

Working with The Dataset

Understanding The Dataset

Team ID	PNT2022TMID15512
Project Name	Estimate The Crop Yield Using Data Analytics

This project is based on a understanding the crop production of India. This dataset has 2,46,092 data points (rows) and 6 features (columns) describing each crop production related details. A state-wise Indian crop production 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.

- Uttar Pradesh– 14 %
- Madhya Pradesh – 9 %
- Other (189842) – 77 %

2. District Name -Different District names.

- 646 unique values
- Most Common - BIJAPUR

3. Crop Year- contains the crop years. (246091 total values)

4. Season – Different seasons for crop production.

- Kharif – 39 %
- Rabi – 27 %
- Others (83153) – 34 %

5. Area- Total number of areas covered.

- Valid – 100 %
- Mismatched – 0
- Missing - 0

6. Production- production of crops.

- Valid – 98 %
- Mismatched – 0
- Missing - 2%