



**Naan mudhalvan**

**Gojan school of business and technology – 1105**

**Course Name: Data analysis in process industries**

**Project title : Agri data explorer - Understanding Indian  
agriculture with EDA**

**Submitted by,**

**Pavithra D – 110523237018**

**Praveena C – 110523237019**

**Radha Krishnan K – 110523237020**

**Rajasri K – 110523237021**

**Revathi E – 110523237022**

# Agri data explorer - Understanding Indian agriculture with EDA

## Abstract:

The agricultural sector in India plays a pivotal role in sustaining the nation's economy and ensuring food security. Despite its importance, Indian agriculture faces challenges including low productivity, unpredictable weather, market instability, and resource mismanagement. This project, *Agridata Explorer*, utilizes **Exploratory Data Analysis (EDA)** techniques to unravel meaningful insights from agricultural datasets. By visualizing and analyzing patterns in crop production, rainfall distribution, land usage, and yield trends, the study aims to assist policymakers, farmers, and researchers in making informed decisions. The outcome highlights hidden correlations and region-specific insights that can support sustainable agricultural development.

**Domain:** Agriculture

## Tools & Skills:

Python, SQL, Power BI, Data Cleaning, Exploratory Data Analysis (EDA), Plotly

## Problem Statement:

India's agricultural landscape is vast and varied, yet often lacks data-driven decision-making. The problem lies in **unorganized data usage**, underutilized historical records, and an absence of visual exploration to detect trends and outliers. Farmers and authorities struggle with inefficiencies due to a lack of actionable insights derived from existing data.

## INTRODUCTION:

Agriculture in India employs over 50% of the workforce and contributes significantly to the GDP. However, it remains vulnerable to climate variability, soil degradation, and market fluctuations. In the era of data science, leveraging structured datasets can transform raw numbers into strategic insights. This project introduces *Agridata Explorer*, an initiative to analyze comprehensive agricultural datasets using EDA to expose meaningful trends, inconsistencies, and relationships. The project emphasizes the potential of data visualization and statistical summaries in redefining how agriculture is understood and improved in India.

## MOTIVATION:

The motivation behind this project stems from the realization that **huge volumes of agricultural data remain underexplored**. While India has centuries-old farming knowledge, today's digital transformation offers tools to enhance that wisdom with precision insights. By converting agricultural data into visual narratives, we can empower decision-makers at every level — from rural farmers to national planners — to act with evidence, not estimates.

# Agri data explorer - Understanding Indian agriculture with EDA

## GOAL :

The primary goal is to explore and interpret agricultural data using EDA techniques to identify:

- Patterns in crop yields and rainfall.
- Regional agricultural strengths and weaknesses.
- Historical trends in production and productivity.
- Anomalies that could indicate data errors or unique circumstances.

The ultimate aim is to develop a visual and statistical dashboard of agricultural health across India.

## METHODOLOGY:

**The methodology follows a structured EDA pipeline:**

1. **Data Collection:** Sourced from government repositories such as the Ministry of Agriculture, ICAR, and weather databases.
2. **Data Cleaning:** Handling missing values, standardizing units, and correcting inconsistencies.
3. **Data Transformation:** Normalizing variables, aggregating state-wise and year-wise metrics.
4. **Visualization:**
  - Line plots for trends over time.
  - Heatmaps for regional analysis.
  - Bar charts for comparing crop outputs.
5. **Correlation Analysis:** Studying the relationship between rainfall and yield, fertilizer use and productivity, etc.
6. **Interpretation:** Summarizing findings and translating them into recommendations.

## ADVANTAGES

- **Insightful Decision Making:** Provides stakeholders with visual insights to back strategies.
- **Data-Driven Planning:** Encourages evidence-based agricultural policies.
- **Identifying Anomalies:** Detects irregularities that could suggest climatic or human impacts.
- **Open-Source Analysis:** Builds a reusable framework for future agricultural datasets.

## DISADVANTAGES

- **Data Quality Issues:** Public datasets may contain inconsistencies or outdated entries.
- **Lack of Real-Time Data:** EDA is more retrospective; it doesn't predict future outcomes without predictive models.
- **Regional Gaps:** Some areas may be underrepresented in the data.
- **Interpretation Dependency:** Visualizations can be misinterpreted without domain expertise.

## POWER BI DASHBOARD FEATURES:

### 1. Data Model Setup:

- Connected multiple related datasets (crop, region, yield, production).
- Established relationships using state/district codes.

### 2. Visualizations:

- **Geo Heatmaps:** Regional yield & production insights.
- **Time-Series Trends:** Production/yield over decades.
- **Bar/Column Charts:** State-wise/year-wise crop stats.
- **KPI Cards:** Total area, yield, and production indicators.
- **Scatter Plots:** Area vs. production correlation.

### 3. Interactivity:

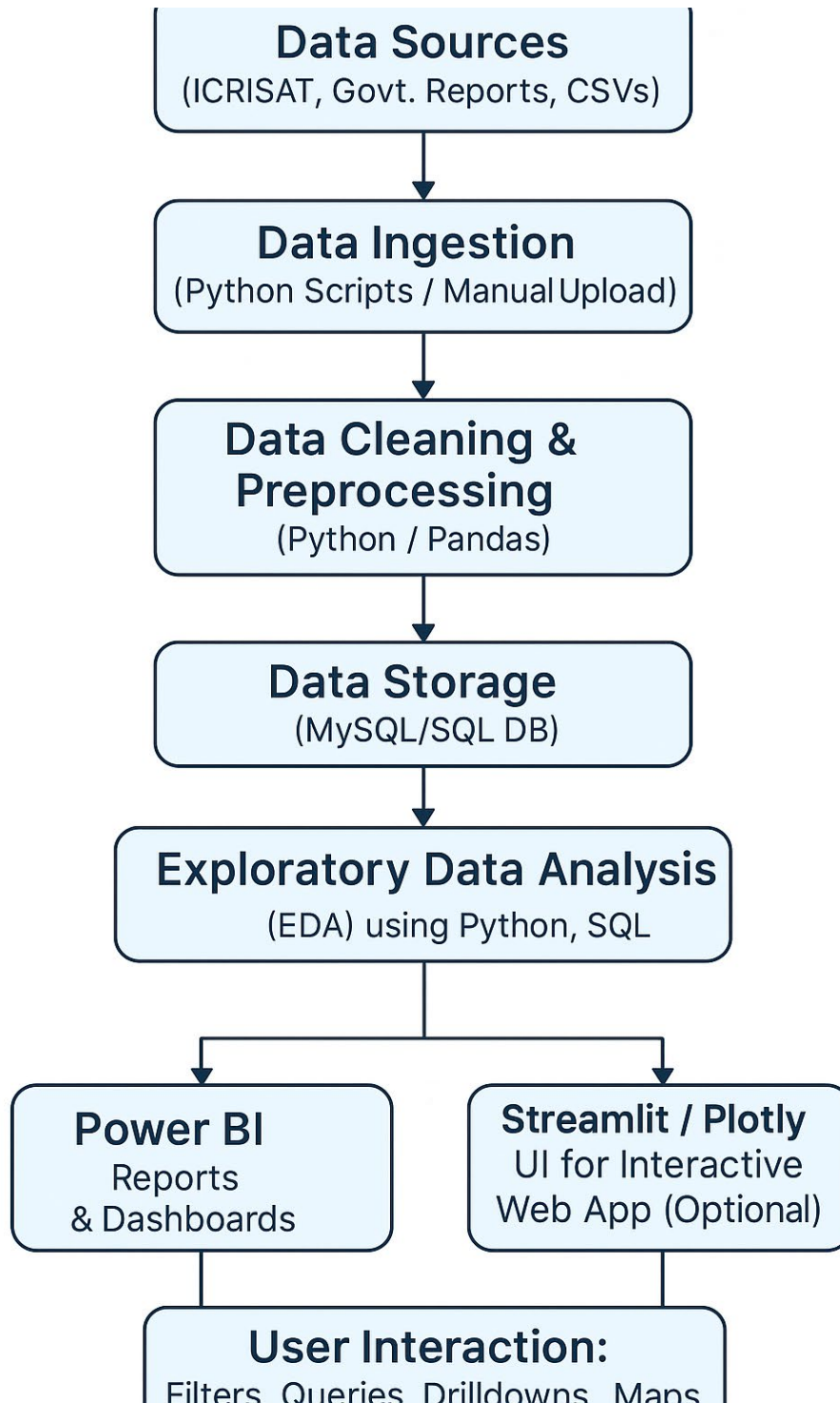
- Slicers for Crop, State, District, and Year.
- Dynamic filters, drill-downs, and tooltips for user experience.

## Evaluation Metrics:

Metric	Description
Accuracy	Verified correctness of visualizations and SQL outputs
Performance	Fast data loading and responsive filters in Power BI
User Engagement	Tracked interactions with slicers, charts, and custom filters
Data Completeness	Covered crops like Rice, Wheat, Maize, Oilseeds, Millets, etc.
User Feedback	Collected qualitative feedback from mock users for improvements

# Agri data explorer - Understanding Indian agriculture with EDA

## System Architecture:



# Agri data explorer - Understanding Indian agriculture with EDA

## RESULTS & DELIVERABLES:

- **Interactive Dashboard:** Easy navigation with crop, region, and year filters.
- **Insight Generation:** Revealed high/low performing crops and districts.
- **Actionable Recommendations:** Provided data-driven suggestions for crop planning and productivity improvements.

## CONCLUSION:

*Agri data Explorer* exemplifies the power of exploratory data analysis in decoding complex, multi-dimensional agricultural datasets. Through systematic visualization and analysis, the project offers a panoramic view of India's agricultural health. While there are limitations in terms of data coverage and timeliness, the approach lays a foundation for more granular, predictive, and prescriptive analytics in agriculture. By enabling smarter decisions rooted in data, *Agri data Explorer* contributes to the broader goal of sustainable and efficient farming in India.