

DATA VISUALISATION (DA 322)

Topic: Agriculture

Shubhi Agarwal
210150016

Table of Contents

- 1) Central Level Analysis
- 2) Case Study of 3 Different Crops: Rice, Wheat, Sugarcane
- 3) Cost of Production of Various Crops
- 4) Case Study on 3 States: Uttar Pradesh, Punjab, Maharashtra

Motivation

Agriculture is the foundation of food production, providing the basic sustenance required for human survival. Analyzing agriculture is crucial because it directly addresses the fundamental need for food security and nutrition.

Agriculture is a significant contributor to India's economy, employing a large portion of the population and contributing a substantial share to the Gross Domestic Product (GDP). Analyzing agricultural trends, productivity, and market dynamics helps in understanding the economic performance of the country.

Data-driven agricultural analysis provides policymakers with insights into the performance of existing policies . This helps in designing targeted policies and programs aimed at addressing challenges such as agricultural productivity, farmer welfare, land use management, and sustainable development.

Central Level Analysis

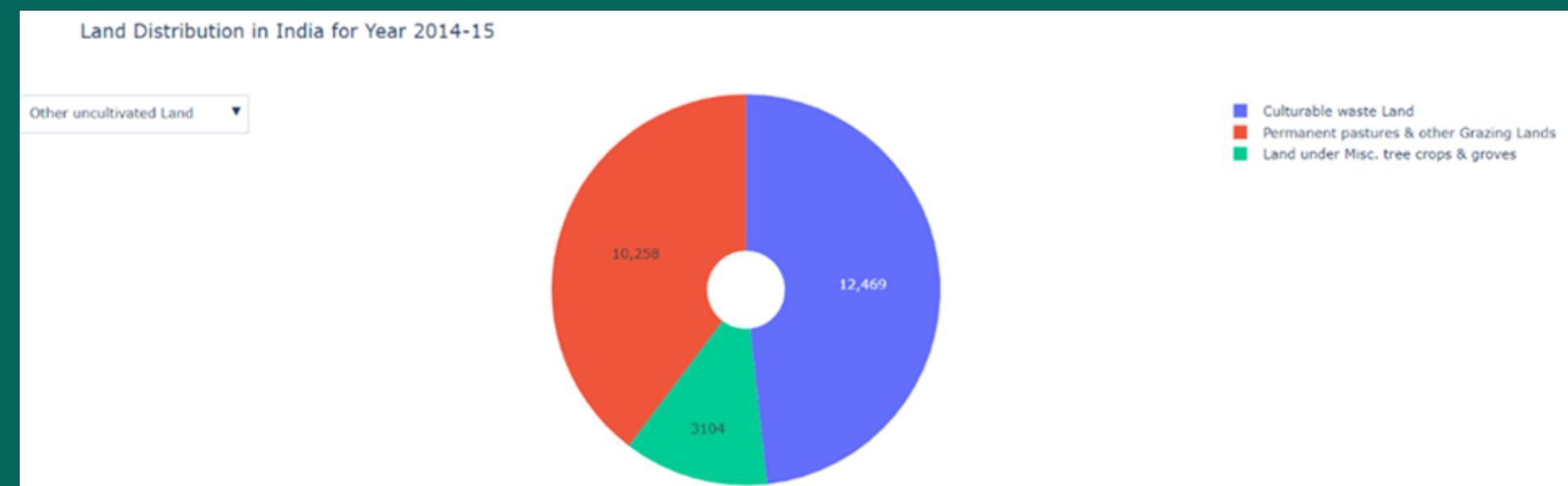
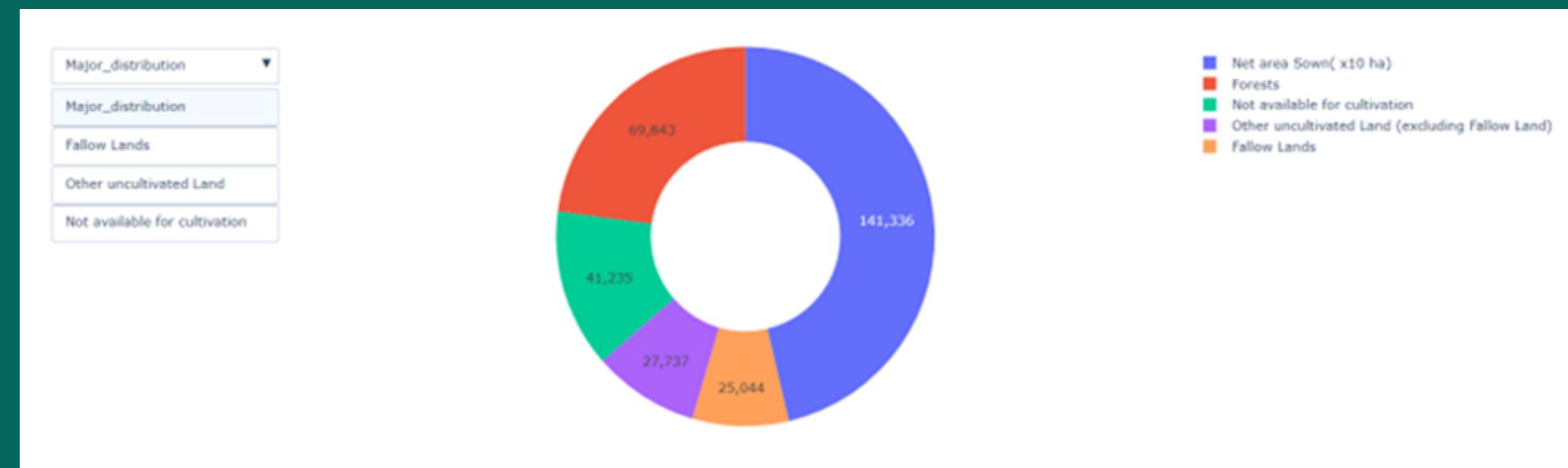
Land distribution of India from 2000 to 2015

India is the 7th largest country in the world in terms of geographical land area. Analyzing land distribution in India holds paramount importance across multiple facets.

By deciphering how land is allocated and utilized, policymakers can unlock critical insights into agricultural productivity, food security, and socioeconomic development

Environmental sustainability also hinges on effective land management practices, which are informed by insights into land distribution.

Plot used : An interactive donut chart



Why the interactive donut chart ?

- 1) Visual Appeal: Donut charts are visually appealing and easy to interpret.
- 2) Users can interact with the chart by clicking on segments to explore details.

Observations:

- 1) India is a vast country and has a large geographical area, out of which almost 23 % are forests, 45% are used for cultivation , around 8-9% is fallow land whereas the rest is not used for crop cultivation while a small part (less than 1%) land is barren.
- 2) The forest area has increased in recent years.

Remark

We know the land is limited, therefore we should move towards sustainable development which reduces burden on land .

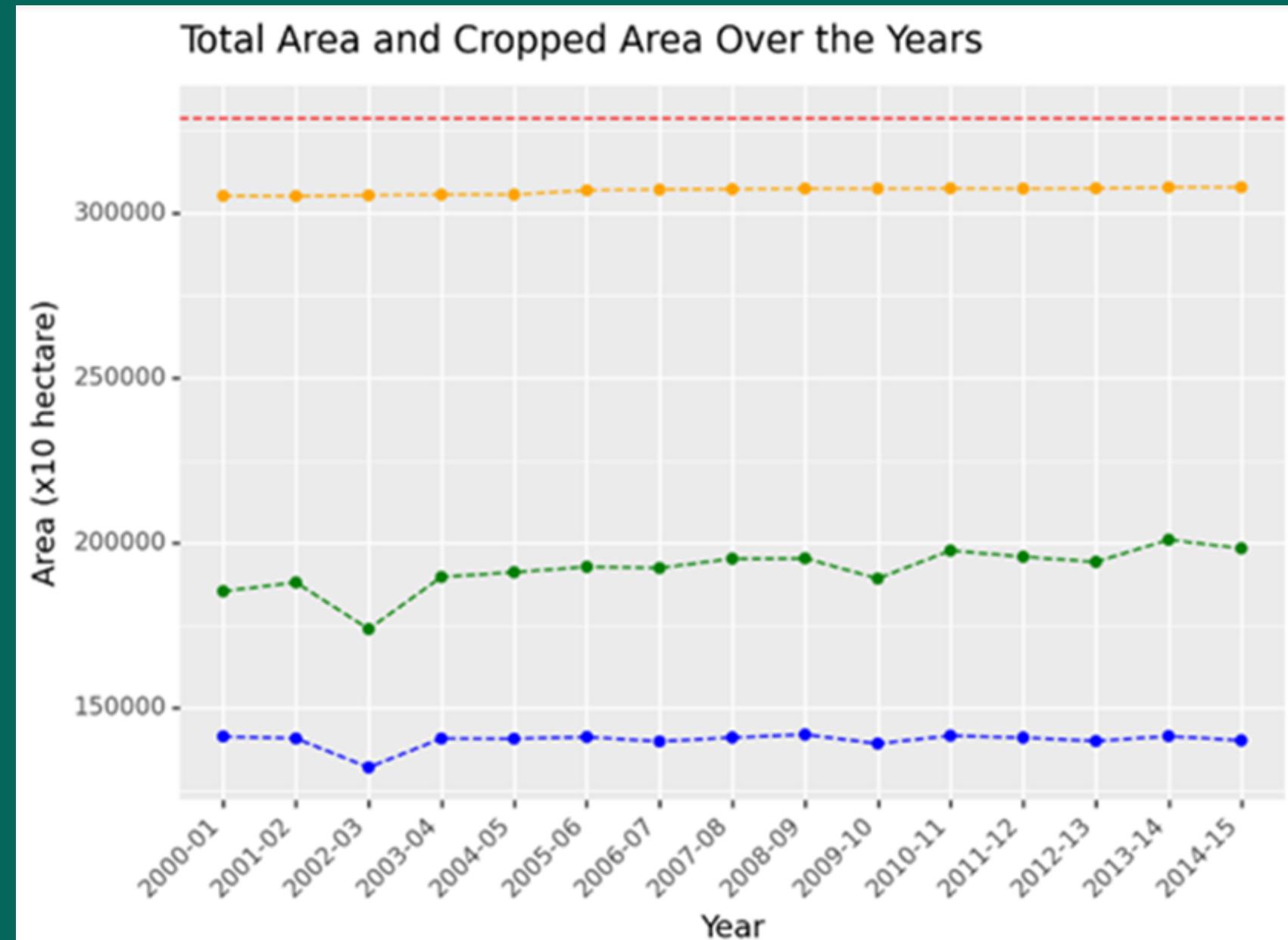
Total cropped area and net sown area over the years

Observations:

- 1) It can be seen that there was a significant decrease in both total cropped area and total area sown in 2002-03.
- 2) The Total cropped area is almost twice the net sown area which seems right as in India, multiple crops are often harvested in a single year.

Remark:

One reason could be the sparse rainfall and droughts that year. As we know, Rajasthan faced one of the worst droughts of the century that year.



Plot used : Line Plot

- Temporal Trends: Line plots are excellent for displaying trends over time.
- Multiple Variables: The plot includes multiple variables (total area sown, total cropped area, total reported area for land utilization statistics), which can be effectively visualized using different lines on the same plot.
- Connection of Data Points: Line plots connect data points, making it easier to see trends and patterns in the data compared to scattered points alone.
- Comparative Analysis: Line plots enable the viewer to compare the changes in different variables (e.g., total area sown, total cropped area) over the same time period, providing valuable insights into how these factors relate to each other.

Line Plot of Area, Production and Yield of Different Crops between (2001-15)

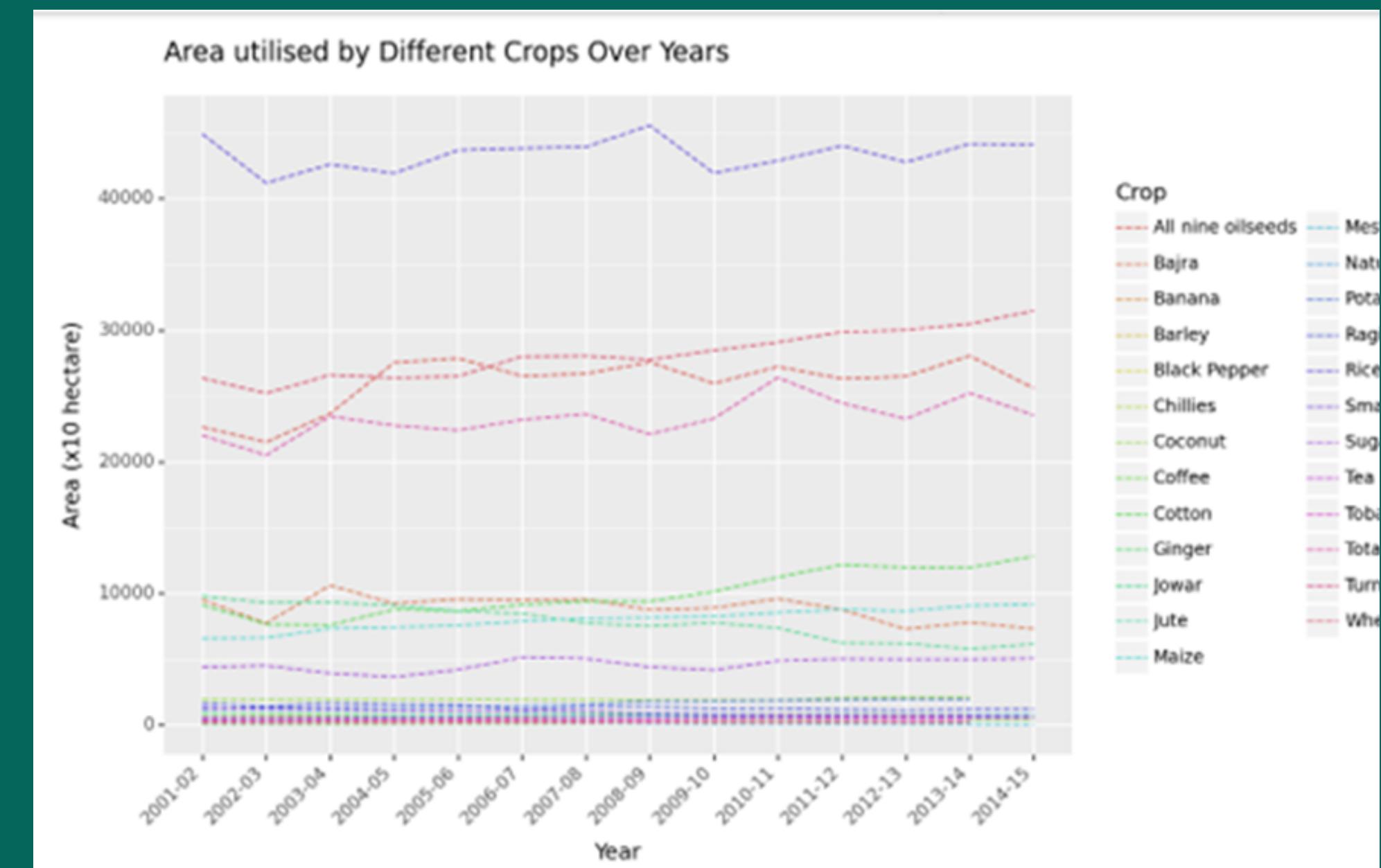
Motivation:

I wanted to see whether there are crops whose land utilization has changed, as this would provide insights into increasing demand of that crop or the arrival of new technology that reduces the need for as much land.

Observation and remarks:

We can see that rice uses the most area followed by wheat followed by grains and then cash crops. This land distribution can be because of 2 reasons:

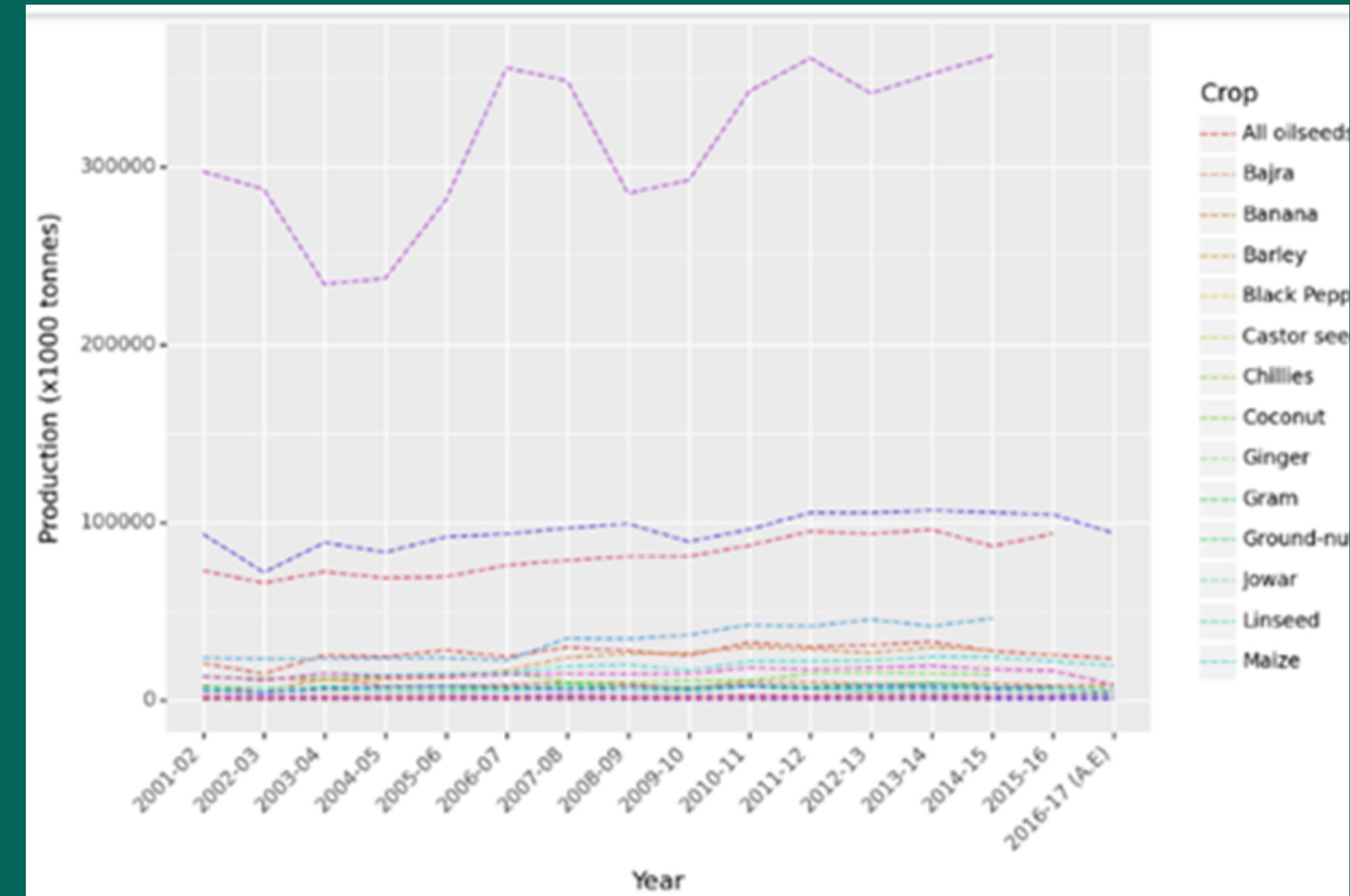
- **Demand and Supply:** More quantity is grown, therefore they require more area or there is less demand, therefore only little amount is grown
- Some crops require less land area to grow. Eg: Tobacco.



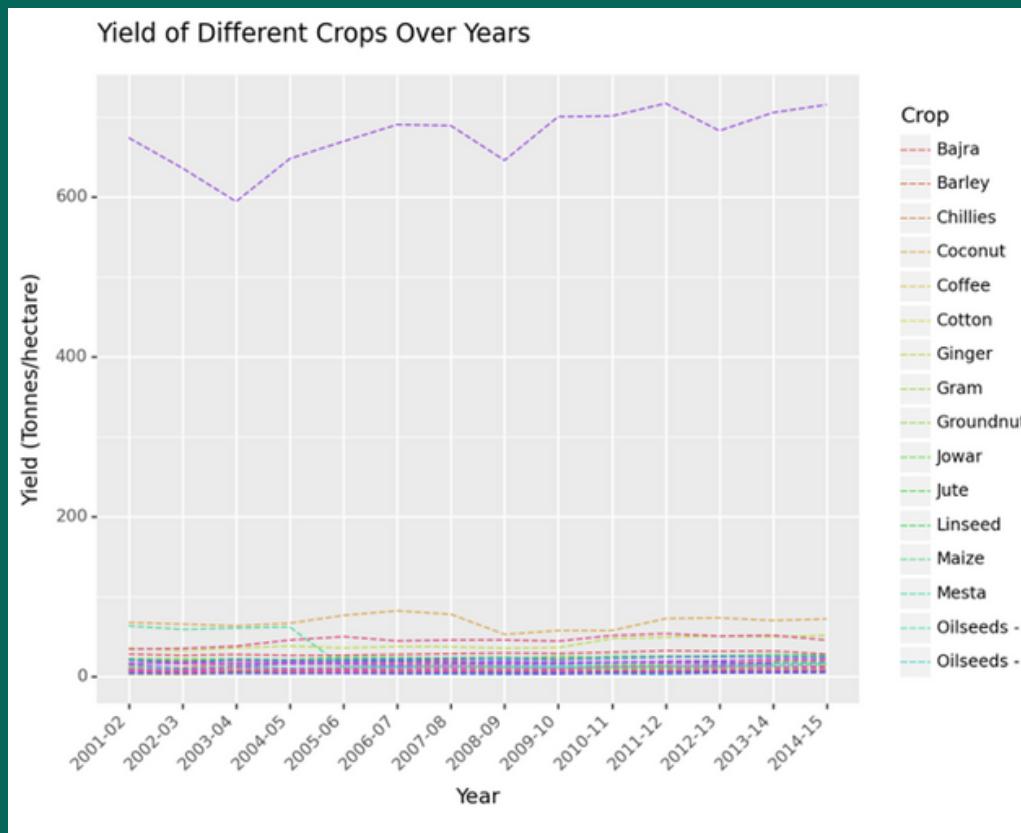
Graph of total production of different crops over years

Observations and Remarks:

- We can see that production of sugarcane is considerably higher than any other crop. It is then followed by rice and wheat.
- India is one of the world's leading producers of sugarcane, wheat and rice. India comes second in terms of sugarcane production only after Brazil, and second in rice and wheat just after China.
- While rice and wheat are primarily used for consumption, (some amount is exported as well), sugarcane is primarily used for sale purposes(Cash Crop)



Yield of different crops over years



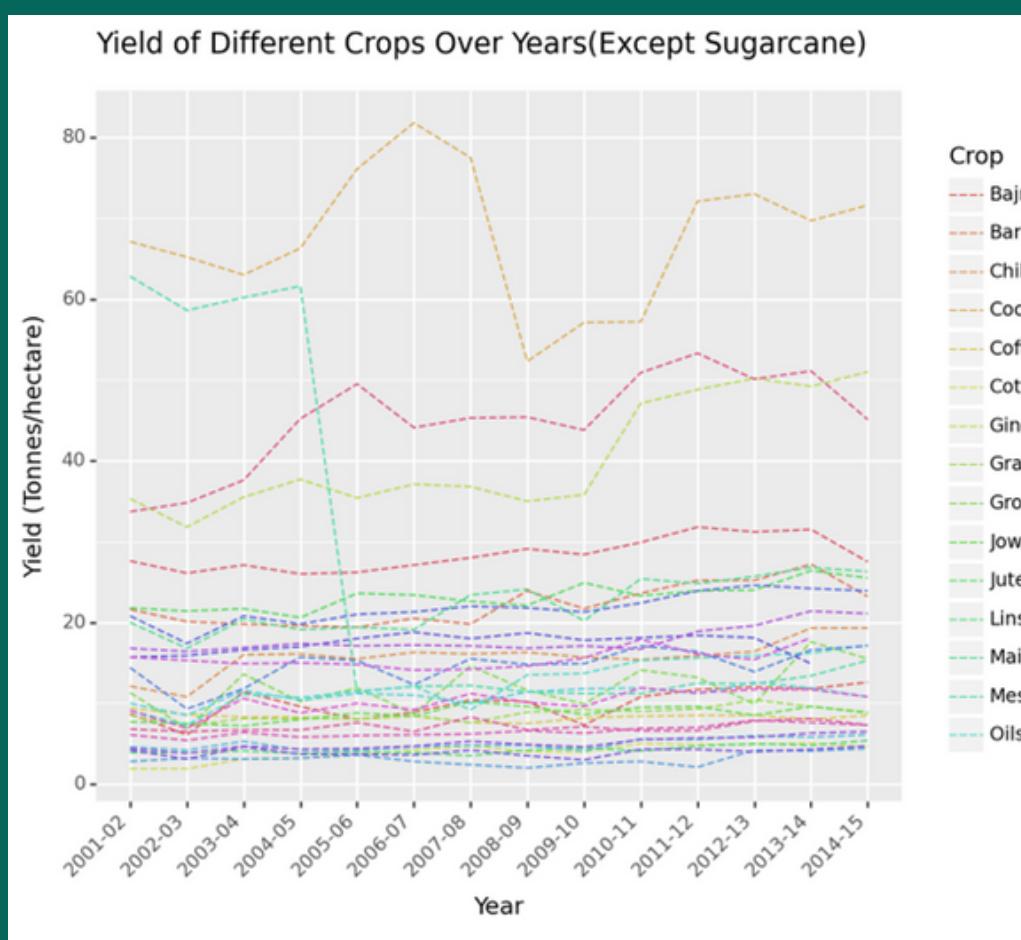
Observations and Remarks:

Yield of sugarcane is very very high.

The overall yield of coconut is high but a lot of fluctuations can be seen.

The potential reasons can be :

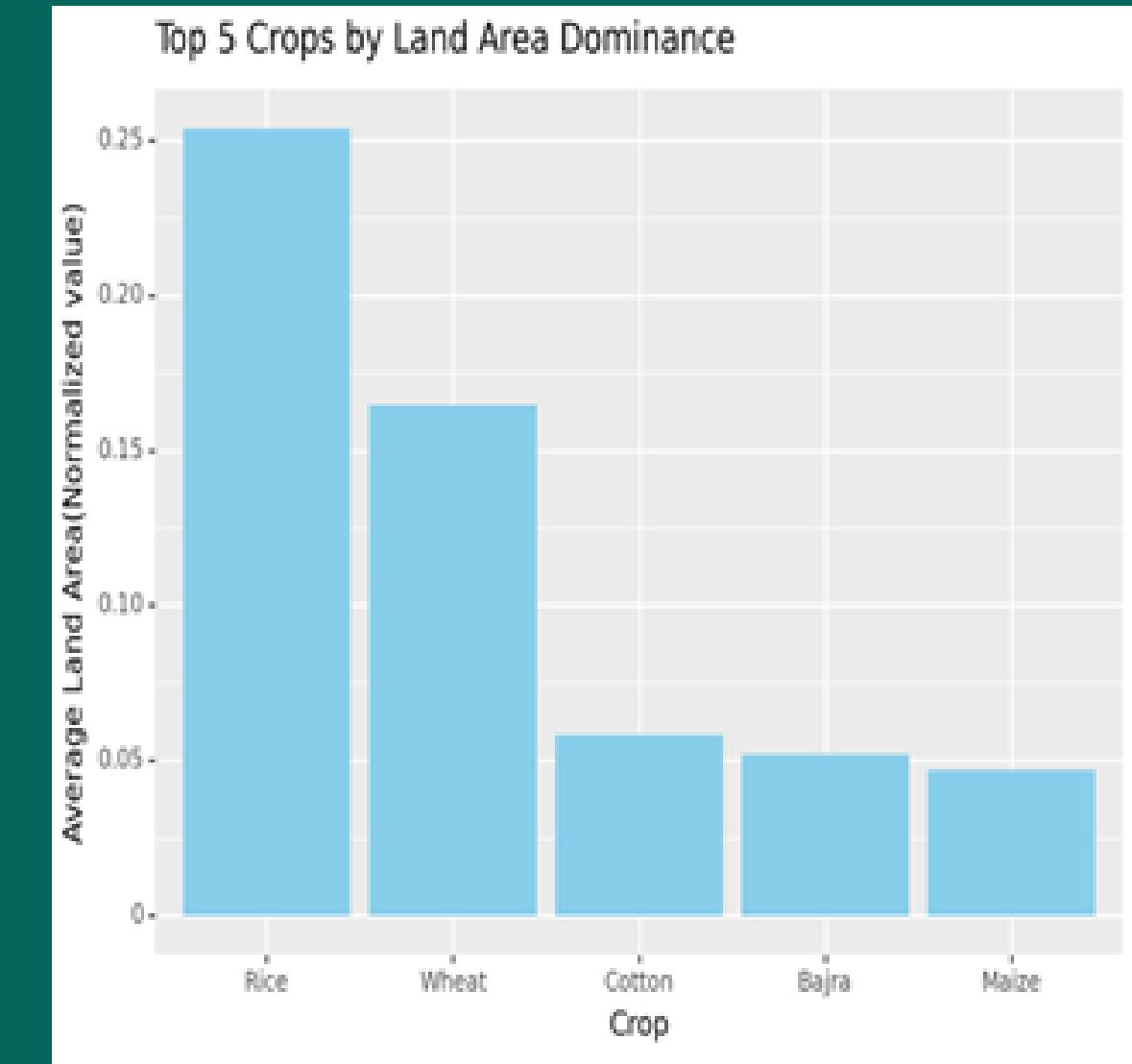
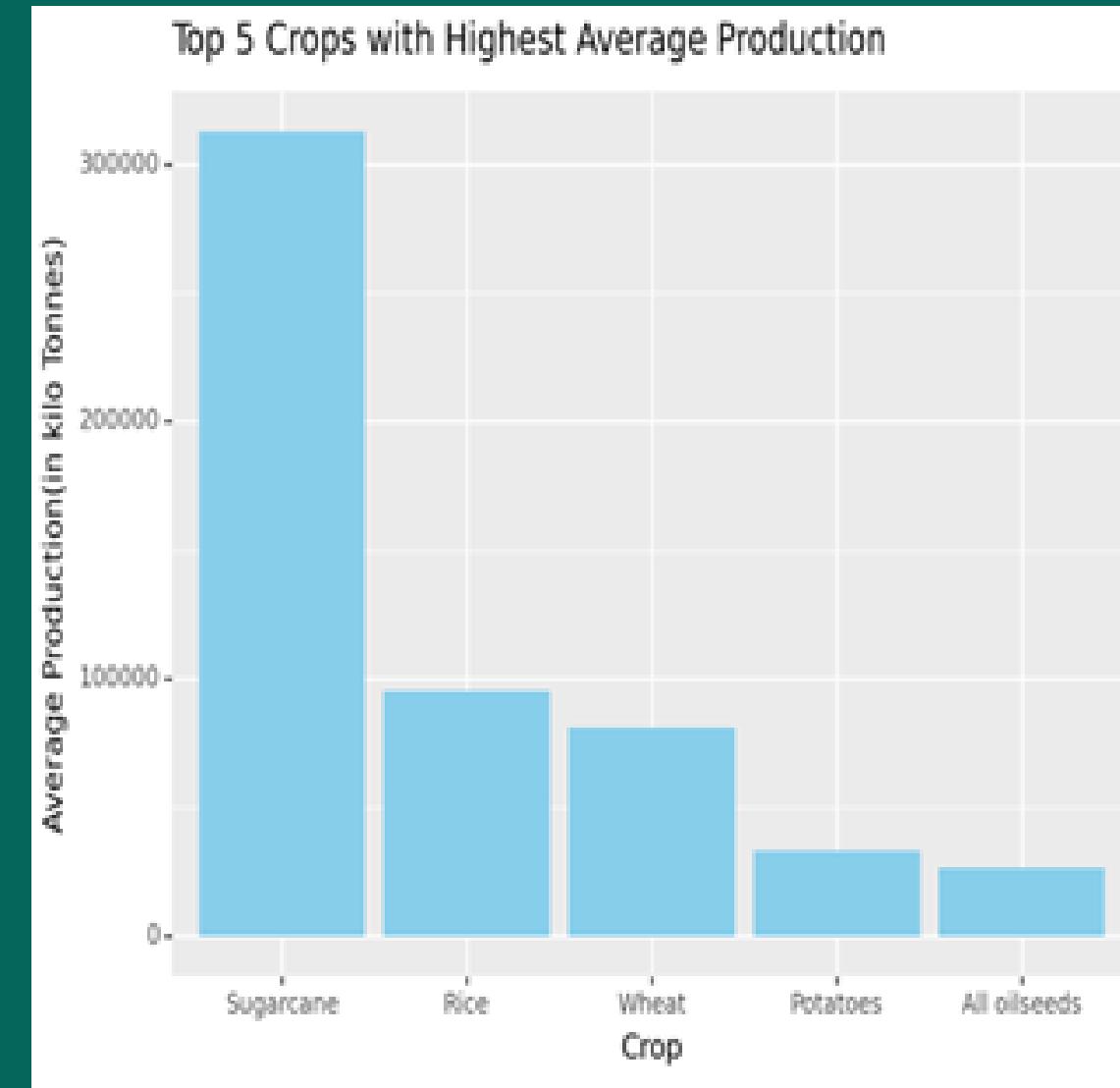
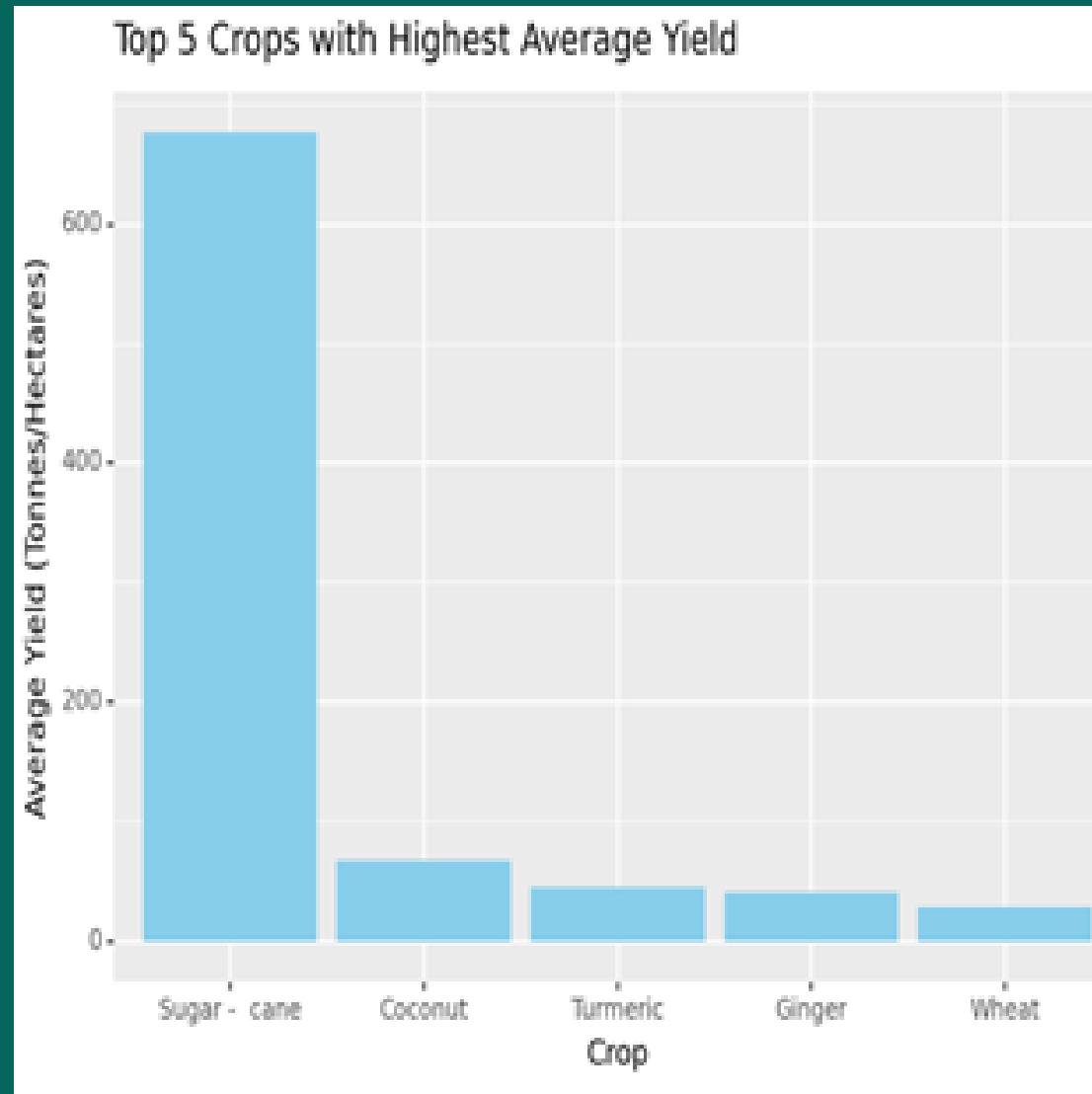
- Coconut cultivation is highly sensitive to weather conditions, including rainfall, temperature, humidity, and wind patterns. Fluctuations in weather patterns can have a significant impact on coconut yield.***
- Pest and disease infestations.***



We can see that the yield of Mesta dropped after 2005. Since it did not increase in the next few years, a possible reason for this drop can be soil degradation, poor agronomic practices like: excessive use of chemicals or not doing crop rotation.

There is a good increase in the yield of wheat and ginger. For other crops, there was a slight increase in yield over the years

Bar Charts of top 5 crops in terms of highest Average Production, Yield and Area covered between 2000 to 2015



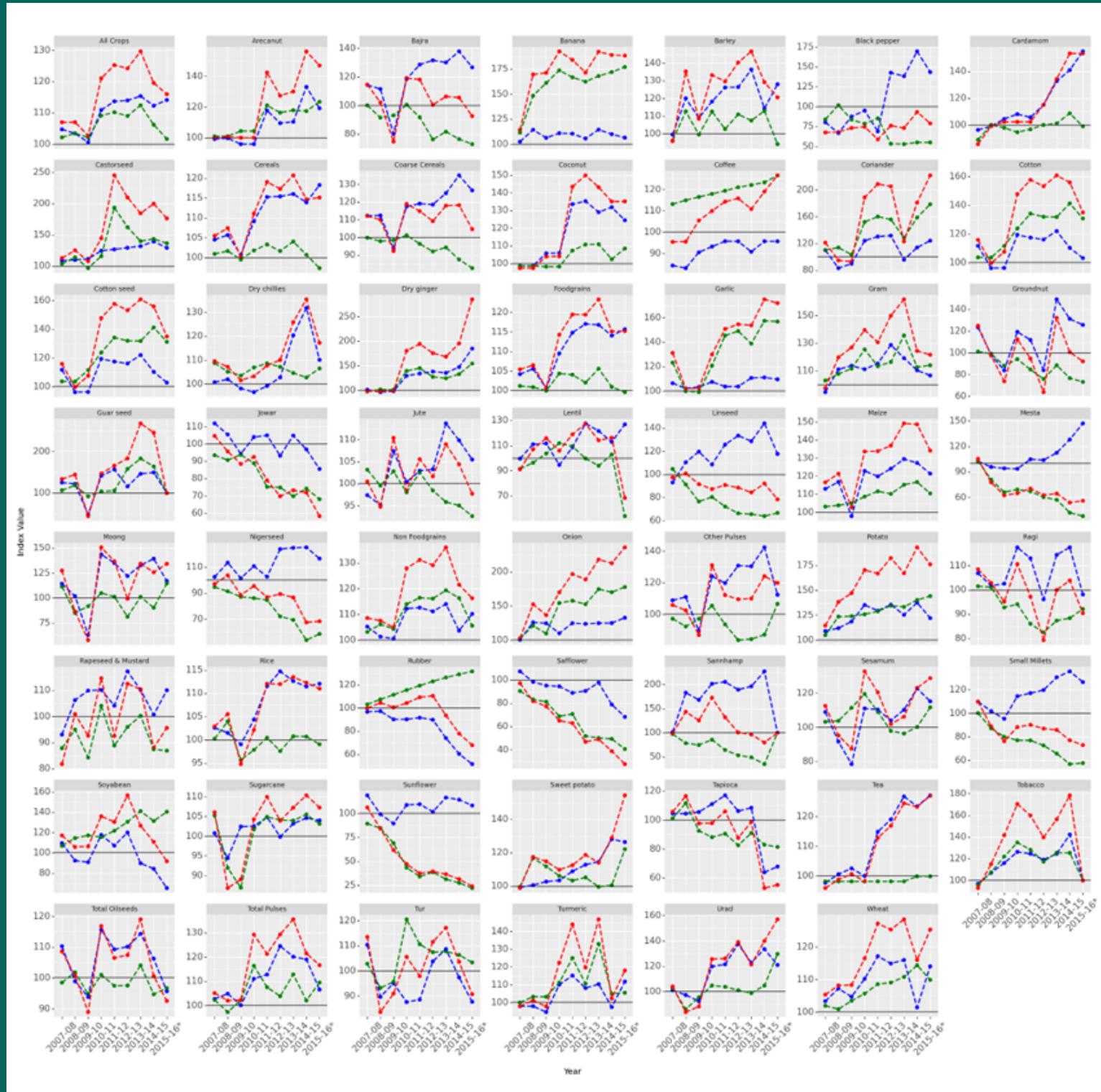
Observations:

1. Sugarcane has the highest Average Production and yield but it does not require much land to grow
2. Rice utilizes the most land.

Reason for Bar Chart:

- 1) *Comparison: Bar charts make it easy to compare the values of different categories (crops) by visually representing their magnitudes with distinct bars.*
- 2) *Ranking: Bar charts naturally lend themselves to ranking data, with the tallest bars indicating the highest values.*
- 3) *Flexibility: Bar graphs can accommodate various types of data, including categorical data and numerical data*

Multiple facet points and line graph showing the index of production, yield and area over the years 2007 to 2017



Observations and Remarks :

- The overall yield and production has increased for most of the crops through the years but there are crops in which a decrease is shown like: jowar, sunflower.
- While most of the crops show erratic behavior for production level there are some crops which show monotonic behavior (like onions, tea).

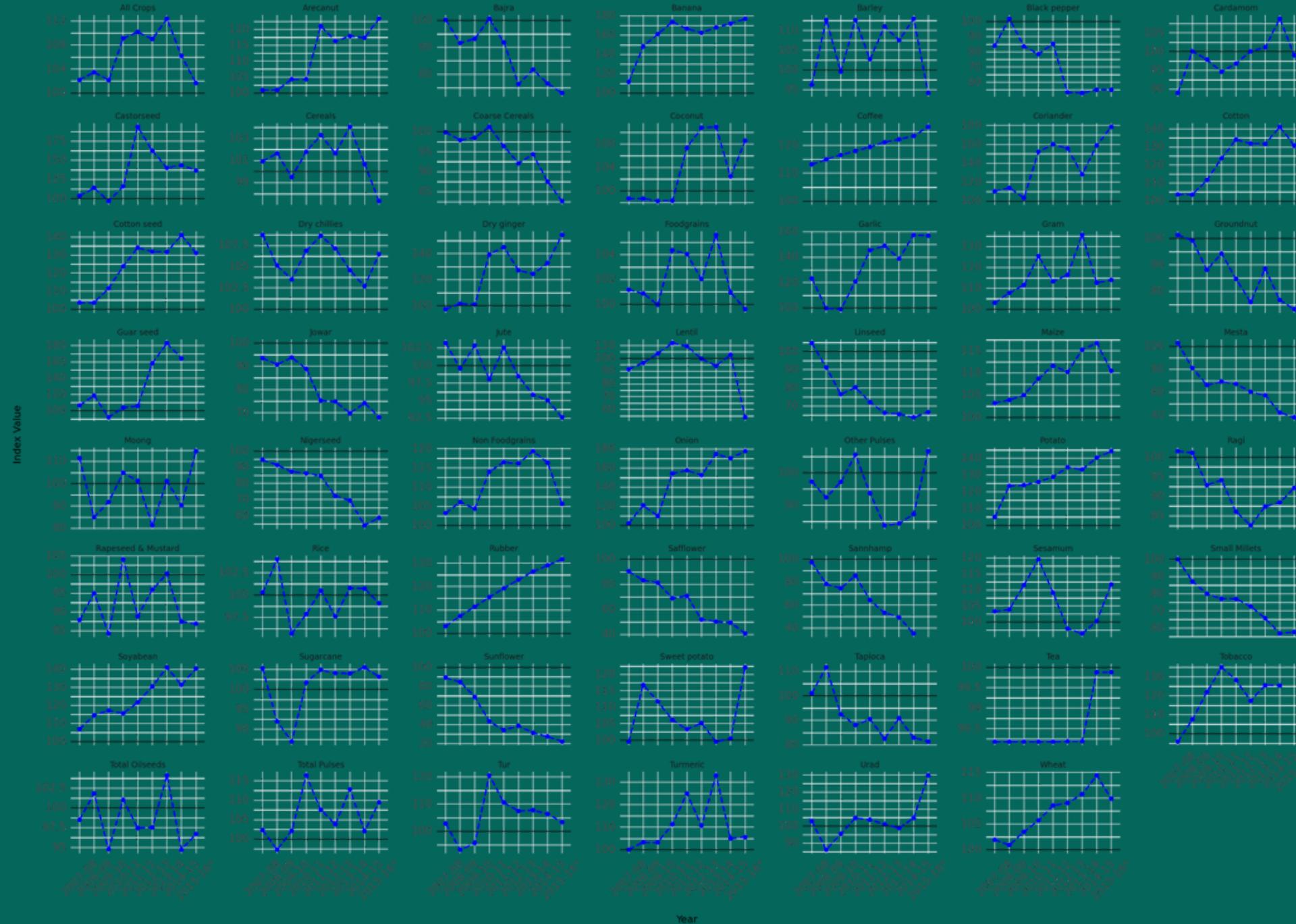
Index of production, yield and Area:

It refers to a relative measure used to compare the production, yield, and area of crops across different time periods or regions. This represents the relative level of production of a crop compared to a reference period or baseline. We use the data from 1993-94 as the baseline.

Why Faceted Line plot ?

- **Comparative Analysis:** *The plot allows for easy comparison of yield, area, and production trends for different crops over time within a single visualization.*
- **Faceted Layout:** *The use of facets (small multiples) organizes the data into individual panels for each crop, making it easier to analyze and compare trends for each crop separately.*
- **Clear Representation:** *By using both points and lines, the plot effectively visualizes the data, providing clear insights into the trends and variations over time.*
- **Space Efficiency:** *The use of facets and a multi-panel layout ensures that each crop's data is displayed compactly, optimizing the use of space and making efficient use of the visualization area.*
- **Scalability:** *The plot can accommodate a large number of crops while still maintaining clarity and readability, thanks to the faceted layout and scalable design.*

Multi facet line plots for price index between 2007 to 2016



I was motivated to create plot to see how the prices of various crops changed over years and what were the trends

Observations:

- 1) Generally the trends are erratic but for some crops they were monotonic
- 2) They were monotonically increasing in rubber, coffee, potato, soyabean
- 3) They were monotonically decreasing for sunflower, wallflower, linseed, mesta

Case Study of 3 Crops :

Rice, Wheat and Sugarcane

My motivation to choose these 3 crops were:

Significance in Agriculture: Rice, wheat, and sugarcane are among the most significant crops in global agriculture, with substantial contributions to food security, economic stability, and livelihoods of millions of people worldwide.

Diverse Cultivation Practices: Each of these crops has unique cultivation practices, requirements, and challenges, providing a rich landscape for analysis. Studying these crops allows for insights into a diverse range of agricultural techniques, including flooded rice paddies, dryland wheat farming, and sugarcane cultivation.

Plot comparing Index values in the years 2007-16

Observations:

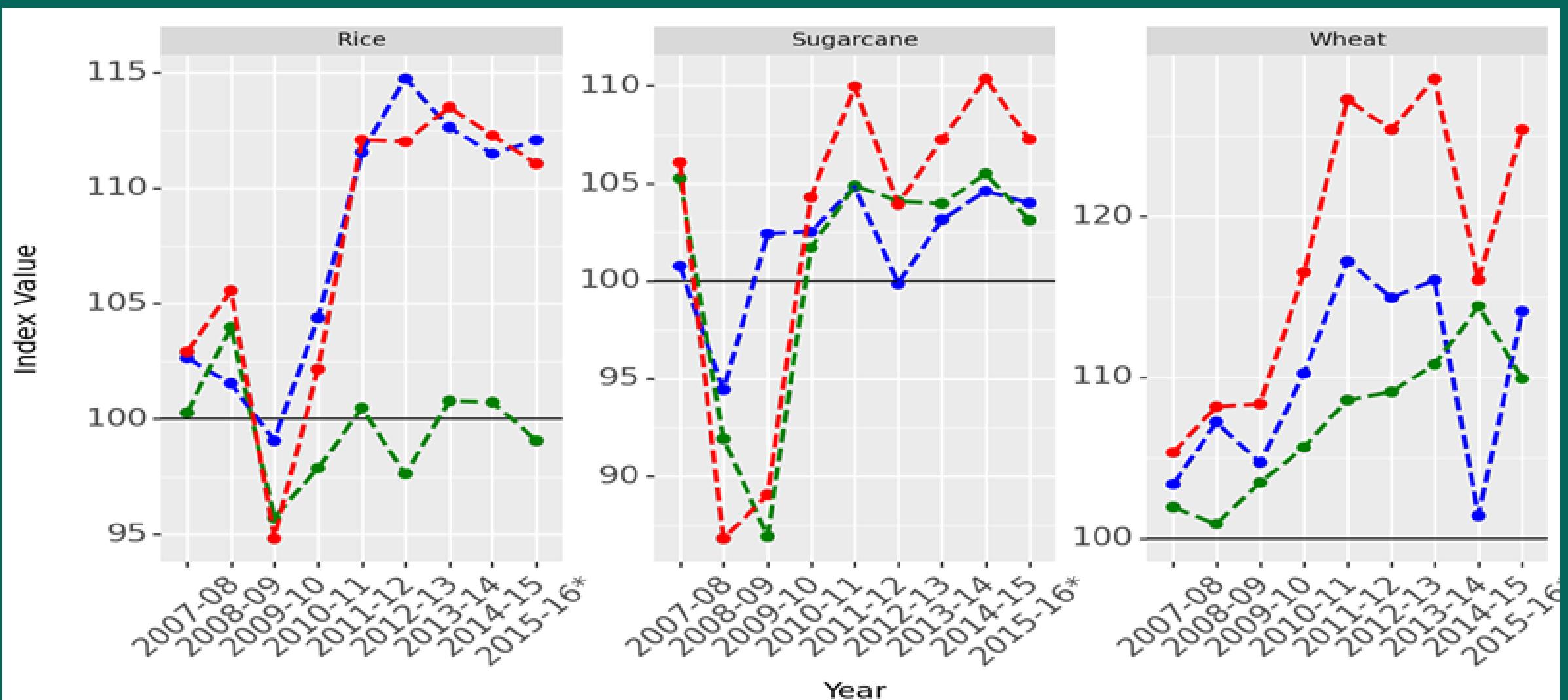
The yield and productivity of all 3 crops increased over time.

The index of wheat was always above 100 i.e. its values were always more than that were in 1993-94 (even at the lowest)

The area allotted to rice cultivation has not changed much since 1993-94

Why did I chose this plot ?

Comparative Analysis
Faceted Layout

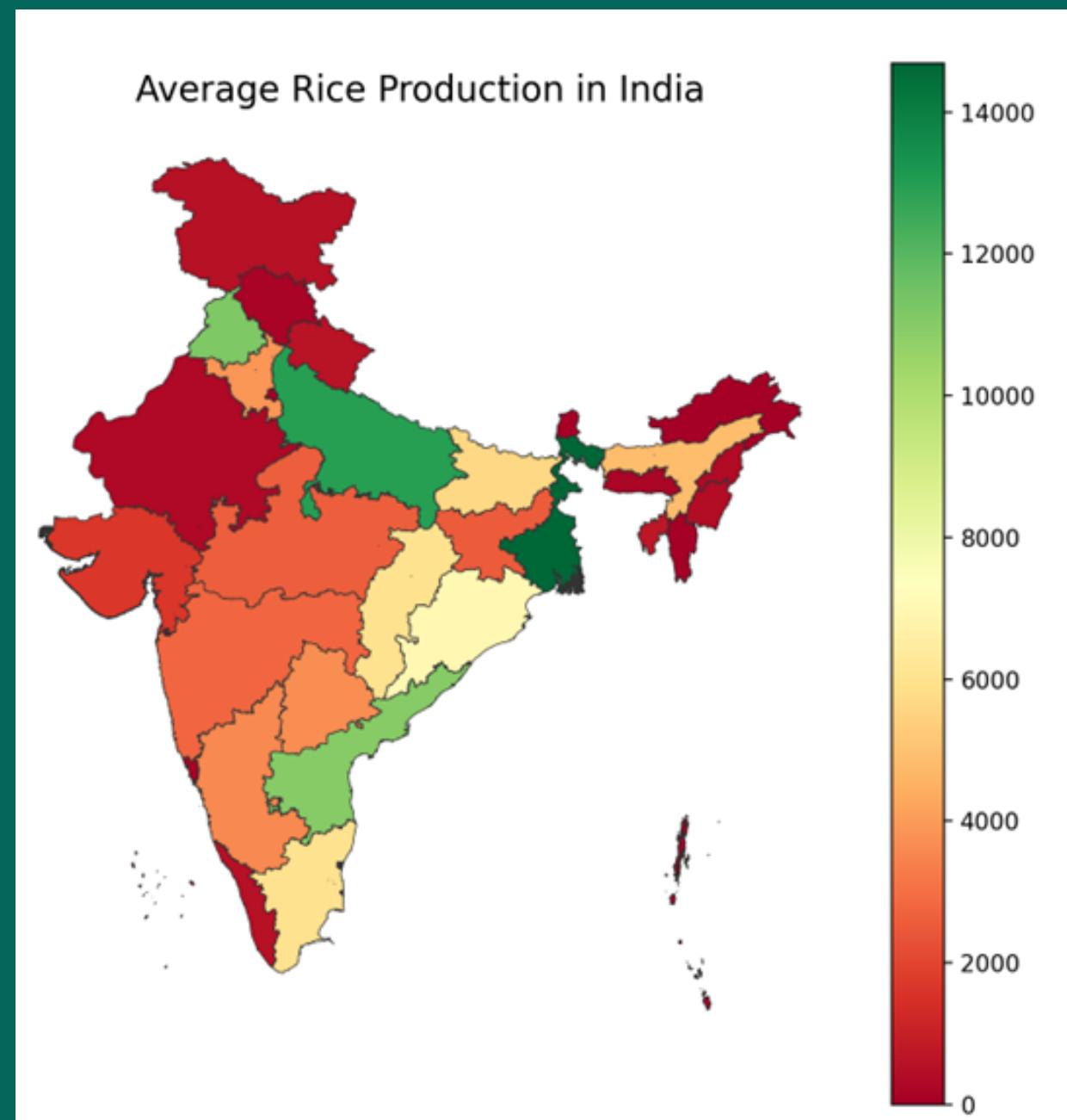


Choropleth map of India to see the production and yield of the 3 crops across different regions of India.

(Made interactive with addition of button to choose which crop to see)

Observations:

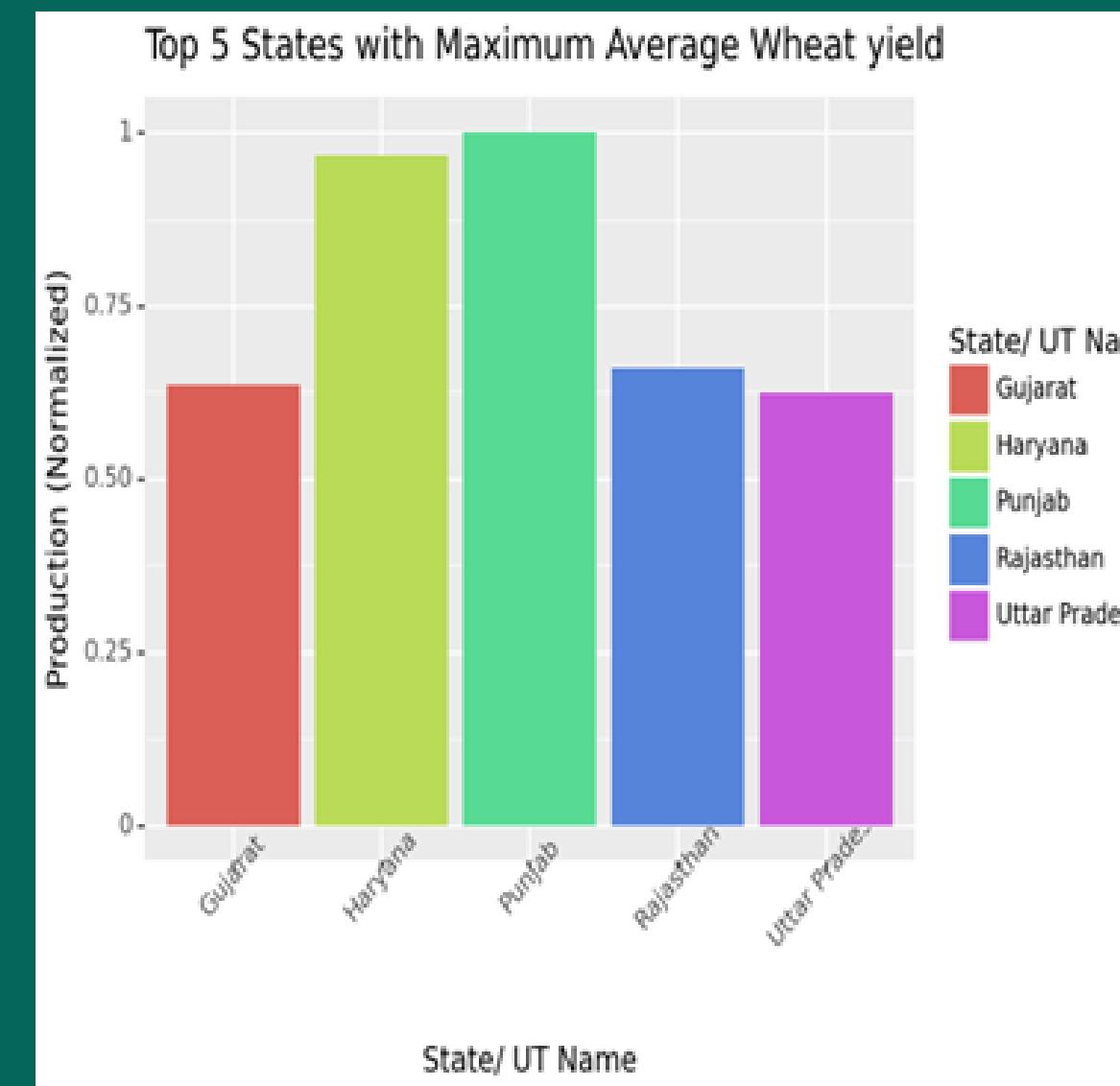
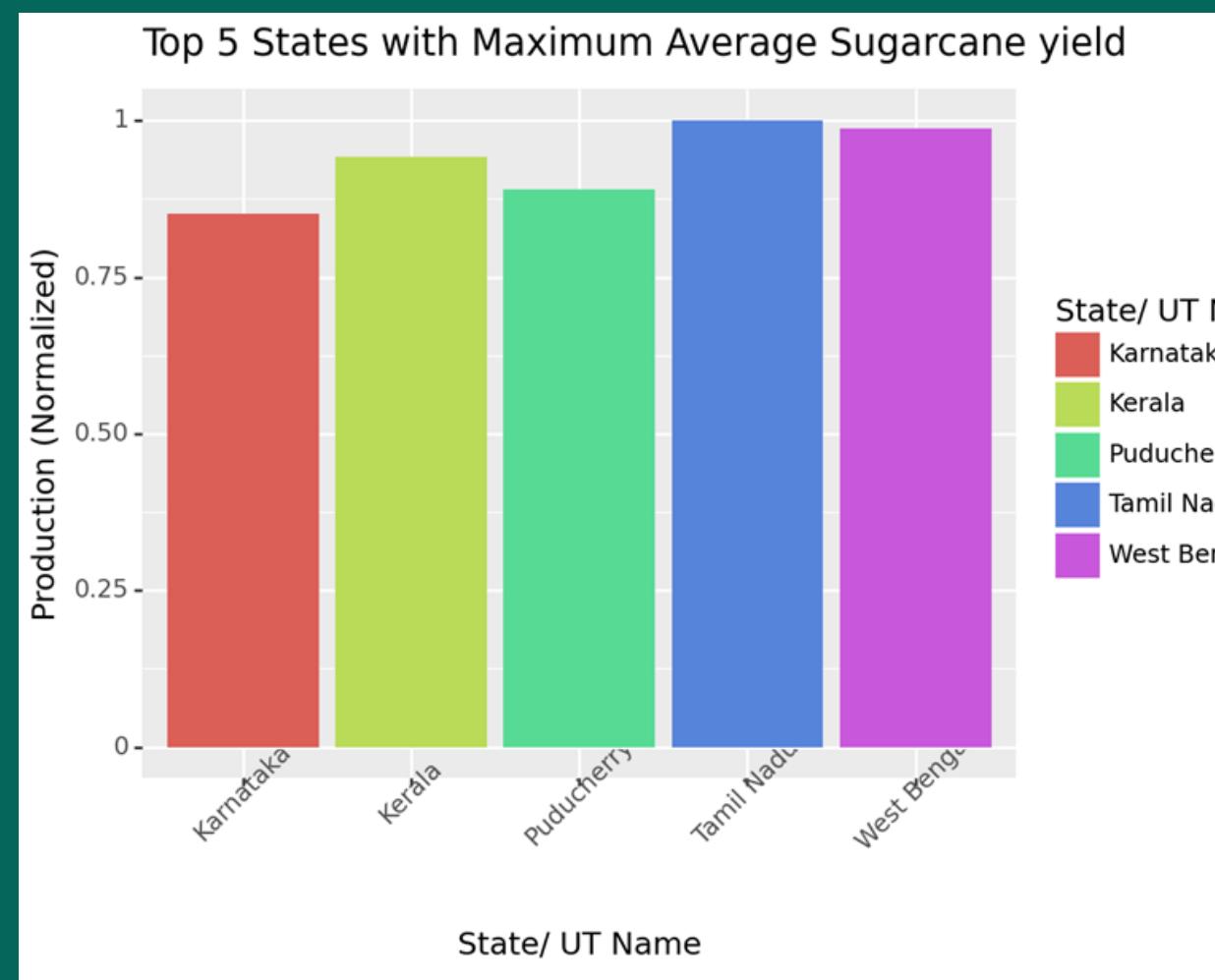
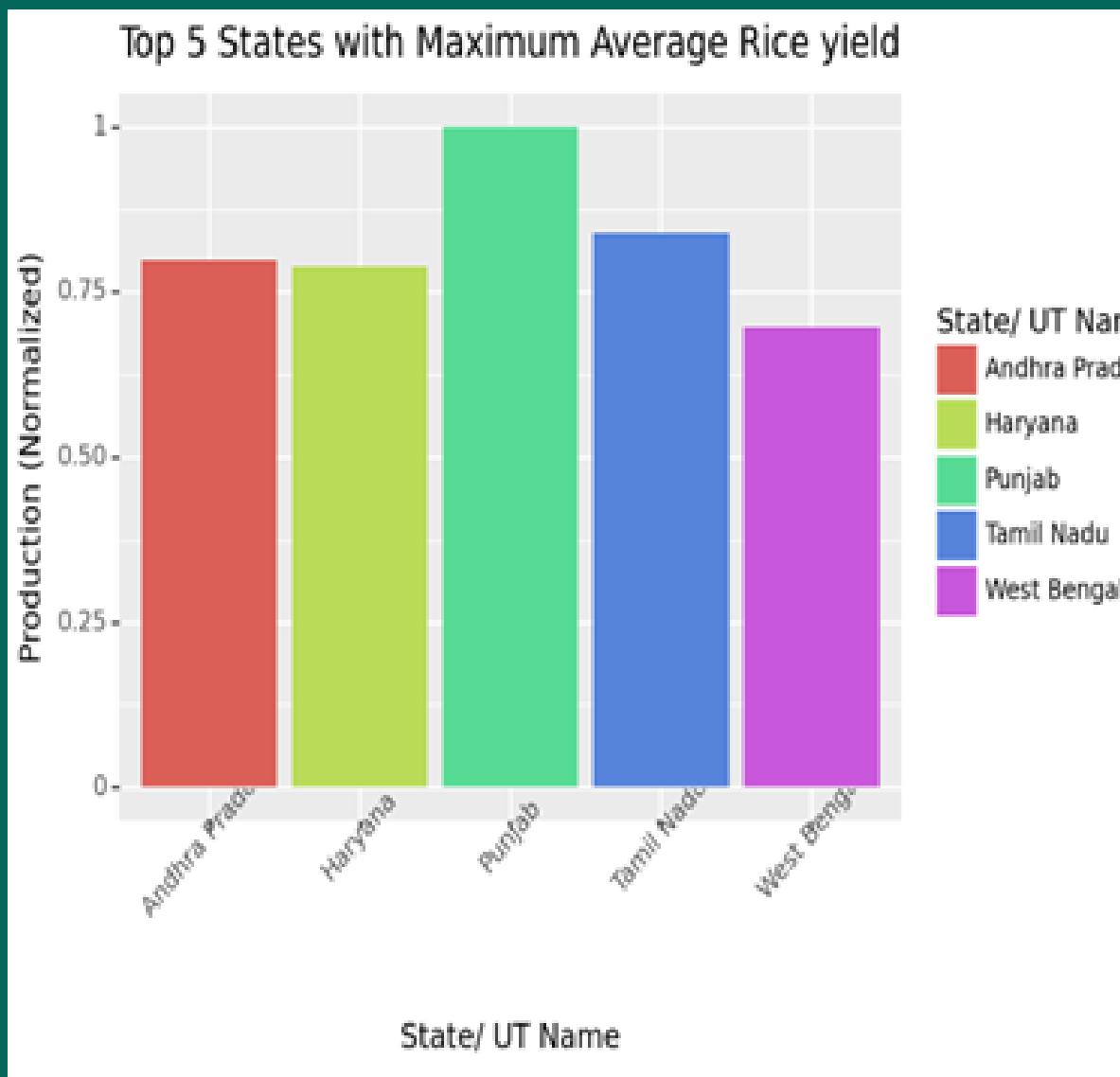
- Sugarcane thrives in tropical and subtropical climates with warm temperatures (20°C to 30°C), abundant sunlight, and adequate moisture. Thus, high production in Maharashtra and Uttar Pradesh.
- Rice is predominantly cultivated in warm, humid climates with abundant rainfall and relatively high temperatures. Therefore, places with a lot of rainfall like West Bengal have high cultivation. Other states like Uttar Pradesh and Punjab are also major producers but they depend more on irrigation for water.
- Wheat is primarily grown in temperate climates with distinct seasons and has a broad temperature range for growth, with optimal temperatures for germination and vegetative growth between 15°C to 24°C. Therefore it is more produced in northern India where the temperature variations are more extremes.



Benefits of Choropleth map:

- **Spatial Analysis:** The map allows for visualizing spatial patterns of crop production across different regions of India, providing insights into geographical variations in production levels.
- **Interactive Exploration:** By selecting different crops from the dropdown, users can dynamically explore and compare the average production of rice, wheat, and sugarcane, facilitating comparative analysis.
- **Data Interpretation:** The choropleth representation makes it easy to interpret the magnitude of production in each region, helping users identify areas of high and low production for the selected crop.
- **User Engagement:** The interactive nature of the map enhances user engagement by allowing users to interactively explore the data and gain deeper insights into crop production patterns.

Top 5 states with maximum average yield or production for a particular Crop.



Observations and Remarks:

We can see that the yield of Sugarcane is quite comparable for many states.

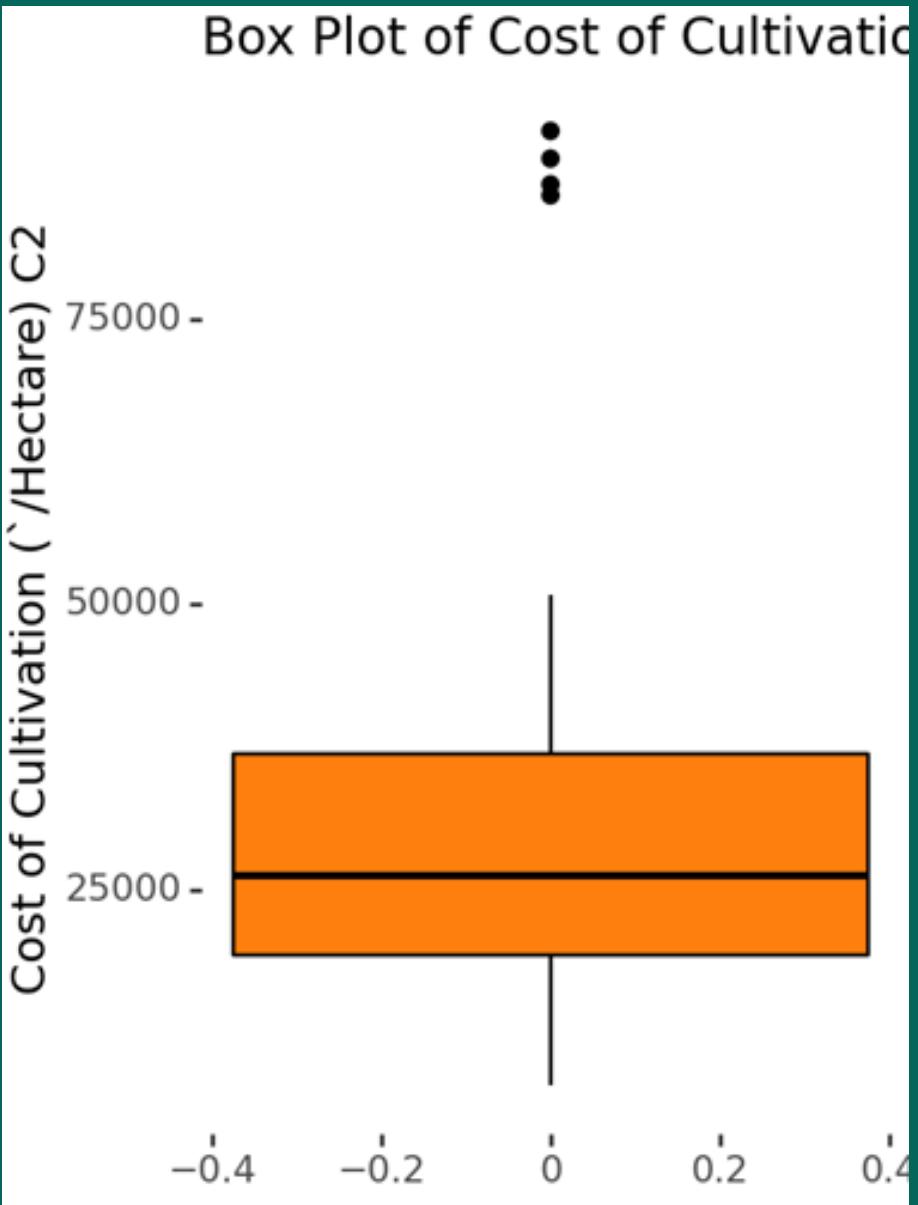
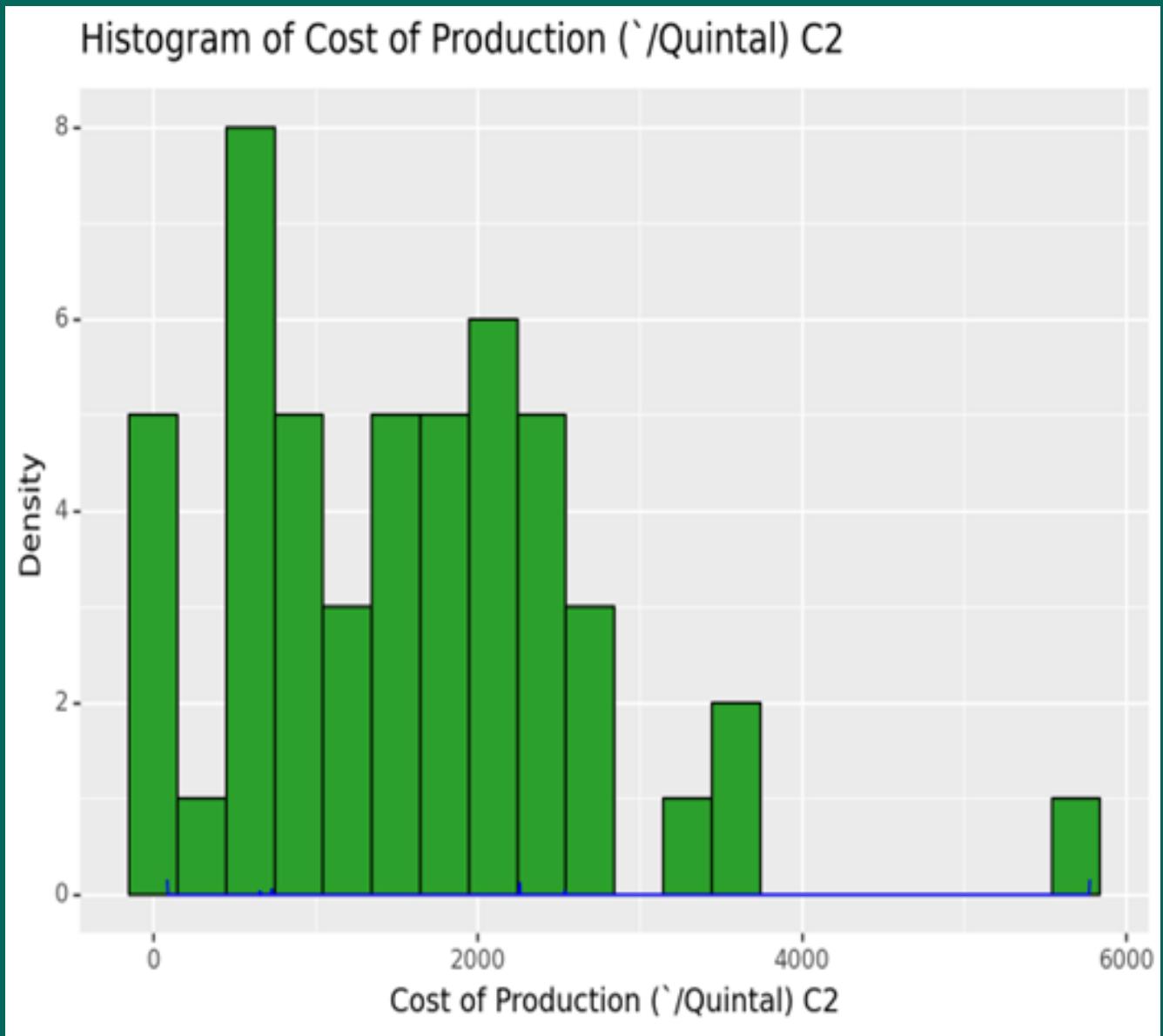
Even though Kerala, Puducherry and West Bengal are not in top 5 producers, they have a very high yield. Few reasons can be :

- They have favorable agro-climatic conditions that are conducive to high-yield.
- These states cultivate a diverse range of crops suited to their respective agro-climatic conditions and Diversification of crops allows farmers to optimize land use and maximize yields.
- Intensive Farming Practices: Farmers in Kerala, Puducherry, and West Bengal employ intensive farming practices such as multiple cropping, intercropping, and integrated pest management. These practices help maximize land productivity and minimize crop losses, leading to higher overall yields
- Smallholder farming is prevalent in Kerala, Puducherry, and West Bengal, where many farmers cultivate crops on relatively small landholdings. Smallholder farmers often have intimate knowledge of local agro-ecosystems and employ traditional knowledge and practices that contribute to sustainable and high-yield agriculture.

I have here used normalized columns as I only needed their relative values and not absolute values of their production of yield

Cost of Production of various crops

Histogram and Box plot of various types of losses



Observations and Remarks:

The above two plots show the histogram of various types of costs encountered. It can be seen that even though majorly the different types of cost lie in a particular bracket it is right skewed. that is there are outliers on the higher value side.

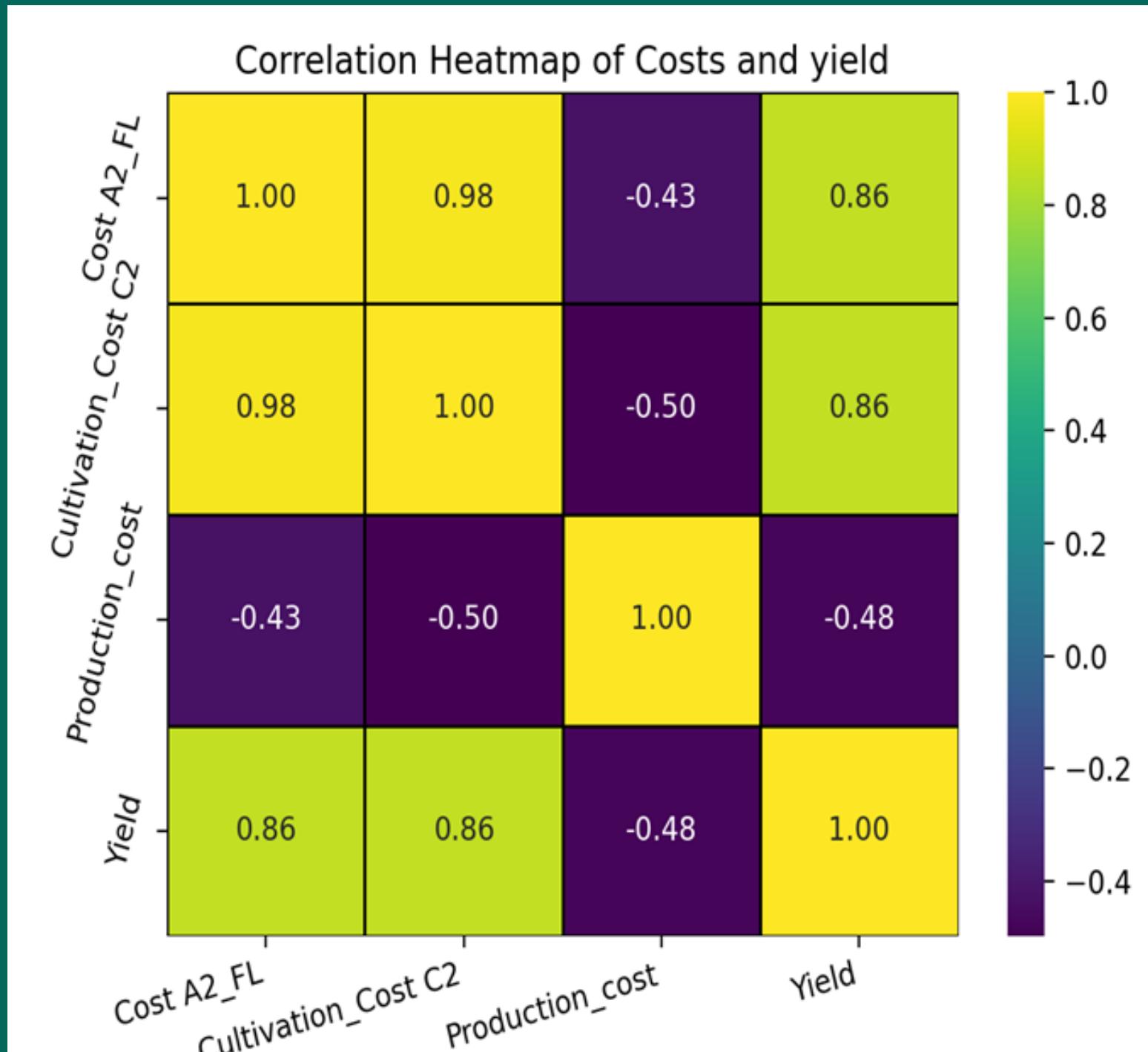
I can see median of each type of losses using the box plot. Median of cost of cultivation (C2) is highest followed by Cost of cultivation (AF_FL) and cost of production which are nearly equal

I tried to plot a density curve showing the distribution of the histogram but I was just getting a horizontal line.

Benefit of Histogram and Bar plot:

- It provide a visual representation of the distribution of data by binning it into intervals and showing the frequency of observations within each bin.
- It helps me to understand the spread and shape of the cost data, allowing me to observe any skewness or outliers in the distribution.
- I tried to plot a density curve which would have helped me see the shape of distribution better. (Benefit of Density Plot)
- The box plot helps to visualize the outliers better also proving with summary like median. Here we can see the median of different types of cost.

Correlation Heatmap for various losses and yield



Benefits of HeatMap

1. It was a good way to see correlation between features.
2. Color Encoding: Heatmaps use color gradients to encode values, making it easy to visualize the strength and direction of correlations.

Observations and Remarks:

We can see that Yield has a negative correlation with production cost but positive correlation with cost of cultivation.

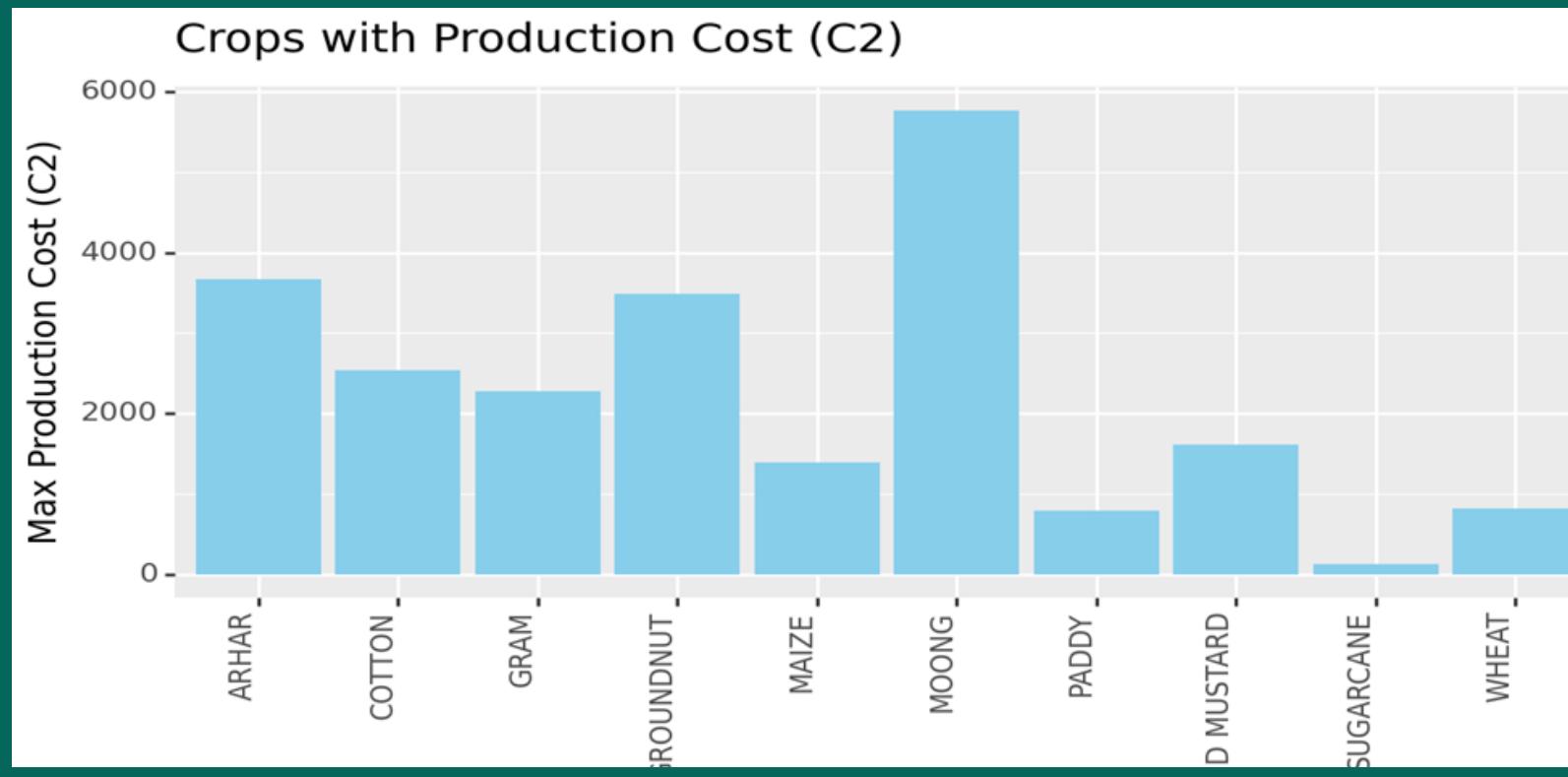
The positive correlation with cultivation cost can be due to :

- Achieving higher yields often requires more intensive cultivation practices, which can drive up the cost of cultivation.
- To obtain higher yields, farmers may invest in premium inputs, such as high-quality seeds, specialized fertilizers, and advanced irrigation systems. While these inputs contribute to increased production, they also add to the overall cost of cultivation.

The negative correlation with production cost can be due to :

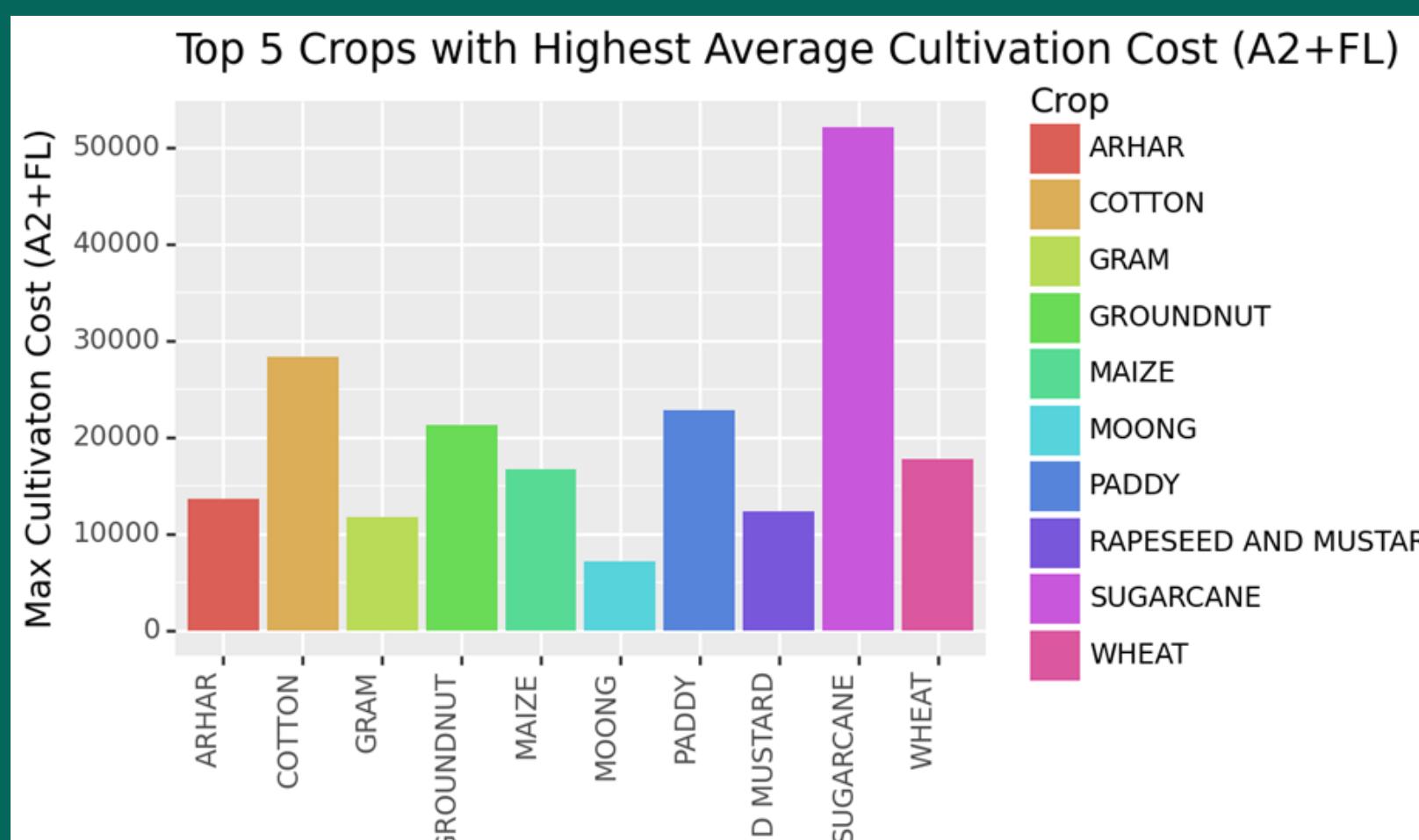
- When farmers are able to produce more crops (higher yield) without significantly increasing their total costs, it means they're becoming more efficient in their operations. This efficiency might come from using resources more wisely or adopting better farming techniques.
- When farmers produce more crops, they can spread their fixed costs (like machinery and land) over a larger output. This means the cost per unit of crop goes down as yield increases.

Bar plot of crops with different types of cost



Observations and Remarks:

We can see that Sugarcane has very low Production Cost and earlier we say that it had the highest yield. Therefore, it shows that a positive correlation between yield and production cost exists.



We can again see that cultivation cost of sugarcaae is very high, showing the positive correlation with yield.

State Level : Case study_3 states :

Uttar Pradesh, Punjab and Maharashtra

Land Distribution in these states using an interactive donut chart

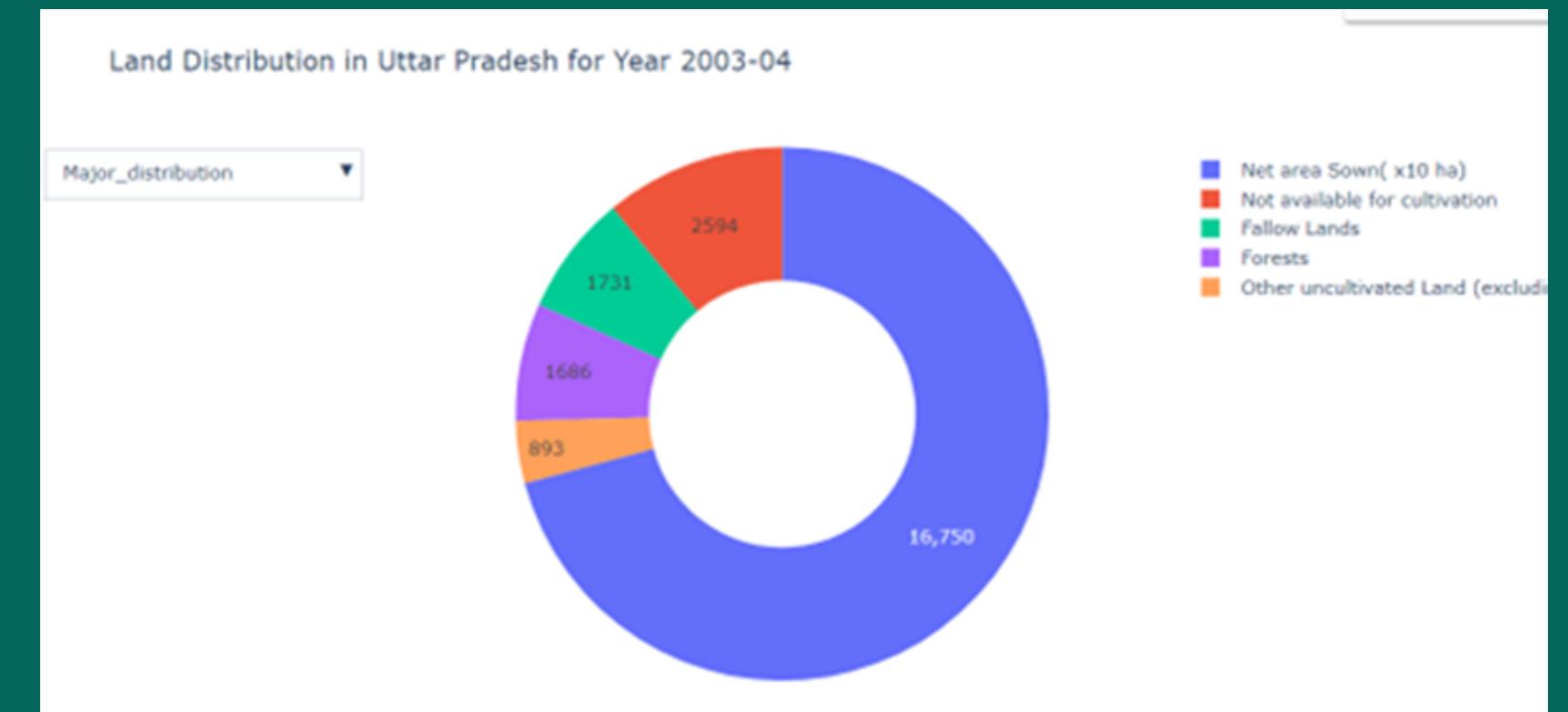
Observations and Remarks:

It can be seen that over 10 years from 2003 to 2013 , there was a decline in forest and net sown area and an increase in barren land in all the 3 states.

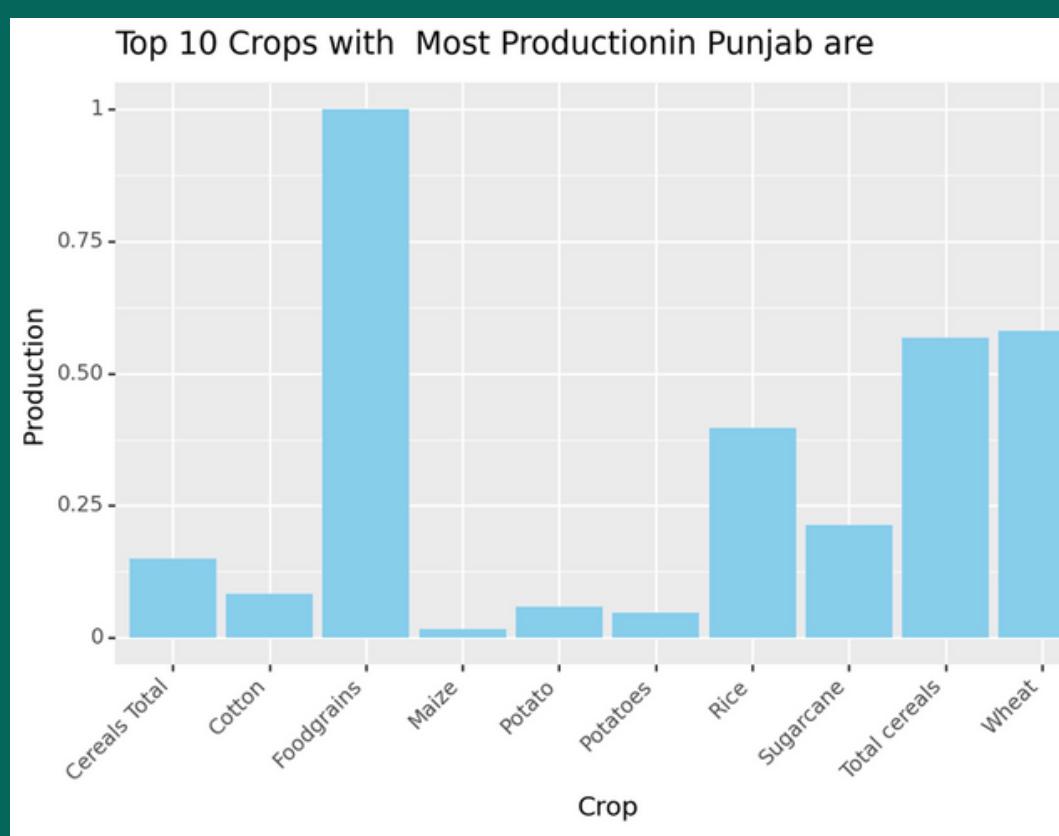
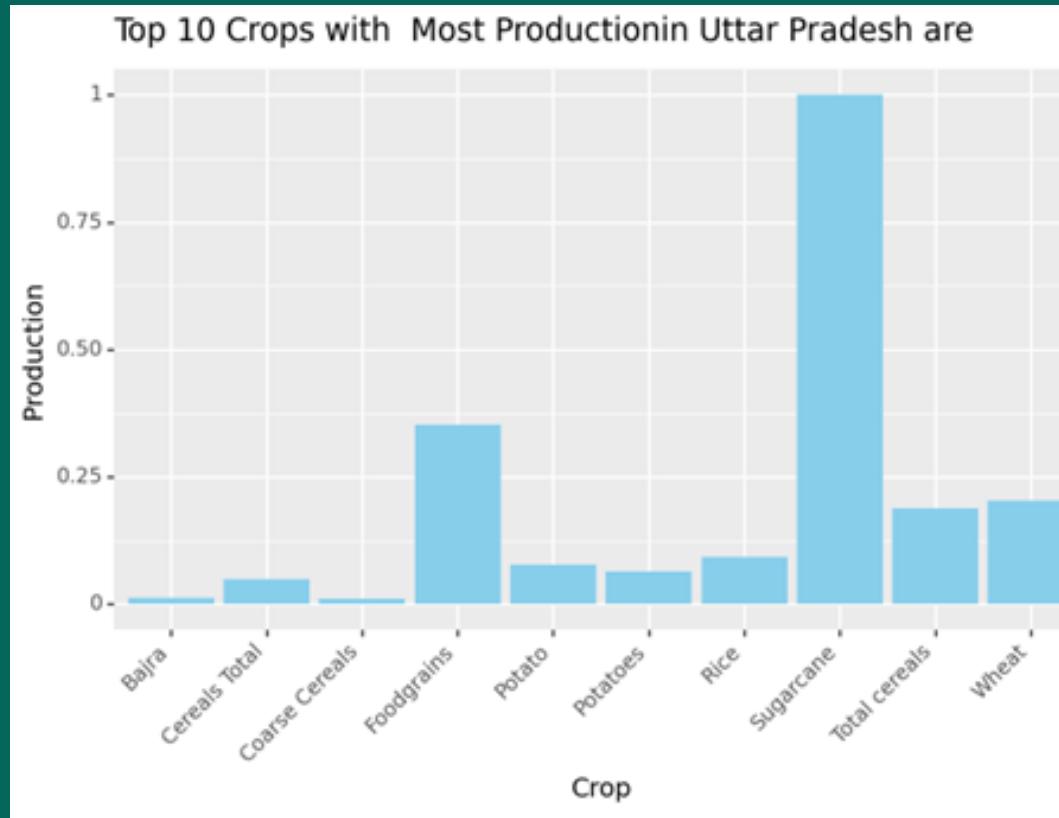
*It shows that there is a need to **promote sustainable agriculture**, and **land use planning** while also prioritizing towards **afforestation and reforestation**.*

More than 80% of land in Punjab is used for agriculture.

Overall, the forest area is highest in Maharashtra followed by Uttar Pradesh and then Punjab

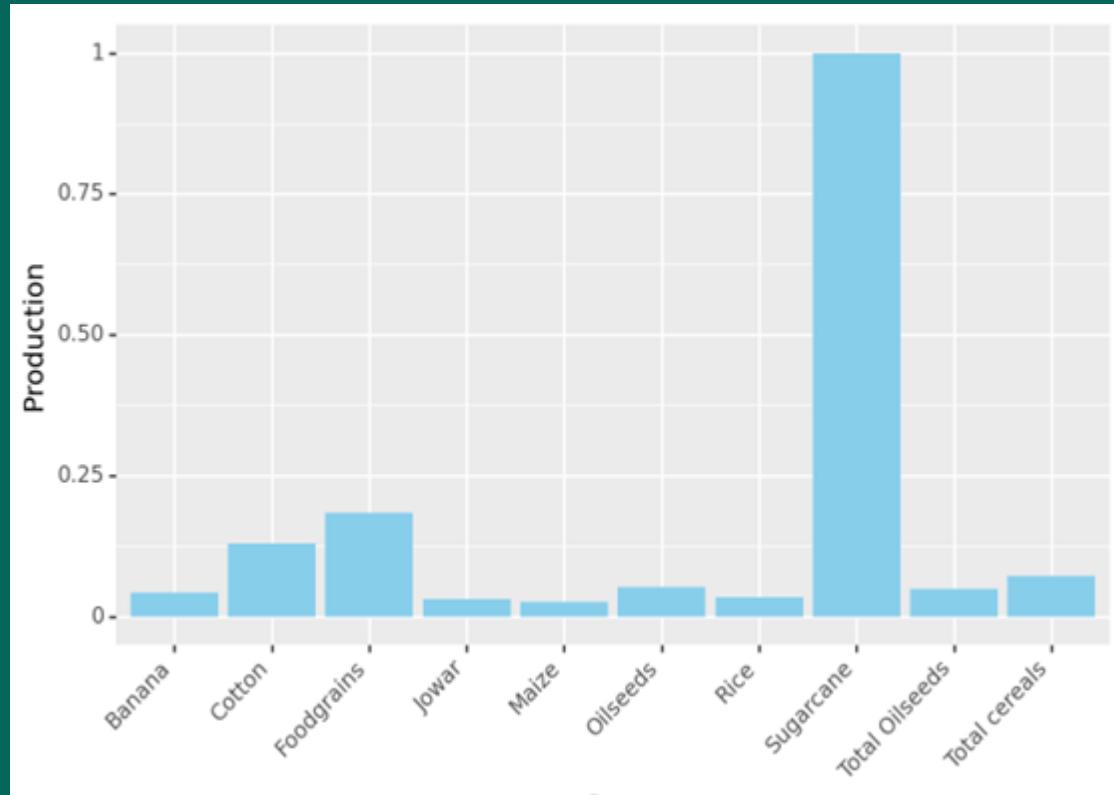


Top 10 crops that major productivity and yield in these states



Observations and Remarks:

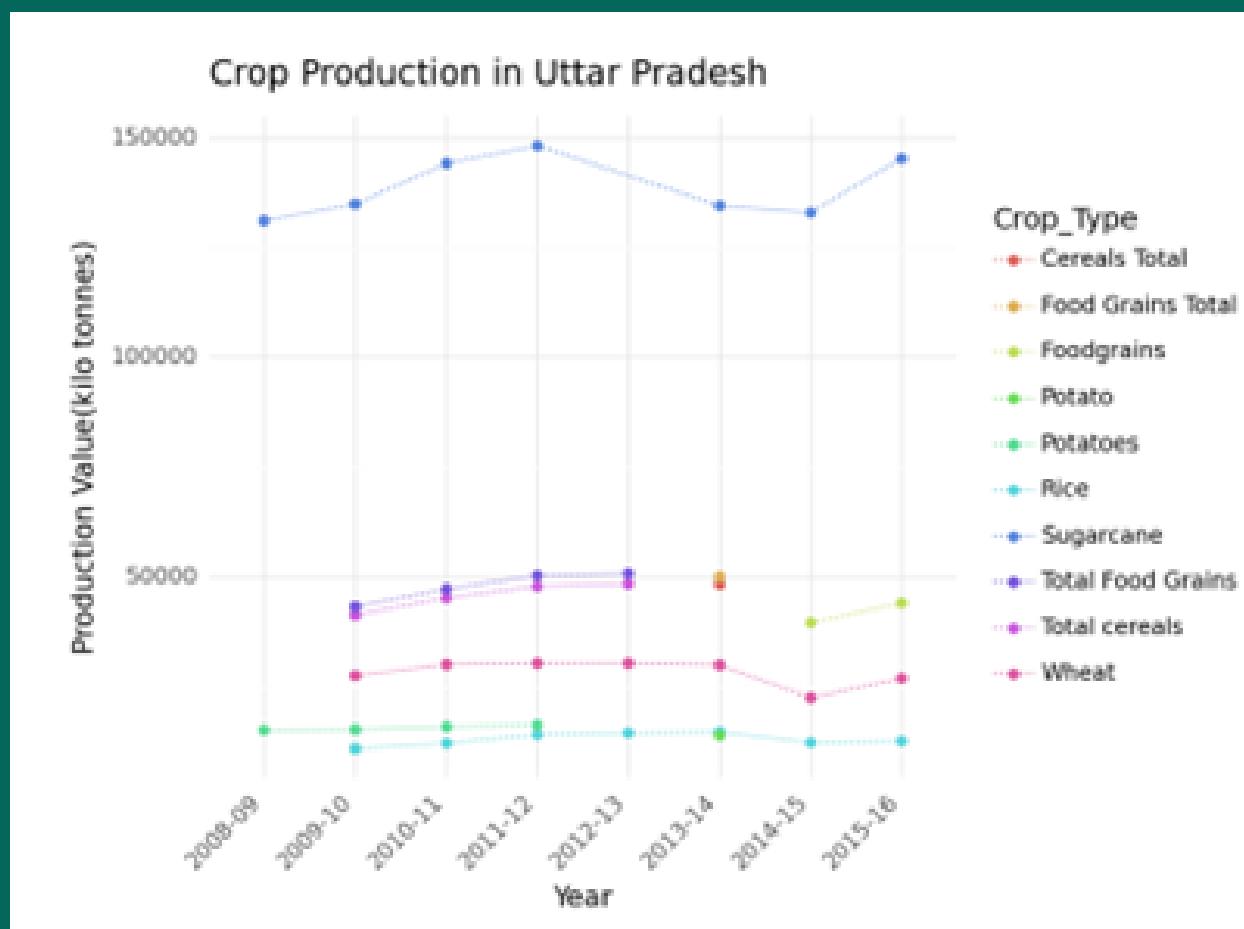
- Sugarcane stands out as the primary crop in Uttar Pradesh, overshadowing the production levels of all other crops, which are notably lower in comparison.
- It also shows that farming here is more focused on one crop.
- All these graphs show that sugarcane have very high yield irrespective of place.
- Maharashtra also focuses on one crop.
- Punjab focuses more on multiple crops harvesting. (As Foodgrains also contain many different types of grains).



I have used normalisation because :

- Without normalization, crops with higher absolute production may visually dominate the graph, making it challenging to discern the relative importance of other crops.
- Highlighting relative contributions: It emphasizes each crop's proportionate contribution to total production, aiding in understanding their importance.

Crop production trend in these states over the past few years



Observations and Remarks:

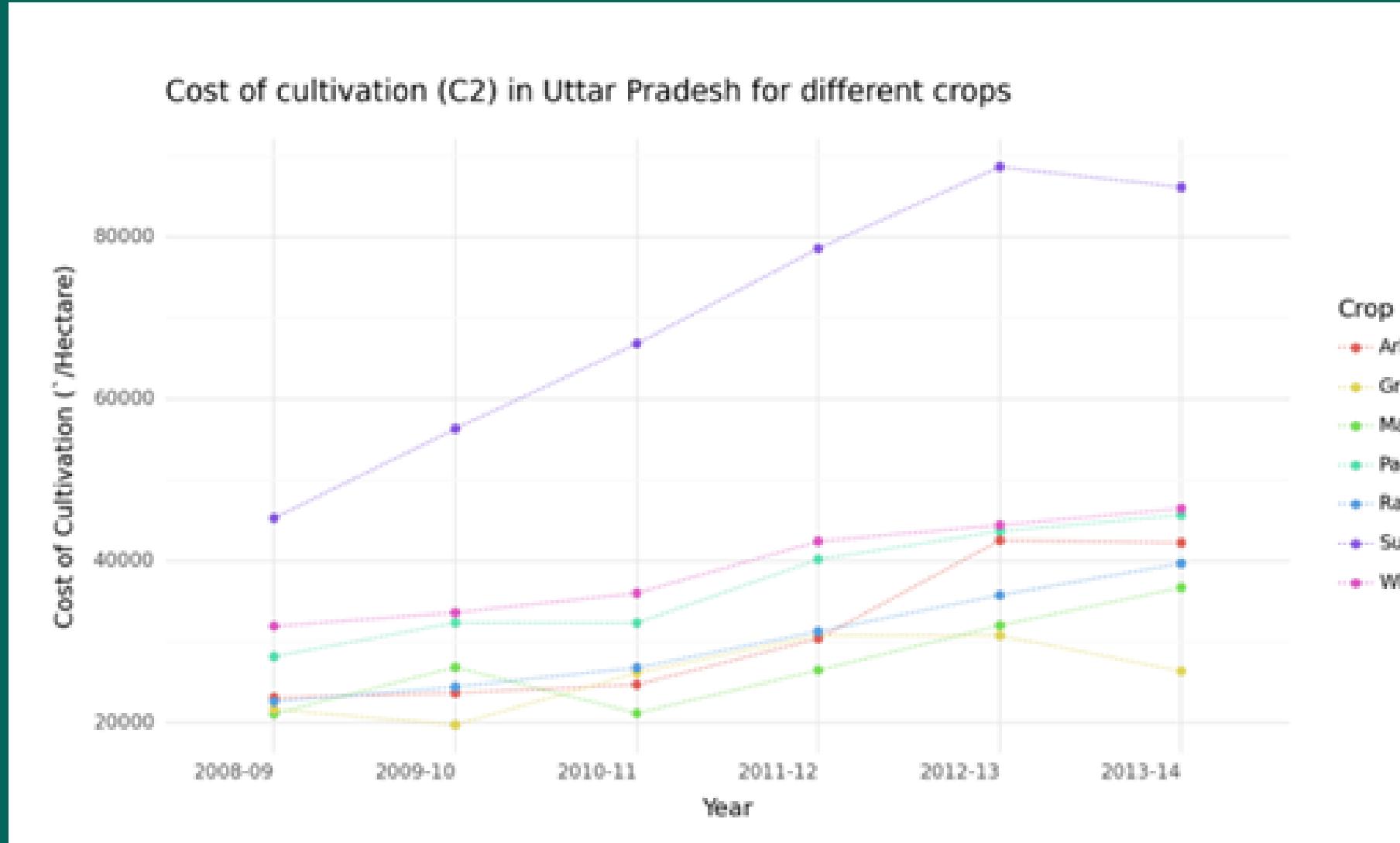
It can be seen that the relative production of crops did not change much in comparison to each other.

I have used line plot for these situation because:

Temporal Trends: Line plots are excellent for displaying trends over time.

Connection of Data Points: it is easier to see trends and patterns in the data compared to scattered points alone.

Observations and Remarks:



In general, the cost(all types) is increasing over the years for all the crops

But the cost of cultivation increased very rapidly for Sugarcane between 2008 to 2012 but then started decreasing again.

The reason could be adoption of some new technology or machinaries.

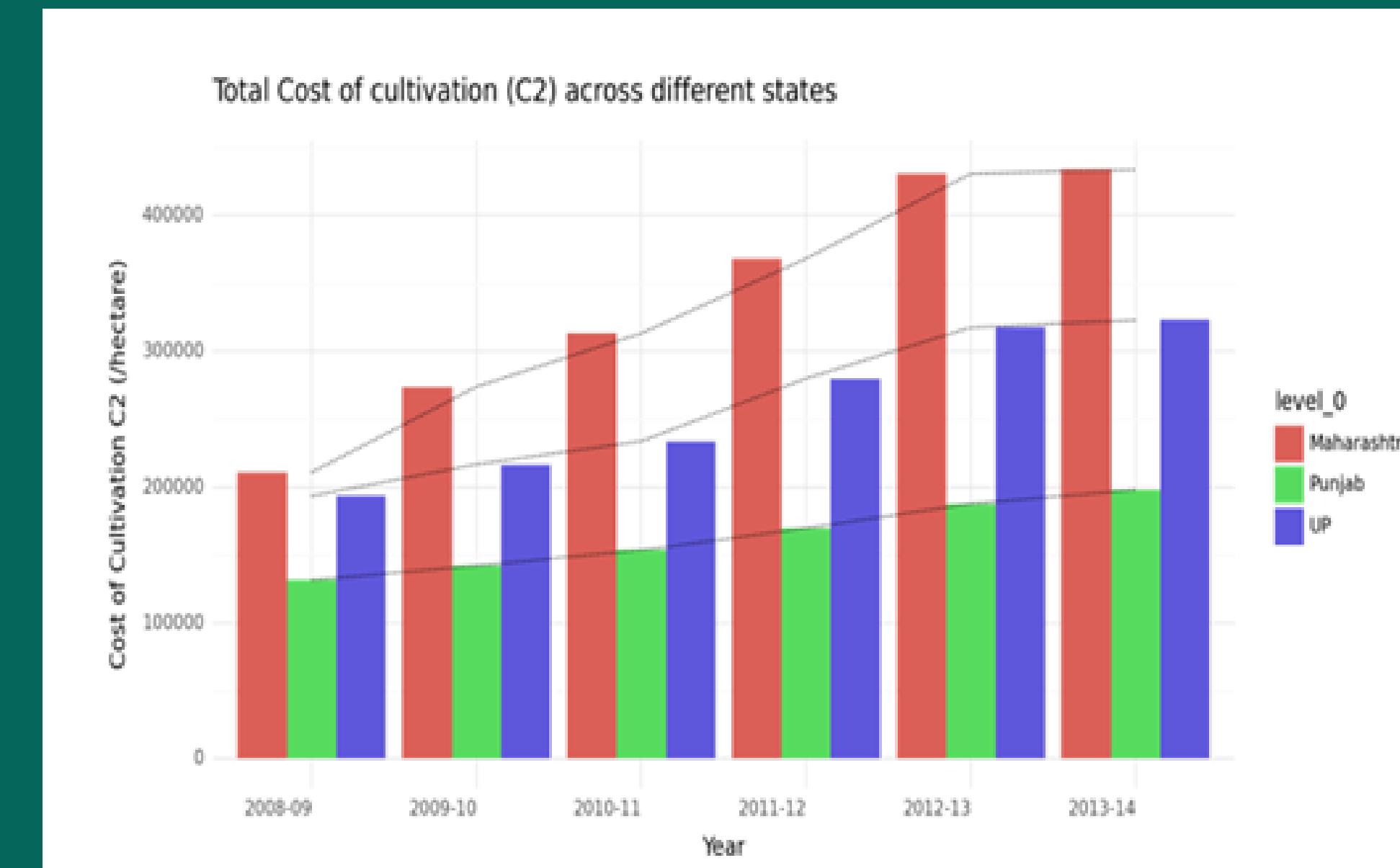
Bar and Line Chart for comparing total cost (for all the types) in these 3 states

Observations and Remarks:

- It can be seen that for all the 3 cases, Maharashtra has the highest cost followed by Uttar Pradesh followed by Punjab.
- Also, these costs for all the 3 states are increasing over time

Motivation Behind the Plot

- The bar chart allows for easy comparison of total costs between different states for each year.
- The line chart provides a trend visualization for one specific state.



Therefore, we can compare the total costs across states while also seeing the trend for one specific state

CONCLUSION POINTS

Land Distribution and Utilization:

- Forest area in India has increased in some places while decreased over other places through the years, indicating positive trends in environmental conservation and people becoming aware and taking necessary steps.
- There is a need for sustainable land management practices to optimize agricultural productivity and minimize environmental degradation.

State-Level Analysis:

- Uttar Pradesh, Maharashtra, and Punjab exhibit diverse agricultural landscapes, each with its own set of challenges and opportunities.
- Sugarcane emerges as a prominent crop in Uttar Pradesh, underscoring its importance in the state's agricultural economy.
- Also, states which do multiple crop agriculture have better yield

Crop Production and Yield:

- Sugarcane emerges as a dominant crop in terms of both production and yield, highlighting its economic significance.
- Regional variations in crop cultivation suggest the importance of tailored interventions to optimize yields and address local agricultural challenges

Cost of Production:

Costs associated with crop cultivation have been increasing over time, emphasizing the financial pressures faced by farmers.

THANKYOU