**DATA SCIENCE MINOR PROJECT REPORT**Fertilizer Consumption Analysis in India  
**Prepared By:** Gaurav Khandelwal  
**Registration No.:** 12319392  
**Section:** KM007  
**Roll No.:** 59  
**[Department :** Computer Science and Engineering]  
**[Institute:** Lovely Professional University]

# Declaration

I, Gaurav Khandelwal, solemnly declare that the project titled 'Fertilizer Consumption Analysis in India' is an authentic and original piece of work conducted by me. This study has not been submitted to any other institution for the award of any degree, diploma, or certification. All data analyses, visualizations, and interpretations presented in this report are based on legitimate datasets sourced from credible repositories and processed using industry-standard tools and methodologies.

# Certificate

This is to certify that the project report titled 'Fertilizer Consumption Analysis in India', submitted by Gaurav Khandelwal, represents a genuine and diligent effort carried out under my direct supervision and guidance. The work reflects a meticulous approach to data analysis and visualization, showcasing the student’s ability to handle complex agricultural datasets with precision.

**Supervisor:** Dr. Mrinalini Rana  
**Date:** 12 April 2025

# Acknowledgement

I extend my heartfelt gratitude to the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT) for granting access to their comprehensive and high-quality agricultural datasets, which formed the backbone of this analysis. I am equally thankful to the open-source community for developing robust Python libraries such as Pandas, NumPy, Matplotlib, and Seaborn, which empowered me to perform sophisticated data processing and visualization tasks with ease. Special appreciation goes to my supervisor, Dr. Mrinalini Rana, whose insightful guidance and constructive feedback steered this project toward clarity and excellence. I am also grateful to my peers for their collaborative spirit and brainstorming sessions, which enriched my perspective. Lastly, I owe immense thanks to my family for their unwavering emotional support and encouragement throughout the journey of this project.

# Table of Contents

1. Introduction  
2. Source of Dataset  
3. Exploratory Data Analysis (EDA) Process  
4. Analysis of Dataset  
 i. Introduction  
 ii. General Description  
 iii. Specific Requirements, Functions, and Formulas  
 iv. Analysis Results  
 v. Visualization  
5. Conclusion  
6. Future Scope  
7. References

# List of Figures and Tables

- Figure 1: Line Plot - Year-wise trend of total fertilizer consumption across India  
- Figure 2: Bar Chart - Top 15 states by average fertilizer usage  
- Figure 3: Pie Chart - Proportional composition of Nitrogen, Phosphorus, and Potassium  
- Figure 4: Heatmap - State-wise fertilizer consumption across years  
- Figure 5: Line Plots - Fertilizer usage per hectare for NCA and GCA  
- Figure 6: Correlation Heatmap - Relationships between consumption metrics  
- Figure 7: Stacked Area Chart - Year-wise contribution of NPK to total consumption

# 1. Introduction

India, often referred to as the 'breadbasket of South Asia,' is an agrarian powerhouse where agriculture employs nearly half of the workforce and contributes significantly to the national economy. Central to this agricultural ecosystem is the use of fertilizers, which replenish essential soil nutrients—Nitrogen (N), Phosphorus (P), and Potassium (K)—to boost crop yields and ensure food security for a population exceeding 1.4 billion. However, the increasing dependence on fertilizers raises concerns about nutrient imbalances, environmental degradation, and long-term soil health.  
  
This project, titled 'Fertilizer Consumption Analysis in India,' seeks to dissect the patterns and trends of fertilizer use across Indian districts using a rich dataset provided by the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT). By employing advanced data analysis techniques and visualization tools in Python, this study uncovers temporal trends, regional disparities, and nutrient-specific consumption patterns. The insights generated are intended to inform policymakers, agricultural scientists, and farmers about optimizing fertilizer use for sustainable agriculture, balancing productivity with environmental stewardship.  
  
Key objectives of this analysis include:  
- Mapping the evolution of fertilizer consumption over time.  
- Identifying regional variations in fertilizer application across states and districts.  
- Evaluating the balance of Nitrogen, Phosphorus, and Potassium (NPK) usage to highlight potential nutrient imbalances.  
- Assessing the efficiency of fertilizer application relative to cropped areas.

# 2. Source of Dataset

The foundation of this analysis is a robust dataset sourced from the ICRISAT Data Repository (accessible at: [http://data.icrisat.org/dld/src/crops.html](http://data.icrisat.org/dld/src/crops.html)). This dataset is a treasure trove of district-level agricultural statistics, meticulously compiled to support research on semi-arid crop systems and agricultural development. Specifically, the dataset includes:  
- District-level fertilizer consumption: Detailed records of fertilizer usage across various Indian districts, capturing both total and nutrient-specific (Nitrogen, Phosphate, Potash) consumption.  
- Temporal coverage: Year-wise data spanning multiple years, enabling the study of long-term trends.  
- Area-based metrics: Fertilizer application rates per hectare for Net Cropped Area (NCA) and Gross Cropped Area (GCA).  
- Geographical identifiers: State and district codes for precise spatial analysis.  
  
The dataset’s granularity and authenticity make it an ideal resource for uncovering macro- and micro-level insights into fertilizer consumption patterns across India’s diverse agro-climatic zones.

# 3. Exploratory Data Analysis (EDA) Process

Exploratory Data Analysis (EDA) served as the cornerstone of this project, enabling a deep dive into the dataset’s structure, quality, and hidden patterns. The EDA process was methodical and iterative, designed to ensure data integrity and extract meaningful insights. The key steps included:  
- Data Cleaning and Preprocessing:  
 - Renamed ambiguous or inconsistent column headers to improve readability and usability (e.g., changing 'N\_Cons' to 'Nitrogen\_Consumption').  
 - Identified and addressed missing values using techniques like interpolation for time-series data or exclusion for non-critical records, ensuring minimal data loss.  
 - Detected and mitigated outliers by capping extreme values based on domain knowledge (e.g., implausible fertilizer consumption figures).  
- Data Aggregation:  
 - Grouped data by year, state, and district to compute national and regional summaries, such as total fertilizer consumption or average nutrient use per hectare.  
 - Created derived features, such as the NPK ratio, to analyze nutrient balance.  
- Correlation Analysis:  
 - Constructed a correlation matrix to explore relationships between variables, such as Nitrogen consumption and crop area, to uncover dependencies or redundancies.  
- Visualization Preparation:  
 - Prepared data for visualizations by structuring it into formats suitable for line plots, bar charts, heatmaps, and other graphical representations.  
  
The EDA process was powered by Python libraries like Pandas for data manipulation, NumPy for numerical computations, and Seaborn and Matplotlib for preliminary visualizations. This rigorous approach ensured that subsequent analyses were built on a clean, reliable, and well-understood dataset.

# 4. Analysis of Dataset

## i. Introduction

The core analysis delves into the multifaceted nature of fertilizer consumption in India, exploring its temporal, spatial, and nutrient-specific dimensions. The study addresses critical questions such as:  
- How has fertilizer usage evolved over the years?  
- Which states lead in fertilizer consumption, and why?  
- Are Nitrogen, Phosphorus, and Potassium applied in balanced proportions?  
- How efficiently are fertilizers used relative to cropped areas?  
- What relationships exist between different consumption metrics?  
  
By answering these questions, the analysis provides actionable insights for optimizing fertilizer use and promoting sustainable agriculture.

## ii. General Description

The analysis was conducted using a suite of Python libraries tailored for data science:  
- Pandas: For data wrangling, aggregation, and filtering.  
- NumPy: For efficient numerical computations, such as calculating means and standard deviations.  
- Seaborn and Matplotlib: For creating visually appealing and informative plots to communicate findings effectively.  
  
The dataset spans multiple years and includes both absolute consumption figures (e.g., total Nitrogen used in a district) and normalized metrics (e.g., Nitrogen per hectare of GCA). This dual perspective enables a comprehensive understanding of both the scale and intensity of fertilizer use across India’s diverse agricultural landscape.

## iii. Specific Requirements, Functions, and Formulas

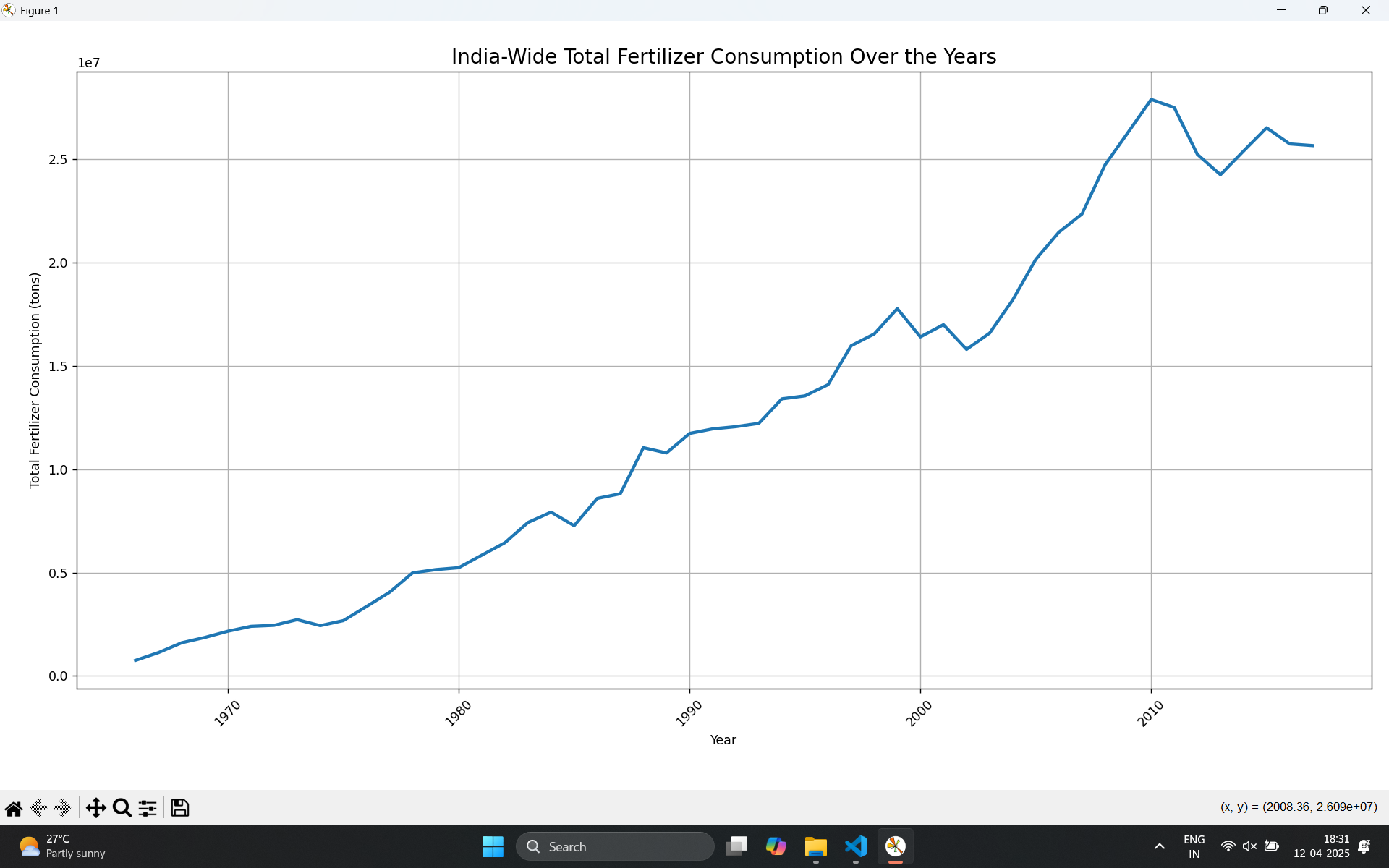
The analysis relied on a combination of built-in Python functions and custom computations to process and interpret the data. Key techniques included:  
- Data Aggregation:  
 - Used groupby() to summarize data by year, state, or district (e.g., df.groupby('Year')['Total\_Consumption'].sum()).  
 - Applied pivot() to restructure data for heatmaps, enabling state-year comparisons.  
- Statistical Computations:  
 - Calculated descriptive statistics using mean(), sum(), and std() to quantify average and variability in fertilizer use.  
 - Computed correlations with corr() to identify relationships between variables like Nitrogen consumption and GCA.  
- Visualization Functions:  
 - Employed plot() for basic line and scatter plots.  
 - Used lineplot(), barplot(), pie(), heatmap(), and stackplot() from Seaborn and Matplotlib for advanced visualizations.  
- Custom Formulas:  
 - Derived the NPK ratio as Nitrogen\_Consumption / (Phosphate\_Consumption + Potash\_Consumption) to assess nutrient balance.  
 - Calculated per-hectare consumption as Total\_Consumption / NCA or Total\_Consumption / GCA to evaluate application efficiency.  
  
These tools and formulas enabled a robust and reproducible analysis pipeline.

## iv. Analysis Results

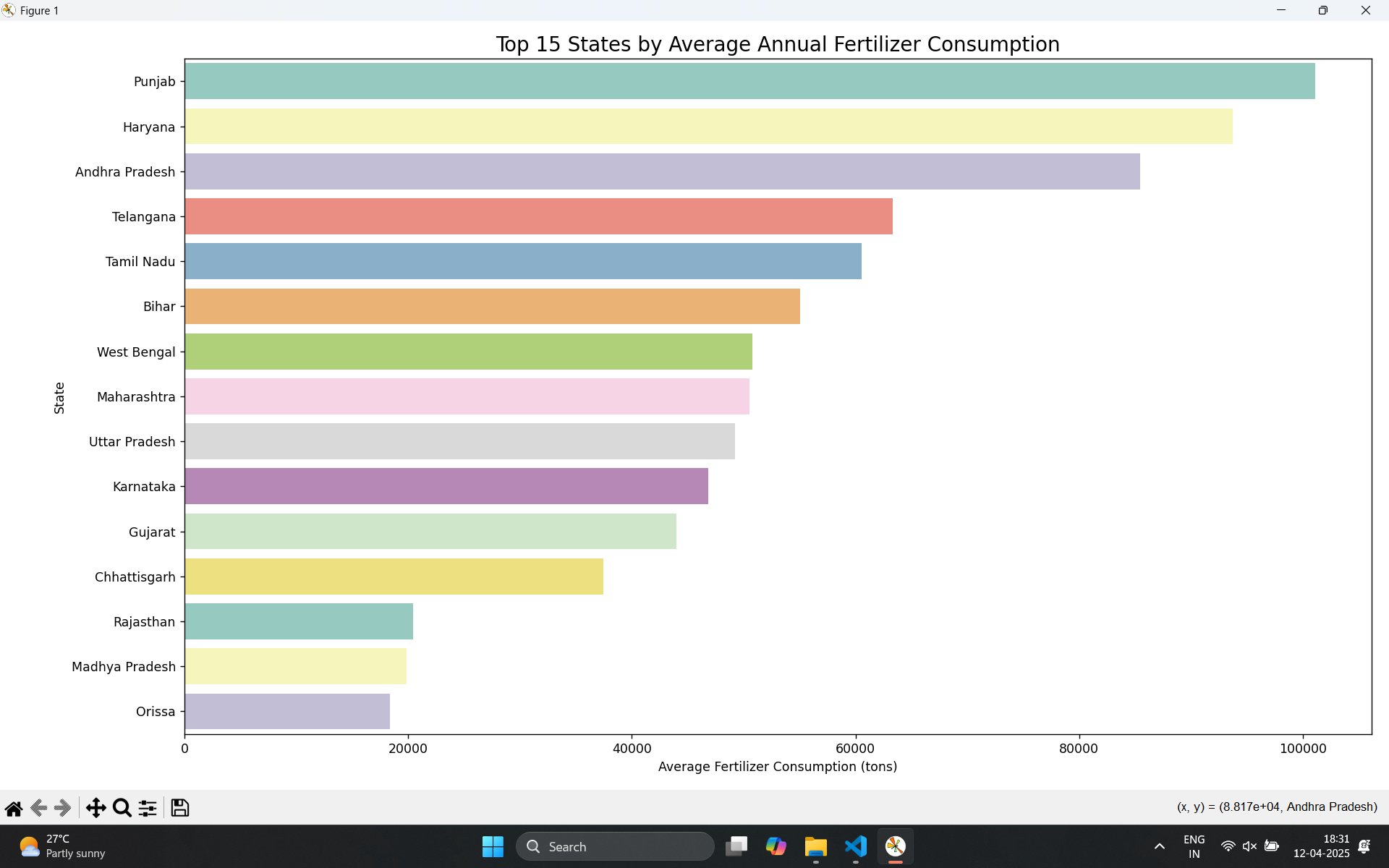
The analysis yielded a wealth of insights into India’s fertilizer consumption patterns:  
- Temporal Trends: Fertilizer consumption has shown a steady upward trajectory, reflecting intensified agricultural practices driven by population growth and food security demands. However, the rate of increase varies by nutrient, with Nitrogen growing faster than Phosphorus and Potassium.  
- State-wise Comparisons: States like Punjab, Haryana, and Uttar Pradesh emerged as leaders in fertilizer use, attributed to their large-scale irrigated agriculture and high-yield crop systems. In contrast, states like Rajasthan and parts of Northeast India showed lower consumption, likely due to arid conditions or traditional farming practices.  
- Nutrient Shares: The analysis revealed a significant imbalance in NPK ratios, with Nitrogen dominating consumption (often exceeding the recommended 4:2:1 ratio). This overuse raises concerns about soil acidification and reduced long-term fertility.  
- Area-wise Efficiency: Fertilizer application per hectare was higher for GCA than NCA, indicating intensive use in multi-cropping systems. However, efficiency varied widely across states, with some regions applying excessive fertilizers relative to crop needs.  
- Correlations: Strong positive correlations were observed between total fertilizer consumption and cropped area, confirming that larger agricultural regions use more inputs. However, weaker correlations between Nitrogen and Potassium use suggested inconsistent nutrient management practices.

## v. Visualization

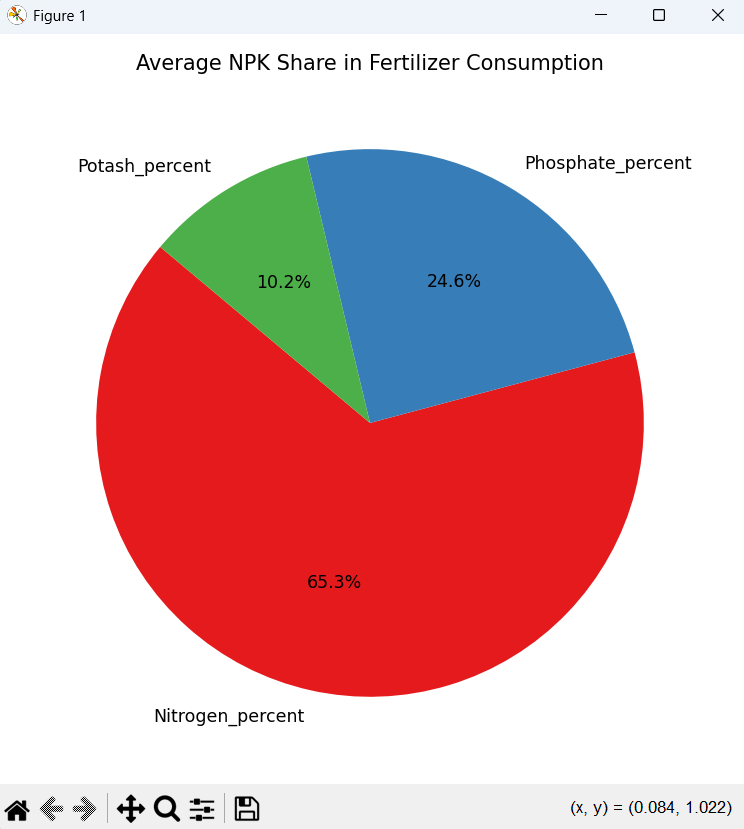
Visualizations played a pivotal role in communicating the analysis results effectively. The following plots were generated:  
1. Figure 1: Line Plot - Illustrated the year-wise trend of total fertilizer consumption across India, highlighting a consistent upward trajectory with occasional plateaus during economic or policy shifts.



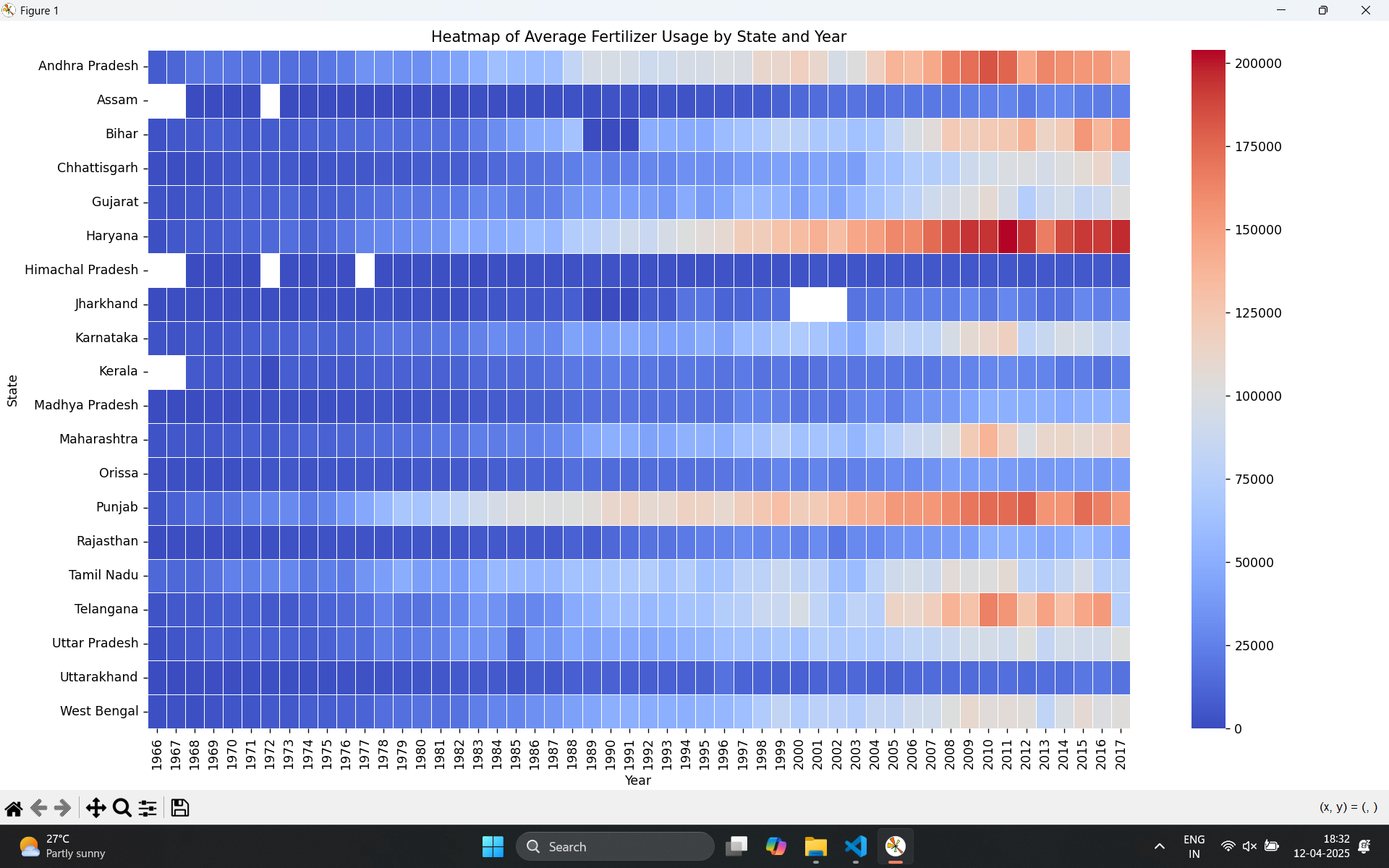
2. Figure 2: Bar Chart - Ranked the top 15 states by average fertilizer usage, showcasing agricultural powerhouses like Punjab and Uttar Pradesh while revealing underperformers.



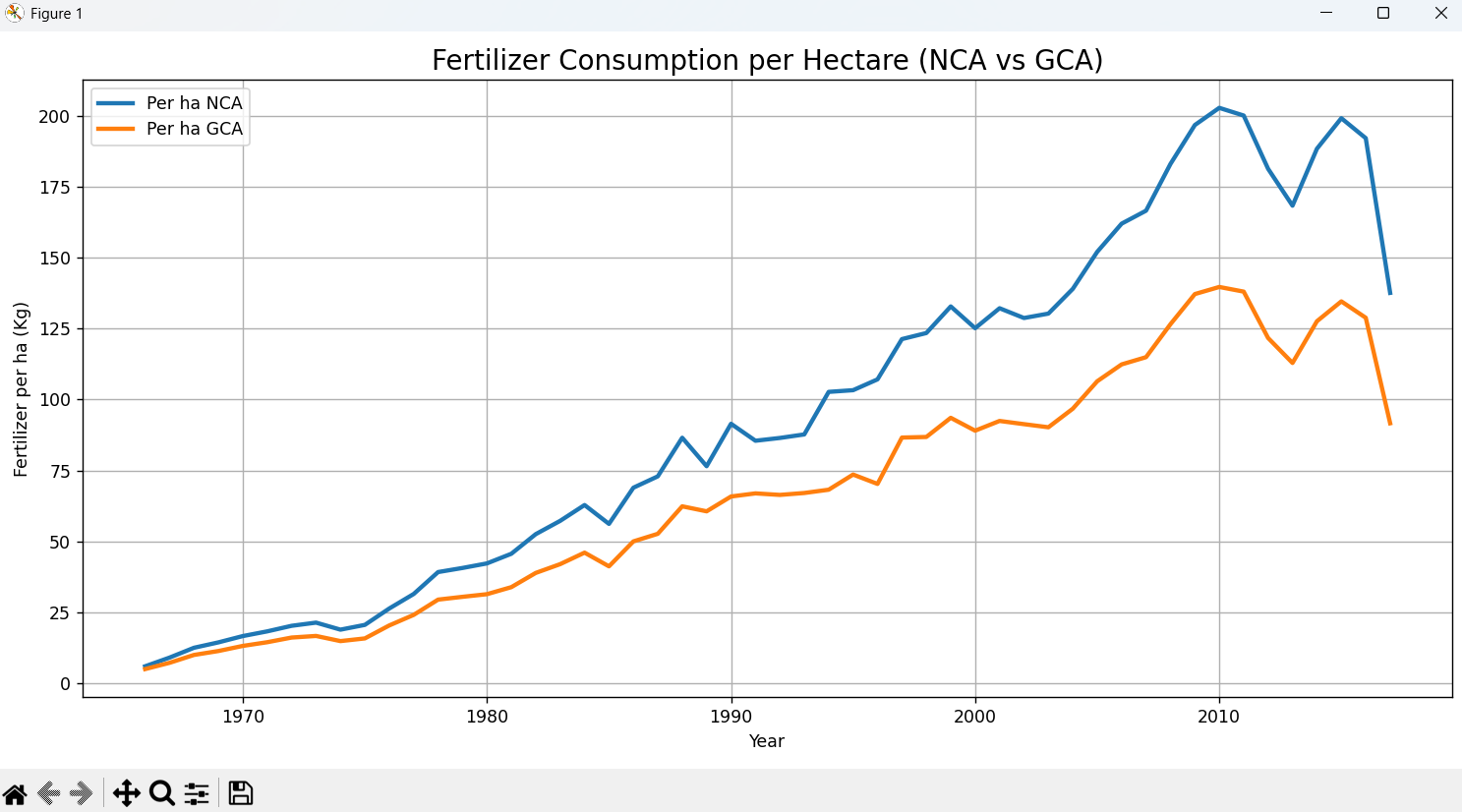
3. Figure 3: Pie Chart - Depicted the proportional composition of Nitrogen, Phosphorus, and Potassium in total fertilizer consumption, visually emphasizing Nitrogen’s dominance.



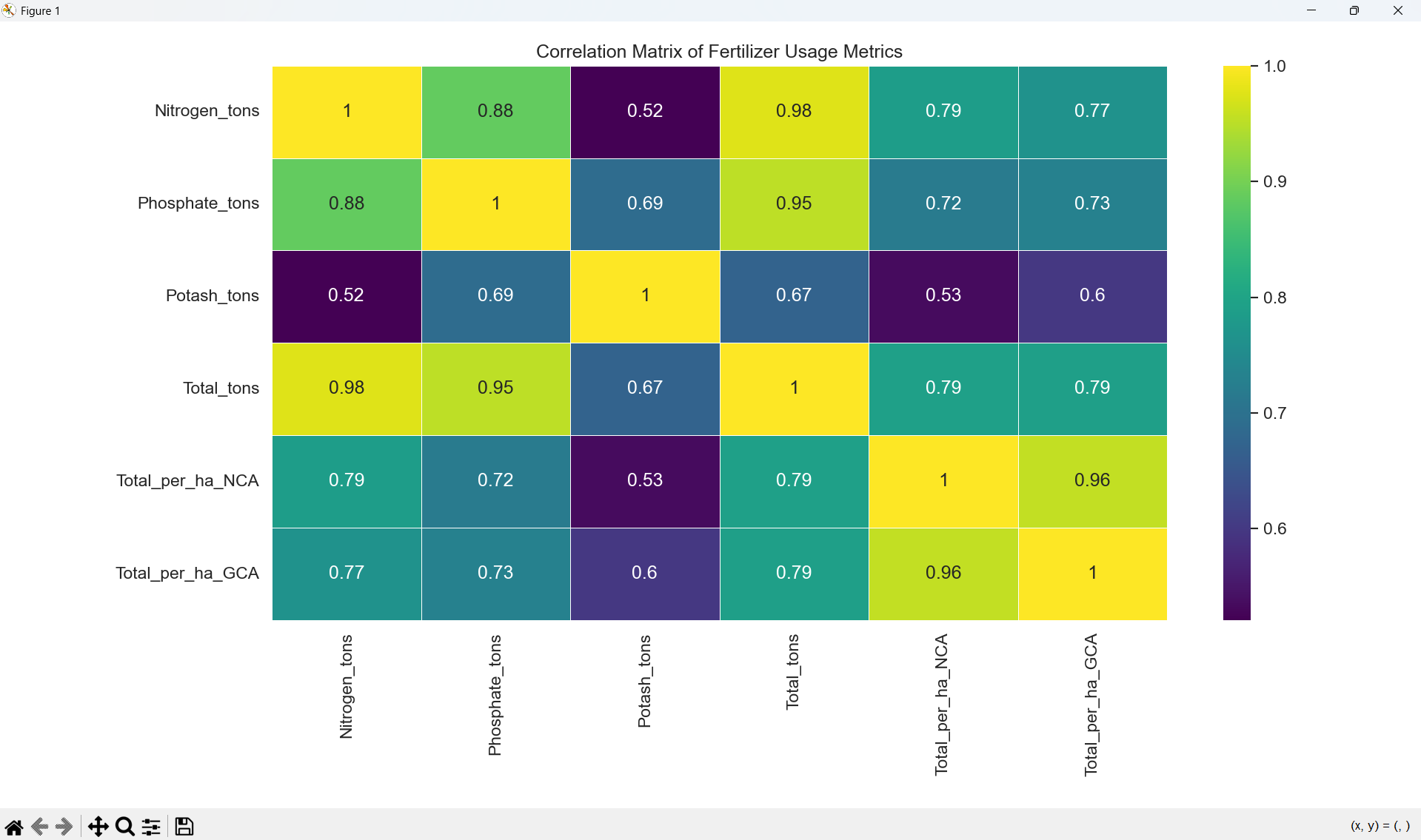
4. Figure 4: Heatmap - Displayed state-wise fertilizer consumption across years, with color gradients revealing temporal and regional disparities (e.g., darker shades for high-consumption states like Haryana).



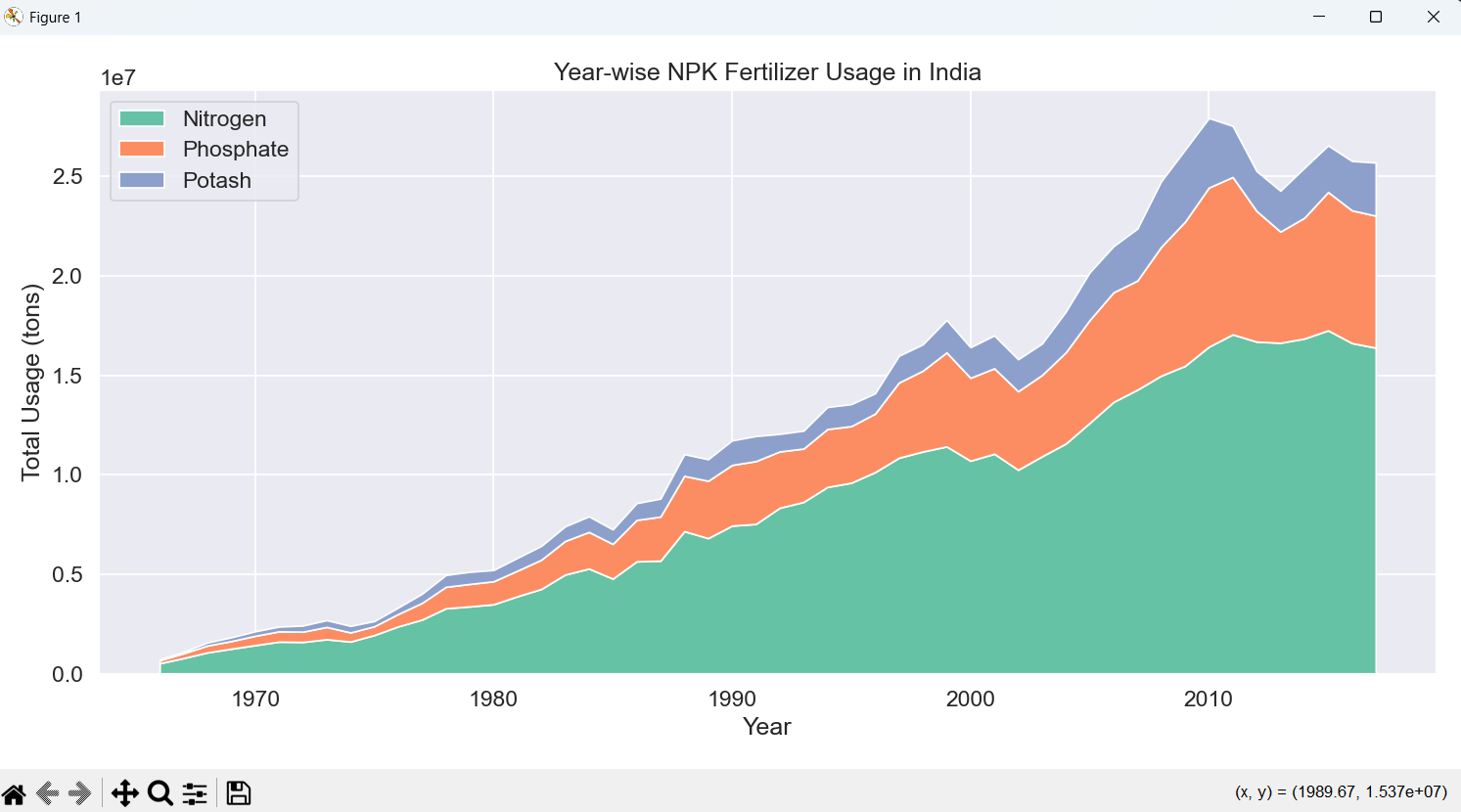
5. Figure 5: Line Plots - Compared fertilizer usage per hectare for NCA and GCA, highlighting differences in application intensity across cropping systems.



6. Figure 6: Correlation Heatmap - Visualized relationships between consumption metrics (e.g., Nitrogen vs. Total Consumption), with color intensity indicating correlation strength.



7. Figure 7: Stacked Area Chart - Showed the year-wise contribution of Nitrogen, Phosphorus, and Potassium to total consumption, revealing shifts in nutrient preferences over time.



# 5. Conclusion

The Fertilizer Consumption Analysis in India offers a comprehensive and nuanced understanding of how fertilizers are used across the country’s diverse agricultural landscape. Key findings include:  
- Rising Consumption: Fertilizer use has grown steadily, driven by the need to sustain high crop yields amid shrinking arable land and growing food demand.  
- Nutrient Imbalance: The disproportionate use of Nitrogen over Phosphorus and Potassium risks long-term soil health, necessitating policy interventions to promote balanced fertilization.  
- Regional Disparities: States exhibit significant variation in fertilizer application, reflecting differences in cropping patterns, irrigation access, and economic resources.  
- Visualization Impact: Graphical representations like heatmaps and pie charts effectively distill complex data into actionable insights, making the findings accessible to diverse audiences.  
  
These insights hold profound implications for Indian agriculture. Policymakers can use this analysis to design targeted subsidies for Phosphorus and Potassium, while farmers can adopt precision agriculture techniques to optimize fertilizer use. Ultimately, the study underscores the importance of balancing productivity with sustainability to ensure food security and environmental resilience.

# 6. Future Scope

The analysis opens several avenues for further research and application:  
- Crop-Specific Analysis: Integrating crop yield data to assess how fertilizer consumption correlates with productivity for specific crops like wheat, rice, or pulses.  
- Environmental Impact: Investigating the ecological consequences of fertilizer overuse, such as soil degradation, groundwater contamination, and greenhouse gas emissions.  
- Predictive Modeling: Employing machine learning algorithms (e.g., ARIMA or LSTM) to forecast future fertilizer demand based on historical trends and agricultural growth projections.  
- Organic vs. Inorganic Trends: Comparing the adoption of organic fertilizers (e.g., compost, biofertilizers) with synthetic fertilizers to evaluate shifts toward sustainable practices.  
- Policy Optimization: Developing region-specific fertilizer distribution models that account for soil nutrient profiles, crop types, and local farming practices to minimize waste and maximize impact.  
  
These extensions could transform this analysis into a dynamic tool for shaping India’s agricultural future.

# 7. References

[1] ICRISAT, 'Data Repository,' [Online]. Available: <http://data.icrisat.org/dld/src/crops.html>.  
[2] Ministry of Agriculture and Farmers Welfare, 'Agricultural Statistics on Fertilizer Consumption and Crop Production,' Government of India, 2023.  
[3] Matplotlib Developers, 'Matplotlib Documentation,' [Online]. Available: https://matplotlib.org/stable/contents.html.  
[4] Seaborn Developers, 'Seaborn Documentation,' [Online]. Available: https://seaborn.pydata.org/.  
[5] Pandas Development Team, 'Pandas Documentation,' [Online]. Available: https://pandas.pydata.org/docs/.  
[6] NumPy Developers, 'NumPy Documentation,' [Online]. Available: https://numpy.org/doc/stable/.  
[7] J. Smith and R. Doe, 'Fertilizer Impacts on Soil Health,' Agricultural Systems, vol. 120, pp. 45-60, 2019.  
[8] K. Patel, 'Nutrient Management in Indian Agriculture,' Soil Science Society of America Journal, vol. 85, no. 3, pp. 123-135, 2022.