

# SPRINT – 1

## WORKING WITH THE DATASET(UNDERSTANDING THE DATASET)

|              |  |
|--------------|--|
| Date         | 12 November 2022                             |
| Team ID      | PNT2022TMID03372                             |
| 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.

Dataset Link : <https://www.kaggle.com/datasets/abhinand05/crop-production-in-india>

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 shows the Kaggle dataset page for 'Crop Production in India' by user ABHINAND, updated 3 years ago. The page has a left sidebar with navigation links: Home, Competitions, Datasets, Code, Discussions, Learn, and More. The main content area features a search bar, a 'Sign In' button, and a 'Register' button. Below these, there's a 'New Notebook' button and a 'Download (2 MB)' button. The dataset title 'Crop Production in India' is prominently displayed, followed by the question 'Can you predict crop production in India?'. A small image of a child in a field is shown to the right. Below the title, there are tabs for 'Data', 'Code (16)', and 'Discussion (0)'. The 'About Dataset' section is visible, showing a 'Usability' score of 10.00 and a 'License' section. At the bottom, a cookie consent banner is present.

- 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:

The screenshot shows an Excel spreadsheet titled 'crop\_production.csv - Excel'. The spreadsheet has columns for State\_Name, District\_Name, Crop\_Year, Season, Crop, Area, and Production. The data is organized into rows, with the first row being the header. The data includes information for various states and districts, such as Andaman and Nicobar Islands, NIOBARS, and various crops like Arecanut, Other Kharif pulses, Rice, Banana, Cashewnut, Coconut, Dry ginger, Sugarcane, Sweet potato, and Tapioca. The production values are listed in the 'Production' column.

| State_Name                  | District_Name | Crop_Year | Season     | Crop                | Area  | Production |
|-----------------------------|---------------|-----------|------------|---------------------|-------|------------|
| Andaman and Nicobar Islands | NIOBARS       | 2000      | Kharif     | Arecanut            | 1254  | 2000       |
| Andaman and Nicobar Islands | NIOBARS       | 2000      | Kharif     | Other Kharif pulses | 2     | 1          |
| Andaman and Nicobar Islands | NIOBARS       | 2000      | Kharif     | Rice                | 102   | 321        |
| Andaman and Nicobar Islands | NIOBARS       | 2000      | Whole Year | Banana              | 176   | 641        |
| Andaman and Nicobar Islands | NIOBARS       | 2000      | Whole Year | Cashewnut           | 720   | 165        |
| Andaman and Nicobar Islands | NIOBARS       | 2000      | Whole Year | Coconut             | 18168 | 65100000   |
| Andaman and Nicobar Islands | NIOBARS       | 2000      | Whole Year | Dry ginger          | 36    | 100        |
| Andaman and Nicobar Islands | NIOBARS       | 2000      | Whole Year | Sugarcane           | 1     | 2          |
| Andaman and Nicobar Islands | NIOBARS       | 2000      | Whole Year | Sweet potato        | 5     | 15         |
| Andaman and Nicobar Islands | NIOBARS       | 2000      | Whole Year | Tapioca             | 40    | 169        |
| Andaman and Nicobar Islands | NIOBARS       | 2001      | Kharif     | Arecanut            | 1254  | 2061       |
| Andaman and Nicobar Islands | NIOBARS       | 2001      | Kharif     | Other Kharif pulses | 2     | 1          |
| Andaman and Nicobar Islands | NIOBARS       | 2001      | Kharif     | Rice                | 83    | 300        |
| Andaman and Nicobar Islands | NIOBARS       | 2001      | Whole Year | Cashewnut           | 719   | 192        |
| Andaman and Nicobar Islands | NIOBARS       | 2001      | Whole Year | Coconut             | 18190 | 64430000   |
| Andaman and Nicobar Islands | NIOBARS       | 2001      | Whole Year | Dry ginger          | 46    | 100        |
| Andaman and Nicobar Islands | NIOBARS       | 2001      | Whole Year | Sugarcane           | 1     | 1          |
| Andaman and Nicobar Islands | NIOBARS       | 2001      | Whole Year | Sweet potato        | 11    | 33         |
| Andaman and Nicobar Islands | NIOBARS       | 2002      | Kharif     | Rice                | 189.2 | 510.84     |
| Andaman and Nicobar Islands | NIOBARS       | 2002      | Whole Year | Arecanut            | 1258  | 2083       |
| Andaman and Nicobar Islands | NIOBARS       | 2002      | Whole Year | Banana              | 213   | 1278       |
| Andaman and Nicobar Islands | NIOBARS       | 2002      | Whole Year | Black pepper        | 63    | 13.5       |
| Andaman and Nicobar Islands | NIOBARS       | 2002      | Whole Year | Cashewnut           | 719   | 208        |
| Andaman and Nicobar Islands | NIOBARS       | 2002      | Whole Year | Coconut             | 18240 | 67490000   |