UNIFIED MENTOR

Project Report

On

Crop-Production Analysis in India

By : Prajakta Patil

PROBLEM STATEMENT

The Agriculture business domain, as a vital part of the overall supply chain, is expected to highly evolve in the upcoming years via the developments, which are taking place on the side of the Future Internet. This paper presents a novel Business-to-Business collaboration platform from the agri-food sector perspective,which aims to facilitate the collaboration of numerous stakeholders belonging to associated business domains, in an effective and flexible manner.

This dataset provides a huge amount of information on crop production in India ranging from several years. Based on the Information the ultimate goal would be to predict crop production and find important insights highlighting key indicators and metrics that influence crop production.

Make views and dashboards first and also make a story out of it.

INTRODUCTION

India is one of the largest producers of agriculture production in the world. It is the second largest producer in the wheat and rice.

Crop production is the process of growing and harvesting plants for food, fiber, fuel, or other purposes. It's the foundation of agriculture and plays a crucial role in feeding the world's population Got a chance to complete a project on crop production of india. Weather plays an important role in agriculture production. Thus there is no aspect of crop culture that is immune to impact of weather. Weather factor contribute to optimal crop growth, development and yield. For rainfall variability needs to be expressed in terms of percentage so that minimum assured rainfall amounts at a certain level of probability. For optimal productivity at a given location crops must be such that their weather requirements match the temporal match of relevant weather elements. A detailed knowledge of rainfall regime at a place is an important prerequisite for agriculture planning and management. Soil fertility refers to the inherent capacity of soil to supply nutrients in adequate amount and in suitable proportion for crop growth and crop yield.

DATASET

```
data
```

	State_Name	District_Name	Crop_Year	Season	Crop	Area	Production
0	Andaman and Nicobar Islands	NICOBARS	2000	Kharif	Arecanut	1254.0	2000.0
1	Andaman and Nicobar Islands	NICOBARS	2000	Kharif	Other Kharif pulses	2.0	1.0
2	Andaman and Nicobar Islands	NICOBARS	2000	Kharif	Rice	102.0	321.0
3	Andaman and Nicobar Islands	NICOBARS	2000	Whole Year	Banana	176.0	641.0
4	Andaman and Nicobar Islands	NICOBARS	2000	Whole Year	Cashewnut	720.0	165.0
...	...	...	...	...	...	...	...
246086	West Bengal	PURULIA	2014	Summer	Rice	306.0	801.0
246087	West Bengal	PURULIA	2014	Summer	Sesamum	627.0	463.0
246088	West Bengal	PURULIA	2014	Whole Year	Sugarcane	324.0	16250.0
246089	West Bengal	PURULIA	2014	Winter	Rice	279151.0	597899.0
246090	West Bengal	PURULIA	2014	Winter	Sesamum	175.0	88.0

246091 rows × 7 columns

CODE DEMONSTRATION

```
import pandas as pd
import numpy as np

# seaborn
import seaborn as sns

# matplotlib
import matplotlib.pyplot as plt

# Load Data

data=pd.read_csv('/content/Crop Production data (1).csv')
data

data.head()
data.shape()

data.info()

# Display missing values
data.isnull().sum()

# Remove null values
data.dropna(subset=["Production"],axis=0,inplace=True)
data.isnull().sum()


data.Season.unique()
data.Crop_Year.unique()
data.Crop.unique()
sns.heatmap(data.isna())
data["Production"].nunique()
```

```
# Descibe all data
data.describe(include="all").T


# # Seasonal Analysis

season_grp=data.groupby("Season")
sns.countplot(data=data,x="Season")
print(data['Season'].unique())

# Season Kharif

kharif = season_grp.get_group('Kharif     ' )
kharif["Crop"].nunique()
kharif["Crop"].unique()
kharif["Crop"].value_counts().head(20)
kharif["Crop"].value_counts().tail(20)
top_crops_kharif = kharif['Crop'].value_counts().head(10)
top_crops_kharif
low_crops_kharif = kharif['Crop'].value_counts().tail(10)
low_crops_kharif =

# Visualization of data
plt.pie(top_crops_kharif, autopct="%.1f%%", labels=top_crops_kharif.index, startangle=90, counterclock=False)
plt.title('Distribution of Top 10 crops')
plt.show()

plt.pie(low_crops_kharif, autopct="%.1f%%", labels=low_crops_kharif.index, startangle=90, counterclock=False)
plt.title('Distribution of low 10 crops')
plt.show()


state_kharif=kharif["State_Name"].value_counts()
state_kharif
```

```
plt.pie(state_kharif, autopct="%.1f%%", labels=state_kharif.index, startangle=90, counterclock=False)
plt.title('Distribution of kharif in states')
plt.show()

sns.countplot(data=kharif,x="Crop_Year")
plt.xticks(rotation=90)

# Season
rabi=season_grp.get_group('Rabi       ')
rabi
rabi["Crop"].nunique()
rabi["Crop"].unique()
rabi['Crop'].value_counts()
rabi_crop_top=rabi['Crop'].value_counts().head(10)

# Visualization of data
plt.pie(rabi_crop_top, autopct="%.1f%%", labels=rabi_crop_top.index, startangle=90, counterclock=False)
plt.title('Distribution of Top 10 crops')
plt.show()

rabi_crop_low=rabi['Crop'].value_counts().tail(10)
rabi_crop_low

plt.pie(rabi_crop_low, autopct="%.1f%%", labels=rabi_crop_low.index, startangle=90, counterclock=False)
plt.title('Distribution of low 10 crops')
plt.show()

state_rabi=rabi["State_Name"].value_counts()
state_rabi

plt.pie(state_rabi, autopct="%.1f%%", labels=state_rabi.index, startangle=90, counterclock=False)
plt.title('Distribution of Rabi in states')
plt.show()
```

```
plt.title('Distribution of Rabi in states')
plt.show()

sns.countplot(data=rabi,x="Crop_Year")
plt.xticks(rotation=90)

# Season autumn

autumn=season_grp.get_group('Autumn     ')
autumn

autumn["State_Name"].value_counts()
sns.countplot(data=autumn,x="State_Name")
plt.xticks(rotation=90)

autumn["Crop"].value_counts()

sns.countplot(data=autumn,x="Crop")
plt.xticks(rotation=90)

sns.countplot(data=autumn,x="Crop_Year")
plt.xticks(rotation=90)

# Season whole_year

whole_year=season_grp.get_group('Whole Year ')
whole_year
whole_year["Crop"].nunique()
whole_year["Crop"].unique()

wholeyear_crop_low=whole_year['Crop'].value_counts().tail(10)
wholeyear_crop_low

wholeyear_crop_top=whole_year['Crop'].value_counts().head(10)
wholeyear_crop_top
```

```
# Visualization of data

plt.pie(wholeyear_crop_low, autopct="%.1f%%", labels=wholeyear_crop_low.index, startangle=90, counterclock=False)
plt.title('Distribution of Top 10 crops')
plt.show()

plt.pie(wholeyear_crop_top, autopct="%.1f%%", labels=wholeyear_crop_top.index, startangle=90, counterclock=False)
plt.title('Distribution of Top 10 crops')
plt.show()


sns.countplot(data=whole_year,x="Crop_Year")
plt.xticks(rotation=90)

sns.countplot(data=data,x="State_Name")
plt.xticks(rotation=90)

sns.histplot(data['Crop_Year'], bins=20, kde=False, color='skyblue')
plt.title('Distribution of Crop_Year')
plt.xlabel('Crop_Year')
plt.ylabel('Count')
plt.show()
```

ANALYSIS APPROACH

1. Data Understanding:

- **Dataset Overview**: Begin by understanding the structure and content of the dataset. This includes examining the columns, data types, and any missing values.
- **Domain Knowledge**: Gain insights into the agricultural domain, including key factors affecting crop production such as climate, soil quality, and farming practices.

2. Data Preprocessing:

- **Handling Missing Values**: Address any missing or incomplete data by either removing rows with missing values or imputing them using appropriate techniques.
- **Data Cleaning**: Check for any inconsistencies or errors in the data and correct them if necessary.
- **Feature Engineering**: Create new features or transform existing ones to extract valuable information for analysis.

3. Exploratory Data Analysis (EDA):

- **Descriptive Statistics**: Compute summary statistics to understand the central tendency, variability, and distribution of key variables such as crop production, area under cultivation, etc.
- **Visualizations**: Create visual representations of the data using plots such as line plots, bar charts, and histograms to identify trends, patterns, and relationships.

Github Link:

https://github.com/Prajakta74/prajakta_7474

CONCLUSION

In conclusion, the analysis of crop production data using Python has provided valuable insights into agricultural practices and productivity in India. By examining trends, patterns, and key factors influencing crop production, we have gained a deeper understanding of the dynamics of the agriculture sector.

Through this analysis, we have identified opportunities for optimizing agricultural practices, improving productivity, and addressing challenges faced by farmers. The findings highlight the importance of data-driven approaches in informing decision-making processes and driving innovation in the agriculture domain.

Moving forward, it is essential to continue leveraging data analysis techniques to monitor crop production trends, identify emerging issues, and implement targeted interventions for sustainable agricultural development. By collaborating with stakeholders across the agricultural value chain, we can work towards enhancing food security, promoting agricultural resilience, and fostering economic growth in rural communities.

THANK YOU