Visualization: Project Report

Name: Anusha R Roll Number: MDS202212

Deadline : 30 Oct 2022

Introduction

The dataset provides information on state-wise crop production in India ranging from 1997 to 2015. We analyze the production of different crops across states in different agricultural season of a year. Various inferences are drawn and the results are visualized using R.

Dataset Description

The dataset is downloaded from Kaggle: https://www.kaggle.com/abhinand05/crop-production-in-india (https://www.kaggle.com/abhinand05/crop-production-in-india)

Show 10 **entries** **Search:**

	State_Name	District_Name	Crop_Year	Season	Crop	Area	P
1	Assam	BAKSA	2005	Rabi	Linseed	728	
2	Goa	SOUTH GOA	2014	Whole Year	Banana	1076	
3	Meghalaya	EAST GARO HILLS	1997	Kharif	Cotton(lint)	2720	
4	Karnataka	BAGALKOT	1998	Kharif	Bajra	48855	
5	Maharashtra	AHMEDNAGAR	1997	Kharif	Arhar/Tur	17600	
6	Odisha	ANUGUL	1997	Autumn	Groundnut	4086	

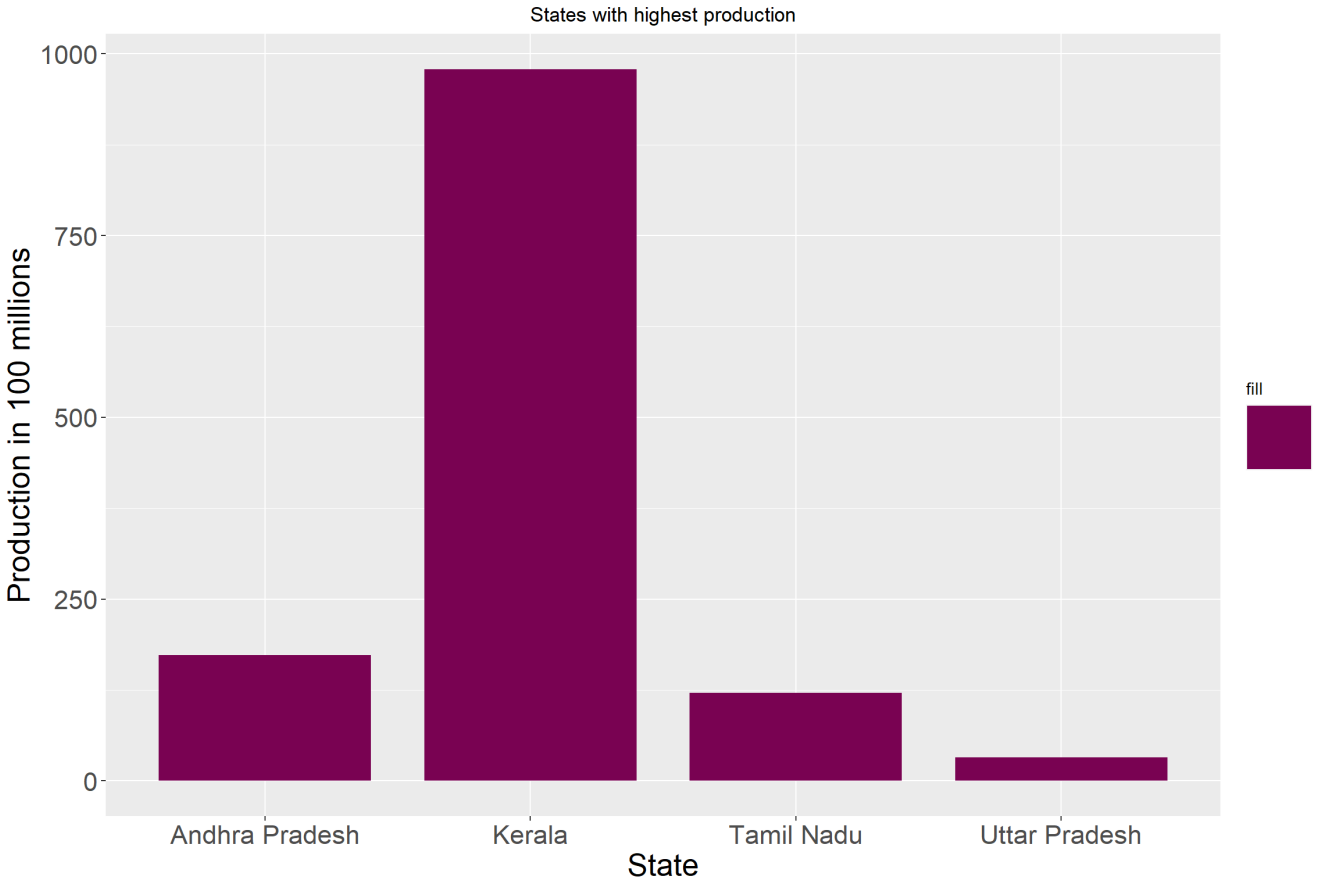
Showing 1 to 6 of 6 entries Previous 1 Next

The data set contains 246091 rows and 7 columns.

Columns with categorical values are ‘State_Name’, ‘District_Name’, ‘Crop_Year’, ‘Season’ and ‘Crop’. ‘Area’ and ‘Production’ are columns with numerical values, with hectares and tonnes respectively. The total number of rows with atleast a null value in any column is 3730. 3730 rows out of 246091 rows is relatively small and can be ignored. Hence these rows are dropped and the new dataframe with 242361 rows and 7 columns obtained is used for further analysis and visualization.

Graphical Presentation of Key Variables

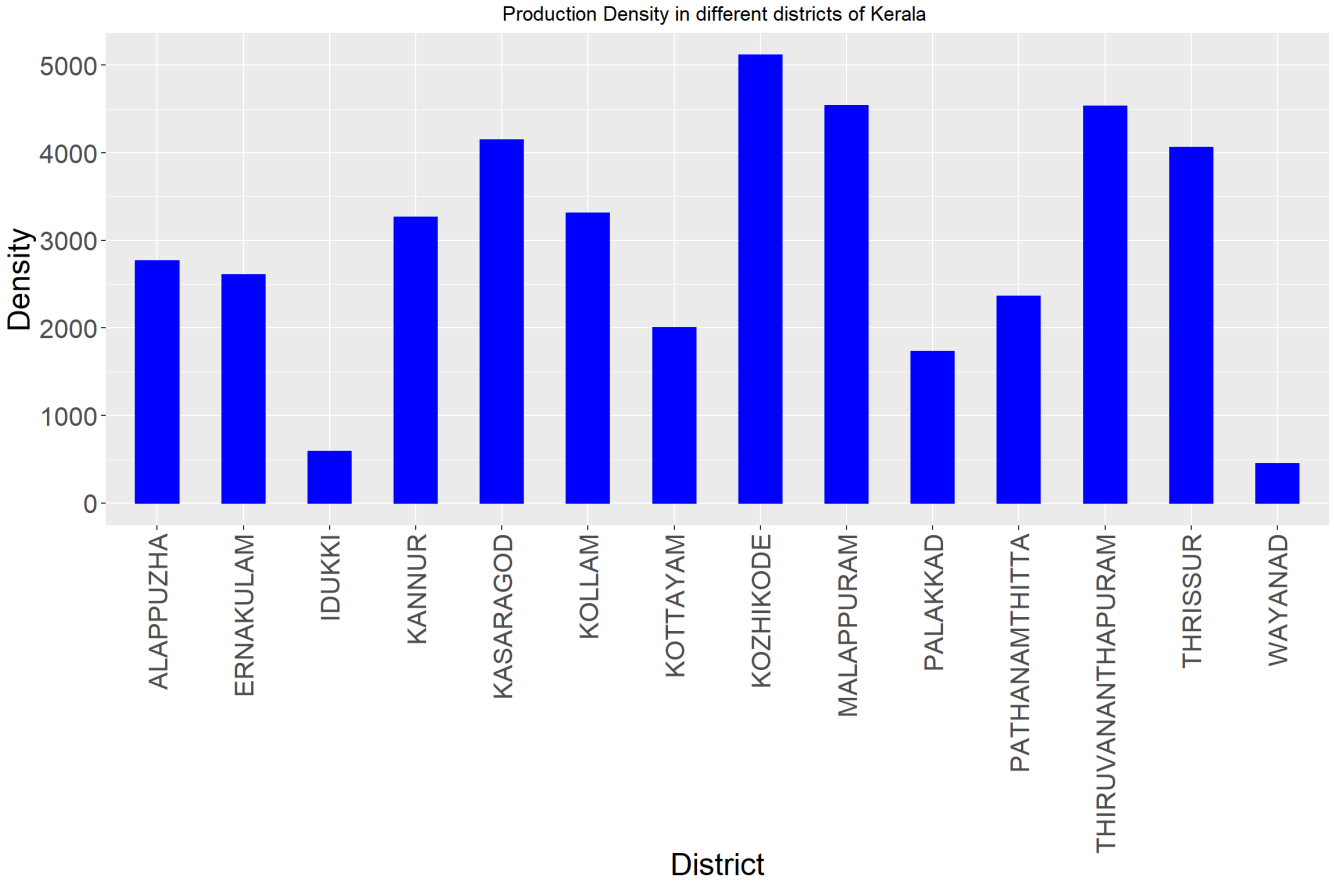
The table is grouped by the name of the state, total production of each state is calculated and top 4 states in terms of production are obtained.

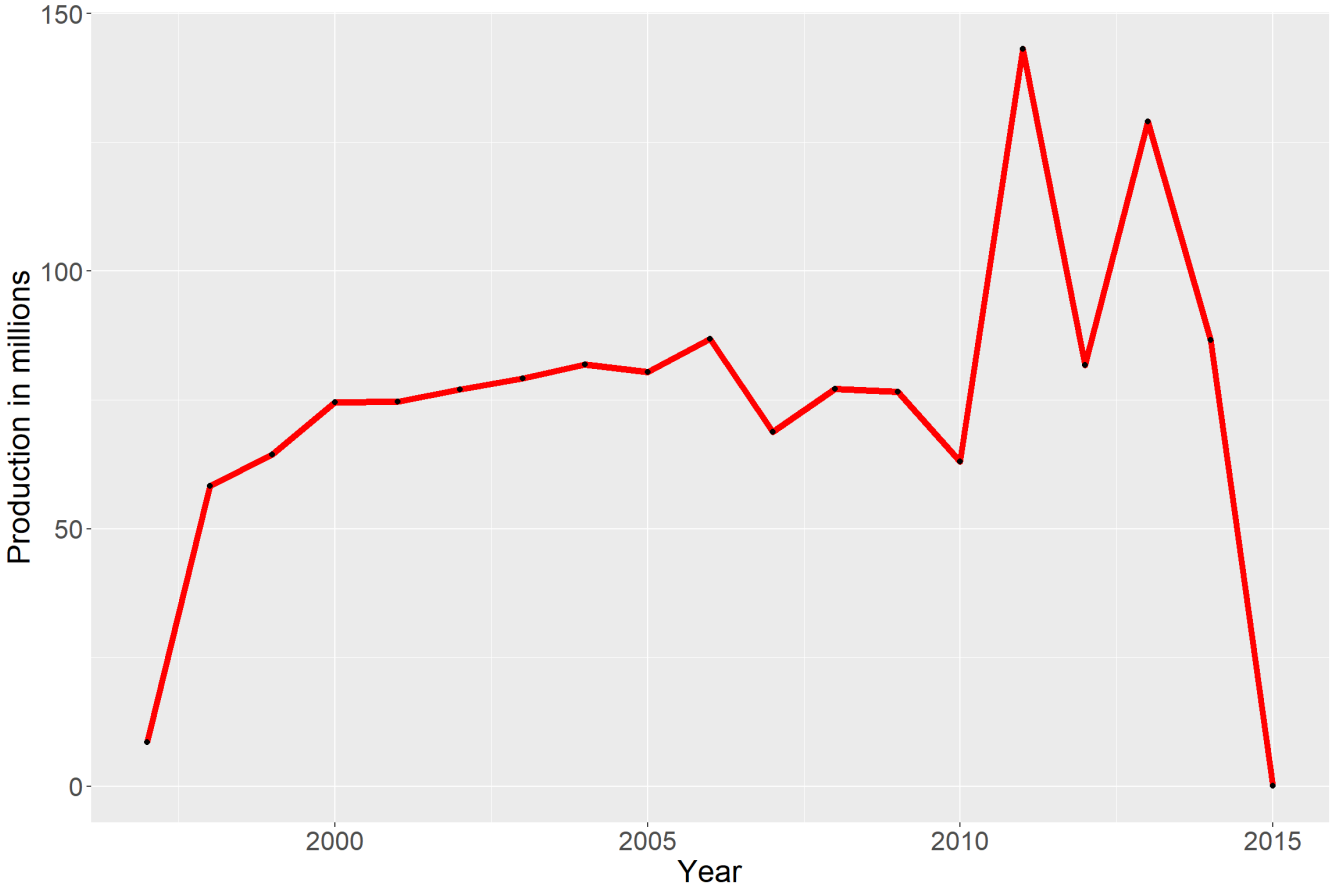


The above bar graph displays the production of top 4 states in terms of their total production. It is evident from the graph that Kerala has the highest production of total crops.

```
## # A tibble: 14 × 4
##    District_Name          tot_prod tot_area Density
##    <chr>                     <dbl>    <dbl>   <dbl>
##  1 ALAPPUZHA           4800461927. 1733208.   2770.
##  2 ERNAKULAM           5021649326. 1923972.   2610.
##  3 IDUKKI              1521358947. 2564981.    593.
##  4 KANNUR              9783432072. 3000090.   3261.
##  5 KASARAGOD           7732216929. 1865369.   4145.
##  6 KOLLAM              7151945342. 2160104.   3311.
##  7 KOTTAYAM            3025203754. 1508852.   2005.
##  8 KOZHIKODE          15280739863. 2988171.   5114.
##  9 MALAPPURAM         14518402342. 3200753.   4536.
## 10 PALAKKAD            6369381579. 3677293.   1732.
## 11 PATHANAMTHITTA      1969921120.  834913.   2359.
## 12 THIRUVANANTHAPURAM 10022714452. 2211755.   4532.
## 13 THRISSUR            9923507680. 2442724.   4062.
## 14 WAYANAD              759110042. 1690068.    449.
```

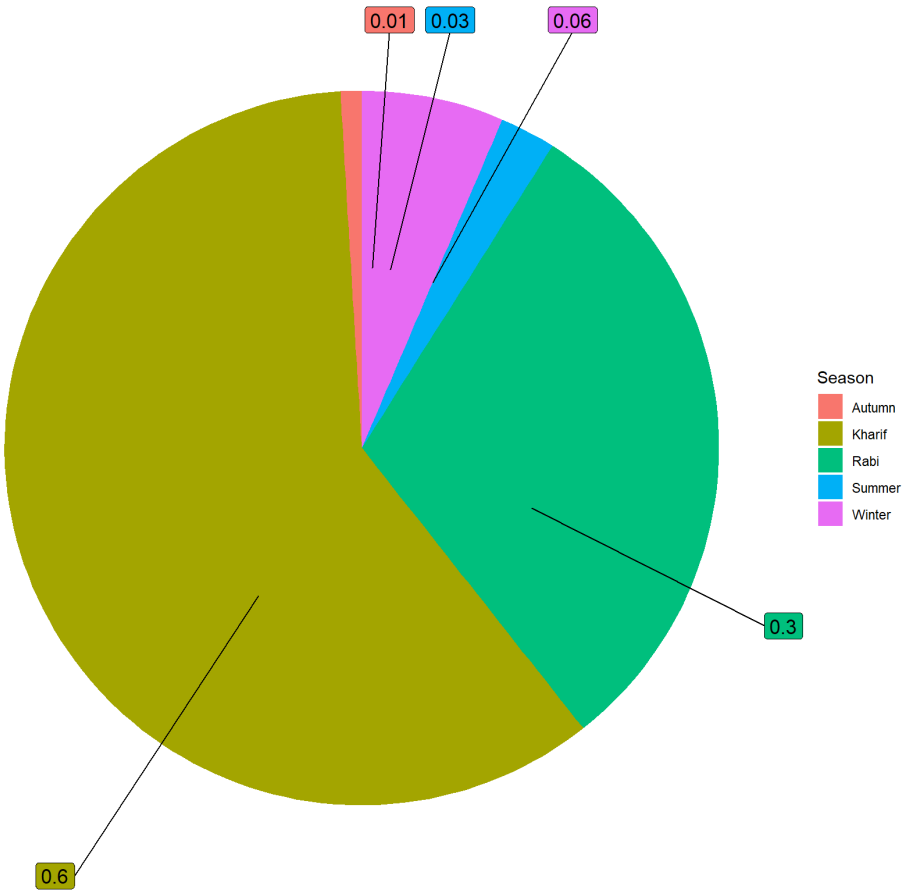
The above table gives the total production and total area of land for agriculture in different districts of Kerala, from which production density (per unit area) is calculated and a bar graph is plotted.





The above graph is a time series plot, which displays the total production of crops in each year from 1997 to 2015.

The proportion of production of seasonal crops in each season is calculated and pie chart is plotted.



The above pie chart shows the proportion of production of crops across different seasons of a year.

Summary

- When state-wise total production across years is plotted, it is seen that Kerala is the state with maximum crop production, followed by Andhra Pradesh, Tamil Nadu and Uttar Pradesh.
- The production density of different districts in a state is plotted as a bar graph.Here, we see that Kozhikode is the district in Kerala with highest crop production density.
- It is visible from the time series plot that highest total production is obtained in the year 2011 and the prodution has been the lowest in the year 2015.
- 60% of production of crops happens in Kharif season, followed by Rabi season (30%).

Conclusion

We analyzed the production of crop varieties in different states over several years and how it varies in each agricultural season. Crop-wise analysis will be made in the interactive dashboard and we can see how the production quantity of each crop varies across seasons depending on each geographical location. State-wise production of each crop in every year can be analyzed to predict its production in future.