

Sprint – 1

Understanding The Dataset

Date	16 November 2022
Team ID	PNT2022TMID35100
Project Name	Estimate the crop yield using Data Analytics

Understanding The Dataset:

This project is based on a understanding the crop production of India. 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.

The data format is as shown in the below image:

Crop Production in India

Data Code (15) Discussion (0) 77 New Notebook

crop_production.csv (15.32 MB)

Detail Compact Column 7 of 7 columns

About this file

A state-wise Indian crop production dataset

State_Name	District_Name	Crop_Year	Season	Crop	Area
Uttar Pradesh	14%	646 unique values	246091 total values	Kharif	39%
Madhya Pradesh	9%		Rabi	27%	6%
Other (189842)	77%		Other (83153)	34%	Other (217040) 88%
Andaman and Nicobar Islands	NICOBARS	2000	Kharif	Areca nut	125
Andaman and Nicobar Islands	NICOBARS	2000	Kharif	Other Kharif pulses	2.6
Andaman and Nicobar Islands	NICOBARS	2000	Kharif	Rice	182