

EMPOWERING SUSTAINABLE AGRICULTURAL GROWTH

EXPLORATORY ANALYSIS ON INDIAN CROP PRODUCTION DATA



Indian Agriculture

- ❑ Agriculture is the backbone of the Indian economy, providing employment to millions of people and contributing significantly to the country's GDP.
- ❑ Sustainable agriculture, in terms of food security, rural employment, and environmentally sustainable technologies such as soil conservation, sustainable natural resource management and biodiversity protection, are essential for holistic rural development.
- ❑ However, the sector is faced with several challenges, including climate change, low productivity, and food security issues. To address these challenges, there is a need for data-driven solutions that can inform policy and decision-making.



About The Dataset

- ❑ This dataset contains comprehensive information on agricultural production statistics in India, sourced from the Indian government's Area Production Statistics (APS) database, maintained by the Ministry of Agriculture and Farmers Welfare
- ❑ The dataset covers four major crop seasons, namely kharif, rabbi, summer, and autumn, from the year 1997 to 2020.
- ❑ The data provides information on the annual production of crops, yield, and area under cultivation across different states and districts in India.

Objective

- ❑ The analysis of this dataset can provide valuable insights into crop yields, area under cultivation, and other metrics that can inform agricultural policies and practices.
- ❑ It can also be used to identify trends and patterns in agricultural production, helping farmers and policymakers make informed decisions about crop selection, irrigation, and other important factors that affect agricultural productivity.

Data Collection

- ❑ Records from 1997-2019 was downloaded as a dataset from Kaggle data community.

<https://www.kaggle.com/datasets/nikhilmahajan29/crop-production-statistics-india>

- ❑ Records of 2020 was collected from the Indian government's Area Production Statistics (APS) database.

https://aps.dac.gov.in/APY/Public_ReportI.aspx

Tools Used

- ❑ MS EXCEL : Data preparation and cleaning was performed and saved as csv file type.
- ❑ PYTHON : Exploratory Data Analysis and Visualization was performed

Overview Of Steps Perform In Python

- ❑ Importing required libraries
- ❑ Dataset loading
- ❑ Retrieval of basic information on the structure of dataset
- ❑ Data cleaning
- ❑ Exploratory Data Analysis

Libraries Imported

```
import pandas as pd
import numpy as np

import matplotlib.pyplot as plt
import seaborn as sns

import plotly.express as px
import plotly as py
import os

import warnings
warnings.filterwarnings('ignore')
```


Dataset Loading And Retrieval Of Basic Information On The Structure Of Dataset

```
df = pd.read_csv("Agri_crop_production.csv")
```

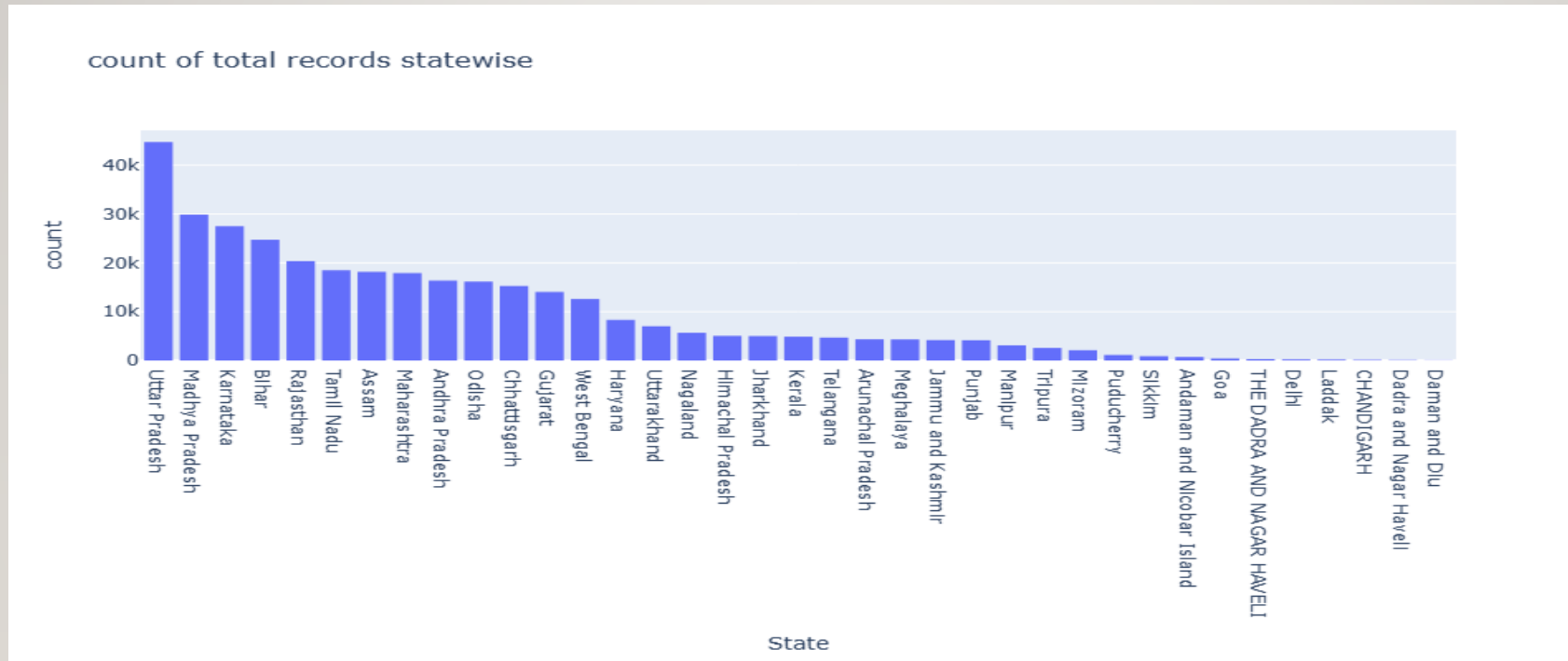
- There are a total of 345646 rows and 8 columns initially in the dataset.
- Null values are observed in crop and production attributes.
- There 3 columns(Area, production and Yield) in float datatype and 4 columns(State, District, Crop, Season) in object datatype and Crop_year in integer datatype.
- The dataset has records from the year 1997 till 2020.
- There are records collected from 37 states and 724 districts in India.
- There are 55 unique crops and 6 seasons from which Rice is the most frequent crop and Kharif is season.

Data Cleaning

As mentioned earlier, there were null values in:

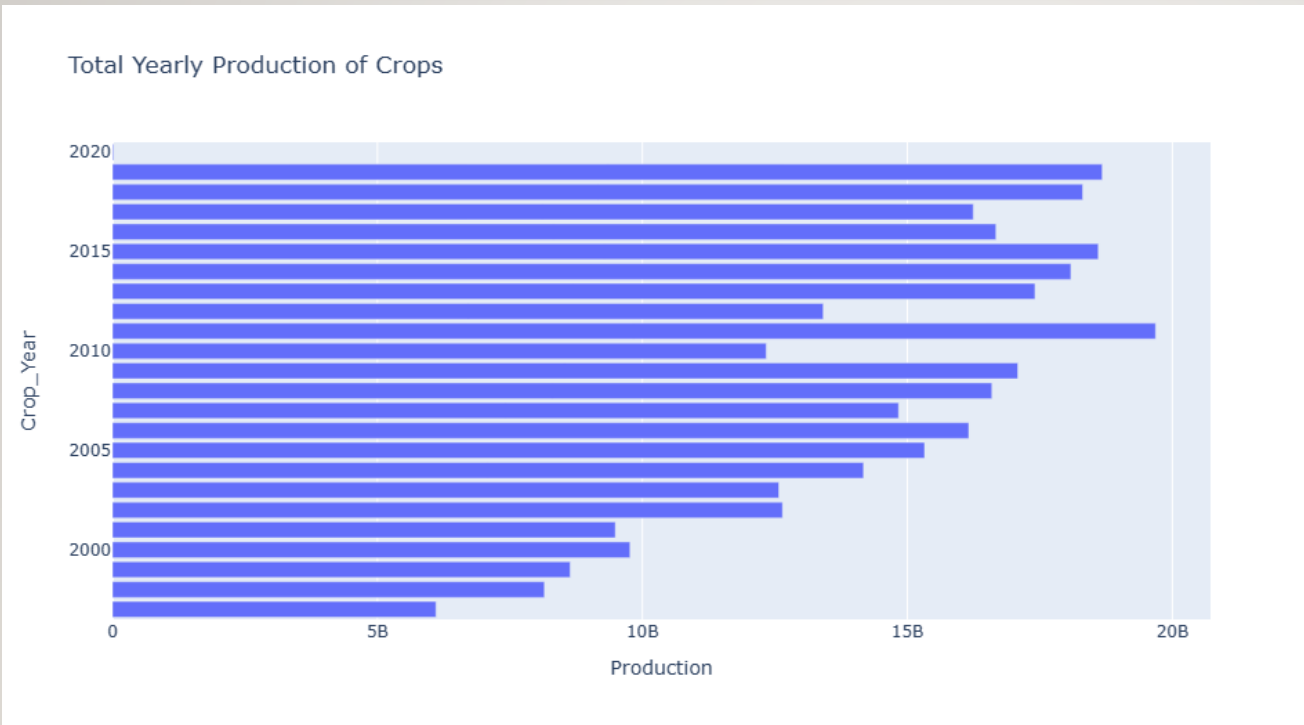
- crop (9) – dropped those 9 rows
- production (4948) – imputed with its overall mean value

The Number Of Records For Each State In The Dataset



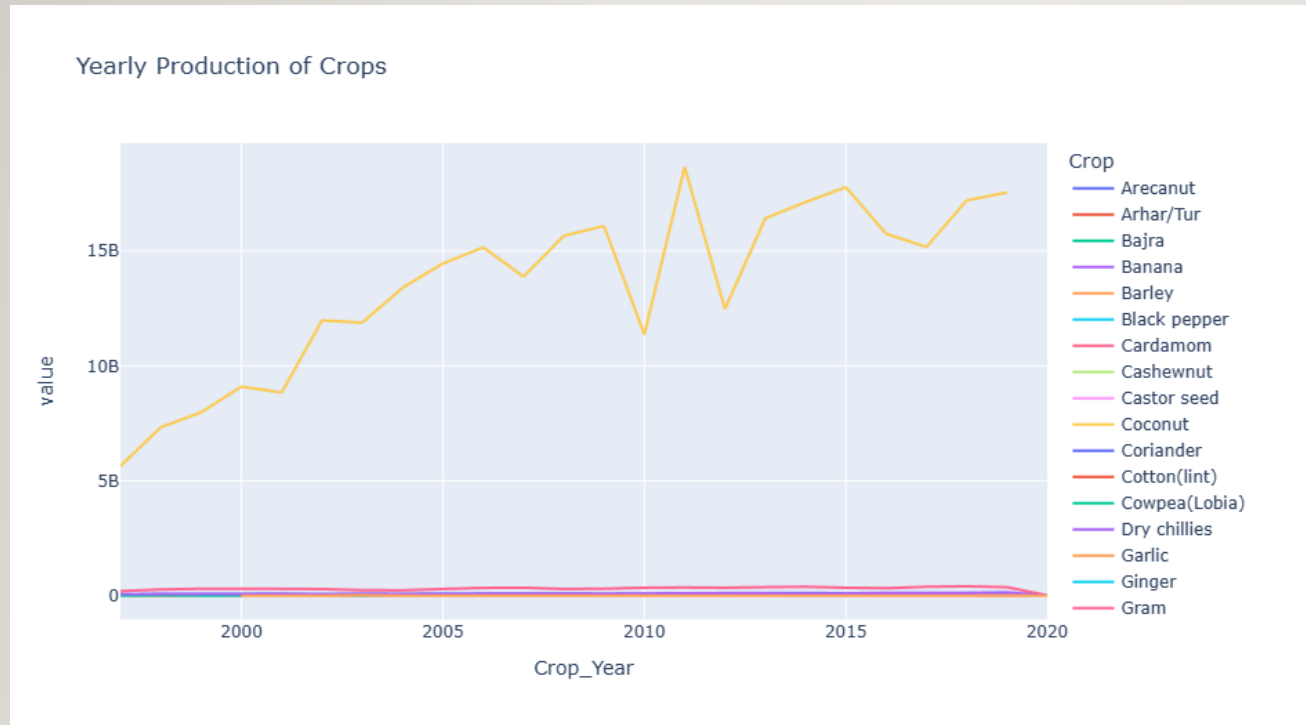
Histogram of the attribute "State" for understanding the total records in each state.

Total Production Of Crops Year-wise



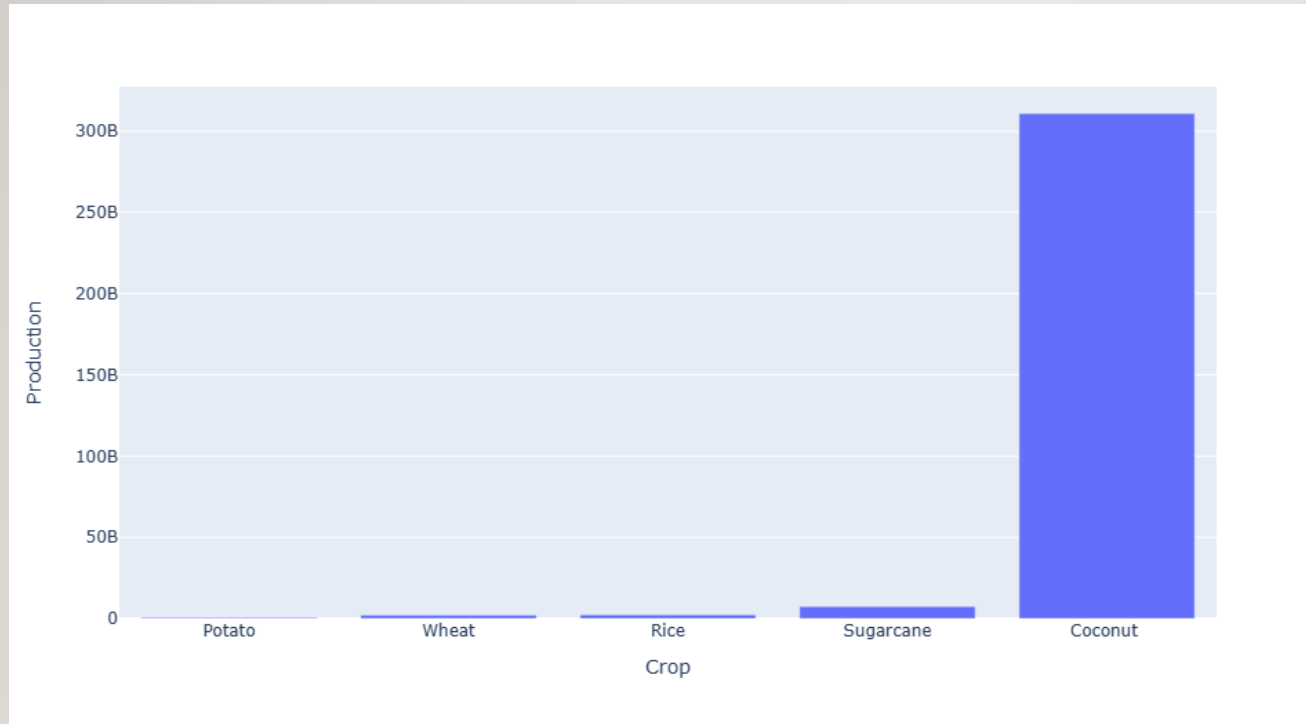
- A group by of “crop_year” and sum of ”production” was performed and viewed as a bar chart.
- Highest production was in the year 2011.

List Of Crops And Their Total Production Every Year



- A group by of “crop_year”, “crop” and sum of “production” was performed and viewed as a line chart.
- The crop coconut shows highest production over the years.

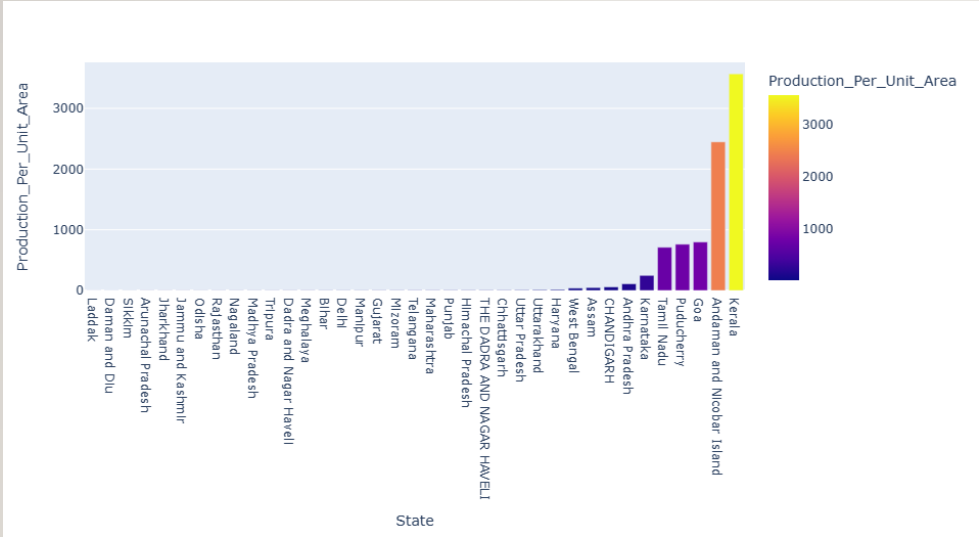
Crop Wise Total Production Data



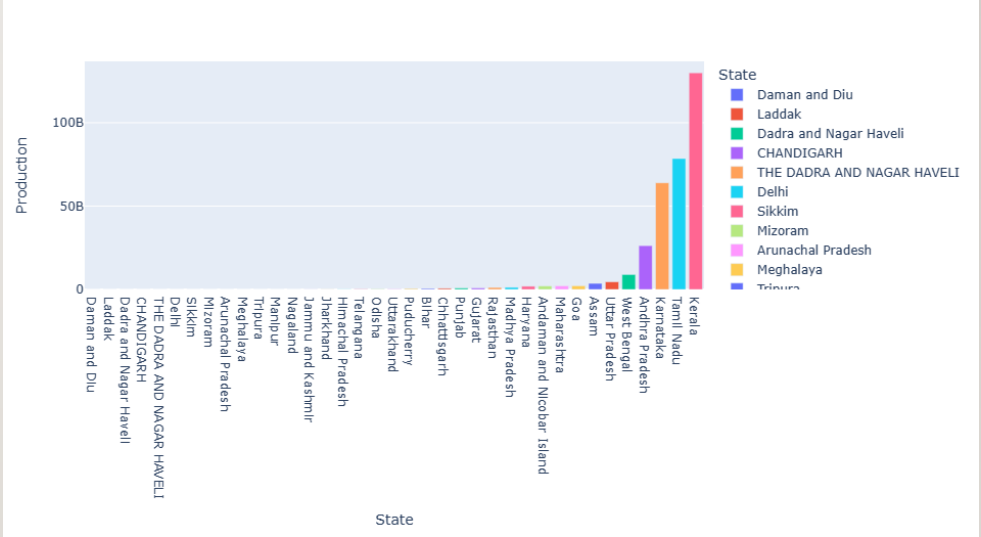
- A group by of "crop" and sum of "production" was performed and viewed as a bar chart only for the tail of the data.
- The crop coconut shows highest production as mentioned in previous slide.

Total Production Of Crops

Production Of Crops Area Wise For Every State

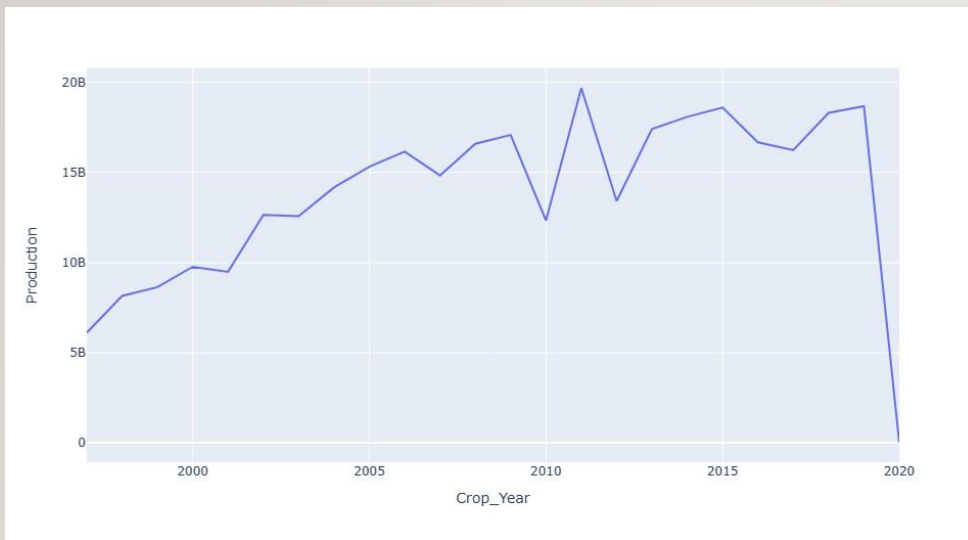


Total Production Of Crops State-wise

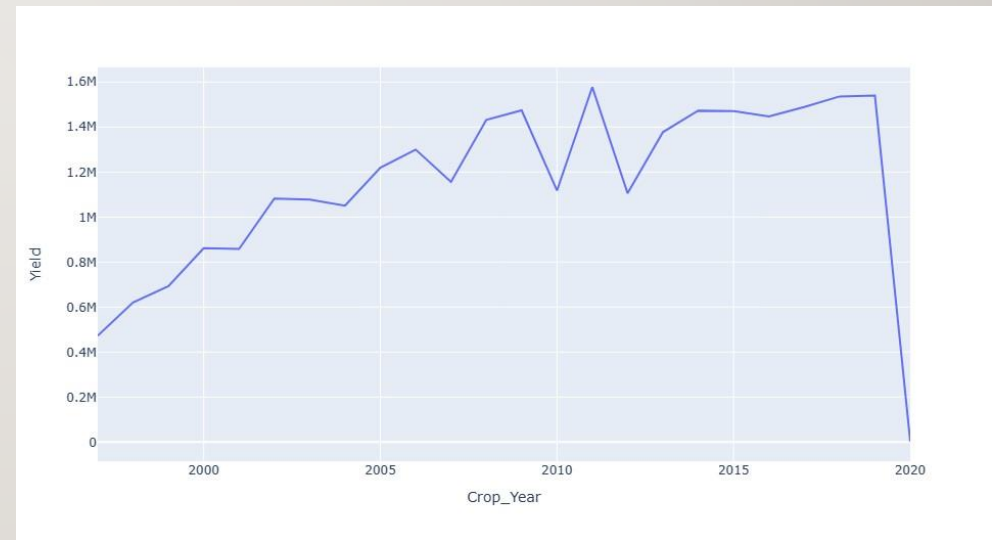


Year Wise Trend Analysis

Production

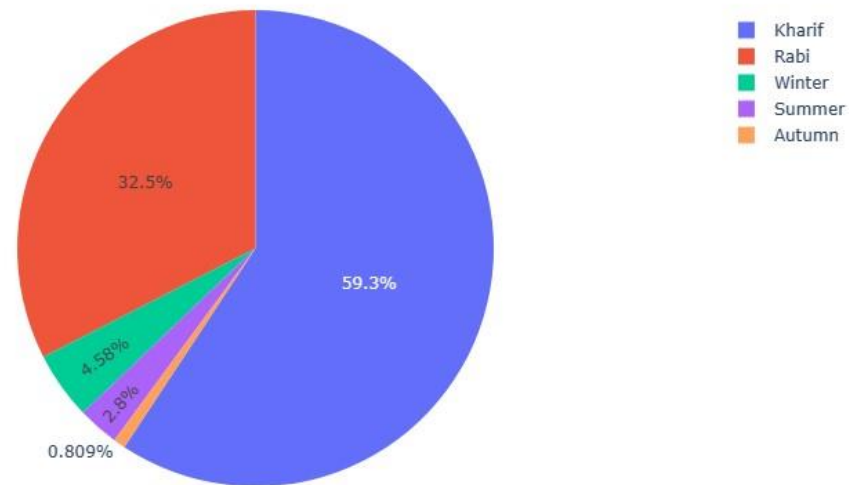


Yield



The Distribution Of Crop Production Across Different Seasons

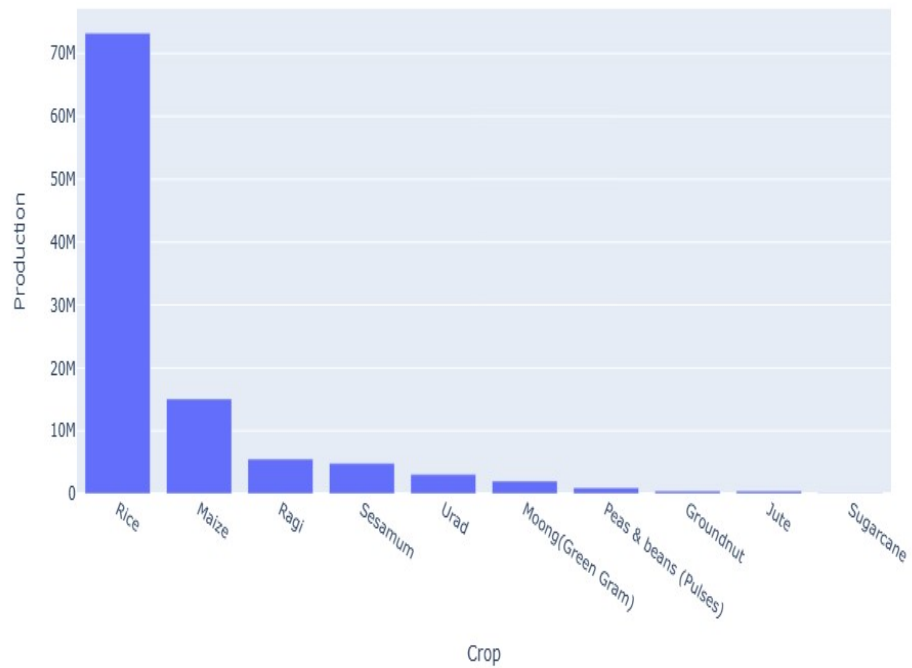
Percentage of Production for Crops in Each Season



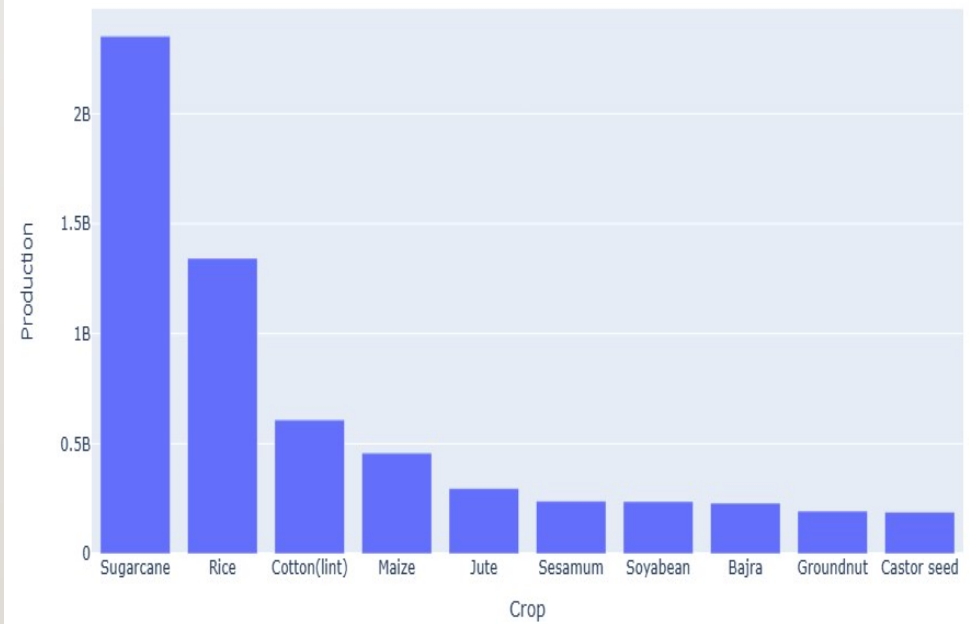
- A group by of “season” and sum of “production” to view as a pie chart.
- More than 50% of the production has taken place in the kharif season i.e., from May to October.

The Highest Crop Production By Season

Highest Crop Production in Autumn

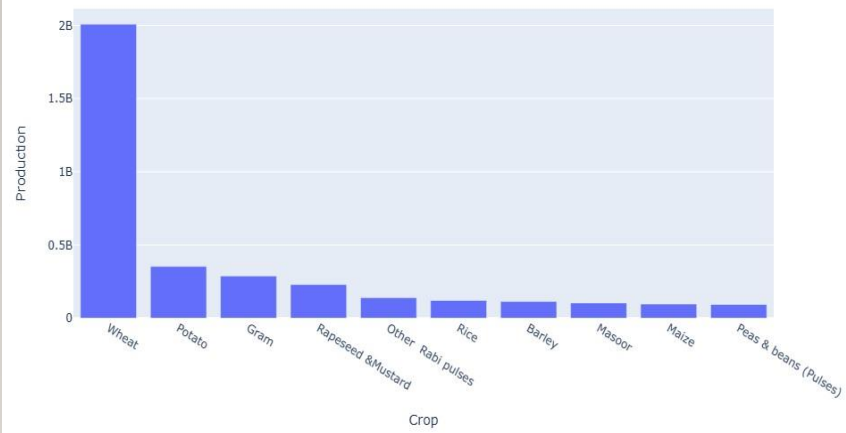


Highest Crop Production in Kharif

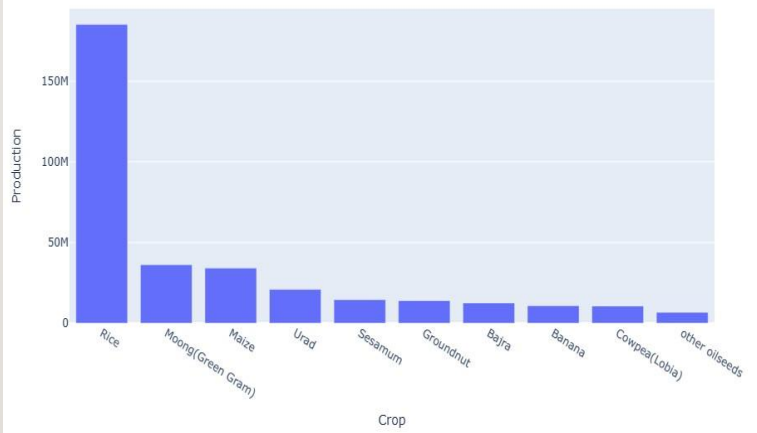


The Highest Crop Production By Season

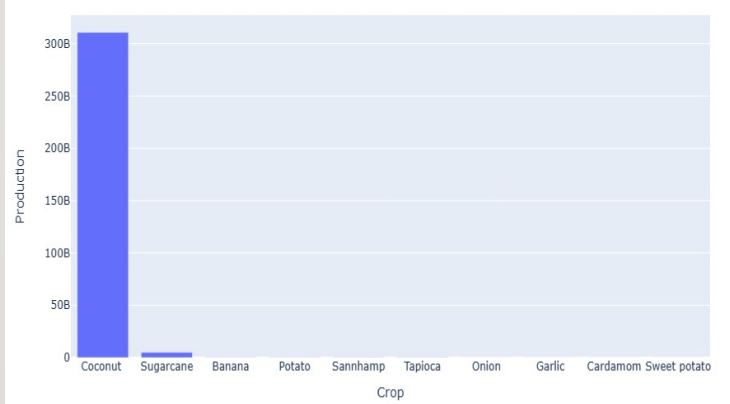
Highest Crop Production in Rabi



Highest Crop Production in Summer



Highest Crop Production in Whole Year



Overall Insights

- ❑ Major crops of India are sugarcane, rice, wheat, coconut.
- ❑ Maximum production was recorded in the kharif season.
- ❑ Kerala has the highest production of crops per unit area (Tonnes/Hectares).
- ❑ Highest production has been recorded in the year 2011.

Conclusion

- ❑ The project provides valuable insights into crop production patterns and trends.
- ❑ By exploring and visualizing the data, we have gained a better understanding of the most commonly grown crops, the distribution of crop production across seasons and states, and the highest production in different years.
- ❑ These insights can be utilized for decision-making, resource allocation, and future planning in the agricultural sector. Additionally, the project highlights the importance of data exploration, visualization, and analytical techniques in understanding and utilizing agricultural data effectively.