

## SPRINT – 1

### WORKING WITH THE DATASET(UNDERSTANDING THE DATASET)

Date	28 October 2022
Team ID	PNT2022TMID41618
Project Name	Estimate The Crop Yield Using Data Analytics

“Estimate The Crop Yield Using Data Analytics” project is based on 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.

**Dataset Link** : [Dataset](#)

Let's understand the data we're 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.

## UNDERSTAND THE DATASET:

The screenshot displays the Kaggle interface for the 'Crop Yield Prediction Dataset'. The dataset is categorized under 'Data' and has 52 versions. The summary on the right indicates it contains 5 files and 33 columns. The table below shows the first few rows of the dataset, with columns for 'Pesticides Use', 'Albania', 'Use', 'Pesticides (total)', and 'Year'.

	1 unique value	168 unique values	1 unique value	1 unique value	
					1990
Pesticides Use	Albania	Use	Pesticides (total)		1990
Pesticides Use	Albania	Use	Pesticides (total)		1991
Pesticides Use	Albania	Use	Pesticides (total)		1992
Pesticides Use	Albania	Use	Pesticides (total)		1993
Pesticides Use	Albania	Use	Pesticides (total)		1994
Pesticides Use	Albania	Use	Pesticides (total)		1995

**Summary**

- 5 files
- 33 columns

We use cookies on Kaggle to deliver our services, analyze web traffic, and improve your experience on the site. By using Kaggle, you agree to our use of cookies. [Got it](#) [Learn more](#)

11:28 AM 11/7/2022

- Download the dataset from the given dataset link.
- Data Given in csv file is understood.
- The data format is as shown in the below image:

State Name	District Name	Crop Year	Season	Crop	Area	Production
Andaman	NICOBARS	2000	Kharif	Areca nut	1254	2000
Andaman	NICOBARS	2000	Kharif	Other Kha	2	1
Andaman	NICOBARS	2000	Kharif	Rice	102	321
Andaman	NICOBARS	2000	Whole Ye	Banana	176	641
Andaman	NICOBARS	2000	Whole Ye	Cashewnu	720	165
Andaman	NICOBARS	2000	Whole Ye	Coconut	18168	65100000
Andaman	NICOBARS	2000	Whole Ye	Dry ginger	36	100
Andaman	NICOBARS	2000	Whole Ye	Sugarcane	1	2
Andaman	NICOBARS	2000	Whole Ye	Sweet pot	5	15
Andaman	NICOBARS	2000	Whole Ye	Tapioca	40	169
Andaman	NICOBARS	2001	Kharif	Areca nut	1254	2061
Andaman	NICOBARS	2001	Kharif	Other Kha	2	1
Andaman	NICOBARS	2001	Kharif	Rice	83	300
Andaman	NICOBARS	2001	Whole Ye	Cashewnu	719	192
Andaman	NICOBARS	2001	Whole Ye	Coconut	18190	64430000
Andaman	NICOBARS	2001	Whole Ye	Dry ginger	46	100
Andaman	NICOBARS	2001	Whole Ye	Sugarcane	1	1
Andaman	NICOBARS	2001	Whole Ye	Sweet pot	11	33
Andaman	NICOBARS	2002	Kharif	Rice	189.2	510.84
Andaman	NICOBARS	2002	Whole Ye	Areca nut	1258	2083
Andaman	NICOBARS	2002	Whole Ye	Banana	213	1278
Andaman	NICOBARS	2002	Whole Ye	Black pepi	63	13.5