

## Project Development phase Sprint-1

### Analysis and Estimation(Working and Loading the dataset)

Date	27 October 2022
Team ID	PNT2022TMID15871
Project Name	Estimate the Crop Yield using Data Analytics

Working with the Dataset:

Understanding the dataset:

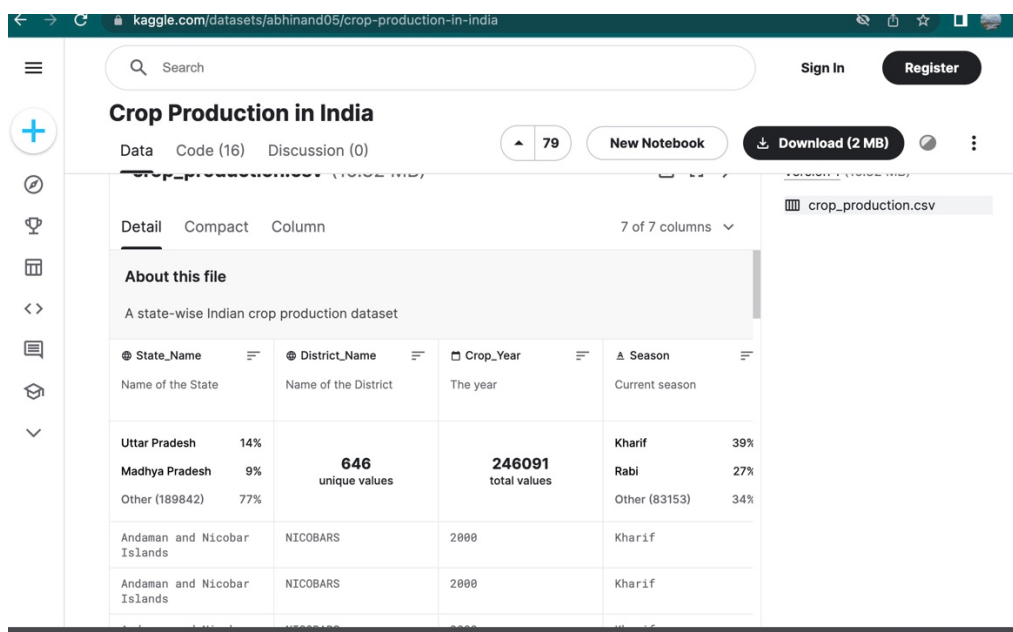
27 October 2022 PNT2022TMID40341

Project - ESTIMATION OF CROP YIELD USING DATA ANALYTICS

This project is based on a understanding the crop production of India .Download the dataset from the below link. It has 2,46,092 data points (rows) and 6 features (columns) describing each crop production related details.

Firstly, download the dataset. Data Given in csv file is understood clearly. In case of understanding the data, we have been working with and give a brief overview of what each feature represents or should represent.

1. State name – All the Indian state names
2. District Name -Different District names.
3. Crop Year- contains the crop years.
4. Season – Different seasons for crop production.
5. Area- Total number of areas covered.
6. Production- production of crops.



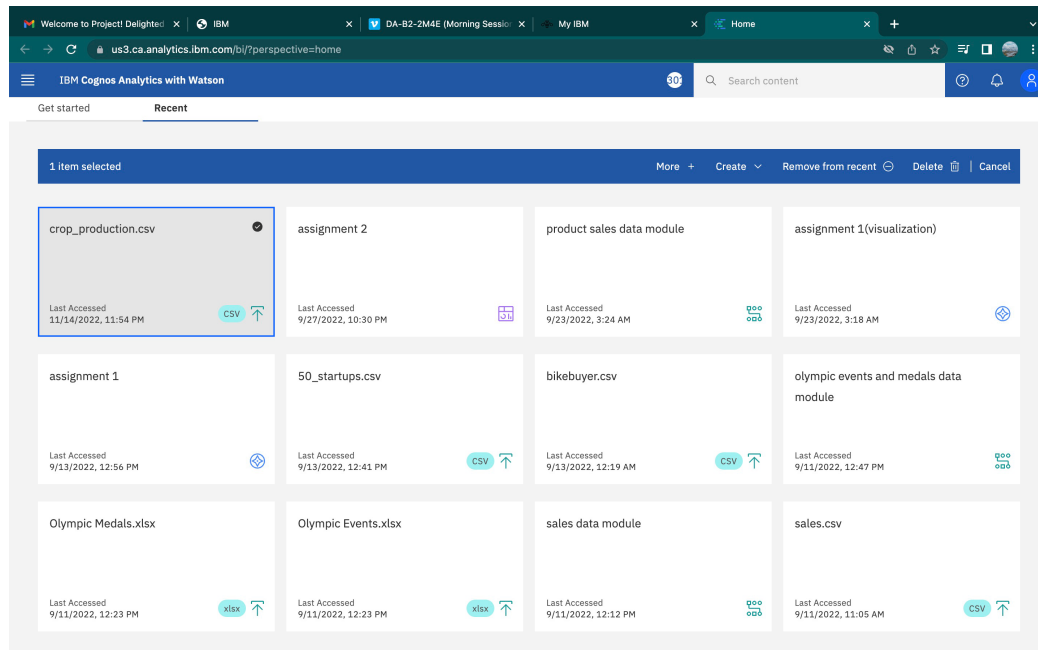
The screenshot shows the Kaggle dataset page for 'Crop Production in India'. The page includes a search bar, 'Sign In' and 'Register' buttons, and a 'Download (2 MB)' button. The dataset is described as 'A state-wise Indian crop production dataset'. The table below provides a summary of the data.

State_Name	District_Name	Crop_Year	Season
Uttar Pradesh	646 unique values	246091 total values	Kharif
Madhya Pradesh			Rabi
Other (189842)			Other (83153)
Andaman and Nicobar Islands	NICOBARS	2000	Kharif
Andaman and Nicobar Islands	NICOBARS	2000	Kharif

## Loading the Dataset:

Before building a view and analyze firstly, we have made a connect of the data to IBM Cognos. Cognos supports connecting to a wide variety of data, stored in a variety of places.

The data might be stored on your computer in a spreadsheet or a text file, or in a big data, relational, or cube (multidimensional) database on a server in your enterprise. In our case, we will be using a spreadsheet or text file for making our analysis.



The screenshot shows the IBM Cognos Analytics 'Grid' view for a data module. The left sidebar shows the 'Data module' structure with a search bar and a list of fields: Row Id, State\_Name, District\_Name, Crop\_Year, Season, Crop, Area, and Production. The main grid displays a table with 15 rows and 8 columns. The columns are: Row Id, State\_Name, District\_Name, Crop\_Year, Season, Crop, and Area. The data represents crop production for Andaman and Nicobar Islands from 2000 to 2001.

Row Id	State_Name	District_Name	Crop_Year	Season	Crop	Area
1	Andaman and Nicobar Islands	NICOBARS	2000	Kharif	Areca nut	125
2	Andaman and Nicobar Islands	NICOBARS	2000	Kharif	Other Kharif pulses	2
3	Andaman and Nicobar Islands	NICOBARS	2000	Kharif	Rice	100
4	Andaman and Nicobar Islands	NICOBARS	2000	Whole Year	Banana	175
5	Andaman and Nicobar Islands	NICOBARS	2000	Whole Year	Cashewnut	725
6	Andaman and Nicobar Islands	NICOBARS	2000	Whole Year	Coconut	185
7	Andaman and Nicobar Islands	NICOBARS	2000	Whole Year	Dry ginger	36
8	Andaman and Nicobar Islands	NICOBARS	2000	Whole Year	Sugarcane	1
9	Andaman and Nicobar Islands	NICOBARS	2000	Whole Year	Sweet potato	5
10	Andaman and Nicobar Islands	NICOBARS	2000	Whole Year	Tapioca	40
11	Andaman and Nicobar Islands	NICOBARS	2001	Kharif	Areca nut	125
12	Andaman and Nicobar Islands	NICOBARS	2001	Kharif	Other Kharif pulses	2
13	Andaman and Nicobar Islands	NICOBARS	2001	Kharif	Rice	83
14	Andaman and Nicobar Islands	NICOBARS	2001	Whole Year	Cashewnut	715
15	Andaman and Nicobar Islands	NICOBARS	2001	Whole Year	Coconut	185

IBM-Cognos-Analytics-with-Watson

Selected sources /

crop\_production.csv

Navigation paths

- crop\_production.csv
  - State\_Name
  - District\_Name
  - Crop\_Year
  - Season
  - Crop
  - Area
  - Production

Explore data relationships

crop\_production.csv

Reset to original

Search Crop

Edit diagram

Relationship diagram

10% 100%

Select single or multiple nodes to see visualizations.

Select a visualization

Explore visualizations related to 'Crop'

Crop

Crop
Apple
Arcanut (Processed)

Add

Area by Crop colored by State\_Name

Wheat Rice Maize Ground... Cotton... Soyabean Rapese... Jowar Gram Bajra

Add

Area by Crop\_Year colored by Crop

.997 1999 2001 2003 2005 2007 2009 2011 2013 2015

Add

IBM-Cognos-Analytics-with-Watson

Selected sources /

crop\_production.csv

Navigation paths

- crop\_production.csv
  - State\_Name
  - District\_Name
  - Crop\_Year
  - Season
  - Crop
  - Area
  - Production

Explore data relationships

crop\_production.csv

Reset to original

Search Crop\_Year

Edit diagram

Relationship diagram

10% 100%

Select single or multiple nodes to see visualizations.

Select a visualization

Explore visualizations related to 'Crop\_Year'

Crop\_Year

Crop_Year
1997
1998

Add

Area by Crop\_Year colored by State\_Name

.997 1999 2001 2003 2005 2007 2009 2011 2013 2015

Add

Area by Crop\_Year colored by District\_Name

.997 1999 2001 2003 2005 2007 2009 2011 2013 2015

Add