## **Understanding the Dataset**

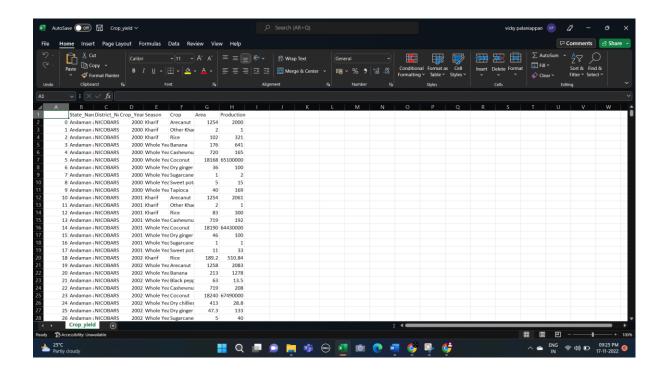
Date	17 November 2022		
Team ID	PNT2022TMID32817		
Project Name	Estimate the crop yield using data analytics		

This project is based on a understanding the crop production of India. The dataset from the below link contains about 2,46,092 data points (rows) and 6 features (columns) describing about the Crop yield in India from the years 1997 to 2015.

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.



The correlation between the dataset is given by,

	Crop_Year	Area	Production
Crop_Year	1.000000	-0.027791	0.006668
Area	-0.027791	1.000000	0.040487
Production	0.006668	0.040487	1.000000