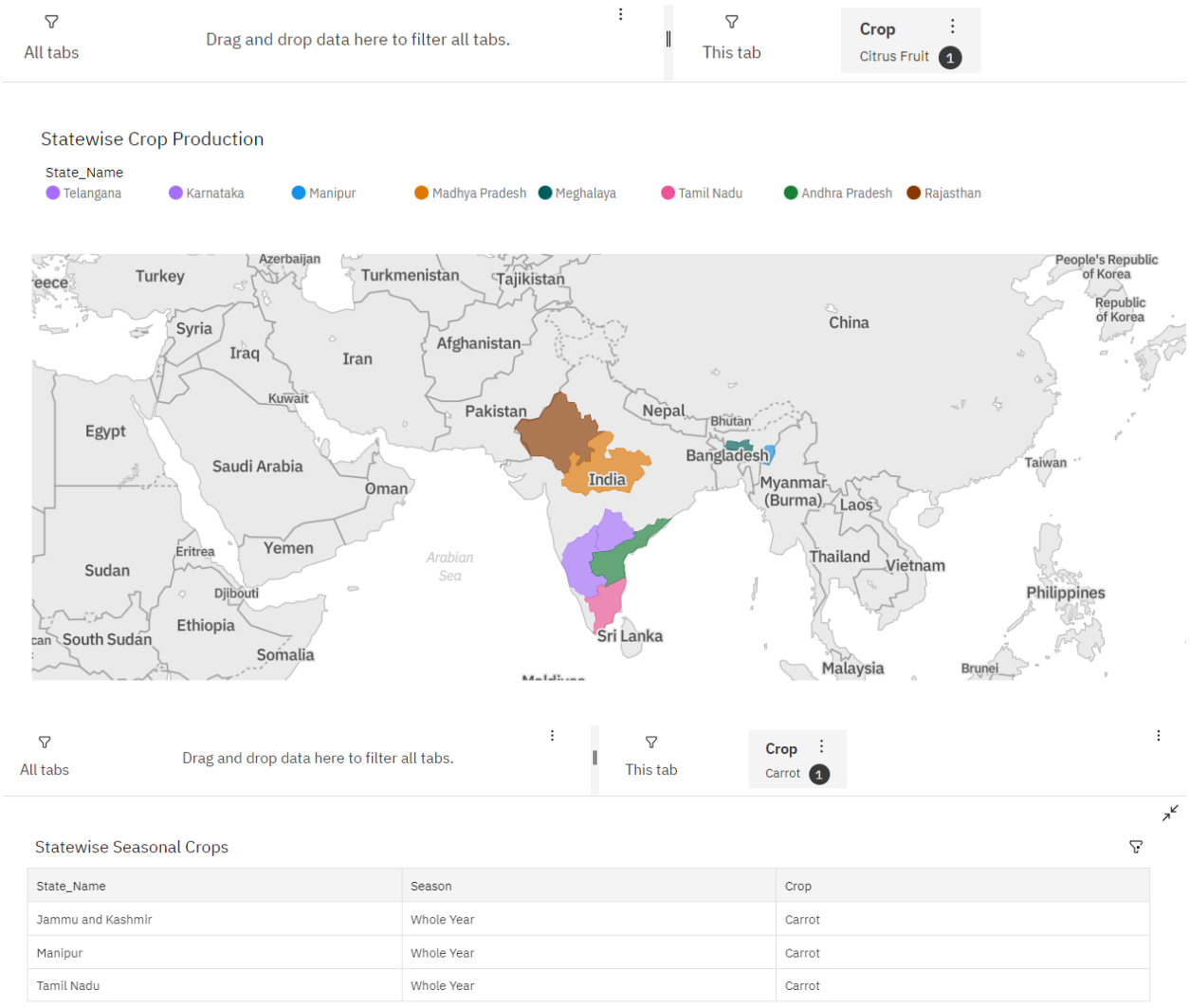
Project Title : Agriculture Data Analytics in Crop Yield Estimation using IBM Cognos



Project Title : Agriculture Data Analytics in Crop Yield Estimation using IBM Cognos



Project Title : Agriculture Data Analytics in Crop Yield Estimation using IBM Cognos



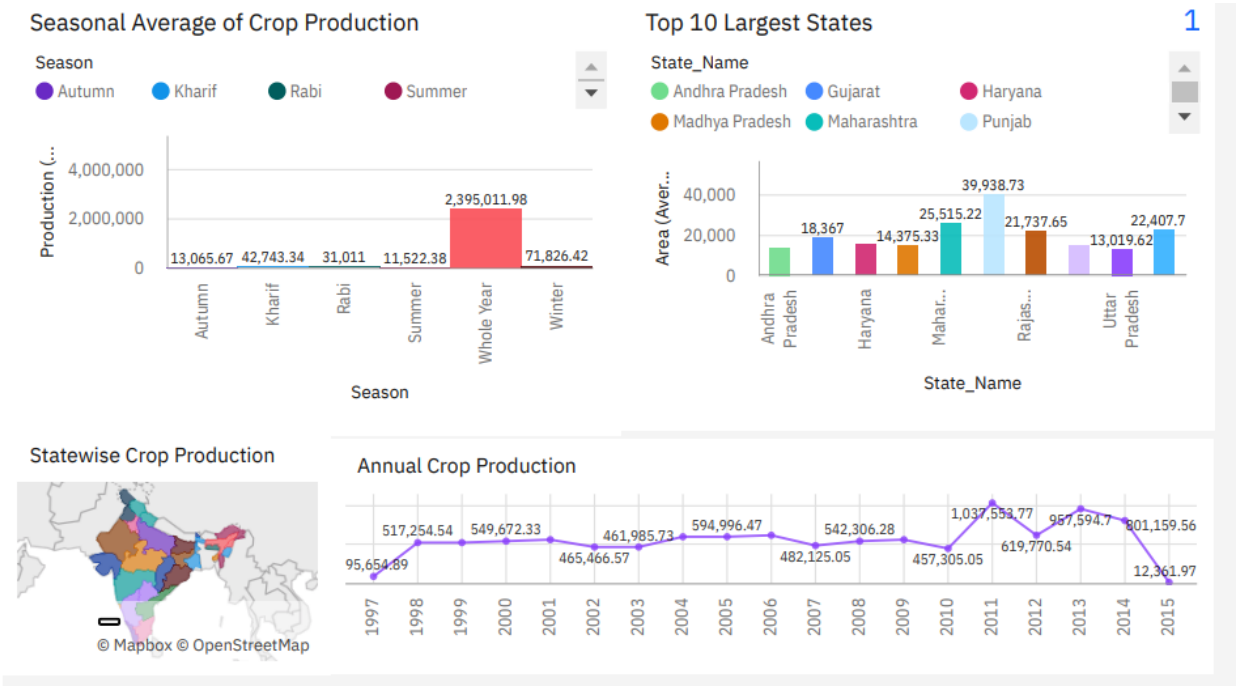
Statewise Seasonal Crops

State_Name	Season	Crop
Jammu and Kashmir	Whole Year	Carrot
Manipur	Whole Year	Carrot
Tamil Nadu	Whole Year	Carrot

6. Result

Final dashboard has been designed to get the insights from the data from various columns.

Project Title : Agriculture Data Analytics in Crop Yield Estimation using IBM Cognos



7. Advantages / Disadvantages

Advantages of the analysis of the agricultural data:

- **Understanding the origin and climate requirements of certain crops**
- **Understanding the topology of various crop cultivation.**
- **Understanding statewise contribution for agriculture both in terms of area as well as production of crops.**
- **Better understanding of seasonal crop production**

Disadvantages of the analysis of the agricultural data:

- **If the data is not genuine, the analysis might not have any meaning. Taking care to understand the authenticity of data source is very important.**

8. Applications

The current study can be carried forward to imporve various sectors including:

- **Agriculture**
- **various agriculture dependent sectors such as transportation, labour, fertilizer manufacturer, imports and exports can have major benefits if the production can be planned ahead.**

9. Conclusion

IBM Cognos is a BI tool which provides variety of functionalities for understanding the data. This tool helps to create reports, analyse the data and monitor the events and metrics for better business decisions.

10. Future Scope

The project can be continued for making prediction of the crops and by which the dependant domains can also plan their functionalities ahead of time.

11. Bibilography

[1]. https://agriculture.vic.gov.au/crops-and-horticulture/grains-pulses-and-cereals/crop-production/general-agronomy/a-brief-guide-to-estimating-crop-yields
[2]. https://www.ibm.com/docs/en/cognos-analytics/10.2.2?topic=SSEP7J_10.2.2/com.ibm.swg.ba.cognos.wig_cr.10.2.2.doc/c_gtstd_c8_bi.html
[3]. https://www.kaggle.com/datasets/abhinand05/crop-production-in-india?datasetId=412860

APPENDIX

A. Source Code

* The exported file of the project visualization is available in the github repository.