



# FARMERS' QUERIES: AN ANALYSIS AND A WEBAPP TO GUIDE TO REGISTER QUERIES





DEPARTMENT OF STATISTICS  
PANJAB  
UNIVERSITY

# **ACKNOWLEDGEMENT**

The successful accomplishment of this project is the result of the contribution of many people to whom we are grateful.

First of all, we would like to express our humble gratitude to the Almighty for giving us the ability to work hard.

We would like to express our gratitude and special thanks to our respected project teacher **Dr. KALPANA K MAHAJAN, Professor, Department of Statistics** for helping us in every possible way to complete this project. Without her guidance, this project would not have seen the light of day. Her guidance and advice played a vital role in the implementation and completion of this project. She always gave us her suggestions which were really very helpful in making this project as flawless as possible.

We are also grateful to the Department of Statistics, Panjab University, for introducing the project work as a part of the course, which has given us the much needed field experience. We thank the Department for providing us the necessary support for completing the project.

Finally, thanks to each other in our group for working on the project as a team.

**AJAY KUMAR, LOBEPREET KAUR, MANJU KUMARI, SONAM RANI**

# **CERTIFICATE**

This is to certify that the project “**FARMERS’ QUERIES: AN ANALYSIS AND A WEBAPP TO GUIDE TO REGISTER QUERIES**” is an academic work done by “**AJAY KUMAR, LOBEPREET KAUR, MANJU KUMARI, SONAM RANI**” submitted in the partial fulfillment of the requirement for the award of the Degree of **MASTER OF SCIENCE(Statistics)** from **DEPARTMENT OF STATISTICS, PANJAB UNIVERSITY, CHANDIGARH (INDIA)**. It has been completed under the guidance of **Dr. KALPANA K MAHAJAN.**

# **CERTIFICATE**

This is Certify that Project “**FARMERS’ QUERIES: AN ANALYSIS AND A WEBAPP TO GUIDE TO REGISTER QUERIES**” was conceived and executed by “**AJAY KUMAR, LOBEPREET KAUR, MANJU KUMARI, SONAM RANI**” under the guidance of undersigned.

**Dr. Kalpana K Mahajan**

**Professor, Department of Statistics,  
Panjab University,  
Chandigarh**

# **INDEX**

## **CHAPTER 1 INTRODUCTION & PROBLEM IDENTIFICATION**

- 1.1 INDIA: An Agriculture Driven Country
- 1.2 Top 10 Crops Producing States of India
- 1.3 Identifying the Problem
- 1.4 Concept of Kisan Call Center

## **CHAPTER 2 RESEARCH METHODOLOGY**

- 2.1 Objectives
- 2.2 Terms
- 2.3 Sample
- 2.4 Data Collection
- 2.5 Statistical Tools

## **CHAPTER 3 DATA PROCESSING AND ANALYSIS**

- 3.1 Analyze overall data
- 3.2 Season-wise analysis of overall data

3.3 Sector-wise analysis of overall data

3.4. Analysis of Null Values

## **CHAPTER 4      DISCUSSION & SUMMARY**

4.1 Essence of Overall Data

4.2 Extract of Season-wise analysis

4.3 Summary of Sector-wise analysis

4.4 Summary of Analysis of Null Values

## **CHAPTER 5      WEBAPP**

5.1 About WebApp

5.2 WebApp framework

5.3 WebApp Deployment

## **CHAPTER 6      CONCLUSION & SUGGESTIONS**

6.1 Findings

6.2 Recommendations

## **REFERENCES**

# **CHAPTER 1**

## **INTRODUCTION & PROBLEM IDENTIFICATION**



### **1.1 INDIA: An Agriculture Driven Country**

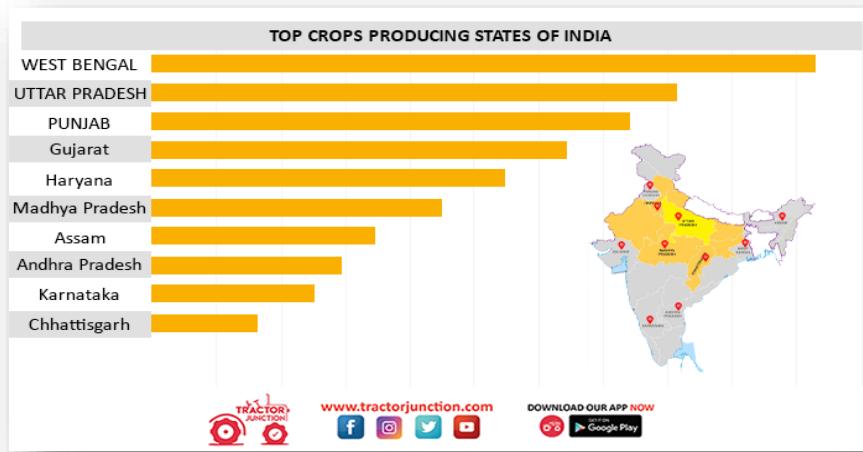
Agriculture in India is the essence of livelihood for around 58% of the population. Agriculture States in India growing with super quality food grains and other food items. The Indian agriculture industry is increasing at a high pace and continuously providing its contribution to world trade. India's grocery and food market is the 6th largest globally, with 70% of the sales.

With developing India, agriculture is evolving continuously at the same pace. The pattern of production and consumption of food is changing. For some years, in India, an increase in population, an increase in Income, rural/urban migration, and extension in rural per capita productivity has been recorded. With the enhancement of all these factors, demand for the food increased too. India's per capita GDP is supposed to rise by 320% in the coming 20 years. With time changing, the trend is remolding too. Indians are now consuming high-quality foods, and focus is switching from plant-based protein to animal-based protein. All thanks to enhancing obtainable incomes and quickly increasing consumer requirements.

## 1.2 Top 10 Crops Producing States of India

A variety of food and non food crops are grown in different parts of the country depending upon the variations in soil, climate and cultivation practices. Major crops grown in India are rice, wheat, millets, pulses, tea, coffee, sugarcane, oil seeds, cotton and jute, etc.

Here we are going to show agricultural production in India state wise. From that you get proper information regarding major crops of India. Let's have a look on the



top 10 agriculture states in India .

### 1. West Bengal

West Bengal is the largest food grain producing state in India. It is known for its rice production, followed by Andhra Pradesh, Punjab, and Uttar Pradesh. With rice production, it is famous for jute, sesamum, tobacco, and tea too. In West Bengal, rice's total production is 146.05 lakh tons on 2600 kilograms per hectare yield. It comes under

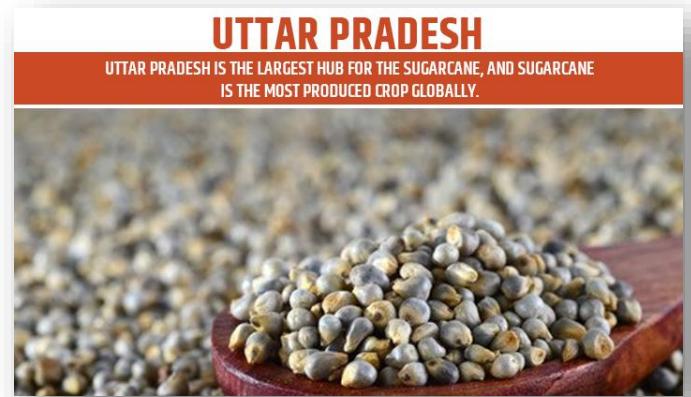


rice producing states in India. West Bengal is engaged in fruit production, including Mango, Litchi, Pineapple, Guava, and Orange.

West Bengal is rich with food commodities, they produced almost all the items, Cauliflower, Tomato, Cucurbits, Cabbage, Okra, and Brinja. Rice, jute, and wheat are the main crops grown in West Bengal. In West Bengal, some spices produced Chilli, Ginger, Garlic, Coriander, and Turmeric. West Bengal comes under the top most agriculture states in India.

## 2. Uttar Pradesh

Uttar Pradesh comes under the top farming state in India and the rank of Uttar Pradesh counted under major state wise crop production in India, bajra, rice, sugarcane, food grains, and many more. It comes under the top wheat producing states in India, followed by Haryana, Punjab, and Madhya Pradesh. Uttar Pradesh has 22.5 million tons of wheat, and the weather condition is appropriate for growing wheat. On 96 lakh hectares, land used to grow wheat in Uttar Pradesh.



Uttar Pradesh is the largest hub for the sugarcane, and sugarcane is the most produced crop globally. Sugarcane grows under hot and humid climate conditions, and Uttar Pradesh is best for that. It has 145.39 million tons of sugarcane and grows on 2.17 million hectares in the state.

### 3. Punjab

Punjab is the most fertile state on earth. It is best to produce wheat, sugarcane, rice, vegetables, and fruits in Punjab. Punjab's other name is the Granary of India and India's breadbasket. Around 93% of the total productive land used to produce food grain. In Punjab, most of the area covered by wheat and paddy cultivation. And this area increases over the year.

Punjab is the 3rd largest farming crop producing state in India. It is known for its irrigation system and these are appropriate for farming. Punjab is the 3rd largest producer of food grains too.



### 4. Gujarat

Gujarat is the fastest growing state in India. This state adopted a wise development pattern. They invested in agriculture, energy, and industry, for that they achieved a double digit growth. Gujarat's weather climate is variable, producing crops there is difficult. One strategy farmers can adopt there is to manipulate crop environments by advanced management for high yield.



Gujarat produced cotton, groundnut, castor, bajra, tur, green gram, sesamum, paddy, maize, and sugarcane. And Gujarat produced cotton in a large scale, followed by Karnataka, Maharashtra, and Telangana. Here groundnut grew too.

## 5. Haryana

And the next farming state is Haryana. Haryana is one of the biggest contributors to agriculture. About 70% of the locals engaged in agriculture. Haryana plays an important role in the Green Revolution in India. With all these, Haryana has a massive irrigation system.

Some top crops produced in Haryana are sugarcane, paddy, wheat, and sunflower. And in India, it is the 2nd largest producer of sunflower. Haryana involved in livestock farming too. There are 99.97 lakh livestock populations in India. India consumes milk products on a daily basis. It is a necessary diet. This comes under one of the top Agriculture States in India.



## 6. Madhya Pradesh

Madhya Pradesh is known for its pulses production, followed by Maharashtra, Rajasthan, and Uttar Pradesh. It is popular for soybean and garlic production too. Madhya Pradesh



earned significant fame in pulse farming. Wheat and maize are the main sources of the earning of farmers in Madhya Pradesh. And other pulses are urad, soybean, and tur.

Madhya Pradesh has the largest area, for that they have different climate and soil conditions best for a wide range of agriculture products. The agriculture sector of Madhya Pradesh is the backbone of the Indian economy. Madhya Pradesh Agriculture sector produced 65% of the employment, and it is  $\frac{1}{4}$  of the GSDP (Gross State Domestic Product).

## 7. Assam

Assam is highly dependent on agriculture in India. In comparison with other states, Assam comes under the least developed states. Assam's almost economy is based on agriculture, and 70% of the population is dependent on agriculture income for livelihood. Tea consumed on a regular basis all over the world after water. Assam is popular for its tea production. It is the largest producer of tea in India, followed by Himachal Pradesh, West Bengal, and other North Indian states. The most famous types of tea in India are Nilgiri tea, Darjeeling tea, Assam tea, and Kangra tea. Assam grew 52% of the total tea production in India.



## 8. Andhra Pradesh

In Andhra Pradesh, 62% of the population engaged in agriculture and more focus is paid on rice production. Andhra Pradesh contributes 77% of crop production in India. And more crops are jowar, bajra, maize, ragi, tobacco, pulses, sugarcane, and others.

In Andhra Pradesh, 1.5 million hectares of land used for horticulture. And from this allotted land nearly 720 thousand hectares used for fruit production.

## ANDHRA PRADESH

ANDHRA PRADESH PRODUCED JOWAR, BAJRA, MAIZE, RAGI, TOBACCO, PULSES, SUGARCANE, AND OTHERS.



## 9. Karnataka

In Karnataka, agriculture is the most important part of the overall economy. Most of the population of the state is engaged in farming. The weather climate of Karnataka greatly supports agriculture. Karnataka's Kharif crops are rice, maize, moong dal, red chili, sugarcane, groundnut, soybean, turmeric, and cotton. Karnataka's rabi crops are mustard, sesamum, barley, wheat, and peas. Karnataka comes under the top Agriculture States in India.

## KARNATAKA

KARNATAKA, AGRICULTURE IS THE MOST IMPORTANT PART OF THE OVERALL ECONOMY.

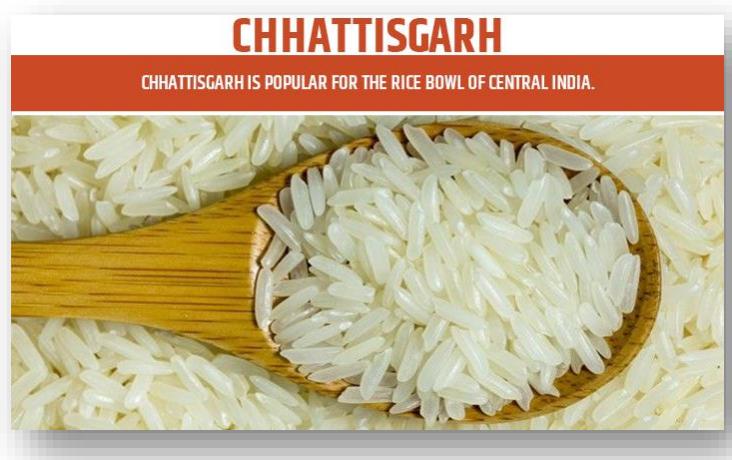


The state is popular for its coffee production, and it contributed 70% of the total production in India. Karnataka grew 222300 metric tons of coffee.

## 10. Chhattisgarh

Chhattisgarh is popular for the Rice Bowl of Central India. Some of the crops that are produced in Chhattisgarh are rice, millets, and maize. In Chhattisgarh, 77% of the area used for rice production. Chhattisgarh is totally dependent on the rain. Only 20% of the total region is under irrigation.

- Fruit Crops – Lime, Cashew-nut, Guava, Cheku, Mango and etc.
- Vegetables Crops – Cucurbits, Cabbage, Beans, Cauliflower etc.
- Spices – Ginger, Turmeric, Methi, Coriander, Chili, Garlic etc.
- Flowers – Marry-gold, Gladiolus, Gaillardia, Gladiolus etc.
- Medicinal Plants – Jamarosa, E.citridora, Pamarosa, Lemongrass etc.





## 1.3 IDENTIFYING THE PROBLEM

In India, the 2020–2021 saw protests from a section of farmers against three farm acts which were passed by the Parliament of India in September 2020.

Legal and political dimensions apart, the issue has drawn the attention of masses towards the problems of farmers, and hence this project.

The title of the project is: “FARMERS’ QUERIES: AN ANALYSIS AND A WEBAPP TO GUIDE TO REGISTER QUERIES”

### 1.3.1. About the Project: In Brief

In this project we have focused in the queries raised by farmers in KISAN CALL CENTRE in 2019. For this analysis we have selected the top 10 crop producing states of India. Secondary data has been used for this analysis which has been downloaded from the government’s website i.edata.gov.in. Data preparation and analysis is done by Python programming language. The name of the major python libraries used in this analysis are pandas and dataprep.

In this project we have also developed a webapp which is for the convenience of the farmers. This web app tells the sector of farmer’s problems. This web app is mostly useful for those farmers who are unaware of technical terminologies of agriculture. If farmers write their queries in this web app then it predicts their sector so that farmers can easily contact the respective department through any means of communication. This web app is built using machine learning algorithms.

## 1.4 Concept of Kisan Call Center

The challenges before Indian Agriculture are immense. This sector needs to grow at a faster rate than in the past to allow for higher per capita income and consumption. It is an accepted fact that the sound agricultural development is essential for the overall economic progress. About two thirds of workforce directly or indirectly dependent on agriculture. This sector generates about 28 percent of its GDP and over 15 percent of exports. Rising consumer prosperity and the search by farmers for higher incomes will simultaneously drive crop diversification. Export opportunities for agricultural products are also expected to continue to grow, provided India could meet the stability, quality and presentation standards demanded by foreign trade and consumers and maintain its comparative advantage as a relatively low cost producer.



Given its range of agro-ecological setting and producers, Indian Agriculture is faced with a great diversity of needs, opportunities and prospects. The well endowed irrigated areas which account for 37 percent of the country's cultivated land currently contribute about 55 percent of agricultural production, whereas,



rainfed agriculture which covers 63 percent accounts for only 45 percent of agricultural production. In these less favorable areas, yields are not only low but also highly unstable and technology transfer gaps are much wider as compared to those in irrigated areas.

If it is to respond successfully to these challenges, greater attention will have to be paid to information-based technologies. Strengthened means of

dissemination will be needed to transmit this information to farmers. Both technology generation and transfer will have to focus more strongly than ever before on the themes of optimization in the management of their available resources by producers, sustainability, coping with diversity by adapting technology more specifically to agro-ecological or social circumstances and raising the economic efficiency of agriculture. To make information transfer more effective, greater use will need to be made of modern information technology and communication among researchers, extensionists and farmers.



Public extension system requires a paradigm shift from top-down, blanket dissemination of technological packages, towards providing producers with the knowledge and understanding with which they solve their own location - specific problems. Continuous two-way interaction among the farmers and agricultural scientists is the most critical component of Agricultural Extension.

At present, the issues have been addressed by the Extension Systems of State Departments of Agriculture, State Agricultural Universities (SAUs), KVKs, NGOs, Private Extension Services through various extension approaches in transfer of technology. A limitation in Transfer of Technology (TOT) model continues to remain a challenge for the public and private extension systems. With the availability of telephone and Internet, it is now possible to bridge this gap to quite a large extent by using an appropriate mix of technologies.

The Department of Agriculture & Cooperation, Ministry of Agriculture, Govt. of India has launched Kisan Call Centers with a view to leverage the extensive telecom infrastructure in the country to deliver extension services to the farming community. The purpose of these Call Centers is mainly to respond to issues raised by farmers instantly in the local language, on continuous basis.

## 1.4.1 About Kisan Call Centre



In order to harness the potential of ICT in Agriculture, Ministry of Agriculture & Farmers Welfare launched the scheme “Kisan Call Centres (KCCs)” on January 21, 2004. Main aim of the project is to answer farmers’ queries on a telephone call in their own dialect. The KCCs are managed by a Service Provider i.e. IFFCO Kisan Sanchar Limited (IKSL). At present there are 454 Farm Tele Advisors. IT Division of DAC&FW from 20/08/2014 has been operationalizing KKMS(Kisan Knowledge Management System) from Shastri Park Centre.

At present these Call Centres are working in 21 different locations covering all the States and UTs. A countrywide common eleven digit Toll Free number 1800-180-1551 or 1551 has been allotted for Kisan Call Centres. This number is accessible through mobile phones and landlines of all telecom networks including private service providers. Replies to the farmers’ queries are given in 22 local languages.



Kisan Call Center services are available from 6.00 am to 10.00 pm on all seven days of the week at each KCC location. Kisan Call Centre agents known as Farm Tele Advisors (FTAs), are graduates or above (i.e. PG or Doctorate) in Agriculture or

allied areas (Horticulture, Animal Husbandry, Fisheries, Poultry, Bee-keeping, Sericulture, Aquaculture, Agricultural Engineering, Agricultural marketing, Bio-technology, Home Science etc.) and possess excellent communication skills in respective local languages.



Queries which cannot be answered by Farm Tele Advisor (FTAs) are transferred to higher level experts in a call conferencing mode. These experts are subject matter specialists of State Agriculture Departments, ICAR and State Agricultural Universities.

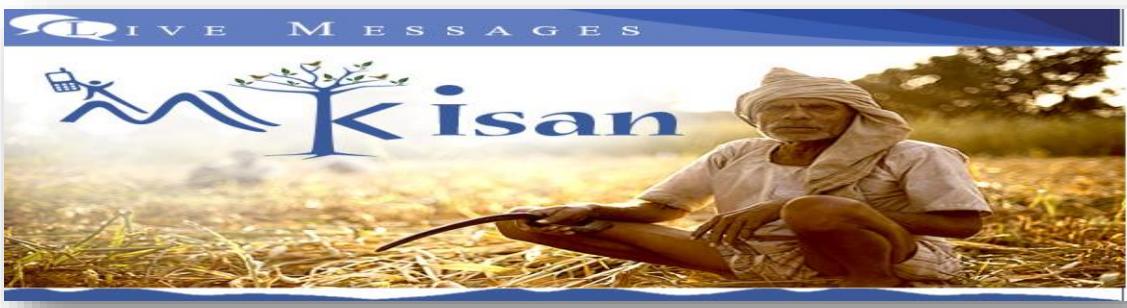
#### 1.4.2. Features of Kisan Call Centre



Restructured KCCs are now providing very reliable and efficient services due to following technological innovations and state of the art hardware/software tools:

- (i) Voice/Media Gateways (IPPBX based decentralized system).
- (ii) Dedicated MPLS leased line network with dedicated bandwidth.

- (iii) Call barge to facilitate listening of call conversation between Farm Tele Advisor and Farmer by the officer monitoring the quality of service provided by KCCs.
- (iv) 100% Call recording and retention of recorded calls for six months so that the call can be listened to in case of a complaint.
- (v) SMS to caller farmers providing a gist of advisories given to them on phone.
- (vi) Voice mail system for recording farmer's queries during idle time of KCC or during call lines busy, with provision for call back to the caller.
- (vii) Soft phones in every personal computer with caller ID facility.
- (viii) Back up through Fixed Wireless Telephone (FWTs) in case of PRI failure.
- (ix) Up scaling the knowledge of FTAs by way of providing latest versions of guide books and booklets issued by the State Agricultural Department or the Agricultural Universities. Facility of video conferencing of each KCC for interaction of KCC agents with the Divisional/Zonal Level Officers of the State Agriculture and allied departments as well as on line monitoring for the working of KCCs.
- (x) Provision for registering the farmers for receiving SMS messages on agriadvisories and mandi prices of different commodities as per their priority.
- (xi) Call holding time substantially reduced to less than 30 seconds.
- (xii) KKMS simplified to enable FTAs to initially reply farmers query.
- (xiii) Farmers rating for FTAs monitoring.
- (xiv) Biometric attendance system of FTAs & Supervisors.



### 1.4.3. Kisan Call Centre Location and Languages

**Kisan Call Centre Location and Languages**

S. No.	Location	States/ UTs Covered	Language
1	Guntur	Andhra Pradesh	Telugu
2	Hyderabad	Telangana	Telugu
3	Patna	Bihar	Hindi
4	Ranchi	Jharkhand	Hindi
5	Jaipur	Delhi Rajasthan	Hindi Hindi
6	Ahmadabad	Gujarat Dadra & Nagar Haveli Daman & Diu	Gujarati Gujarati Gujarati/Konkani
7	Chandigarh	Haryana Punjab Chandigarh	Hindi Punjabi Punjabi
8	Solan	Himachal Pradesh	Hindi
9	Jammu	Jammu & Kashmir	Dogri/Kashmiri/Ladakhi
10	Bengaluru	Karnataka	Kannada
11	Trivandrum	Kerala Lakshadweep	Malayalam Malayalam
12	Jabalpur	Madhya Pradesh	Hindi
13	Raipur	Chhattisgarh	Hindi
14	Pune	Maharashtra Goa	Marathi Konkani/Marathi
15	Coimbatore	Tamil Nadu Puducherry	Tamil Tamil
16	Kanpur	Uttar Pradesh	Hindi
17	Pant Nagar	Uttarakhand	Hindi
18	Kolkata	West Bengal Sikkim Andaman & Nicobar	Bengali Sikkimese/Nepali/Hindi Bengali, Tamil, Hindi
19	Bhubaneswar	Odisha	Oriya
20	Guwahati	Arunachal Pradesh Assam Manipur Nagaland	Hindi/Adi Assamese Manipuri Nagamese
21	Agartala	Tripura Mizoram Meghalaya	Bengali Mizo Khasi/Garo/Jayantia



# CHAPTER 2

# RESEARCH METHODOLOGY

## 2.1 Objectives

The project “FARMERS’ QUERIES: AN ANALYSIS AND A WEBAPP TO GUIDE TO REGISTER QUERIES”, focused on the following objectives.

- To list State-wise, Region-wise and Sector-wise Farmers’ most common queries.
- To find out the districts which registered maximum and minimum queries on various counts.
- To find out month-wise (season-wise) effect on the queries.
- To find out the crop for which maximum and minimum queries are raised.
- To find out the category of crops in a particular month for which most of the questions are asked.
- To develop a WebApp to facilitate registering of queries.



## 2.2 Terms

The project: “FARMERS’ QUERIES: AN ANALYSIS AND A WEBAPP TO GUIDE TO REGISTER QUERIES” has following key terms:

Farmers, Queries, Analysis and WEB APP

- **Farmer:** A farmer (also called an agriculturist) is a person engaged in agriculture, raising living organisms for food or raw materials. The term

usually applies to people who do some combination of raising field crops, orchards, vineyards, poultry, or other livestock. A farmer might own the farmed land or might work as a laborer on land owned by others, but in advanced economies, a farmer is usually a farm owner, while employees of the farm are known as farm workers, or farmhands. However, in the not so distant past, a farmer was a person who promotes or improves the growth of (a plant, crop, etc.) by labor and attention, land or crops or raises animals (as livestock or fish).

- **Queries**: The project focuses on queries as raised by farmers from time to time. These queries are reflecting the problem faced by the farmers, question about something especially in order to express one's doubts about it or to check its validity or accuracy and they seek solution by raising queries.
- **Analysis**: Analysis is the process of breaking a complex topic or substance into smaller parts in order to gain a better understanding of it.

Types of Analysis:

A) Qualitative Analysis: It is concerned with which components are in a given sample or compound. Example: Precipitation reaction

B) Quantitative Analysis: It is to determine the quantity of individual component present in a given sample or compound. Example: To find concentration by uv-spectrophotometer.

- **WEB APP** : A web application (or web app) is application software that runs on a web server, unlike computer-based software programs that are run locally on the operating system (OS) of the device. Web applications are accessed by the user through a web browser with an active network connection. These applications are programmed using a client–server modeled structure—the user ("client") is provided services through an off-site server that is hosted by a third-party. Examples of commonly-used web applications include: web-mail, online retail sales, online banking, and online auctions.

## 2.3 Sample

SAMPLE is a group of people, objects or items that are taken from a larger population for measurement. The sample should be representative of the population to ensure that we can generalize the findings from the research sample to the population as a whole.

For the project we have selected 10 states based on the crops production in India.

- West Bengal
- Uttar Pradesh
- Punjab
- Gujarat
- Haryana
- Madhya Pradesh
- Assam
- Andhra Pradesh
- Karnataka
- Chhattisgarh



## 2.4 Data Collection

A screenshot of the data.gov.in website. The header features the 'data.gov.in' logo and a search bar. Below the header, a large banner reads 'HIGH VALUE DATASETS' with a graphic of three screens displaying data visualizations like charts and graphs. The main content area is divided into four sections: 'Recently Added Datasets', 'Most Viewed Datasets', 'Ministry/Dept. Contributed New Datasets', and 'High Value Datasets'. Each section contains a thumbnail icon and a brief description.

For the project we had collected secondary data from the website:

<https://data.gov.in/>

This site is designed, hosted and maintained by National Informatics Centre (NIC), Ministry of Electronics & Information Technology, Government of India.

Open Government Data (OGD) Platform India or data.gov.in is a platform for supporting open data initiative of Government of India. This portal is a single-point access to datasets, documents, services, tools and applications and organizations of the Government of India.

Data for the following districts was not available on the website:-  
Jhargram(WB), Kalimpong(WB), Purba Burdwan(WB), Bengaluru\_urban(KA), Bolado Bazar(CG), Bastar(CG), Bemetara(CG), Kabirdham(CG).

## 2.4.1 About Data



- In this data we have range index: 2835387 entries, 0 to 2835386 and total 11 data columns.
- These columns have Season, Sector , Category, Crop, Query Type, Query Text, Kcc Ans, State Name, District Name, Block Name, Created On variables. The data type of all the variables except Created on are qualitative.
- There were 39953 duplicates in this data which we have removed.
- This data set have following null values:-

```
    Season          2788558
    Sector           11
    Category         13
    Crop             16
    QueryType        17
    QueryText        17
    KccAns          290451
    StateName        18
    DistrictName     19
    BlockName        20
    CreatedOn        19
dtype: int64
```

Null values of Season and KccAns are not removed from the data set.

## VARIABLES

### Season:

A **season** is a period of the year that is distinguished by special climate conditions.

However, the growing **season** in a particular country or area is the period in each year when the weather and temperature is right for plants and crops to grow.

- **India** has three **cropping seasons** — *Rabi, Kharif and Zaid*.
- Rabi crops are sown in winter from October to December and harvested in summer from April to June.
- Kharif crops are grown with the onset of monsoon in different parts of the country and these are harvested in September to October.
- In between the Rabi and the Kharif seasons, there is a short season during the summer months known as the Zaid season.

The data we have collected have the queries asked by the callers which have been sorted out according to different seasons also.

### Sector:

The **Agriculture sectors** comprise establishments primarily engaged in growing crops, raising animals, and harvesting fish and other animals from a farm, ranch, or their natural habitats.

There are mainly four sectors listed here

- **Agriculture(crop farming)** Crop farming is the cultivation of plants for food, animal foodstuffs, or other commercial uses. A variety of techniques including organic production methods can be used to manage crops.
- **Horticulture** is the art of cultivating plants in gardens to produce food and medicinal ingredients, or for comfort and ornamental purposes.

Horticulturists are agriculturists who grow flowers, fruits and nuts, vegetables and herbs, as well as ornamental trees and lawns.

- Animal husbandry is the branch of science deals with the practice of breeding, farming and care of farm animals such as cattle, dogs, sheep and horses by humans for advantages. Animal husbandry refers to livestock raising and selective breeding.
- **Fishery** is the enterprise of raising or harvesting fish and other aquatic life. Commercial fisheries include wild fisheries and fish farms, both in fresh water and the oceans.

The queries belong to different sectors and they been labeled in the data table. Some queries has not been defined in any sectors and are null values as shown below:

```
[ ] df[df['Sector'].isnull()]
```

	Season	Sector	Category	Crop	QueryType	QueryText	KccAns	StateName	DistrictName	BlockName	CreatedOn
920605	to the Weather	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
920606		NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
920610	Information about weather forecast of block- i...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
920615	52700163824 ganna	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
920616	Localo app	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
920618	7409132588 mentha ki kheti	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
920624	Manjeet@1234 new	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
920625	KCC601672 Pass@123456 purana	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
920629	Information about weather forecast of block- K...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
920630	in district- Agra	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
2835386	+	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN

These values are dropped from the data.

## Categories:

Crops are divided into different classes according to their economic purposes.

But some categories are not defined as shown below

```
[ ] df[df['Category'].isnull()]
```

	Season	Sector	Category	Crop	QueryType	QueryText	KccAns	StateName	DistrictName	BlockName	CreatedOn
920619	7503456760) job -850	5963057	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
920620	DIRECTORATE OF PLANT PROTECTION QUARANTINE & STORAGE		NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN

These values are dropped from the data. Also, Some of the values in this column are 0. So, we have replaced these values with Govt.

schemes/Weather/Fertilizers/Weed (as these queries do not belong to a particular crop, so they have not been put in any category and were given value 0).

## Crops:

The different queries that have been asked belong to different crops.

Like Paddy, Wheat, Potato etc.

And those which were not related to crops or which cannot be related to any particular crop are listed as others.

Some queries have null value in crop column as shown below

```
[ ] df[df['Crop'].isnull()]
```

Season	Sector	Category	Crop	QueryType	QueryText	KccAns	StateName	DistrictName	BlockName	CreatedOn
920609	????? 3-8 ??????? ?? ?? ?????? ?????? ??? ... ?????? ?? ???? ?????? ?? ?? ?????? ... NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
920612	Sir moderate to heavy rains are expected between ... partly cloudy. Today's minimum temperature is...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
920613	????? 24-29 ??????? ?? ?? ?????? ??? ?????? ??? ... ?????? ?? ???? ?????? ?? ?? ?????? ... NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN

These values are then removed from the data. Also, some of the values in crop column are 0. So, we have replaced these values with Not related to crop.

## Query Type:

Query Type here refers to the type of question asked like if some questions are asked by pardhanmantri kisan samman nidhi yojna then it fits the query type 'Govt. schemes' and if some questions are asked about weather then it belongs to query type 'Weather' and so on.

But some queries have null values as shown below

```
[ ] df[df['QueryType'].isnull()]
```

Season	Sector	Category	Crop	QueryType	QueryText	KccAns	StateName	DistrictName	BlockName	CreatedOn
920623	Village Name : Picnic Spot Road	Khurram Nagar	Kamla Nehru Nagar	Vikas Nagar ( ?????? ?????? ??? ?????? ?? )		NaN	NaN	NaN	NaN	NaN

It is then removed from the data.

Also, some of the values in *Query Type* are 0. So, we have replaced these values with Irrelevant Calls, as some of the Query Text says these were Irrelavant calls.

## Query Text:

In query text, the question which has been asked by caller has been written.

## KccAns: (kcc stands for Kisan Call Centre)

In this column the questions asked by the caller has been answered

## State Name:

It refers to the different states from where we have been receiving the queries.

Like Punjab, Gujarat, Uttar pradesh etc.

The State Name also have null values

```
[ ] df[df['StateName'].isnull()]
```

Season	Sector	Category	Crop	QueryType	QueryText	KccAns	StateName	DistrictName	BlockName	CreatedOn
920604	NaN	AGRICULTURE	Others	Others	Weather \nInformation about weather forecast of block-	NaN	NaN	NaN	NaN	NaN

It have been removed from the data

## District Name:

District Name refers to the different districts of different states from where we have been receiving the queries.

Like Sirsa, Bathinda, Banas kantha etc.

The District Name also have null values as shown

```
[ ] df[df['StateName'].isnull()]
```

Season	Sector	Category	Crop	QueryType	QueryText	KccAns	StateName	DistrictName	BlockName	CreatedOn
920604	NaN	AGRICULTURE	Others	Others	Weather \nInformation about weather forecast of block-...	NaN	NaN	NaN	NaN	NaN

It have been removed from the data.

## Block Name:

Block Name refers to the different blocks of different districts from where we have been receiving the queries.

Like Balod, Itihar, Karandighi etc.

Block Names also have some null values as given below

```
[ ] df[df['BlockName'].isnull()]
```

Season	Sector	Category	Crop	QueryType	QueryText	KccAns	StateName	DistrictName	BlockName	CreatedOn
1932733	NaN	AGRICULTURE	Others	Others	Crop Insurance ask about crop insurance ??????????? ?????? ?????? ?????? ?????? ?????? ?....	GUJARAT	AMRELI	NaN	2019-11-09T21:42:57.55	

This value has been dropped from the data. Also, some of the values in Block Name are 0. So, We have replaced these values with Not Mentioned (these blocks were not mentioned in the data).

### Created on:

It tells us the date on which the following queries have been asked on the call.

## 2.5 Statistical Tools

In the light of the objectives of the project and nature of data following tools have been used for the preparation and analysis of the data:

- **Frequency distribution tables**:A frequency distribution is a list, table (i.e.: frequency table) that displays the frequency of various outcomes in a sample. Each entry in the table contains the frequency or count of the occurrences of values within a particular group or interval.

A frequency distribution shows us a summarized grouping of data divided into mutually exclusive classes and the number of occurrences in a class. Some of the graphs that can be used with frequency distributions are histograms, line charts, bar charts and pie charts. Frequency distributions are used for both qualitative and quantitative data.

In this project we used frequency distribution tables to represent the number of questions in the different bases of analysis.

- **Diagrams and graphs**:A diagram is a visual form for presentation of statistical data, highlighting their basic facts and relationship. If we draw diagrams on the basis of the data collected they will easily be understood and

appreciated by all. It is readily intelligible and save a considerable amount of time and energy. Significance of Diagrams and Graphs:

Diagrams and graphs are extremely useful because of the following reasons.

1. They are attractive and impressive.
2. They make data simple and intelligible.
3. They make comparison possible
4. They save time and labour.
5. They have universal utility.
6. They give more information.
7. They have a great memorizing effect.

In this analysis we used pie graphs and bar graphs for visual interpretation of the analysis.

- **Chi-Square Test of Independence:** The Chi-Square Test of Independence determines whether there is an association between categorical variables (i.e., whether the variables are independent or related). It is a nonparametric test.

We used Chi Square Test of Independence in testing the association between sectors of queries in different states.

- **Chi-Square Goodness of Fit Test:** Chi-Square goodness of fit test is a non-parametric test that is used to find out how the observed value of a given phenomena is significantly different from the expected value. In Chi-Square goodness of fit test, the term goodness of fit is used to compare the observed sample distribution with the expected probability distribution. Chi-Square goodness of fit test determines how well theoretical distribution (such as normal, binomial, or Poisson) fits the empirical distribution.

We used this test to check the uniformity of the data on different basis.

- Python: Python is a multi-paradigm programming language. Object-oriented programming and structured programming are fully supported, and many of its features support functional programming and aspect-oriented programming (including by metaprogramming and metaobjects (magic methods)). Many other paradigms are supported via extensions, including design by contract and logic programming.

We used Python programming language to make analysis easy and fast.

## CHAPTER 3

# DATA PROCESSING AND ANALYSIS

We have used Python programming language to analyze this data.

Python is an interpreted high-level general-purpose programming language. Python's design philosophy emphasizes code readability with its notable use of significant indentation. Its language constructs as well as its object-oriented approach aim to help programmers write clear, logical code for small and large-scale projects.



Libraries used for analysis:

- Pandas: Pandas is a software library written for the Python programming language for data manipulation and analysis. In particular, it offers data structures and operations for manipulating numerical tables and time series.
- DataPrep: Dataprep is an open-source python library that allows you to prepare your data and that too with just a few lines of code. By preparing data it means that we can analyze the properties of the attributes that are there in the data.

## 3.1 Analyze overall data

### **3.1.1 SECTOR-WISE**

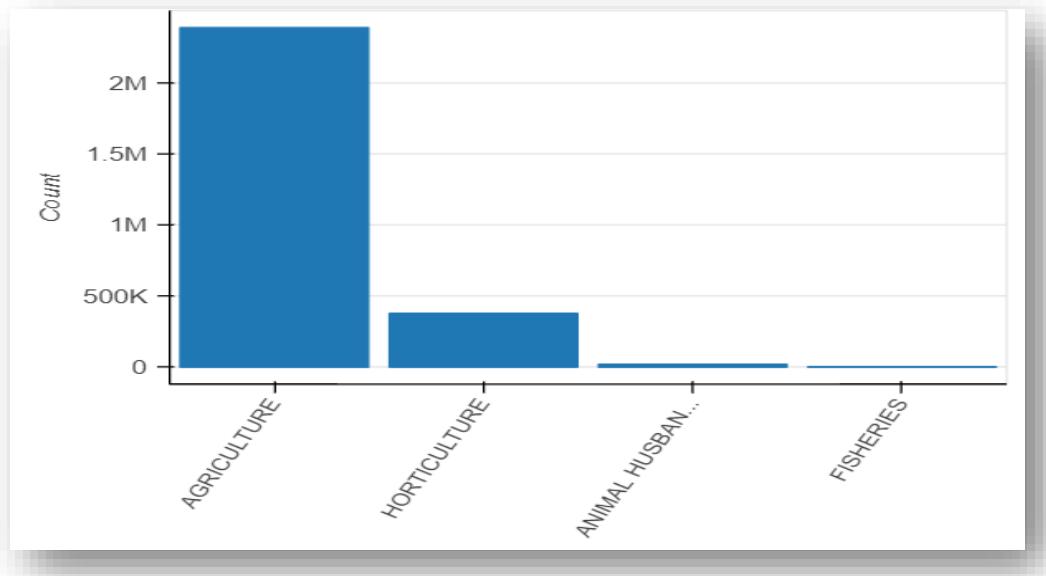
In the given data, we have four different sectors i.e. AGRICULTURE, HORTICULTURE, ANIMAL HUSBANDRY & FISHERIES. The number of queries in these four sectors with respect to 10 selected states can be represented using below table :-

	Agriculture	Horticulture	Animal Husbandry	Fisheries
West Bengal	80741	38391	859	628
UP	813940	133847	8442	1471
Punjab	326815	21719	1791	72
Gujarat	194341	41578	1477	65
Haryana	333401	20792	2059	175
MP	334974	48603	2143	241
Assam	6364	5178	2024	560
AP	103787	28569	776	79
Karnataka	157282	26972	933	125
Chhattisgarh	39379	13922	695	201

**INTERPRETATION:-** From the above table, we can see that we have 4 different sectors in the data & most of the data entries are from AGRICULTURE & HORTICULTURE . We have only few number of queries from other two sectors.

- UP have the maximum number of queries for each sector.
- Assam have minimum number of queries for AGRICULTURE & HORTICULTURE.

Now, by combining the data of all the 10 states , we have results which can be represented using bar chart as follows :-



**INTERPRETATION :-** The largest value (AGRICULTURE) is over 6.3 times larger than the second largest value (HORTICULTURE). Maximum queries are from AGRICULTURE sector & minimum queries are from FISHERIES sector.

#### Chi-Square Test of Independence :

Null Hypothesis: There is no sector wise variation in the states

Alternate Hypothesis: There is sector wise variation in the states

p-value = 0.0

**Result:** p-value < 0.05. Therefore, we conclude that there is significant sector wise variation in the states.

### Chi-Square Goodness of Fit Test:

Null Hypothesis: There is no sector wise variation in different states.

Alternate Hypothesis: There is sector wise variation in different states.

p-value= 0.0

Result: p-value < 0.05. Therefore, we conclude that state wise variation in different sectors has been significantly at variance.

### **3.1.2 CATEGORY-WISE**

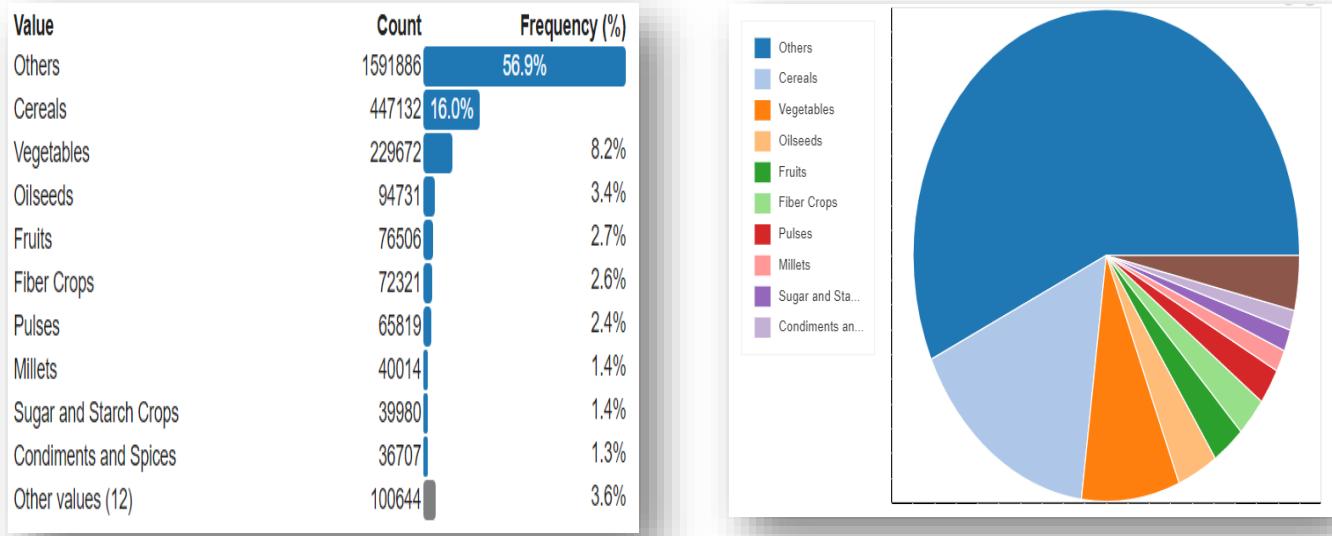
In the given data, we have 22 different categories of which we selected 4 categories which are common to all the 10 states. The number of queries in these four categories with respect to 10 selected states can be represented using below table :-

	Cereals	Vegetables	Oilseeds	Fruits
West Bengal	19556	27891	2224	6429
UP	176251	80649	23402	29178
Punjab	121835	14760	2219	5728
Gujarat	9448	14514	26874	7502
Haryana	42597	12700	7609	6256
MP	45811	30160	25147	6821
Assam	1164	3458	190	906
AP	11132	20001	3814	5952
Karnataka	2268	14141	2320	6202
Chhattisgarh	17070	11398	932	1532

From the above table, it can clearly seen that:-

- All the states have maximum number of queries in mainly two categories :- CEREALS & VEGETABLES.
- Only Gujarat have maximum number of queries in 'OILSEEDS'.

Now, by combining the data of all the 10 states , we have results which can be represented using pie chart as follows:-



**INTERPRETATION:-**Below Pie chart shows the proportion of farmer queries in each category. It can be clearly seen that 16% of the total queries are from 'CEREALS' category . Rest of the 27% queries are from other categories.

\*Others – Particular Category is not mentioned in the given data.[More than 50% of total queries are from Others]

### Chi-Square Goodness of Fit Test:

Null Hypothesis: There is no categories wise variation in different states.

Alternate Hypothesis: There is categories wise variation in different states.

p-value= 0.0

**Result:** p-value < 0.05. Therefore, we conclude that state wise variation in different categories has been significantly at variance.

### **3.1.3 CROP-WISE**

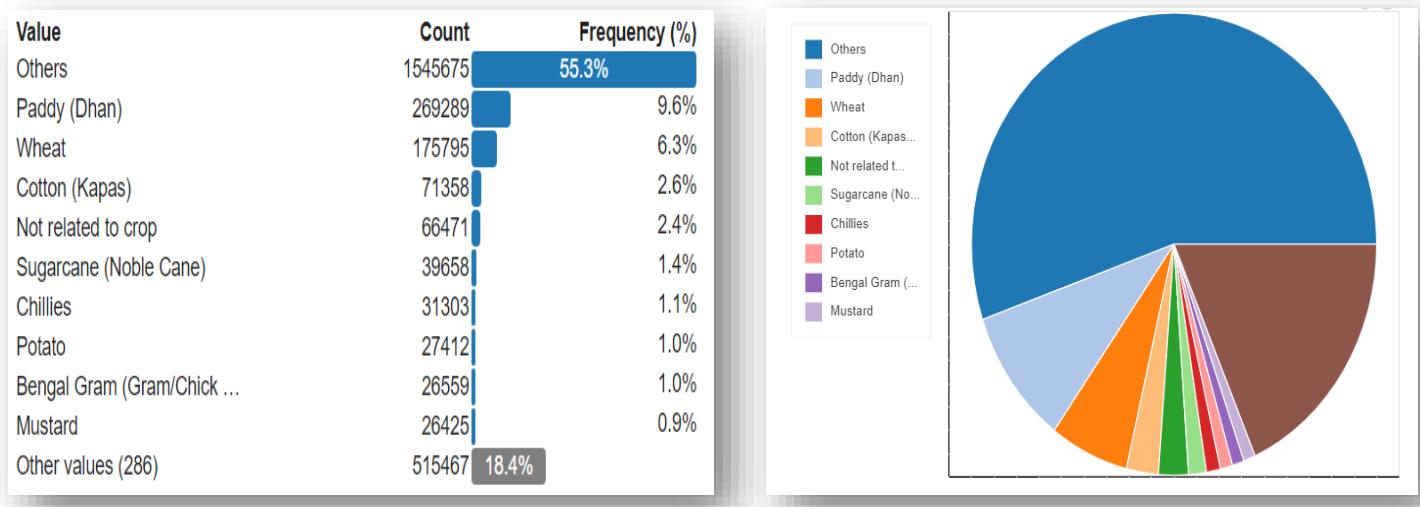
In the given data, we have 296 different crops of which we selected 4 crops which are common to all the 10 states. The number of queries in these four crops with respect to 10 selected states can be represented using the below table :-

	Paddy	Wheat	Cotton	Chilies
West Bengal	19428	-	-	2686
UP	104745	70794	-	8395
Punjab	79636	42368	10396	-
Gujarat	-	6397	21507	2312
Haryana	19737	22631	24237	-
MP	14659	31137	-	-
Assam	1126	-	-	653
AP	11056	-	7353	9141
Karnataka	-	-	4057	2453
Chhattisgarh	15382	1622	-	864

From the above table, it can clearly seen that:-

- Most of the above mentioned states have maximum number of queries in Paddy (Dhan) crop.
- UP have highest number of queries in Paddy crop whereas Assam have minimum number of queries in Paddy crop in comparison to all other states.

Now, by combining the data of all the 10 states , we have results which can be represented using pie chart as follows:-



**INTERPRETATION:-**Below Pie chart shows the proportion of farmer queries in each crop. It can be clearly seen that around 20% of the total queries are from Paddy, Wheat ,Cotton & Sugarcane crops. Rest of the crops constitute remaining 25% proportion.

\*Others – Particular Crop is not mentioned in the given data.[More than 50% of total queries are from Others]

### Chi-Square Goodness of Fit Test:

Null Hypothesis: There is no crops wise variation in different states.

Alternate Hypothesis: There is crops wise variation in different states.

p-value= 0.0

**Result:** p-value < 0.05. Therefore, we conclude that state wise variation in different crops has been significantly at variance.

### **3.1.4 QUERY TYPE**

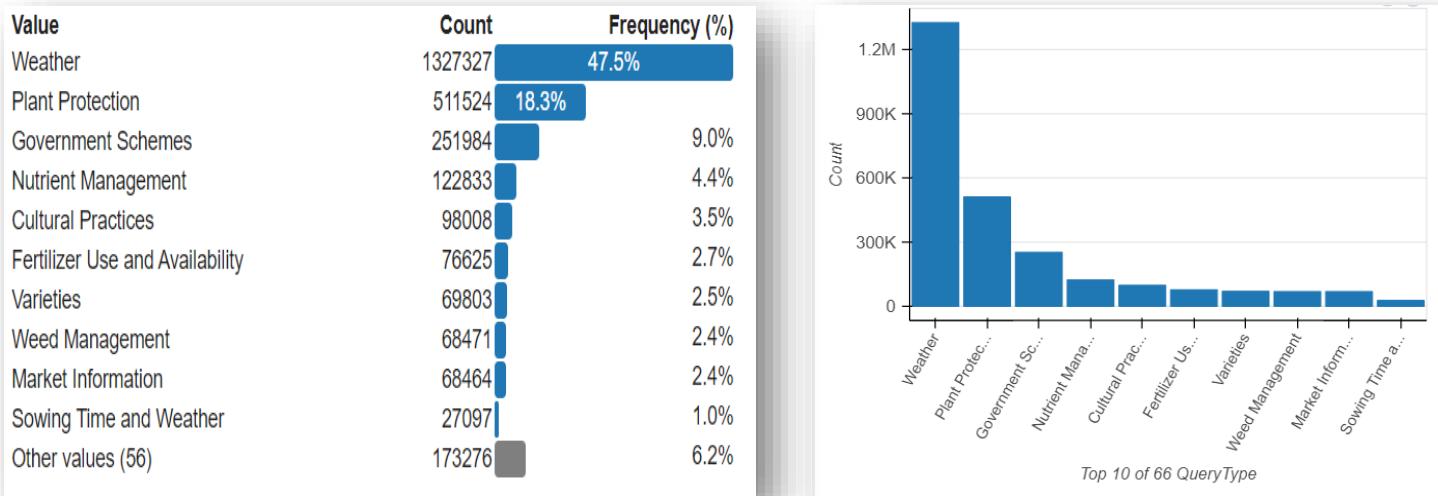
In the given data, we have 66 different query types of which we selected 4 query types which are common to all the 10 states. The number of queries in these four query types with respect to 10 selected states can be represented using the below table :-

	Weather	Plant Protection	Government Schemes	Cultural Practices
West Bengal	53093	40091	4799	4084
UP	489627	146055	104423	30527
Punjab	160983	66388	6742	11756
Gujarat	70133	59016	35596	21462
Haryana	221726	44429	15211	7745
MP	168824	76269	49657	8790
Assam	2008	3704	2068	908
AP	56383	33705	9846	2744
Karnataka	90594	24073	20810	5209
Chhattisgarh	13956	17794	2832	4782

From the above table, it can clearly seen that:-

- All the states have maximum number of queries in mainly two Query types :- WEATHER & PLANT PROTECTION.
- UP have maximum number of queries for WEATHER than other states.

Now, by combining the data of all the 10 states , we have results which can be represented using bar chart as follows:-



**INTERPRETATION:-** As represented in the bar chart, maximum queries 47.5% are for weather , followed by plant production with 18.3 % ,nutrient management with 4.4 % and govt. schemes with 9.0% . These 4 categories accounts for total of 79.2% queries originated. The remaining 20.8% queries comprises of other 62 categories.

### Chi-Square Goodness of Fit Test:

Null Hypothesis: There is no query-types wise variation in different states.

Alternate Hypothesis: There is query-types wise variation in different states.

p-value= 0.0

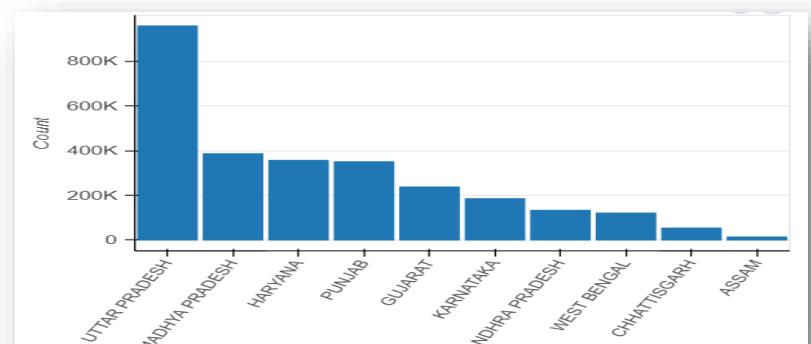
**Result:** p-value < 0.05. Therefore, we conclude that state wise variation in different query-types has been significantly at variance.

### **3.1.5 STATE-WISE COMPARISON**

The different states along with their count of queries can be shown by using the **Value Table** as :-

<b>Value</b>	<b>Count</b>
UTTAR PRADESH	957700
MADHYA PRADESH	385962
HARYANA	356427
PUNJAB	350397
GUJARAT	237461
KARNATAKA	185312
ANDHRA PRADESH	133211
WEST BENGAL	120995
CHHATTISGARH	53821
ASSAM	14126

Above mentioned values can be represented in the form of bar chart as follows:-



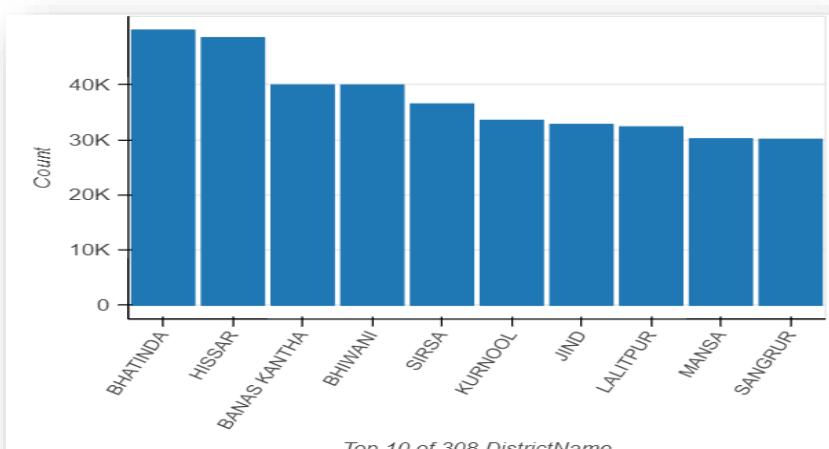
**INTERPRETATION**:-Out of all 10states, as represented in bar chart ,UP has maximum queries with count 957700 , followed by MADHYA PRADESH with count 2.48 times smaller than UP's count. Further , Punjab & Haryana faced nearly equal number of queries in year 2019. ASSAM have minimum number of queries in comparison to all other states.

### **3.1.6 DISTRICT-WISE COMPARISON**

By combining the data of all the 10 states , the district names mentioned in the data along with their count of queries can be shown using **Value Table** as :-

<b>Value</b>	<b>Count</b>
BHATINDA	49906
HISSAR	48530
BANAS KANTHA	39973
BHIWANI	39956
SIRSA	36522
KURNOOL	33552
JIND	32812
LALITPUR	32365
MANSA	30199
SANGRUR	30118
Other values (298)	2421479

In the above table , we have shown top 10 districts of all 308 districts present in the given dataand these values can be represented using **Bar chart** as-

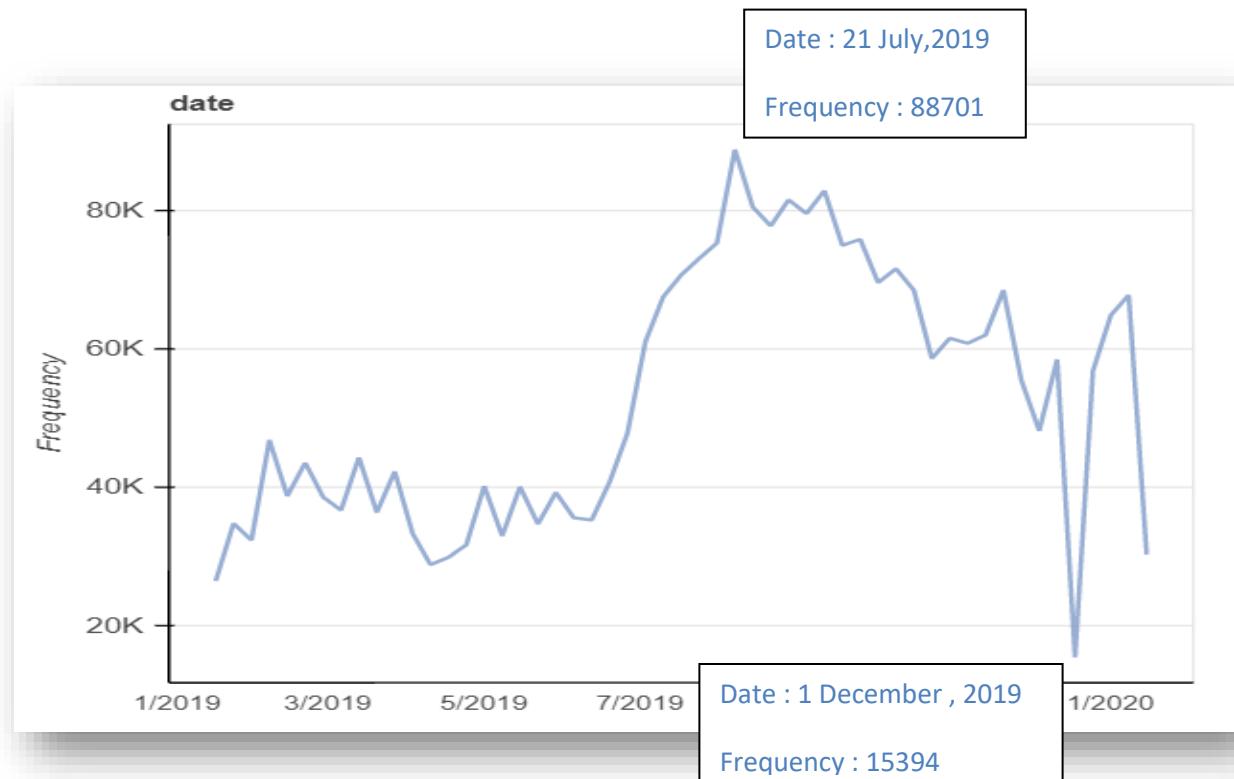


**INTERPRETATION:-**Out of all 308districts, as represented in bar chart ,BHATINDA(PB) has maximum queries with count 49906 , followed by HISSAR(HR)

with slighter less count than BHATINDA's count. Further , BANAS KANTHA(GJ) &BHIWANI(HR) faced equal number of queries in year 2019.

### **3.1.7 DATE-WISE QUERIES**

The change in the number of queries according as different dates of the year 2019 can be represented using a **LINE CHART** as:-

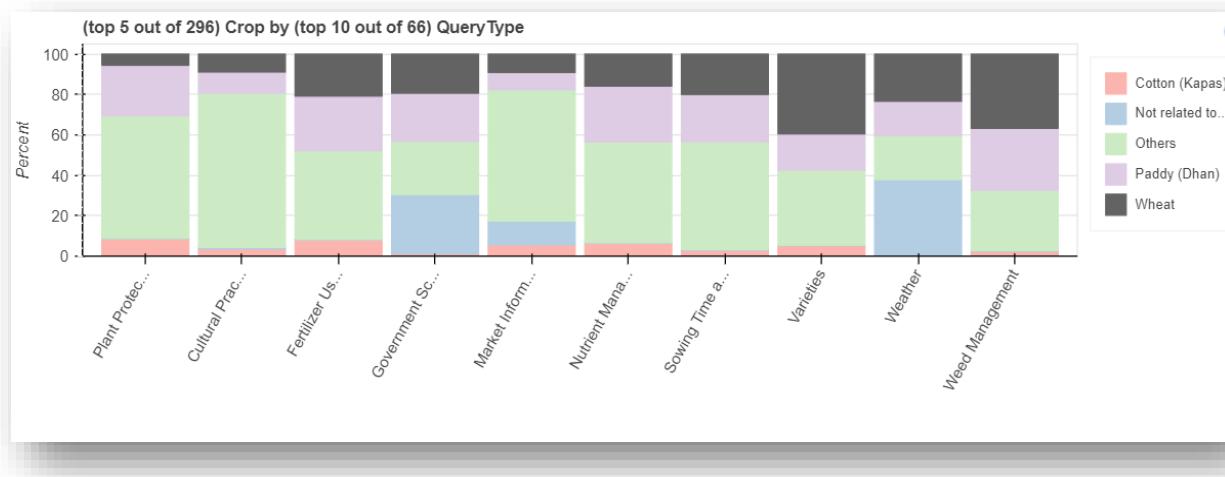


**INTERPRETATION :-** It can be clearly seen from the above line chart that the highest queries are in the middle of July , whereas lowest queries are in the

starting of December month which are approximately 6 times less than that in the month of July. A decreasing trend can be seen in the number of queries from July to December.

### **3.1.8 COMBINATION OF QUERY TYPE & CROP**

The percent of queries across two categorical variables : Query type and Crop type



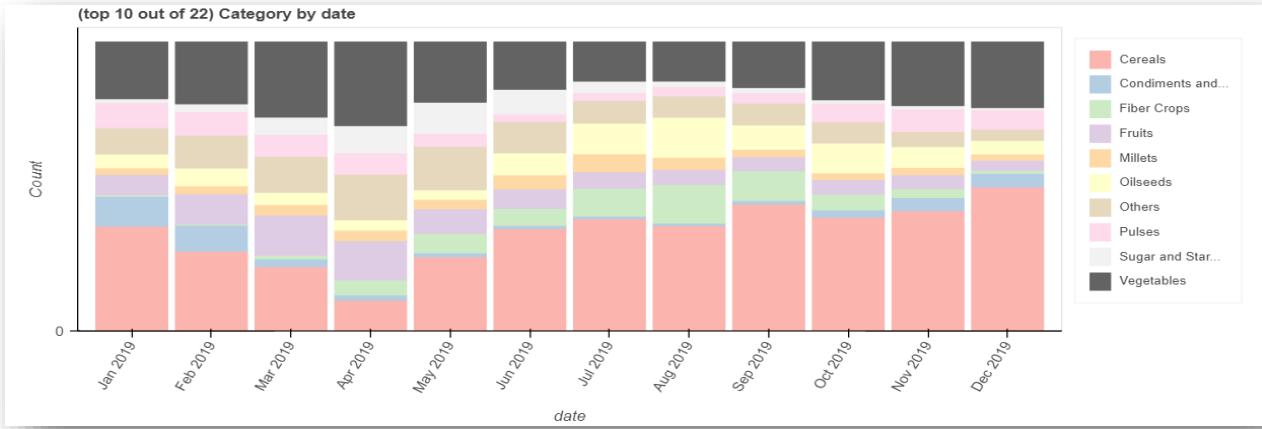
can be shown using the below **Stacked Bar Chart** as:-

In the above chart , we have shown top 5 crops of all 296 crops by top 10 query types of all 66 query types present in the given data .

**INTERPRETATION :-** The maximum queries in Wheat crop are related to Varieties & weed management, whereas in Paddy & Cotton , these are related to Plant protection & Nutrient management. The negligible queries about Weather & Government schemes in Cotton crop .

### **3.1.9 COMBINATION OF DATE & CATEGORY**

The count of queries across two categorical variables : Date and Category type can



be shown using the below **Stacked Bar Chart** as:-

**INTERPRETATION :-** During the year 2019, the maximum number of queries are related to Cereals, followed by Vegetables , Fruits & Oilseeds. However, there are very less queries about Millets during the whole year.

## 3.2 Season-wise analysis of overall data

India has three cropping seasons — *Rabi*, *Kharif* and *Zaid*.

In this section, we are analyzing the farmer queries on the basis of three different seasons.

We are using Pie charts or Bar graphs to represent the percentage or no. of queries in different categories. Also, for state-wise analysis we are giving a table of no. of queries of different states with respect to different categories. We've built this table in taking account the top 5 no. of queries from each state.

### **3.2.1 KHARIF SEASON**

- **Category :-** The no. of queries from different types of categories present in the Kharif season are :-

**Overall analysis (Combined data of all the states)**

Value represents different Categories

Value	Count
Others	888851
Cereals	244361
Vegetables	98749
Oilseeds	65241
Fiber Crops	57993
Fruits	33654
Millets	24235
Pulses	23995
Sugar and Starch Crops	18800
Govt.schemes/weather/ferti...	12184
Other values (12)	46080

Count represents no. of queries

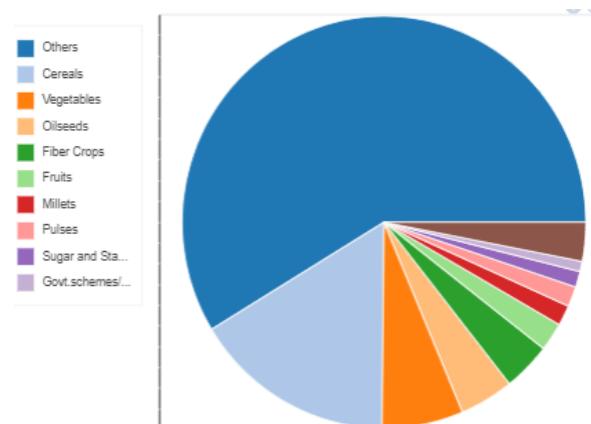


Figure (1): Top 10 of all 22 categories present in Kharif season

\*Others – Particular Category is not mentioned in the given data.

**State-wise analysis:** In the following table, no. of queries in 10 different states (in rows) of different categories (in columns) is given.

(This table is made by taking in account only top 5 no. of queries from each state)

	Cereals	Oilseeds	Vegetables	Fiber Crops	Fruits
West Bengal	11,191		11,429	–	2,787
Uttar Pradesh	99,049	13,859	34,729	–	–
Punjab	73,391	–	7,606	8,689	–
Gujarat	–	20,228	5,942	17,845	3,285
Haryana	20,809	4,325	6,084	19,729	–
Madhya Pradesh	18,144	22,139	11,669	–	–
Assam	709	–	1,380	–	416
Andhra Pradesh	5,305	–	7,859	5,204	2,191
Karnataka	–	–	6,743	3,114	–
Chhattisgarh	12,376	–	5,108	–	606

Interpretation:

Overall analysis of Category: Highest numbers of queries are from CEREALS category. (Ref Figure (1))

State-wise analysis of Category: We see that the three main categories that have the highest queries in these ten states were Cereals, vegetables and Oilseeds. The highest no. of queries about Cereals and Vegetables are from Uttar Pradesh (99,049 and 34,729 respectively) and Oilseeds from West Bengal (22,139).

- Crop :- The no. of queries from different types of crops present in the Kharif season are :-

### Overall analysis

Value represents different Crops

Count represents no. of queries

Value	Count
Others	857499
Paddy (Dhan)	224673
Cotton (Kapas)	57526
Not related to crop	46840
Soybean (bhat)	21533
Sugarcane (Noble Cane)	18597
Wheat	18259
Groundnut (pea nut/mung ...	17792
Chillies	15723
Maize (Makka)	14919
Other values (281)	220782

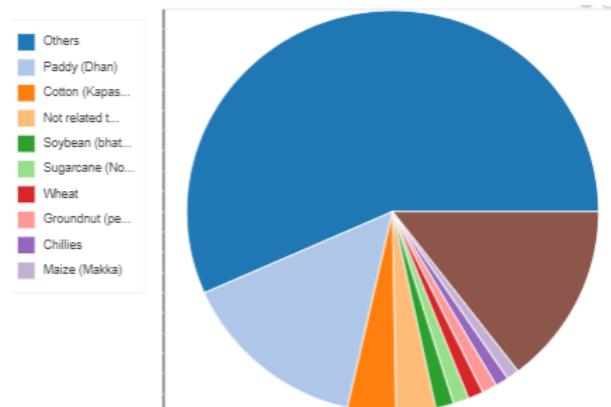


Figure (2): Top 10 of all 291 crops present in Kharif season

\*Others – Particular Crop is not mentioned in the given data.

**State-wise analysis:** In the following table, no. of queries in 10 different states (in rows) of different crops (in columns) is given.

(This table is made by taking in account only top 5 no. of queries from each state)

	Paddy	Cotton	Soybean	Wheat	Maize
West Bengal	11,130	—	—	—	—
Uttar Pradesh	94,270	—	—	—	—
Punjab	67,088	8,686	—	6,025	—
Gujarat	—	17,828	—	—	—
Haryana	17,912	19,559	—	—	—
Madhya Pradesh	13,629	—	19,815	4,483	5,074
Assam	698	—	—	—	—
Andhra Pradesh	5,265	5,195	—	—	—
Karnataka	—	3,109	—	—	2,411
Chhattisgarh	12,217	—	—	—	—

#### Interpretation:

Overall analysis of Crops: Paddy has the highest queries among crops with 14.8% (no. of queries=224673) & then cotton with 3.8% (no. of queries= 57526), whereas all other crops have less than 2% out of total queries.(Ref Figure (2))

State-wise analysis of crops: Paddy, Cotton and Soybean have most no. of queries in these states. The highest no. of queries about Paddy is from Uttar Pradesh (94,270), Cotton from Haryana (19,559) and Soybean from Madhya Pradesh (19,815).

- **Query Type :-** The no. of queries from different query types present in the Kharif season are :-

## Overall analysis

Value represents different Query type

Count represents no. of queries

Value	Count
Weather	762109
Plant Protection	295309
Government Schemes	121381
Nutrient Management	64982
Cultural Practices	43367
Fertilizer Use and Availability	41663
Weed Management	33853
Varieties	30792
Market Information	27546
Sowing Time and Weather	13491
Other values (56)	79650

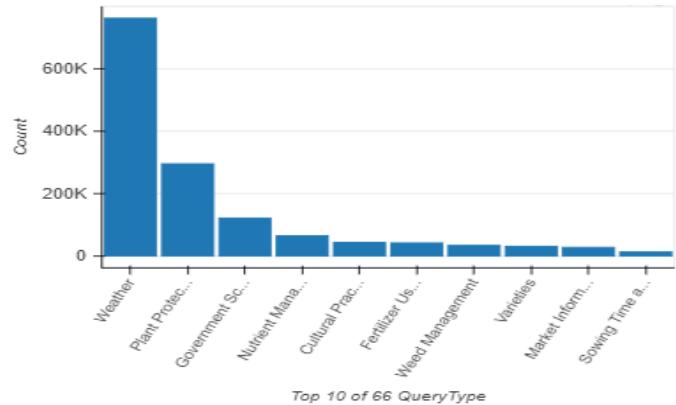


Figure (3): Top 10 of all 66 query types present in Kharif season

**State-wise analysis:** In the following table, no. of queries in 10 different states (in rows) of different query types (in columns) is given.

(This table is made by taking in account only top 5 no. of queries from each state)

	Weather	Plant Protection	Govt. schemes	Nutrient Management	Cultural Practices
West Bengal	22,794	18,340	3,090	3,083	1,649
Uttar Pradesh	269,902	86,169	49,979	19,131	—
Punjab	82,375	50,261	—	16,361	5,250
Gujarat	55,128	30,640	16,456	—	8,704
Haryana	119,610	32,443	6,380	5,954	—

Madhya Pradesh	116,193	37,783	22,248	6,602	—
Assam	—	1,631	1,668	571	429
Andhra Pradesh	34,282	13,815	5,052	5,285	—
Karnataka	52,498	13,003	11,425	5,500	—
Chhattisgarh	8,909	11,224	1,693	1,741	2,005

### Interpretation:

Overall analysis of Query Type: Maximum queries are related to WEATHER, followed by Plant protection, followed by Govt. schemes. These 3 categories accounts for 77.8% queries originated.(Ref Figure (3))

State-wise analysis of Query Type: Mostly all the states have top queries about Weather except for Assam and Chhattisgarh which have highest no. of queries about Govt. schemes and Plant Protection respectively. The highest no. of queries about Weather is from Uttar Pradesh (2, 69,902)

- **Date :-** The change in the number of queries according as different dates of Kharif season can be represented as :-



Figure (4): Line Chart of queries in Kharif season

Interpretation:In Kharif Season, it can be clearly seen that:

The highest no. of queries is in the month of September, 2019 and the lowest no. of queries is at the end of August, 2019. (Ref Figure (4))

- **State Name :-** The no. of queries from different states in the Kharif season are :-

Value represents different States

Count represents no. of queries

Value	Count
UTTAR PRADESH	514757
MADHYA PRADESH	217264
HARYANA	194004
PUNJAB	193034
GUJARAT	134211
KARNATAKA	99235
ANDHRA PRADESH	68847
WEST BENGAL	54989
CHHATTISGARH	30906
ASSAM	6896

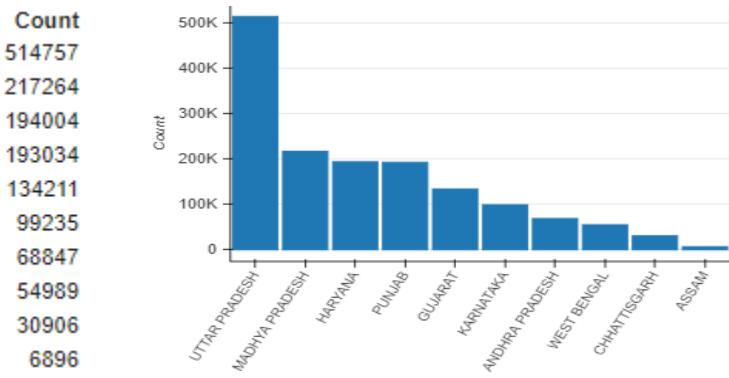


Figure (5): All the 10 states

Interpretation: Uttar Pradesh state has the maximum no. of queries (34.4%), followed by Madhya Pradesh (14.3%), followed by Haryana (12.8%) in Kharif season. (Ref Figure (5))

- **District Name :-** The no. of queries from different districts of the ten states in the Kharif season are :-

Value represents different districts

Count represents no. of queries

Value	Count
BHATINDA	30186
HISSAR	29395
BANAS KANTHA	24655
KURNOOL	22954
BHIWANI	22556
SIRSA	21579
LALITPUR	19436
JIND	19217
MANSA	18964
JUNAGADH	16842
Other values (298)	1288359

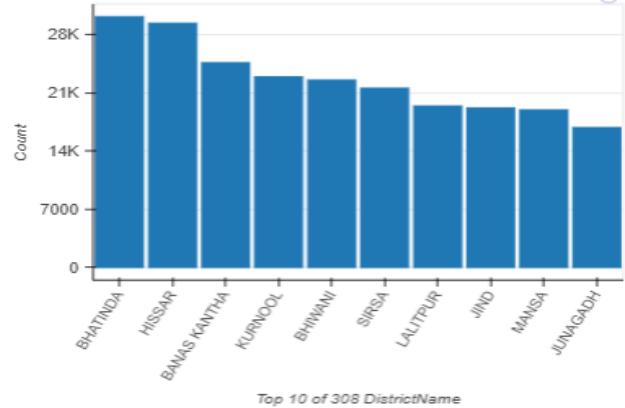


Figure (6): Top 10 of all 308 districts of the ten states

**Interpretation:** BHATINDA district (Punjab) had maximum queries 30186 (2.0%) in Kharif season, followed by HISSAR district (Haryana) with 29395 (1.9%) and BANAS KANTHA (Gujarat) with 25655 (1.6%) no. of queries. (Ref Figure (6))

- **Query Type & Crop :-** The percent of queries across two categorical variables : Query type and Crop type in Kharif season are

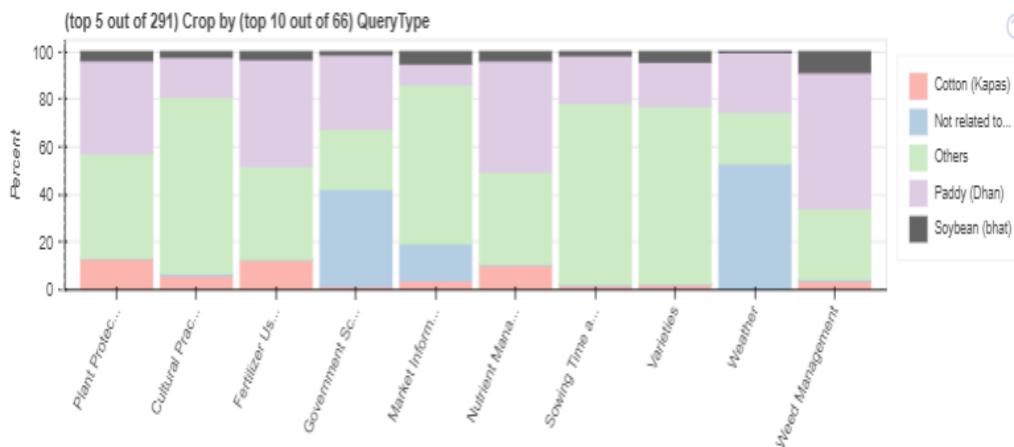


Figure (7): Top (5 out of 291) Crops by ( top 10 out of 66) Query Types

- **Interpretation:** Maximum queries are about Weed Management, Nutrient Management & Fertilizer use and availability in Paddy crop, whereas in case of Soybean most queries are about Weed Management, Market Information and Plant Protection and in case of cotton these are Fertilizer use and availability, Plant Protection and Nutrient Management in Kharif season. Negligible queries about Government schemes, Varieties and Weather in Cotton crop. (Ref (7))
- **Date & Category :-** The count of queries across two categorical variables : Date and Category type in Kharif season are:-

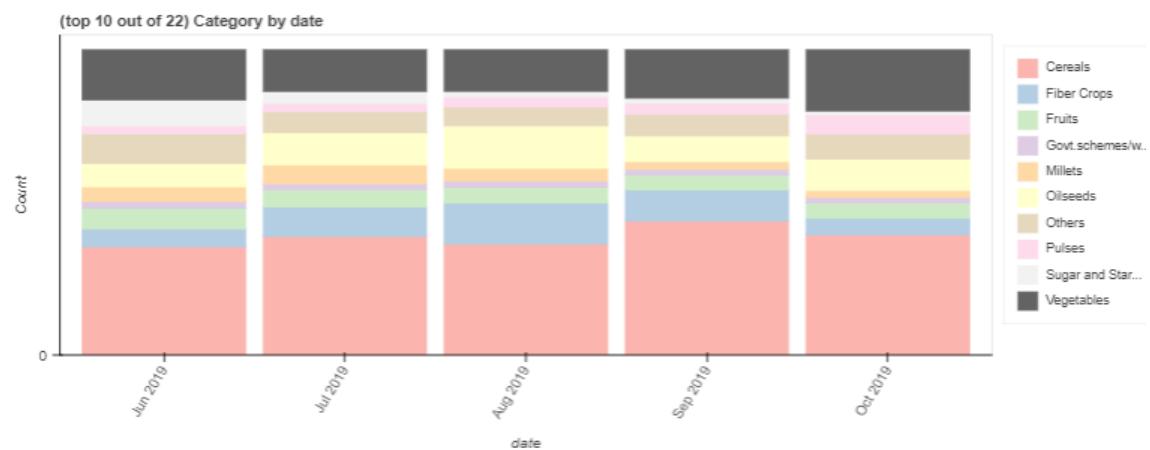


Figure (8): Top (10 out of all 22) categories by different months of Kharif season

**Interpretation:** In Kharif Season, it can be clearly seen that:

Maximum numbers of queries are related to Cereals are highest in September, 2019, followed by Vegetables which are highest in October, 2019 and Oilseeds which are highest in August. However, there are very less queries about Millets in the Kharif season. (Ref Figure (8))

## 3.2.2 RABI SEASON

- Category :-** The percent of queries from different types of categories present in the Rabi season are :-

### Overall analysis

Value represents different Categories

Count represents no. of queries

Value	Count
Others	518591
Cereals	178455
Vegetables	100785
Pulses	34740
Fruits	29641
Condiments and Spices	26281
Oilseeds	25198
Millets	11508
Sugar and Starch Crops	8416
Animal	7647
Other values (12)	30795

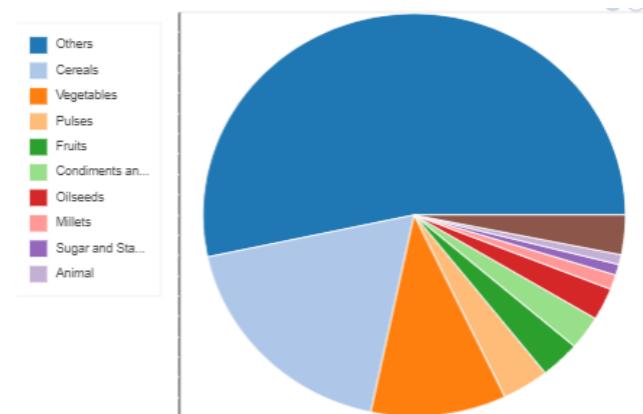


Figure (1): Top 10 of all 22 categories present in Rabi season

\*Others – Particular Category is not mentioned in the given data.

**State-wise analysis:** In the following table, no. of queries in 10 different states (in rows) of different categories (in columns) is given.

(This table is made by taking in account only top 5 no. of queries from each state)

	Cereals	Oilseeds	Vegetables	Pulses	Condiment and spices
West Bengal	7,421	1,596	13,296	–	–
Uttar Pradesh	70,777	8,927	35,232	–	–
Punjab	35,746	–	4,562	–	–

Gujarat	7,014	5,271	7,043	—	14,847
Haryana	19932	2835	4361	—	—
Madhya Pradesh	26,689	—	13,873	12,503	6,446
Assam	288	—	1,478	—	—
Andhra Pradesh	5,547	—	10,589	3,399	—
Karnataka	—	—	5,770	2,533	—
Chhattisgarh	3,969	—	4,584	1,669	—

### Interpretation:

Overall analysis of Category: Highest numbers of queries are from CEREALS category. (Ref Figure (1))

State-wise analysis of Category: We see that the three main categories that have the highest queries in these ten states are Cereals, Vegetables and Condiment and spices. The highest no. of queries about Cereals and Vegetables are from Uttar Pradesh (70,777 and 35,232 respectively), Vegetables from Uttar Pradesh (20,116), Condiment and spices from Gujarat (14,847).

- **Crop :-** The percent of queries from different types of crops present in the Rabi season are :-

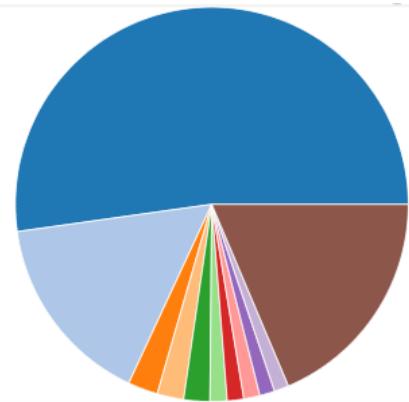
### **Overall analysis**

Value represents different Crop

Count represents no. of queries

Value	Count
Others	507547
Wheat	154001
Paddy (Dhan)	23965
Potato	21248
Bengal Gram (Gram/Chick ...	20877
Mustard	13614
Onion	13220
Not related to crop	12888
Chillies	12026
Cumin	11670
Other values (285)	181001

Fig  
ur  
e  
(2)  
:  
To  
p  
10



of all 295 crops present in Rabi season

\*Others – Particular Crop is not mentioned in the given data.

**State-wise analysis:** In the following table, no. of queries in 10 different states (in rows) of different crops (in columns) is given.

(This table is made by taking in account only top 5 no. of queries from each state)

	Paddy	Wheat	Potato	Cumin	Maize
West Bengal	7,335	–	3,591	–	–
Uttar Pradesh	–	64,756	13,665	–	–
Punjab	–	35,577	1,106	–	–
Gujarat	–	5,677	–	11,397	–
Haryana	–	19,790	–	–	–
Madhya Pradesh	–	25,995	–	–	–
Assam	265	–	254	–	–
Andhra Pradesh	5,514	–	–	–	–
Karnataka	–	–	–	–	1,655
Chhattisgarh	2,479	1,477	–	–	–

## Interpretation:

Overall analysis of Crops: Wheat has the highest queries among crops with 15.8% ( no. of queries= 1,54,001) & then Paddy with 2.5% (no. of queries= 23,965), followed by Potato (2.2%), followed by Bengal Gram (2.1%), whereas all other crops have less than 2% out of total queries.(Ref Figure (2))

State-wise analysis of crops: Paddy, Wheat, Potato have most no. of queries in these states. The highest no. of queries about Paddy is from West Bengal (7,355), Wheat from Uttar Pradesh (64,756) and Potato from Uttar Pradesh (13,665). However, Maximum no. of queries from Gujarat and Karnataka are for Cumin and Maize respectively.

- **Query Type :-** The count of queries from different query types present in the Rabi season are :-

### Overall analysis

Value represents different Query type

Count represents no. of queries

Value	Count
Weather	417367
Plant Protection	167427
Government Schemes	106230
Nutrient Management	46371
Cultural Practices	41713
Market Information	30321
Weed Management	29864
Fertilizer Use and Availability	28800
Varieties	26545
Seeds and Planting Material	9051
Other values (56)	68368

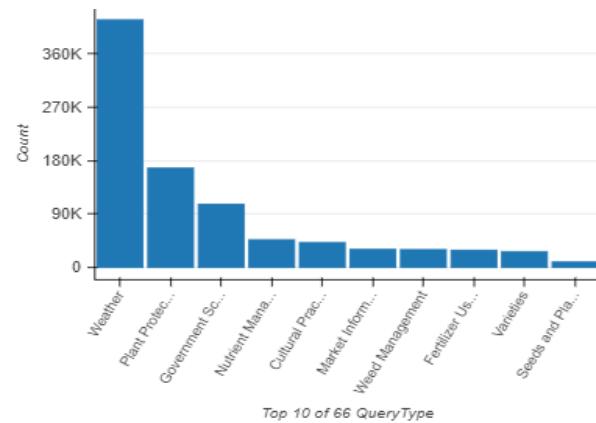


Figure (3): Top 10 of all 66 query types present in Rabi season

- **State-wise analysis:** In the following table, no. of queries in 10 different states (in rows) of different query types (in columns) is given.

(This table is made by taking in account only top 5 no. of queries from each state)

	Weather	Plant Protection	Govt. schemes	Nutrient Management	Cultural Practices
West Bengal	20,363	17,108	1,581	3,472	2,141
Uttar Pradesh	1,60,760	37,999	44,125	11,164	11,813
Punjab	58,975	11,182	—	8,225	—
Gujarat	12,968	26,074	13,806	—	10,833
Haryana	73,874	8,027	6,953	4,175	—
Madhya Pradesh	43,414	33,360	23,104	7,512	—
Assam	844	1,359	—	536	386
Andhra Pradesh	17,516	18,239	3,529	5,745	—
Karnataka	24,995	9,125	9,060	3,591	—
Chhattisgarh	3,657	4,954	939	1,306	2,158

#### Interpretation:

Overall analysis of Query Type: Maximum queries are related to WEATHER, followed by Plant protection, followed by Govt. schemes. These 3 categories accounts for 71% queries originated. (Ref Figure (3))

State-wise analysis of Query Type: Mostly all the states have top queries about Weather except for Gujarat, Andhra Pradesh, Assam and Chhattisgarh which have highest no. of queries about Plant Protection. The highest no. of queries about Weather is Uttar Pradesh (1, 60,760).

- **Date :-** The change in the number of queries according as different dates of Rabi season can be represented as :-



Figure (4): Line Chart of queries in Rabi season

### Interpretation:

We study only the months of Rabi season so we don't consider any query between June to October. The highest no. of queries is in the month of November, 2019 and the lowest no. of queries is at the end of December, 2019. (Ref Figure (4))

- **State Name :-** The count of queries from different states in the Rabi season are :-

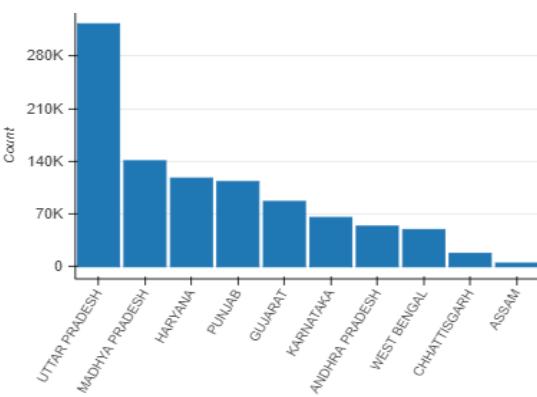
Value represents different States

Value
UTTAR PRADESH
MADHYA PRADESH
HARYANA
PUNJAB
GUJARAT
KARNATAKA
ANDHRA PRADESH
WEST BENGAL
CHHATTISGARH
ASSAM

Count represents no. of queries

Count
322749
140716
117722
113012
86731
65499
53869
49370
17616
4773

Figure (5): All the 10 states



### Interpretation

**Uttar Pradesh** state has the maximum no. of queries (33.2%), followed by **Madhya Pradesh** (14.5%), followed by **Haryana** (12.1%) in Rabi season. (Ref Figure (5))

- **District Name :-** The count of queries from different districts of the ten states in the Rabi season are :-

Value represents different districts

Count represents no. of queries

Value	Count
HISSAR	13299
BHATINDA	13146
BANAS KANTHA	12809
BHIWANI	12363
JUNAGADH	11055
SANGRUR	10524
LALITPUR	10400
SITAPUR	10174
SIRSA	9925
JIND	9712
Other values (298)	858650

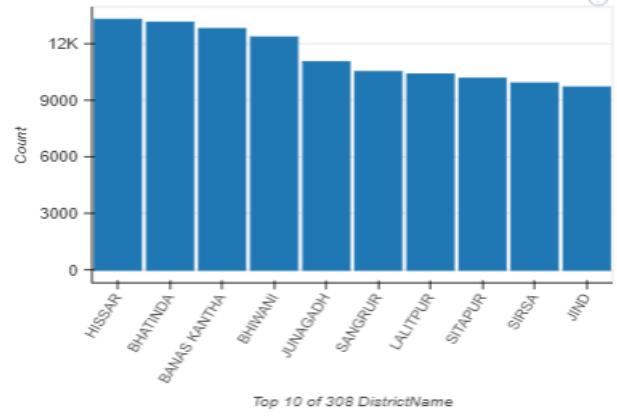


Figure (6): Top 10 of all 308 districts of the ten states

### Interpretation:

HISSAR district (Haryana) and BHATINDA district (Punjab) has approximate same no. of queries (1.4%) in Rabi season, followed by BANAS KANTHA (Gujarat) with 1.3% of total no. of queries. (Ref Figure (6))

- **Query Type & Crop :-** The percent of queries across two categorical variables : Query type and Crop type in Rabi season are :-

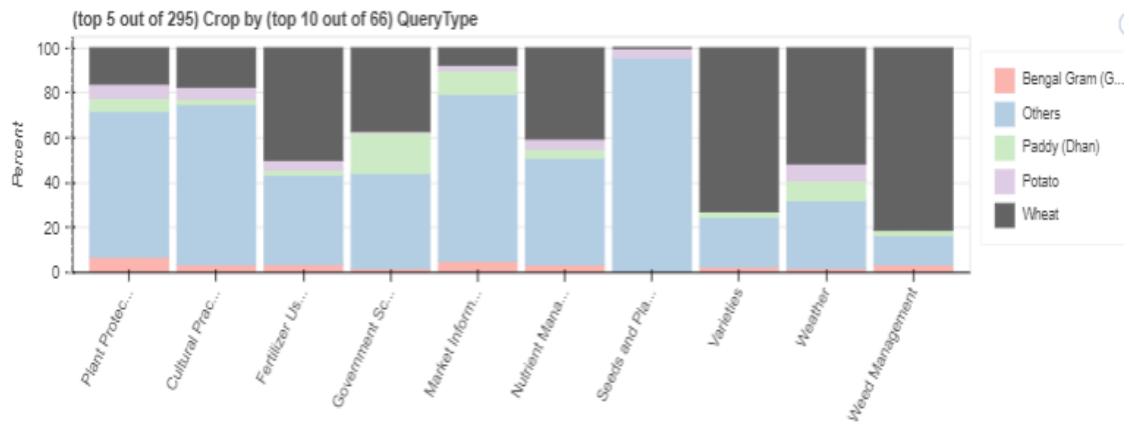


Figure (7): Top (5 out of 295) Crops by (top 10 out of 66) Query Types

### Interpretation:

Maximum queries are about Weed Management, Varieties & Fertilizer use and availability in Wheat crop, whereas in case of Paddy most queries are about Govt. schemes, Market Information and Weather and in case of Potato these are Weather, Plant Protection and Cultural Practices in Rabi season. Negligible queries are about Government schemes, Varieties and Weather in Bengal gram. (Ref Figure (7))

- **Date & Category :-** The count of queries across two categorical variables : Date and Category type in Rabi season are:-

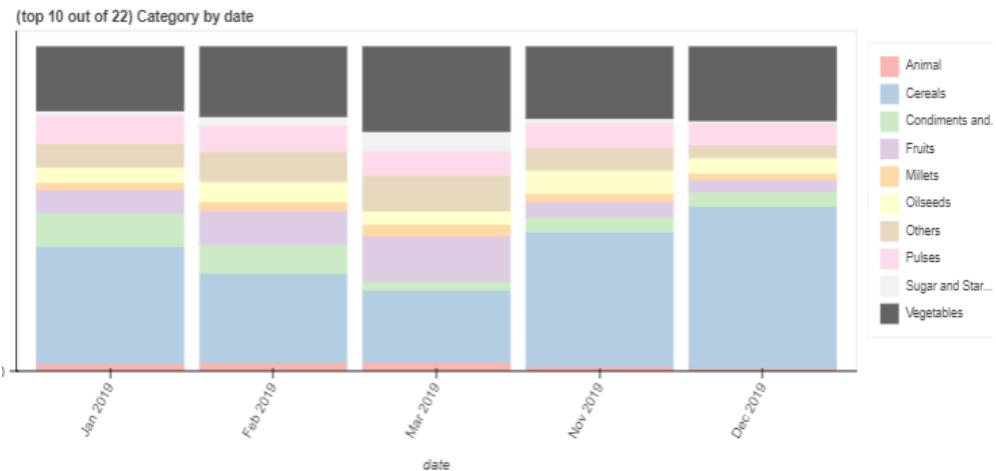


Figure (8): Top (10 out of all 22) categories by different months of Rabi season

### Interpretation:

Maximum numbers of queries are related to Cereals which are highest in December, followed by Vegetables which are highest in May. However, there are very less queries about Millets in the Rabi season. (Ref Figure (8))

### 3.2.3 ZAID SEASON

- **Category :-** The percent of queries from different types of categories present in the Zaid season are :-

#### Overall analysis

Value represents different Categories

Count represents no. of queries

Value	Count
Others	432524
Cereals	66760
Vegetables	58510
Fruits	26746
Sugar and Starch Crops	23317
Pulses	13624
Fiber Crops	13283
Oilseeds	13088
Medicinal and Aromatic Pla...	10674
Millets	10415
Other values (12)	28946

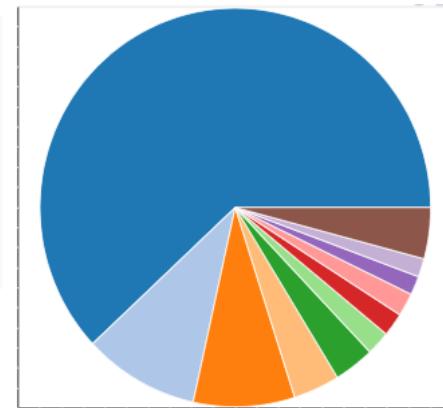


Figure (1): Top 10 of all 22 categories present in Zaid season

\*Others – Particular Category is not mentioned in the given data.

**State-wise analysis:** In the following table, no. of queries in 10 different states (in rows) of different categories (in columns) is given.

(This table is made by taking in account only top 5 no. of queries from each state)

	Cereals	Oilseeds	Vegetables	Fiber Crops	Fruits
West Bengal	3,425	576	6,435	—	2,069
Uttar Pradesh	21,066	—	20,116	—	11,546
Punjab	29,335	—	4,529	2,338	—
Gujarat	—	4,760	3,527	2,370	2,203
Haryana	—	20,228	5,942	17,845	3,285
Madhya Pradesh	2,644	3,586	8,515	—	—
Assam	400	—	1,103	—	379
Andhra Pradesh	1,343	710	3,594	—	1,524
Karnataka	—	—	3,250	—	1,486
Chhattisgarh	2,225	—	3,250	—	557

### Interpretation:

Overall analysis of Category: Highest numbers of queries are from CEREALS category. (Ref Figure (1))

State-wise analysis of Category: We see that the three main categories that have the highest queries in these ten states are Cereals, Vegetables, Oilseeds and Fiber crops. The highest no. of queries about Cereals is from Punjab (29,335), Vegetables from Uttar Pradesh (20,116), Oilseeds from Gujarat (4,760) and Fiber Crops from Haryana (6,412).

- **Crop :-** The percent of queries from different types of crops present in the Zaid season are :-

## Overall analysis

Value represents different Crops

Count represents no. of queries

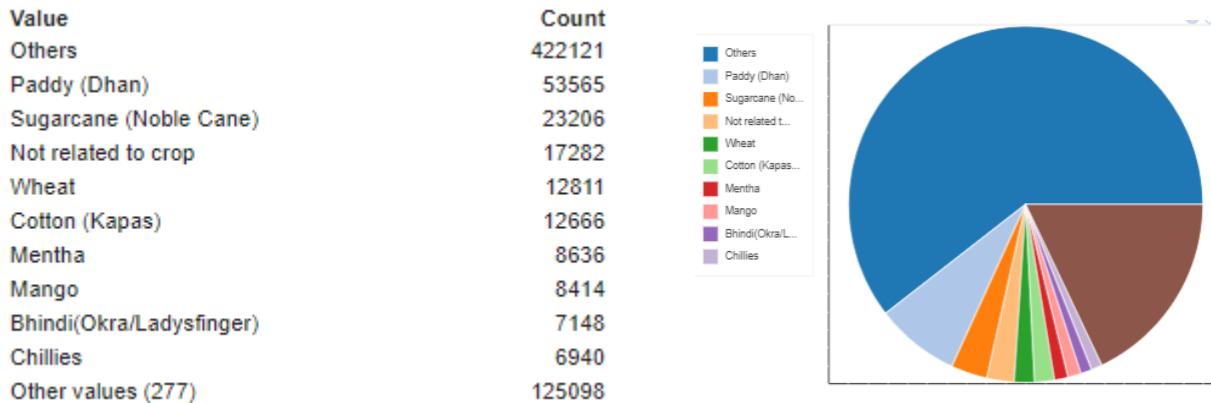


Figure (2): Top 10 of all 287 crops present in Zaid season

\*Others – Particular Crop is not mentioned in the given data.

**State-wise analysis:** In the following table, no. of queries in 10 different states (in rows) of different crops (in columns) is given.

(This table is made by taking in account only top 5 no. of queries from each state)

	Paddy	Cotton	Sugarcane	Chillies	Wheat
West Bengal	3,407	–	–	–	–
Uttar Pradesh	16,630	–	20,009	–	–
Punjab	23,945	2,339	–	–	5,333
Gujarat	–	2,362	–	–	–
Haryana	3,824	6,389	–	–	1,590
Madhya Pradesh	–	–	–	–	–
Assam	389	–	–	297	–
Andhra	1,333	–	–	615	–

Pradesh					
Karnataka	-	-	579	749	-
Chhattisgarh	2095	-	-	-	-

### Interpretation:

Overall analysis of Crops: Paddy has the highest queries among crops with 7.7% (no. of queries= 53565) & then Sugarcane with 3.3% (no. of queries= 23206), followed by Potato (2.5%), whereas all other crops have less than 2% out of total queries.(Ref Figure (2))

State-wise analysis of crops: Paddy, Cotton, Sugarcane, and Chillies have most no. of queries in these states. The highest no. of queries about Paddy is from Punjab (23,945), Cotton from Haryana (6,389), Sugarcane from Uttar Pradesh (20,009) and Chillies from Karnataka (749).

- **Query Type :-** The count of queries from different query types present in the Zaid season are :-

### Overall analysis

Value represents different Query type

Count represents no. of queries

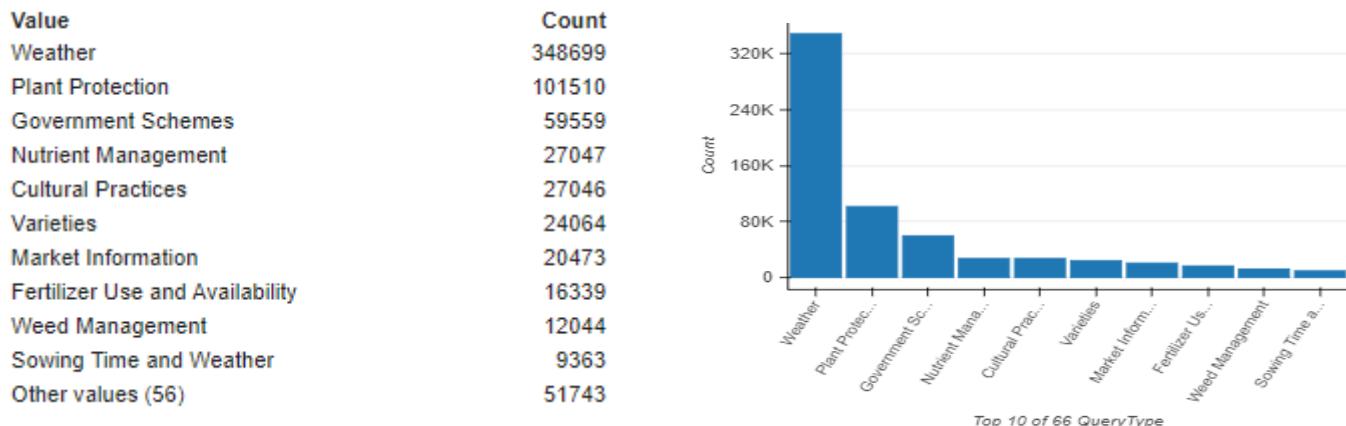


Figure (3): Top 10 of all 66 query types present in Zaid season

**State-wise analysis:** In the following table, no. of queries in 10 different states (in rows) of different query types (in columns) is given.

(This table is made by taking in account only top 5 no. of queries from each state)

	Weather	Plant Protection	Govt. schemes	Nutrient Management	Cultural Practices
West Bengal	19,430	9,783	504	1,569	724
Uttar Pradesh	1,37,134	42,396	24,534	9,906	10,683
Punjab	46,955	11,893	—	6,723	—
Gujarat	10,723	6,226	12,125	—	4,743
Haryana	62,865	8,209	5,767	2,135	—
Madhya Pradesh	31,140	9,937	10,424	—	—
Assam	1,181	1,371	349	357	—
Andhra Pradesh	10,208	4,494	3,009	1,886	—

Karnataka	25,973	4,026	–	1,686	1,476
Chhattisgarh	3,090	3,175	–	628	1,323

### Interpretation:

Overall analysis of Query Type: Maximum queries are related to WEATHER, followed by Plant protection, followed by Govt. schemes. These 3 categories accounts for 73% queries originated.(Ref Figure (3))

State-wise analysis of Query Type: Mostly all the states have top queries about Weather except for Gujarat, Assam and Chhattisgarh which have highest no. of queries about Govt. schemes and Plant Protection. The highest no. of queries about Weather is Uttar Pradesh (1, 37,134).

- **Date :-** The change in the number of queries according as different dates of Zaid season can be represented as :-



Figure (4): Line Chart of queries in Zaid season

### Interpretation:

The highest no. of queries was in the month of July, 2019 and the lowest no. of queries was in the month of April, 2019. An increasing trend can be seen from June to July, 2019. (Ref Figure (4))

- **State Name :-** The count of queries from different states in the Zaid season are :-

Value represents different States

Count represents no. of queries

Value	Count
UTTAR PRADESH	267558
PUNJAB	98541
HARYANA	96917
MADHYA PRADESH	71987
GUJARAT	46267
KARNATAKA	41285
WEST BENGAL	34358
ANDHRA PRADESH	24500
CHHATTISGARH	11630
ASSAM	4844

Value	Count
UTTAR PRADESH	267558
PUNJAB	98541
HARYANA	96917
MADHYA PRADESH	71987
GUJARAT	46267
KARNATAKA	41285
WEST BENGAL	34358
ANDHRA PRADESH	24500
CHHATTISGARH	11630
ASSAM	4844

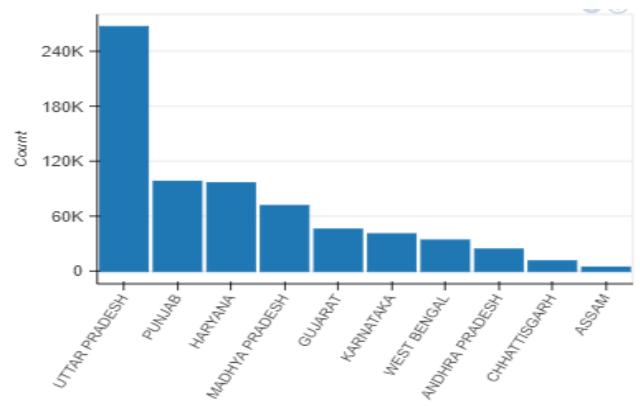


Figure (5): All the 10 States

#### Interpretation:

**Uttar Pradesh** state has the maximum no. of queries (38.3%), followed by **Punjab** (14.1%), followed by **Haryana** (13.9%) in Zaid season. (Ref Figure (5))

- **District Name :-** The count of queries from different districts of all the ten states in the Zaid season are :-

Value represents different districts

Count represents no. of queries

Value	Count
BHATINDA	14992
HISSAR	12478
SIRSA	10747
SITAPUR	9857
BHIWANI	9846
BAREILLY	9386
JIND	8767
BADAUN	8335
MANSA	8258
SANGRUR	7769
Other values (298)	597452

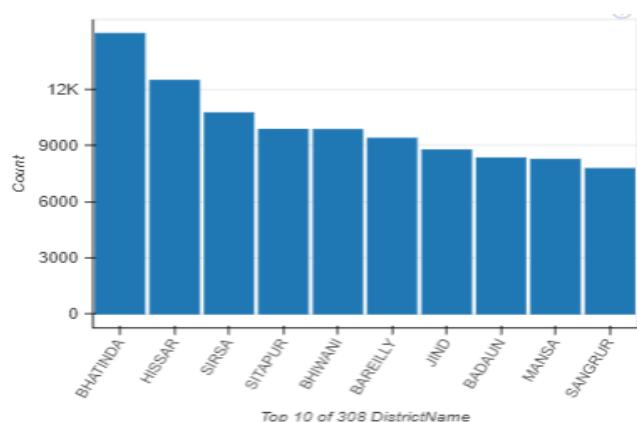


Figure (6): Top 10 of all 308 districts of the ten states

### Interpretation:

BHATINDA district (Punjab) has maximum no. of queries in Zaid season (2.1%), followed by HISSAR district (Haryana) with 1.8%, and followed by SIRSA (Haryana) with 1.5% of total no. of queries. (Ref Figure (6))

- **Query Type & Crop :-** The percent of queries across two categorical variables : Query type and Crop type in Zaid season are :-

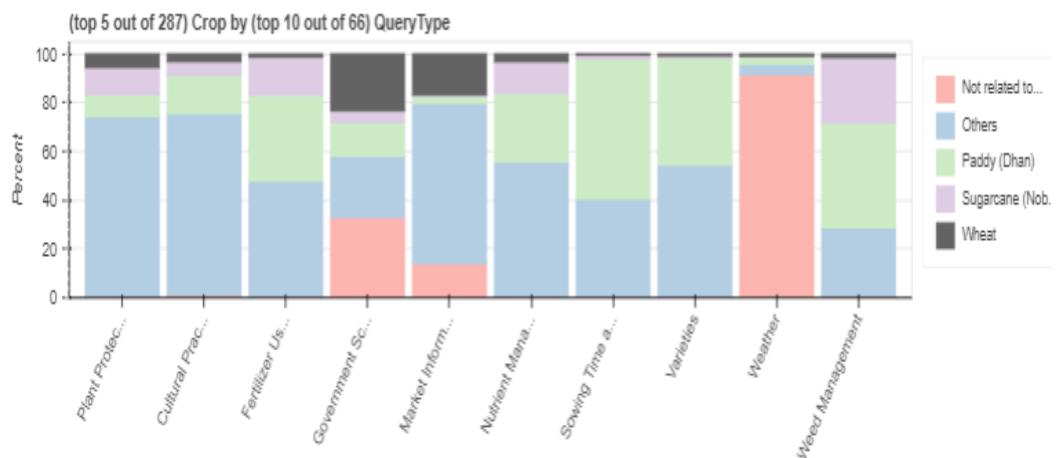
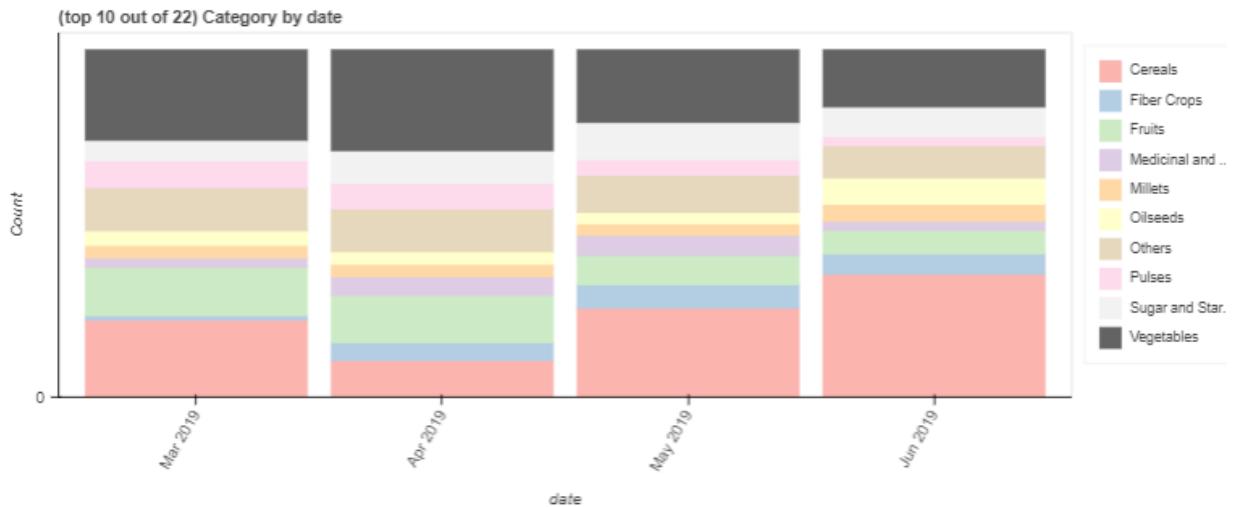


Figure (7): Top (5 out of 287) Crops by (top 10 out of 66) Query Types

### Interpretation:

Maximum queries are about Weed Management, Sowing time & Varieties in Paddy crop, whereas in case of Wheat most queries are about Govt. schemes, Market Information and Plant Protection and in case of Sugarcane these are Weed Management, Nutrient Management, Fertilizer use and availability in Zaid season. (Ref Figure (7))

- **Date & Category :-** The count of queries across two categorical variables : Date and Category type in Zaid season are:-



**Figure (8):** Top (10 out of all 22) categories by different months of Zaid season

### Interpretation:

Maximum numbers of queries are related to Cereals which are highest in June, 2019, followed by Vegetables which are highest in August, 2019. However, there are very less queries about Millets in the Zaid season. (Ref Figure (8))

### 3.3 Sector-wise analysis of overall data

In this section, we are analyzing the farmer queries on the basis of four different sectors.

We are using Pie charts or Bar graphs to represent the percentage or no. of queries in different categories. Also, for state-wise analysis we are giving a table of no. of queries of different states with respect to different categories. We've built this table in taking account the top 5 no. of queries from each state.

#### 3.3.1 AGRICULTURE

- Category :-The percent of queries from different types of categories present in the Agriculture sector are :-

**State-wise analysis:** In the following table, no. of queries in 10 different states (in rows) of different categories (in columns) is given.

(This table is made by taking in account only top 5 no. of queries from each state)

	Cereal	Oilseed	Fiber crop	Pulses	Millets
Assam	1164	190	-----	----	97
Andhra Pradesh	11132	3814	7349	4853	-----
Chhattisgarh	17070	932	-----	2342	730
PUNJAB	121835	----	10391	-----	2888
HARYANA	42579	7609	24420	-----	-----

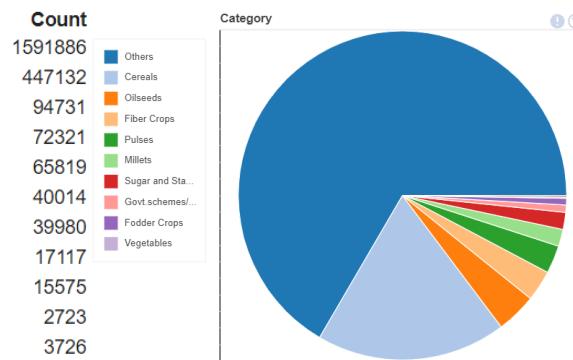
GUJURAT	9448	26874	21516	6902	-----
KARNATAKA		2320	4058	5501	6065
MADHAYA PRADESH	45811	25147	-----	19895	6689
UTTAR PRADESH	176251	23402	-----	19936	-----
WEST BENGAL	19556	2224	-----	1087	-----

### Overall analysis (Combined data of all the states)

Value represents different Categories

Count represents no. of queries

Value	Count
Others	1591886
Cereals	447132
Oilseeds	94731
Fiber Crops	72321
Pulses	65819
Millets	40014
Sugar and Starch Crops	39980
Govt.schemes/weather/ferti...	17117
Fodder Crops	15575
Vegetables	2723
Other values (11)	3726



\*Others – Particular Category is not mentioned in the given data.

- Crop :-The percent of queries from different types of crops present in the Agriculture sector are :-

**State-wise analysis:** In the following table, no. of queries in 10 different states (in rows) of different crops (in columns) is given.

(This table is made by taking in account only top 5 no. of queries from each state)

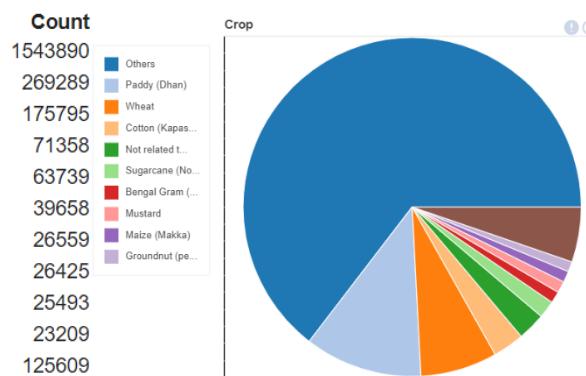
	Paddy	Wheat	Cotton	Mustrad	Bengal Gram
Assam	1126	-----	-----	99	-----
Andhra Pradesh	11056	-----	7353	-----	-----
Chhattisgarh	15382	1622	-----	-----	1276
PUNJAB	79636	42368	10396	-----	-----
HARYANA	19737	22631	24237	-----	-----
GUJURAT	-----	6397	21507	-----	-----
KARNATAKA	-----	-----	4057	-----	-----
MADHAYA PRADESH	14659	31137	-----	-----	10927
UTTAR PRADESH	10745	33050	-----	1351	-----
WEST BENGAL	19428	-----	-----	-----	-----

### Overall analysis (Combined data of all the states)

Value represents different Query type

Count represents no. of queries

Value	Count
Others	1543890
Paddy (Dhan)	269289
Wheat	175795
Cotton (Kapas)	71358
Not related to crop	63739
Sugarcane (Noble Cane)	39658
Bengal Gram (Gram/Chick ...	26559
Mustard	26425
Maize (Makka)	25493
Groundnut (pea nut/mung ...	23209
Other values (229)	125609



\*Others – Particular Crop is not mentioned in the given data.

- **Query Type** :-The count of queries from different query types present in the Agriculture sector are :-

**State-wise analysis:** In the following table, no. of queries in 10 different states (in rows) of different query type (in columns) is given.

(This table is made by taking in account only top 5 no. of queries from each state)

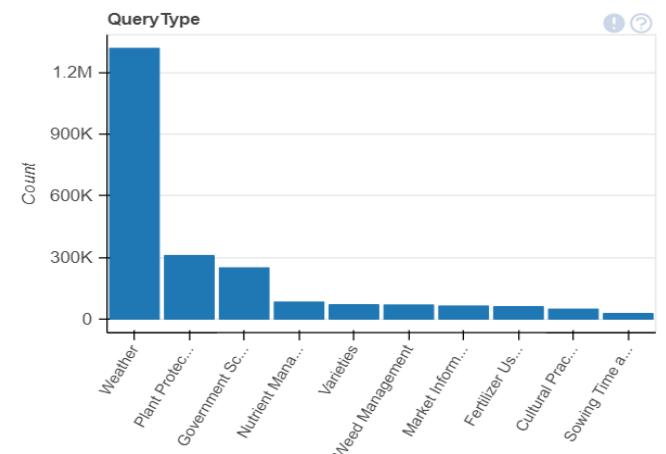
	Weather	Plant Protection	Nutrient Management	Govt.Schemes	Varieties
Assam	1991	745	290	2060	-----
Andhra Pradesh	56177	16936	5282	9683	-----
Chhattisgarh	13880	11008	1963	2763	-----
PUNJAB	160880	56354	24376	-----	13810
HARYANA	221312	35077	9508	15017	11295
GUJURAT	69987	36090	-----	35307	-----
KARNATAKA	90498	11312	-----	20346	-----
MADHAYA PRADESH	168759	48182	-----	49225	12708
UTTAR PRADESH	480099	79503	23345	103229	-----
WEST BENGAL	19917	26492	3794	16584	-----

### Overall analysis (Combined data of all the states)

Value represents different Query type

Count represents no. of queries

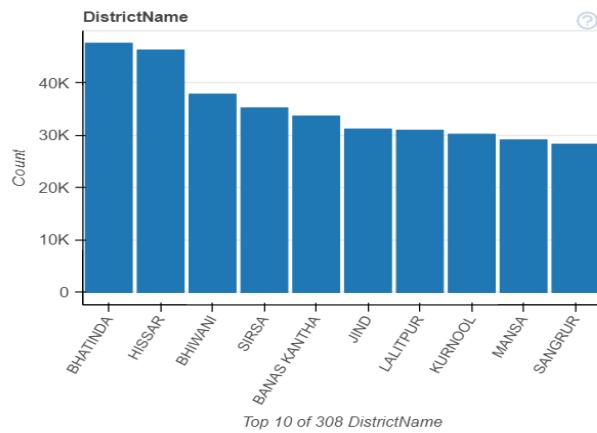
Value	Count
Weather	1316093
Plant Protection	308915
Government Schemes	248982
Nutrient Management	82656
Varieties	69741
Weed Management	68447
Market Information	63449
Fertilizer Use and Availability	60243
Cultural Practices	48218
Sowing Time and Weather	27046
Other values (20)	97234



- District Name :-The count of queries from different districts of the states in the Agriculturesector are :-

Value represents different Query type

Count represents no. of queries



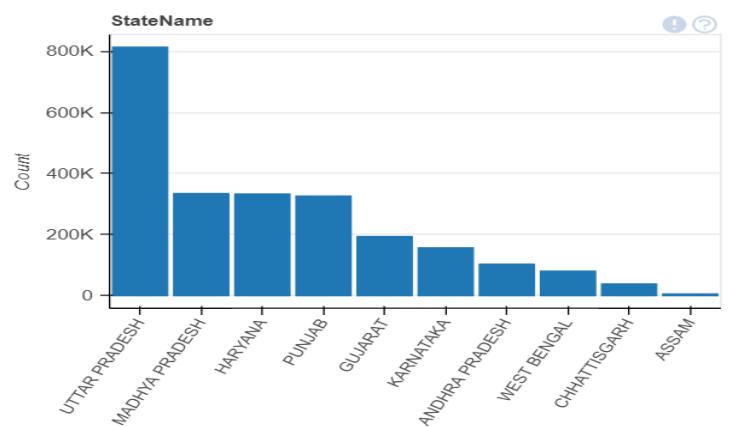
Value	Count
BHATINDA	47620
HISsar	46328
BHIWANI	37893
SIRSA	35267
BANAS KANTHA	33713
JIND	31220
LALITPUR	31002
KURNOOL	30244
MANSA	29155
SANGRUR	28330
Other values (298)	2040252

- State Name :-The count of queries from different states in the Agriculturesector are :-

Value represents different Query type

Count represents no. of queries

Value	Count
UTTAR PRADESH	813940
MADHYA PRADESH	334974
HARYANA	333401
PUNJAB	326815
GUJARAT	194341
KARNATAKA	157282
ANDHRA PRADESH	103787
WEST BENGAL	80935
CHHATTISGARH	39185
ASSAM	6364



- Date :-The change in the number of queries according as different months of year 2019 in Agriculture sector can be represented as :-



## Results :

The maximum queries are for the paddy crop which belong to Cereals category. As Paddy crop requires abundant water supply, so most of the queries are related to Weather & an increasing trend can be seen in the monsoon season from June to October .The maximum queries are from UTTAR PRADESH followed by MADHAYA PRADESH and HARYANA & if we see district wise then maximum queries are from BHATINDA(PB)followed by HISSAR(HR),BHIWANI(HR) and SIRSA(HR).

The summary is the result of above graphs.

### **3.3.2 HORTICULTURE**

- Category :-The percent of queries from different types of categories present in the Horticulture sector are :-

**State-wise analysis:** In the following table, no. of queries in 10 different states (in rows) of different categories (in columns) is given.

(This table is made by taking in account only top 5 no. of queries from each state)

	Vegetables	Fruits	Condiments and spices	Medicinal and Aromatic Plant	Plantation Crops
Assam	3455	905	65	80	582
Andhra Pradesh	19742	5930	377	-----	1031
Chhattisgarh	11359	1526	606	122	-----
PUNJAB	14720	5720	779	-----	191
HARYANA	12667	6246	1032	183	205
GUJURAT	13110	6017	2065	-----	3946
KARNATAKA	14405	7471	16609	-----	1281
MADHAYA PRADESH	29851	6792	8682	1527	-----
UTTAR PRADESH	79995	29157	5080	14728	-----
WEST BENGAL	27645	6418	-----	1206	1199

### Overall analysis (Combined data of all the states)

Value represents different Query type

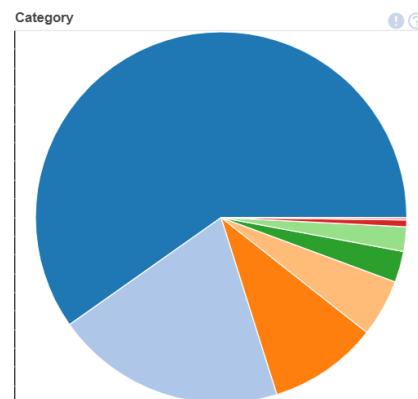
Count represents no. of queries

#### Value

Vegetables  
Fruits  
Condiments and Spices  
Medicinal and Aromatic Pla...  
Plantation Crops  
Flowers  
Drug and Narcotics  
Govt.schemes/weather/ferti...

#### Count

226949	Vegetables
76182	Fruits
36100	Condiments an...
18931	Medicinal and...
10276	Plantation Cr...
8108	Flowers
2236	Drug and Narc...
790	Govt.schemes/...



- Crop :-The percent of queries from different types of crops present in the Horticulture sector are :-

**State-wise analysis:** In the following table, no. of queries in 10 different states (in rows) of different crop (in columns) is given.

(This table is made by taking in account only top 5 no. of queries from each state)

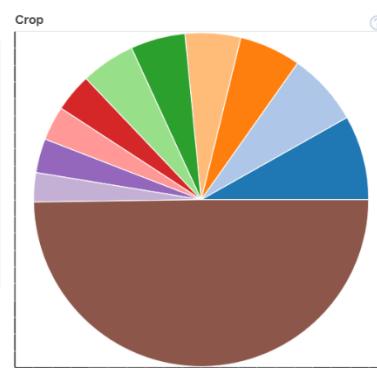
	Chillies	Tomato	Potato	Brinjal	Onion
Assam	653	345	304	299	-----
Andhra Pradesh	9054	2627	-----	1488	-----
Chhattisgarh	863	2747	-----	1177	-----
PUNJAB	1203	1600	-----	-----	-----
HARYANA	-----	1336	-----	-----	1118
GUJURAT	2298	-----	-----	1830	1884
KARNATAKA	2397	-----	-----	-----	2578
MADHAYA PRADESH	2619	4216	-----	-----	6822
UTTAR PRADESH	8381	6465	17557	-----	-----
WEST BENGAL	2678	-----	3629	4646	-----

**Overall analysis (Combined data of all the states)**

Value represents different Query type

Count represents no. of queries

Value	Count
Chillies	31108
Potato	26742
Tomato	22604
Brinjal	20309
Onion	20069
Mango	19891
Bhindi(Okra/Ladysfinger)	14066
Mentha	12591
Cumin	12542
Bottle Gourd	10735
Other values (186)	188915



\*Others – Particular Crop is not mentioned in the given data.

- Query Type :-The count of queries from different query types present in the Horticulture sector are :-

**State-wise analysis:** In the following table, no. of queries in 10 different states (in rows) of different query type (in columns) is given.

(This table is made by taking in account only top 5 no. of queries from each state)

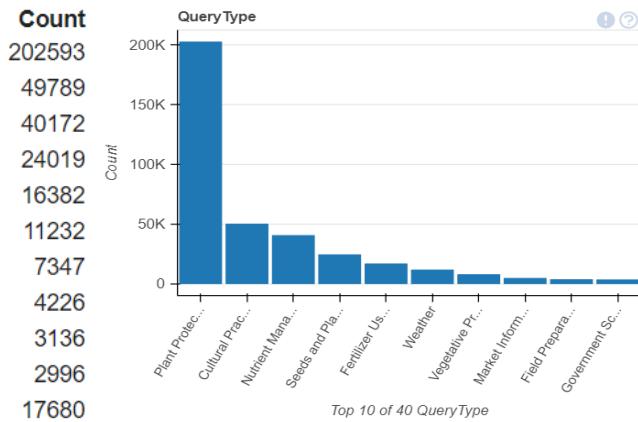
	Seeds And Planting Material	Plant Protection	Nutrient Management	Cultural Practices	Fertilizer Use And Availability
Assam	136	2959	975	665	150
Andhra Pradesh	1031	16742	6611	1965	619
Chhattisgarh	1149	6786	1381	3134	381
PUNJAB	2211	10034	2135	4058	1627
HARYANA	2499	9352	1383	2870	-----
GUJURAT	1405	22926	586	10129	3863
KARNATAKA	1963	12761	5660	2811	-----
MADHAYA PRADESH	4297	28086	4621	4187	2705
UTTAR PRADESH	8516	66537	11742	17058	-----
WEST BENGAL	812	26410	5077	2911	1126

**Overall analysis** (Combined data of all the states)

Value represents different Query type

Count represents no. of queries

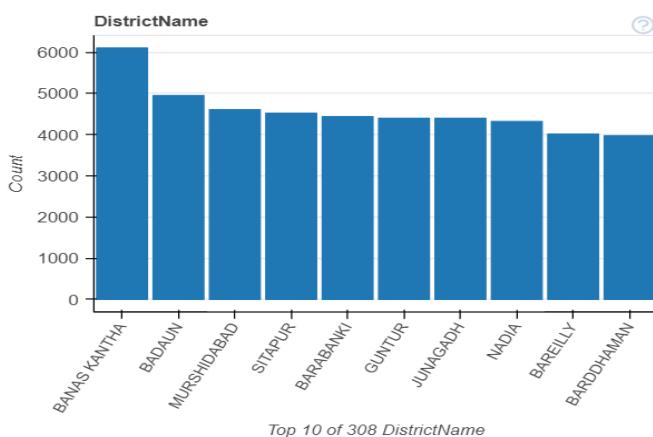
Value
Plant Protection
Cultural Practices
Nutrient Management
Seeds and Planting Material
Fertilizer Use and Availability
Weather
Vegetative Propagation an...
Market Information
Field Preparation
Government Schemes
Other values (30)



- District Name :-The count of queries from different districts of states in the Horticulture sector are :-

Value represents different Query type

Count represents no. of queries

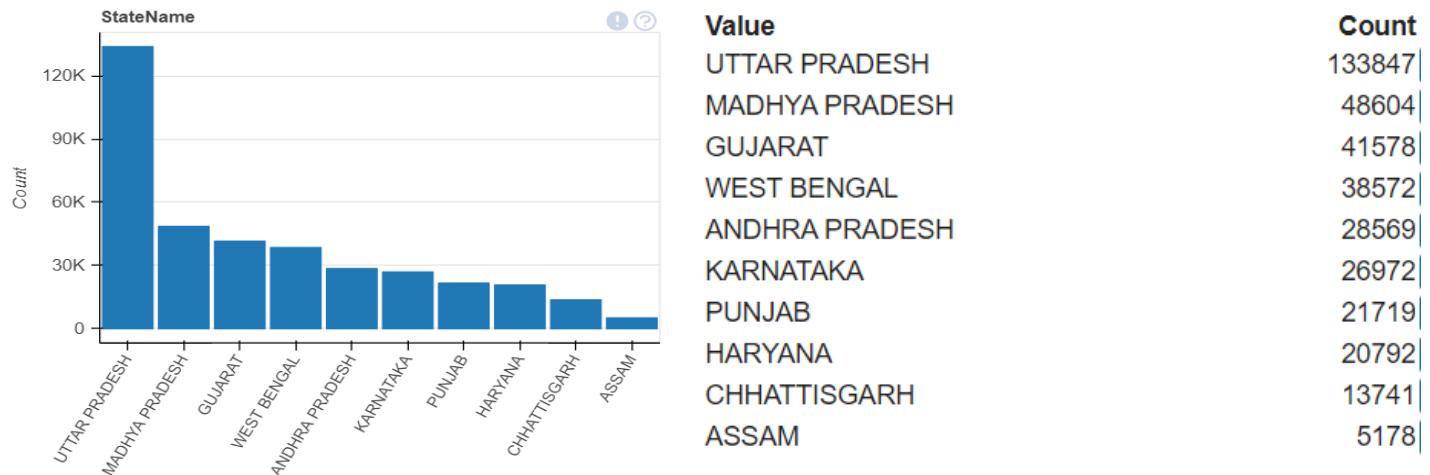


Value	Count
BANAS KANTHA	6105
BADAUN	4950
MURSHIDABAD	4608
SITAPUR	4522
BARABANKI	4439
GUNTUR	4399
JUNAGADH	4399
NADIA	4321
BAREILLY	4016
BARDDHAMAN	3976
Other values (298)	333837

- State Name :-The count of queries from different states in the Horticulture sector are :-

Value represents different Query type

Count represents no. of queries



- Date :-The change in the number of queries according as different months of year 2019 in Horticulture sector can be represented as :-



## Results :

The maximum queries are for the CHILLES crop which belongs to Vegetable category. . maximum type of queries are for plant protection. The maximum queries are from UTTAR PRADESH followed by MADHYA PRADESH and GUJARAT

& if we see district wise then maximum queries are from BANAS KANTHA(GJ) followed by BADAUN(UP), MURSHIDABAD(WB) and SITAPUR(UP).

Above summary is the result of above graphs.

### **3.3.3FISHERIES**

- Category : The percent of queries from different types of categories present in the Fisheries sector are:-

**State-wise analysis:** In the following table, no. of queries in 10 different states (in rows) of different categories (in columns) is given.

(This table is made by taking in account only top 5 no. of queries from each state)

	Inland	Marine
Assam	542	14
Andhra Pradesh	59	20
Chhattisgarh	180	19
PUNJAB	61	11
HARYANA	140	31
GUJURAT	43	19
KARNATAKA	103	20
MADHAYA PRADESH	208	33
UTTAR PRADESH	1119	345
WEST BENGAL	795	51

**Overall analysis** (Combined data of all the states)

Value represents different Query type

Count represents no. of queries

**Value**

Inland

**Count**

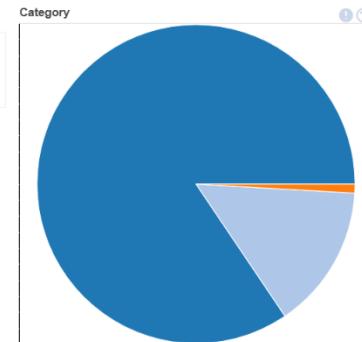
3250

Marine

563

Govt.schemes/weather/ferti...

35



- Crop :-The percent of queries from different types of crops present in the Fisheries sector are :-

**State-wise analysis:** In the following table, no. of queries in 10 different states (in rows) of different crop (in columns) is given.

(This table is made by taking in account only top 5 no. of queries from each state)

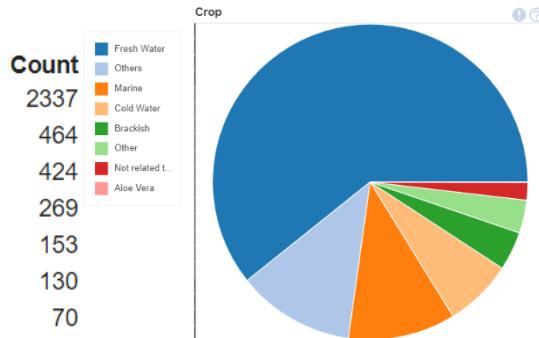
	Fresh Water	Cold Water	Marine	Brackish
Assam	506	7	-----	-----
Andhra Pradesh	34	-----	13	4
Chhattisgarh	110	35	15	8
PUNJAB	30	-----	5	5
HARYANA	57	-----	18	29
GUJURAT	26	-----	11	-----
KARNATAKA	44	6	16	18
MADHAYA PRADESH	116	14	19	17
UTTAR PRADESH	738	122	285	62
WEST BENGAL	693	66	39	-----

## Overall analysis (Combined data of all the states)

Value represents different Query type

Count represents no. of queries

Value	Count
Fresh Water	2337
Others	464
Marine	424
Cold Water	269
Brackish	153
Other	130
Not related to crop	70



\*Others – Particular Crop is not mentioned in the given data.

- Query Type :-The count of queries from different query types present in the Fisheries sector are :-

**State-wise analysis:** In the following table, no. of queries in 10 different states (in rows) of different query type (in columns) is given.

(This table is made by taking in account only top 5 no. of queries from each state)

	Tank And Pond	Fish Fingerling Production	Fish Nutrition	Fish Marketing	Water Testing
Assam	450	38	28	-----	16
Andhra Pradesh	33	6	15	-----	-----
Chhattisgarh	122	19	18	-----	17
PUNJAB	42	8	-----	5	-----
HARYANA	39	14	24	38	-----
GUJURAT	14	8	-----	6	6
KARNATAKA	59	18	8	-----	11

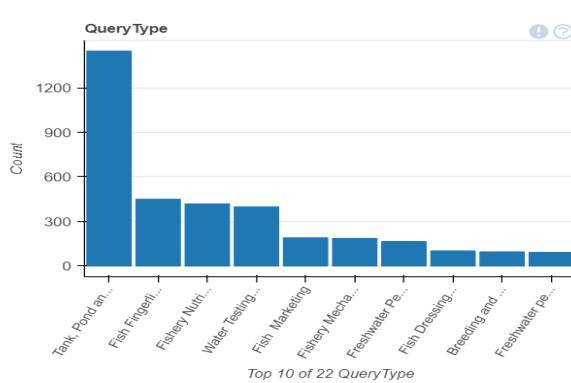
MADHAYA PRADESH	84	45	15	-----	26
UTTAR PRADESH	263	250	190	-----	130
WEST BENGAL	341	44	115	-----	175

## Overall analysis (Combined data of all the states)

Value represents different Query type

Count represents no. of queries

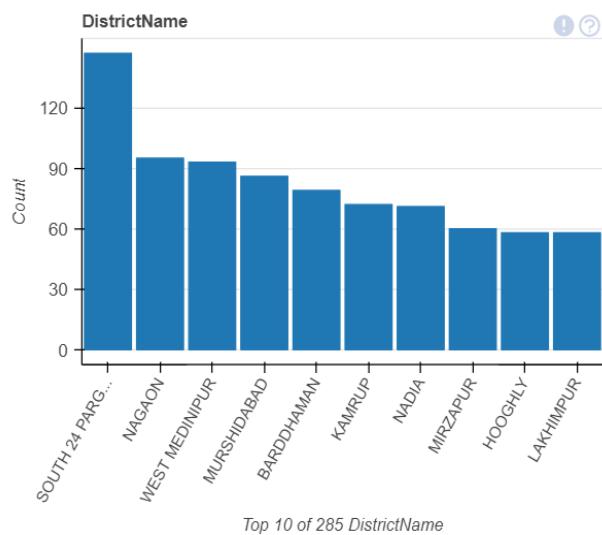
Value	Count
Tank, Pond and Reservoir ...	1447
Fish Fingerling Production	450
Fishery Nutrition	418
Water Testing for Fish Prod...	398
Fish Marketing	190
Fishery Mechanization	186
Freshwater Pearl Farming.	165
Fish Dressing Drying	102
Breeding and culture of orn...	95
Freshwater pearl culture	92
Other values (12)	305



- District Name :-The count of queries from different districts of the states in the **Fisheries** sector are :-

Value represents different Query type

Count represents no. of queries

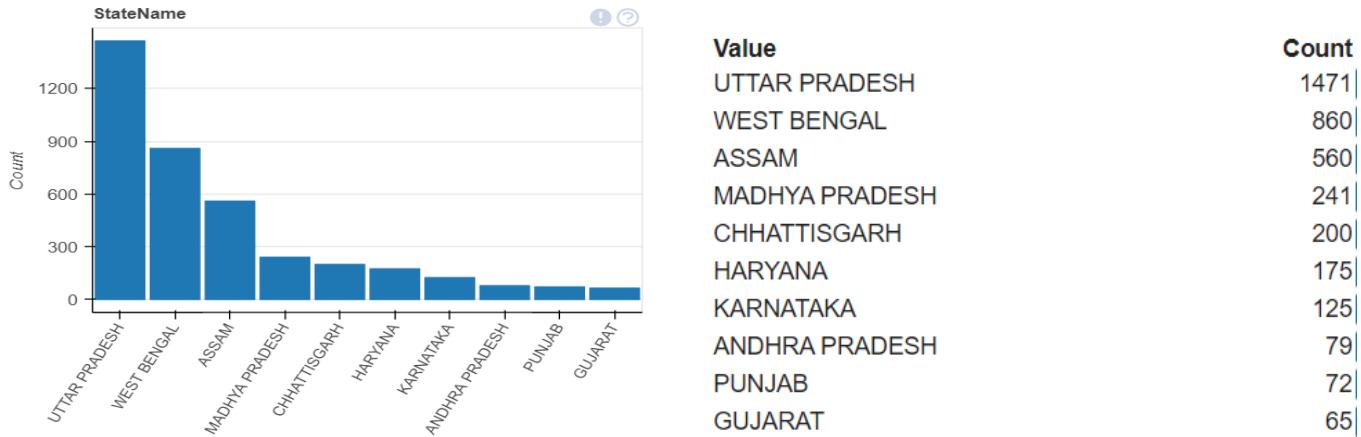


Value	Count
SOUTH 24 PARGANAS	147
NAGAON	95
WEST MEDINIPUR	93
MURSHIDABAD	86
BARDDHAMAN	79
KAMRUP	72
NADIA	71
MIRZAPUR	60
HOOGHLY	58
LAKHIMPUR	58
Other values (275)	3029

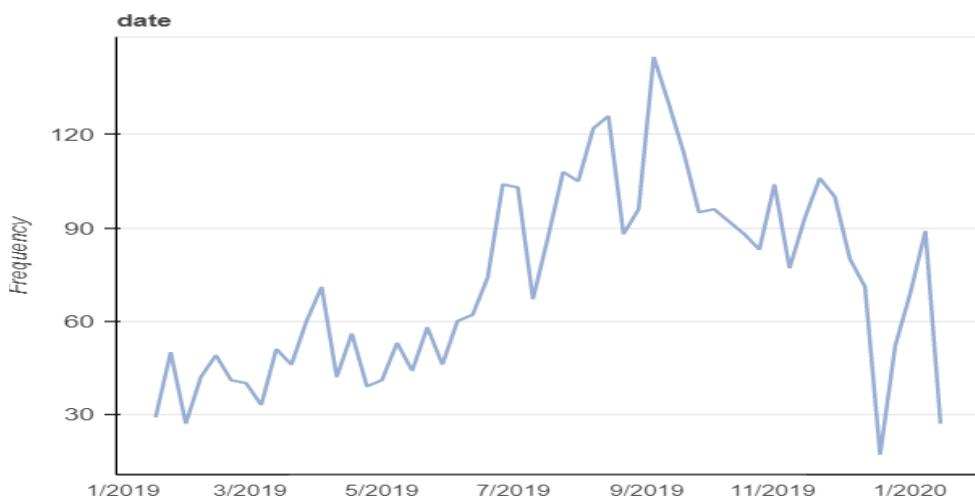
- **State Name** :-The count of queries from different states in the **Fisheries** sector are :-

Value represents different Query type

Count represents no. of queries



- **Date** :-The change in the number of queries according as different months of year 2019 in **Fisheries** sector can be represented as :-



## **Results :**

The maximum queries are for the fresh water fisheries which belongs to inland category. The maximum queries are from UTTAR PRADESH followed by WEST BENGAL and ASSAM & if we see district wise then maximum queries are from SOUTH 24 PARGANAS followed by NAGAON(ASSAM),WEST MEDINIPUR(WB) and MURSHIDABAD(WB).

Above summary is the result of above graphs.

### **3.3.4 ANIMAL HUSBANDRY**

- **Category:**The percent of queries from different types of categories present in the Animal Husbandry sectors are:-

**State-wise analysis:** In the following table, no. of queries in 10 different states (in rows) of different categories (in columns) is given.

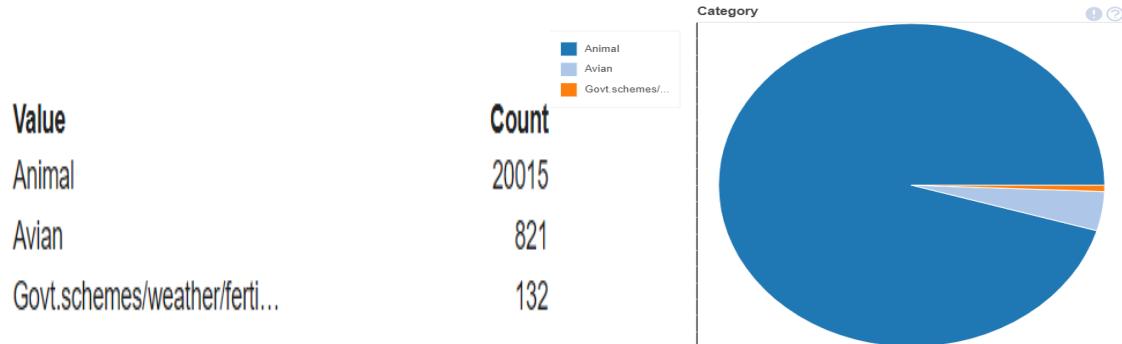
(This table is made by taking in account only top 5 no. of queries from each state)

	Animal	Avian
Assam	1771	245
Andhra Pradesh	695	63
CHHATTISGARH	662	29
PUNJAB	1741	49
HARYANA	2015	35
GUJURAT	1387	18
KARNATAKA	2099	41
MADHAYA PRADESH	1358	357
UTTAR PRADESH	8288	144
WEST BENGAL	533	91

## Overall analysis (Combined data of all the states)

Value represents different Query type

Count represents no. of queries



- Crop :-The percent of queries from different types of crops present in the Animal Husbandry sector are :-

**State-wise analysis:** In the following table, no. of queries in 10 different states (in rows) of different crop (in columns) is given.

(This table is made by taking in account only top 5 no. of queries from each state)

	Bovine	Poultry Farm	Pig	Goat	Duck
Assam	723	438	367	172	204
Andhra Pradesh	405	128	-----	34	-----
Chhattisgarh	317	231	30	60	20
PUNJAB	1146	225	91	113	-----
HARYANA	1681	131	58	39	-----
GUJURAT	1188	35	-----	-----	-----
KARNATAKA	436	123	-----	85	38
MADHAYA PRADESH	1541	282	-----	200	-----
UTTAR PRADESH	6461	660	212	663	-----

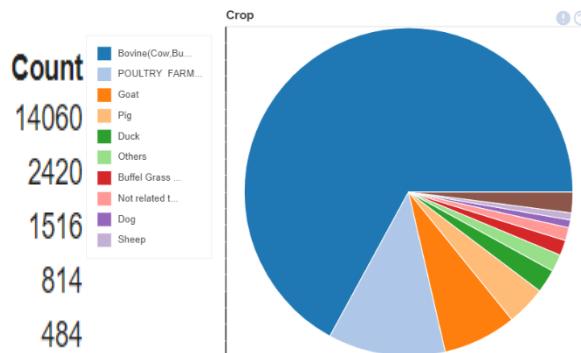
WEST BENGAL	189	167	-----	124	73
----------------	-----	-----	-------	-----	----

### Overall analysis (Combined data of all the states)

Value represents different Query type

Count represents no. of queries

Value	Count
Bovine(Cow,Buffalo)	14060
POULTRY FARM	2420
Goat	1516
Pig	814
Duck	484



\*Others – Particular Crop is not mentioned in the given data.

- Query Type :-The count of queries from different query types present in the Animal Husbandry sector are :-

**State-wise analysis:** In the following table, no. of queries in 10 different states (in rows) of different query type (in columns) is given.

(This table is made by taking in account only top 5 no. of queries from each state)

	Disease Management	Animal Production	Cattle Shed Planning	Dairy Production	Disease Reporting
Assam	1114	238	157	-----	179
Andhra Pradesh	165	92	155	-----	-----
Chhattisgarh	123	155	116	58	-----
PUNJAB	430	409	244	134	-----

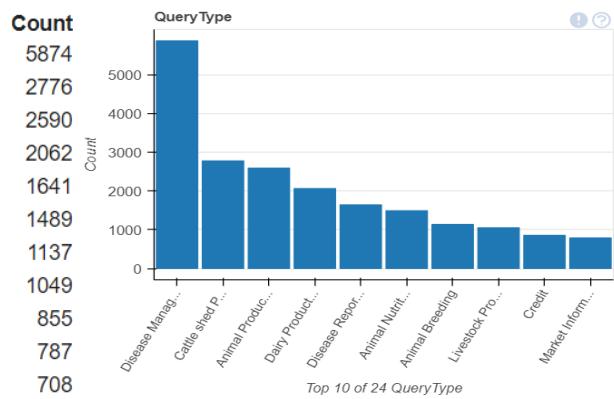
HARYANA	528	-----	240	243	198
GUJURAT	472	-----	162	-----	-----
KARNATAKA	89	249	207	-----	-----
MADHAYA PRADESH	526	208	300	388	-----
UTTAR PRADESH	2308	905	1138	1001	741
WEST BENGAL	119	90	57	-----	-----

## Overall analysis (Combined data of all the states)

Value represents different Query type

Count represents no. of queries

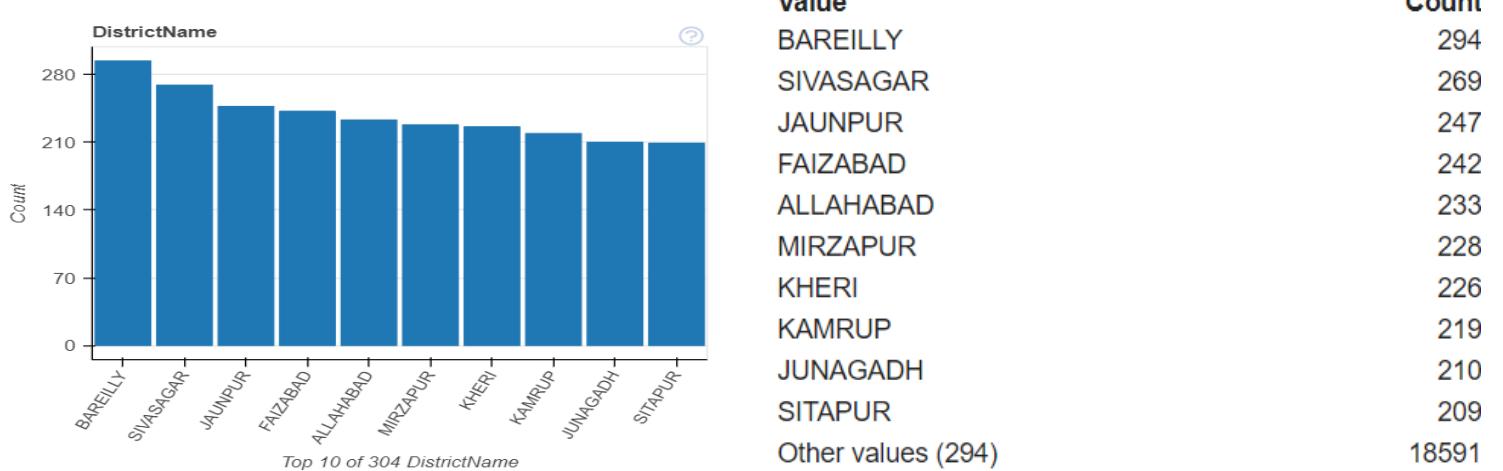
Value	Count
Disease Management	5874
Cattle shed Planning and ...	2776
Animal Production (Piggery...	2590
Dairy Production	2062
Disease Reporting	1641
Animal Nutrition	1489
Animal Breeding	1137
Livestock Products Proces...	1049
Credit	855
Market Information	787
Other values (14)	708



- District Name :-The count of queries from different districts of the states in the Animal Husbandry sector are :-

Value represents different Query type

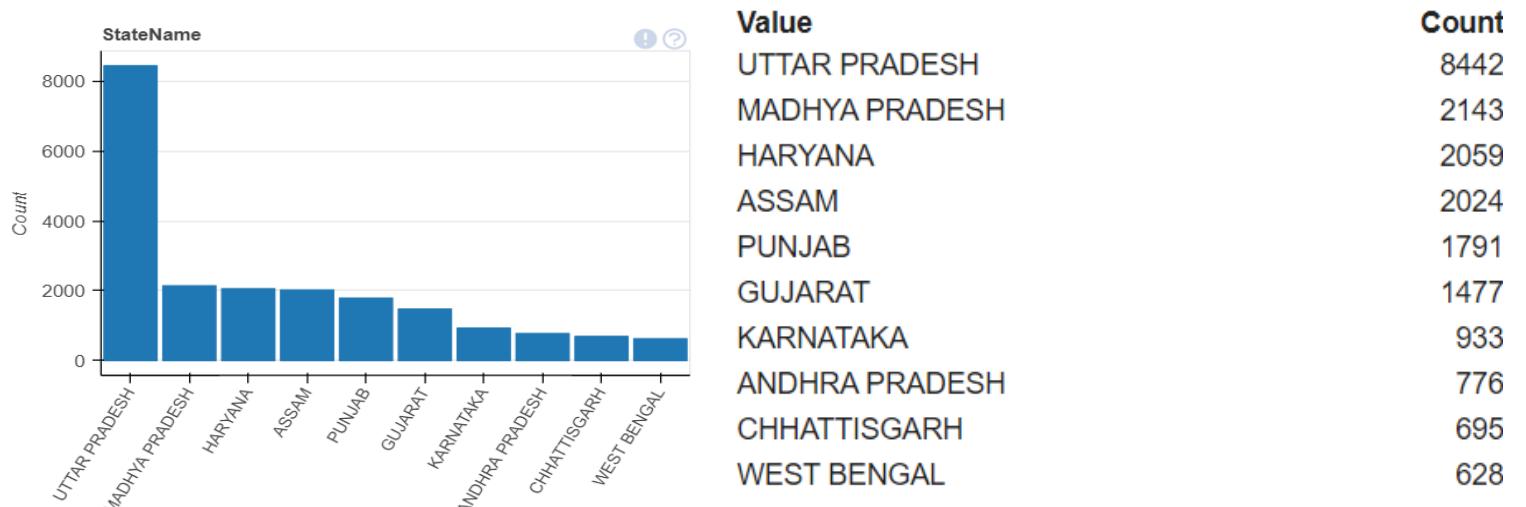
Count represents no. of queries



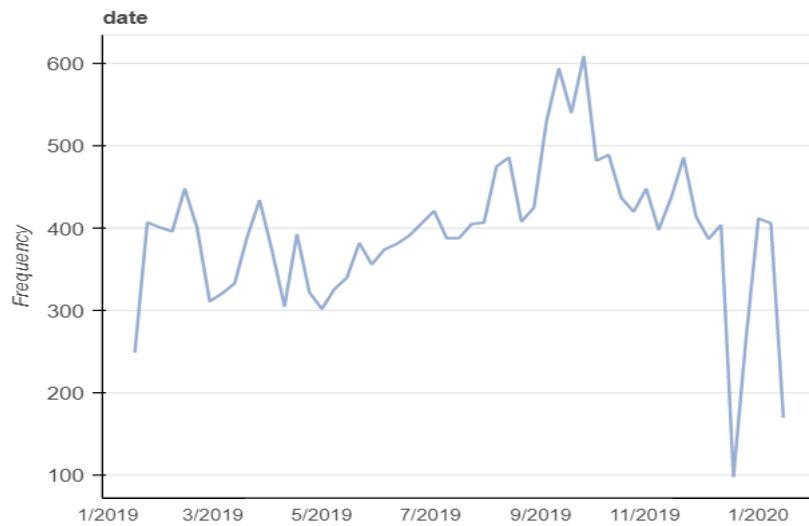
- **State Name** :-The count of queries from different states in the Animal Husbandry sector are :-

Value represents different Query type

Count represents no. of queries



- Date :-The change in the number of queries according as different months of year 2019 in Animal Husbandry sector can be represented as :-



## Results :

The maximum queries are for the Bovine which belongs to the Animal category. And the most queries are regarding disease management. The maximum queries are from UTTAR PRADESH followed by MADHAYA PRADESH and HARYANA & if we see district wise then maximum queries are from BAREILLY(UP), SIVASAGAR(ASSAM), JAUNPUR(UP) and FAIZABAD(UP).

Above summary is the result of above graphs.

## 3.4. Analysis of Null Values

As we can see in our dataset, season and KccAns has lot of null values which we can't remove as it will reduce useful information.

So let's analyze them :-

### 3.4.1. Null values in Season

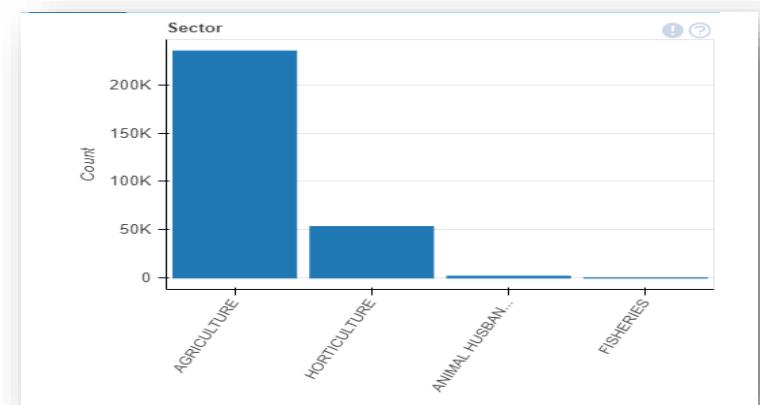
A season is a period of the year that is distinguished by special climate conditions.

Summary: In our dataset, approximately 2788558 season variable have null values. Most of the time, the queries farmers ask don't mention the season. This shows that Kisan Call Centre experts don't care much about the season of the query is about.

### 3.4.2. Null values in KccAns

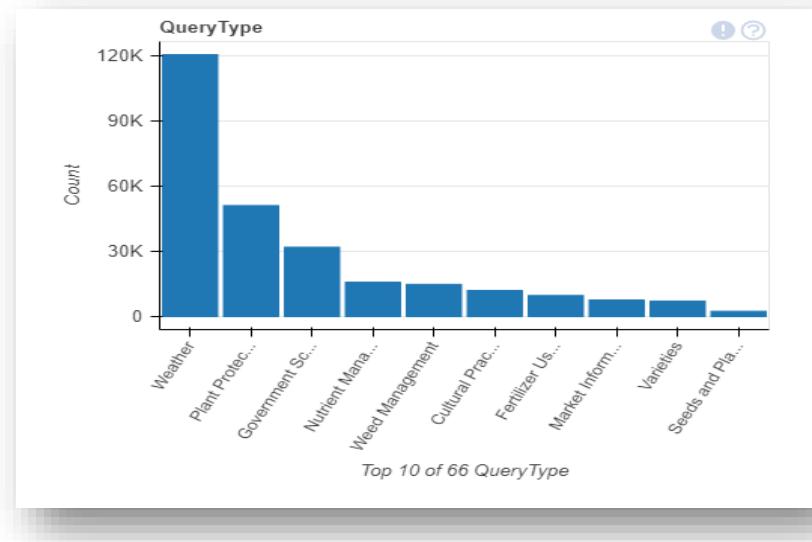
In this column the questions asked by the caller has been answered.

- **Sector wise:** The no. of queries from different sector which have not been answered by Kisan call centre.



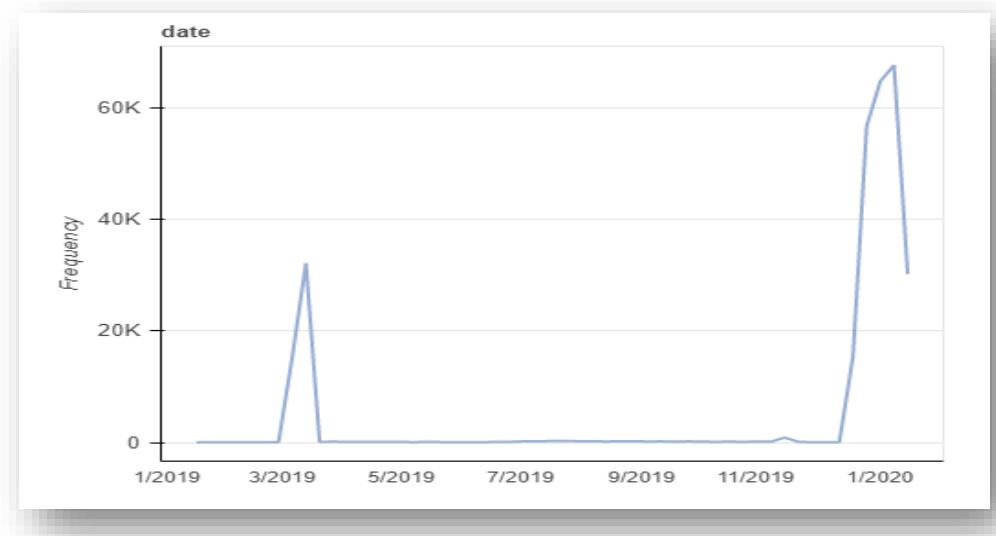
Bar chart of number of queries not answered by Kisan call centre in different sector.

- **Query Type wise** : The no. of queries from different query types which have not been answered by Kisan call centre.



Bar chart of top 10 of 66 query type of queries not answered by Kisan call centre

- **Date wise:** The change in number of queries according as different date of year 2019 which have not been answered by Kisan call centre can be representing by using line chart as



Line chart of number of queries not answered by Kisan call centre

## Results:

The above analysis reveals that most of the queries which have not been answered by Kisan call centre are related to agriculture sector and most of these questions are of weather query type. The important result that we can see is from the line chart of dates. This shows that only in the month of February and December, the peak of the frequency curve is higher. So we conclude that the queries which have not been answered by kisan call centre are only in the months of February and December.

# **CHAPTER 4**

## **DISCUSSION & SUMMARY**

### **4.1 Essence of Overall Data**

#### *State-wise analysis of Sector*

- we have 4 different sectors in the data & most of the data entries are from AGRICULTURE & HORTICULTURE .
- We have only few number of queries from other two sectors.UP have the maximum number of queries for each sector.
- Assam have minimum number of queries for AGRICULTURE & HORTICULTURE

#### *Overall analysis of Sector*

The largest value (AGRICULTURE) is over 6.3 times larger than the second largest value (HORTICULTURE ). Maximum queries are from AGRICULTURE sector & minimum queries are from FISHERIES sector.

- we conclude that there is significant sector wise variation in the states.
- we conclude that state wise variation in different sectors has been significantly at variance.

#### *State-wise analysis of Category*

- All the states have maximum number of queries in mainly two categories :- CEREALS & VEGETABLES.
- Only Gujarat have maximum number of queries in ‘OILSEEDS’.

### Overall analysis of Category

the proportion of farmer queries in each category. It can be clearly seen that 16% of the total queries are from 'CEREALS' category . Rest of the 27% queries are from other categories.

- we conclude that state wise variation in different categories has been significantly at variance.

### State-wise analysis of Crops

- Most of the above mentioned states have maximum number of queries in Paddy (Dhan) crop.
- UP have highest number of queries in Paddy crop whereas Assam have minimum number of queries in Paddy crop in comparison to all other states.

### Overall analysis of Crop

the proportion of farmer queries in each crop. It can be clearly seen that around 20% of the total queries are from Paddy, Wheat ,Cotton & Sugarcane crops. Rest of the crops constitute remaining 25% proportion.

- we conclude that state wise variation in different crops has been significantly at variance.

### State-wise analysis of Query Type

- All the states have maximum number of queries in mainly two Query types :- WEATHER & PLANT PROTECTION.
- UP have maximum number of queries for WEATHER than other states.

### Overall analysis of Query Type

Maximum queries 47.5% are for weather , followed by plant production with 18.3 % ,nutrient management with 4.4 % and govt. schemes with 9.0% . These 4

categories accounts for total of 79.2% queries originated. The remaining 20.8% queries comprises of other 62 categories.

- we conclude that state wise variation in different Query Type has been significantly at variance.

### Overall analysis of States

Out of all 10 states, as represented in bar chart ,UP has maximum queries with count 957700 , followed by MADHYA PRADESH with count 2.48 times smaller than UP's count. Further , Punjab & Haryana faced nearly equal number of queries in year 2019. ASSAM have minimum number of queries in comparison to all other states.

### Overall analysis of Districts

Out of all 308 districts, as represented in bar chart ,BHATINDA(PB) has maximum queries with count 49906 , followed by HISSAR(HR) with slighter less count than BHATINDA's count. Further , BANAS KANTHA(GJ) & BHIWANI(HR) faced equal number of queries in year 2019.

### Date wise analysis of Overall Data

It can be clearly seen from the above line chart that the highest queries are in the middle of July , whereas lowest queries are in the starting of December month which are approximately 6 times less than that in the month of July. A decreasing trend can be seen in the number of queries from July to December.

### Combined analysis of Query Type and Crop

The maximum queries in Wheat crop are related to Varieties & weed management, whereas in Paddy & Cotton , these are related to Plant protection & Nutrient management. The negligible queries about Weather & Government schemes in Cotton crop .

### Combined analysis of Date and Category

During the year 2019,the maximum number of queries are related to Cereals, followed by Vegetables , Fruits & Oilseeds. However, there are very less queries about Millets during the whole year.

## 4.2 Extract of Season-wise analysis

### 4.2.1 KHARIF SEASON

#### Overall analysis of Category:

Highest numbers of queries are from CEREALS category.

#### State-wise analysis of Category:

We see that the three main categories that have the highest queries in these ten states were Cereals, vegetables and Oilseeds. The highest no. of queries about Cereals and Vegetables are from Uttar Pradesh (99,049 and 34,729 respectively) and Oilseeds from West Bengal (22,139).

### Overall analysis of Crops:

Paddy has the highest queries among crops with 14.8% (no. of queries=224673) & then cotton with 3.8% (no. of queries= 57526), whereas all other crops have less than 2% out of total queries.

### State-wise analysis of crops:

Paddy, Cotton and Soybean have most no. of queries in these states. The highest no. of queries about Paddy is from Uttar Pradesh (94,270), Cotton from Haryana (19,559) and Soybean from Madhya Pradesh (19,815).

### Overall analysis of Query Type:

Maximum queries are related to WEATHER, followed by Plant protection, followed by Govt. schemes. These 3 categories accounts for 77.8% queries originated.

### State-wise analysis of Query Type:

Mostly all the states have top queries about Weather except for Assam and Chhattisgarh which have highest no. of queries about Govt. schemes and Plant Protection respectively. The highest no. of queries about Weather is from Uttar Pradesh (2,69,902)

### Date-wise analysis:

In Kharif Season, The highest no. of queries is in the month of September, 2019 and the lowest no. of queries is at the end of August, 2019.

### States-wise analysis:

Uttar Pradesh state has the maximum no. of queries (34.4%), followed by Madhya Pradesh(14.3%), followed by Haryana(12.8%) in Kharif season.

### Districts-wise analysis:

BHATINDA district (Punjab) had maximum queries 30186 (2.0%) in Kharif season, followed by HISSAR district (Haryana) with 29395 (1.9%) and BANAS KANTHA (Gujarat) with 25655 (1.6%) no. of queries.

#### Combined analysis of Query Type and Crop:

Maximum queries are about Weed Management, Nutrient Management & Fertilizer use and availability in Paddy crop, whereas in case of Soybean most queries are about Weed Management, Market Information and Plant Protection and in case of cotton these are Fertilizer use and availability, Plant Protection and Nutrient Management in Kharif season. Negligible queries about Government schemes, Varieties and Weather in Cotton crop.

#### Combined analysis of Date and Category:

In Kharif Season, it can be clearly seen that:

Maximum numbers of queries are related to Cereals are highest in September, 2019, followed by Vegetables which are highest in October, 2019 and Oilseeds which are highest in August. However, there are very less queries about Millets in the Kharif season.

## **4.2.2 RABI SEASON**

#### Overall analysis of Category:

Highest numbers of queries are from CEREALS category. (Ref Figure (1))

#### State-wise analysis of Category:

We see that the three main categories that have the highest queries in these ten states are Cereals, Vegetables and Condiment and spices. The highest no. of queries about Cereals and Vegetables are from Uttar Pradesh (70,777 and 35,232

respectively), Vegetables from Uttar Pradesh (20,116), Condiment and spices from Gujarat (14,847).

#### Overall analysis of Crops:

Wheat has the highest queries among crops with 15.8% ( no. of queries= 1,54,001) & then Paddy with 2.5% (no. of queries= 23,965), followed by Potato (2.2%), followed by Bengal Gram (2.1%), whereas all other crops have less than 2% out of total queries.

#### State-wise analysis of crops:

Paddy, Wheat, Potato have most no. of queries in these states. The highest no. of queries about Paddy is from West Bengal (7,355), Wheat from Uttar Pradesh (64,756) and Potato from Uttar Pradesh (13,665). However, Maximum no. of queries from Gujarat and Karnataka are for Cumin and Maize respectively.

#### Overall analysis of Query Type:

Maximum queries are related to WEATHER, followed by Plant protection, followed by Govt. schemes. These 3 categories accounts for 71% queries originated.

#### State-wise analysis of Query Type:

Mostly all the states have top queries about Weather except for Gujarat, Andhra Pradesh, Assam and Chhattisgarh which have highest no. of queries about Plant Protection. The highest no. of queries about Weather is Uttar Pradesh (1, 60,760).

#### Date-wise analysis:

The highest no. of queries is in the month of November, 2019 and the lowest no. of queries is at the end of December, 2019.

#### States-wise analysis:

Uttar Pradesh state has the maximum no. of queries (33.2%), followed by Madhya Pradesh(14.5%), followed by Haryana(12.1%) in Rabi season.

#### Districts-wise analysis:

HISSAR district (Haryana) and BHATINDA district (Punjab) has approximate same no. of queries (1.4%) in Rabi season, followed by BANAS KANTHA (Gujarat) with 1.3% of total no. of queries.

#### Combined analysis of Query Type and Crop:

Maximum queries are about Weed Management, Varieties & Fertilizer use and availability in Wheat crop, whereas in case of Paddy most queries are about Govt. schemes, Market Information and Weather and in case of Potato these are Weather, Plant Protection and Cultural Practices in Rabi season. Negligible queries are about Government schemes, Varieties and Weather in Bengal gram.

#### Combined analysis of Date and Category:

Maximum numbers of queries are related to Cereals which are highest in December, followed by Vegetables which are highest in May. However, there are very less queries about Millets in the Rabi season.

### **4.2.3 ZAID SEASON**

#### Overall analysis of Category:

Highest numbers of queries are from CEREALS category.

#### State-wise analysis of Category:

We see that the three main categories that have the highest queries in these ten states are Cereals, Vegetables, Oilseeds and Fiber crops. The highest no. of queries about Cereals is from Punjab (29,335), Vegetables from Uttar Pradesh (20,116), Oilseeds from Gujarat (4,760) and Fiber Crops from Haryana (6,412).

### Overall analysis of Crops:

Paddy has the highest queries among crops with 7.7% (no. of queries= 53565) & then Sugarcane with 3.3% (no. of queries= 23206), followed by Potato (2.5%), whereas all other crops have less than 2% out of total queries.(Ref Figure (2))

### State-wise analysis of crops:

Paddy, Cotton, Sugarcane, and Chillies have most no. of queries in these states. The highest no. of queries about Paddy is from Punjab (23,945), Cotton from Haryana (6,389), Sugarcane from Uttar Pradesh (20,009) and Chillies from Karnataka (749).

### Overall analysis of Query Type:

Maximum queries are related to WEATHER, followed by Plant protection, followed by Govt. schemes. These 3 categories accounts for 73% queries originated.

### State-wise analysis of Query Type:

Mostly all the states have top queries about Weather except for Gujarat, Assam and Chhattisgarh which have highest no. of queries about Govt. schemes and Plant Protection. The highest no. of queries about Weather is Uttar Pradesh (1, 37,134).

### Date-wise analysis:

The highest no. of queries was in the month of July, 2019 and the lowest no. of queries was in the month of April, 2019. An increasing trend can be seen from June to July, 2019.

### States-wise analysis:

Uttar Pradesh state has the maximum no. of queries (38.3%), followed by Punjab (14.1%), followed by Haryana(13.9%) in Zaid season.

#### Districts-wise analysis:

BHATINDA district (Punjab) has maximum no. of queries in Zaid season (2.1%), followed by HISSAR district (Haryana) with 1.8%, and followed by SIRSA (Haryana) with 1.5% of total no. of queries.

#### Combined analysis of Query Type and Crop:

Maximum queries are about Weed Management, Sowing time & Varieties in Paddy crop, whereas in case of Wheat most queries are about Govt. schemes, Market Information and Plant Protection and in case of Sugarcane these are Weed Management, Nutrient Management, Fertilizer use and availability in Zaid season.

#### Combined analysis of Date and Category:

Maximum numbers of queries are related to Cereals which are highest in June, 2019, followed by Vegetables which are highest in August, 2019. However, there are very less queries about Millets in the Zaid season.

## **4.3 Summary of Sector-wise analysis**

### **4.3.1 AGRICULTURE**

The maximum queries are for the paddy crop which belong to Cereals category. As Paddy crop requires abundant water supply, so most of the queries are related to Weather & an increasing trend can be seen in the monsoon season from June to October .The maximum queries are from UTTAR PRADESH followed by MADHAYA PRADESH and HARYANA & if we see district wise then maximum queries are from BHATINDA(PB)followed by HISSAR(HR),BHIWANI(HR) and SIRSA(HR).

### **4.3.2 HORTICULTURE**

The maximum queries are for the CHILLES crop which belongs to Vegetable category. . maximum type of queries are for plant protection.The maximum queries are from UTTAR PRADESH followed by MADHAYA PRADESH and GUJARAT & if we see district wise then maximum queries are from BANAS KANTHA(GJ) followed by BADAUN(UP),MURSHIDABAD(WB) and SITAPUR(UP).

### **4.3.3 FISHERIES**

The maximum queries are for the fresh water fisheries which belongs to inland category. The maximum queries are from UTTAR PRADESH followed by WEST BENGAL and ASSAM & if we see district wise then maximum queries are froM SOUTH 24 PARGANAS followed by NAGAON(ASSAM),WEST MEDINIPUR(WB) and MURSHIDABAD(WB).

### **4.3.4 ANIMAL HUSBANDRY**

The maximum queries are for the Bovine which belongs to the Animal category.And the most queries are regarding disease management.The maximum queries are from UTTAR PRADESH followed by MADHAYA PRADESH and HARYANA & if we see district wise then maximum queries are from BAREILLY(UP),SIVASAGAR(ASSAM),JAUNPUR(UP) and FAIZABAD(UP).

## 4.4 Summary of Analysis of Null Values

In our dataset, approximately 2788558 season variable have null values. Most of the time, the queries farmers ask don't mention the season. This shows that Kisan Call Centre experts don't care much about the season of the query is about.

The analysis reveals that most of the queries which have not been answered by Kisan call centre are related to agriculture sector and most of these questions are of weather query type. The important result that we can see is from the line chart of dates. This shows that only in the month of February and December, the peak of the frequency curve is higher. So we conclude that the queries which have not been answered by kisan call centre are only in the months of February and December.

# **CHAPTER-5**

## **WEBAPP**

### **5.1 About WebApp**

As we know 30 per cent of marginal and small farmers are still illiterate. Farmer's do some combination work such as crops, orchards, vineyards, poultry, or other livestock. Because of this, their problems may be related to any of these work. But the government has made separate departments for all these. There are many such farmers who do not even know whom to talk to for their problems. In such situations, farmers keep roaming in the many government departments for their problems. That's why we have created a web app which is for the convenience of the farmers. This web app is mostly useful for those farmers who are unaware of technical terminologies of agriculture.

This web app tells the sector of farmer's problems. If farmers write their queries in this web app then it predicts their sector so that farmers can easily contact the respective department through any means of communication. This web app is built using machine learning algorithms. Which we have explained later on.

LINK: <https://farmer-query.herokuapp.com>

# WELCOME TO KISAN CALL CENTER

Here we will help you to find the sector of your queries

Please enter your name:

Type Here



Your name

Please enter your state

Type Here



Your state

Please enter your query:

Type Here



Your query

NEXT



This is the Machine Learning webapp of accuracy level 90%

THANKYOU



Please enter your query:

WATER MANAGEMENT OF POND

NEXT

## HELLO Manju!

**You should consult to FISHERIES  
Department for your query**

please contact to toll free number:1800-180-1551 and for BSNL landline  
Toll-Free number: 1551

This is the Machine Learning webapp of accuracy level 90%

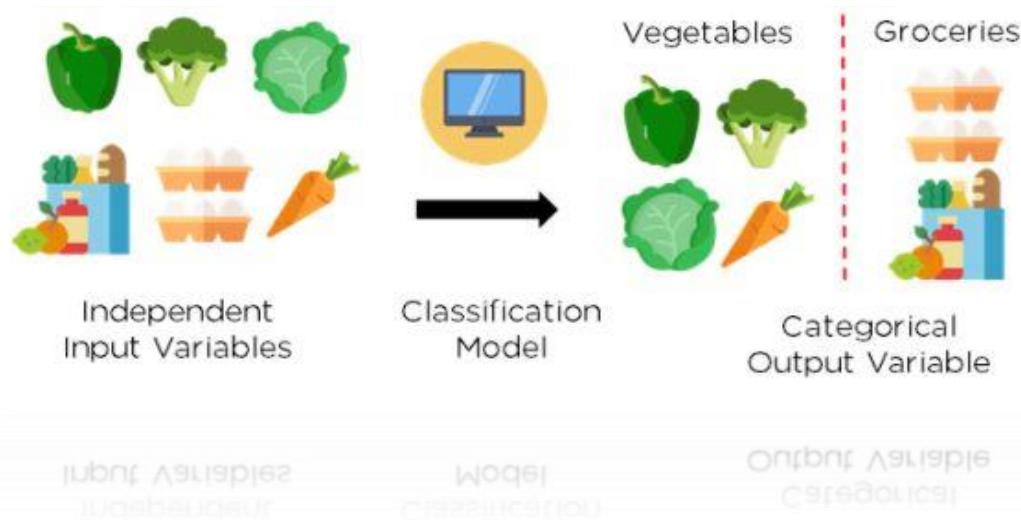
THANKYOU

## **5.1.1 Classification Predictive Modeling**

Machine learning is a field of study and is concerned with algorithms that learn from examples.

From a modeling perspective, classification requires a training dataset with many examples of inputs and outputs from which to learn.

Classification refers to a predictive modeling problem where a class label is predicted for a given example of input data.



## **5.1.2 Text Classification**

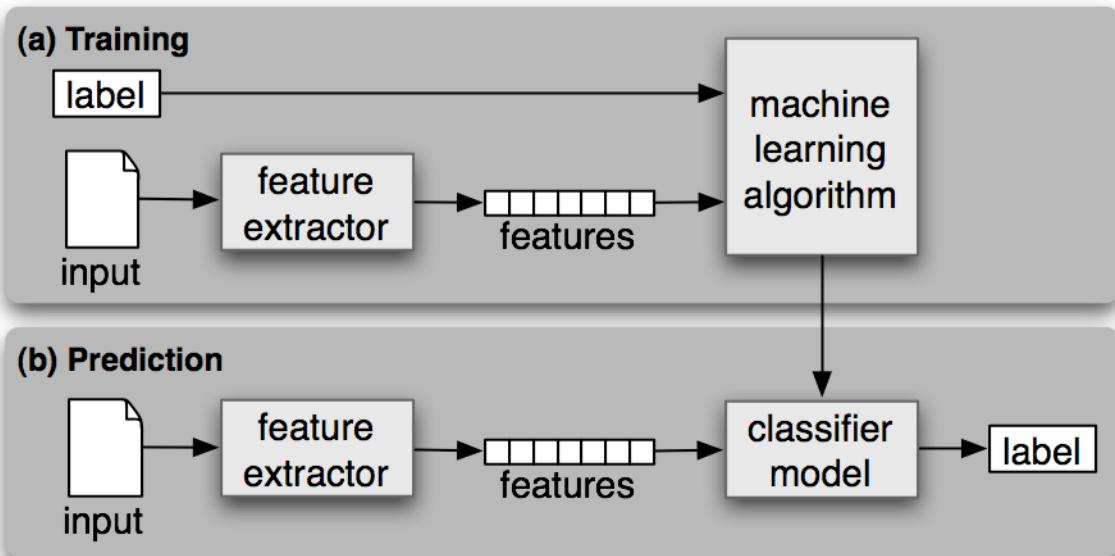
The goal of text classification is to automatically classify the text documents into one or more defined categories.

**1. Dataset Preparation:** The first step is the Dataset Preparation step which includes the process of loading a dataset and performing basic pre-processing. The dataset is then split into train and validation sets.

**2. Feature Engineering:** The next step is the Feature Engineering in which the raw dataset is transformed into flat features which can be used in a machine learning

model. This step also includes the process of creating new features from the existing data.

**3. Model Training:** The final step is the Model Building step in which a machine learning model is trained on a labelled dataset.



## ➤ Dataset Preparation

### Scikit-Learn (Sklearn)

Scikit-learn (Sklearn) is the most useful and robust library for machine learning in Python. It provides a selection of efficient tools for machine learning and statistical modeling including classification, regression, clustering and dimensionality reduction via a consistent interface in Python. This library, which is largely written in Python, is built upon NumPy, SciPy and Matplotlib.



- Model Selection

Model selection is the process of selecting one final machine learning model from among a collection of candidate machine learning models for a training dataset.

Model selection is a process that can be applied both across different types of models (e.g. logistic regression, SVM, KNN, etc.) and across models of the same type configured with different model hyper parameters (e.g. different kernels in an SVM).

## ➤ Feature Engineering

### TF-IDF Vectors as features

TF-IDF score represents the relative importance of a term in the document and the entire corpus. TF-IDF score is composed by two terms: the first computes the normalized Term Frequency (TF), the second term is the Inverse Document Frequency (IDF), computed as the logarithm of the number of the documents in the corpus divided by the number of documents where the specific term appears.

TF(t) = (Number of times term t appears in a document) / (Total number of terms in the document)

IDF(t) =  $\log_e(\text{Total number of documents} / \text{Number of documents with term } t \text{ in it})$

TF-IDF Vectors can be generated at different levels of input tokens (words, characters, n-grams)

$$w_{i,j} = tf_{i,j} \times \log \left( \frac{N}{df_i} \right)$$

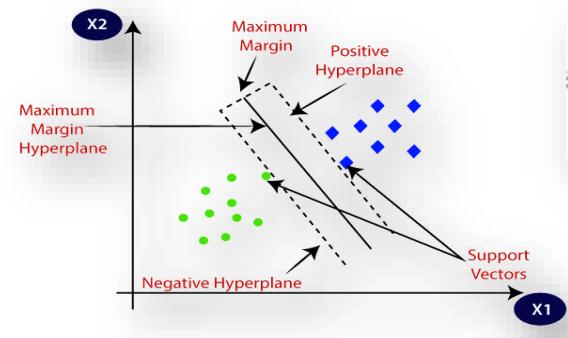
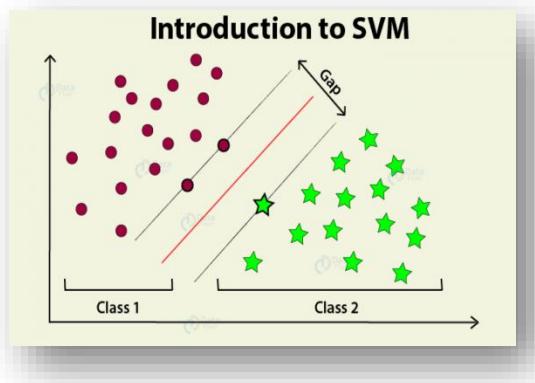
$tf_{i,j}$  = number of occurrences of  $i$  in  $j$   
 $df_i$  = number of documents containing  $i$   
 $N$  = total number of documents

## ➤ Model Training

### SVM Model

“Support Vector Machine” (SVM) is a supervised machine learning algorithm which can be used for both classification or regression challenges. However, it is mostly used in classification problems. In the SVM algorithm, we plot each data item as a point in n-dimensional space (where n is number of features you have) with the value of each feature being the value of a particular coordinate. Then, we perform classification by finding the hyper-plane that differentiates the two classes very well (look at the below snapshot).

Support Vectors are simply the co-ordinates of individual observation. The SVM classifier is a frontier which best segregates the two classes (hyper-plane/ line).



## **5.2 App framework**

### **Streamlit**

Streamlit is an open source app framework in python language. It helps us create beautiful web-apps for data science and machine learning in a little time. It is compatible with major python libraries such as scikit-learn, keras, pytorch, latex, numpy, pandas, matplotlib, etc..



## **5.3 App Deployment**

### **Heroku**

Heroku is a container-based cloud Platform as a Service (PaaS). Developers use Heroku to deploy, manage, and scale modern apps. Our platform is elegant, flexible, and easy to use, offering developers the simplest path to getting their apps to market.



# **CHAPTER 6**

## **CONCLUSION & SUGGESTIONS**

### **6.1 FINDINGS**

It can be clearly seen that:-

-  Out of all the 10 selected states, UP has maximum number of queries. This may be because UTTAR PRADESH is the state with highest population in INDIA. Whereas, ASSAM has minimum number of queries in comparison to other states. This may be because ASSAM is at last position in production than other selected states.
-  Data for the top crops in 10 selected states shows that Paddy has the largest share which belong to Cereals category.. Examining the broad reasons for calling, it is found that the highest number of calls recorded are for weather information, followed after a margin by plant protection, government schemes and then nutrient management, cultural practices, fertilizer use/ availability, and then market information. Weather and plant protection are major in all of them – indicating that concerns of risk are a major reason for calling
-  An increasing trend in the number of queries can be seen before starting of monsoon season. This may be due to sowing of Paddy crop in these months.

## **6.2RECOMMENDATIONS**

- For enhancing the use of the KCC system further& to help the farmers, for this strong publicity to the farming community should be done - to increase awareness about KCCs, how they can help, and how to reach them, so that the user base and the call frequency can be greatly increased.
- There is great need to regularly monitor the call efficiency statistics of the KCC and seek to reduce the waiting time, the calls not answered, the call drops, and to increase the percentage of calls effectively answered.
- Weather information is a significant reason for calling and should be substantially strengthened and kept up to date. But very importantly, there is great dissatisfaction with respect to the price and market information and so a substantial need to strengthen the price and market information database.
- Frequent and good training programmes for the **Farmer Tele-advisors(FTAs)**are a must to regularly enhance their skills and knowledge.
- The FTAs play the most important role in the KCC system and need to be well compensated and supported. There is need to provide good office infrastructure facilities and create a good working environment for them, and the terms and compensation of FTAs need to be enhanced to attract the best talent, motivate them, get the good performance, and retain them. They play the most critical role in helping the farmers and delivering the KCC service.

# REFERENCES

- Book name: Fundamentals of Statistics  
Author: A M Goon; M K Gupta; B Dasgupta  
Publisher: Calcutta- The World Press Private Ltd, 1968
- Book name: Research Methodology: Methods And Techniques.  
Author: C R Kothari; Gaurav Garg.  
Publisher: New Age International Publishers.
- Kisan Call Centre Official website:  
<https://dackkms.gov.in/account/login.aspx>
- Department of Agriculture Cooperation & Farmers Welfare Official website:  
<https://agricoop.nic.in/en>
- Official Python website:  
<https://www.python.org/>
- Wikipedia  
[https://en.wikipedia.org/wiki/Main\\_Page](https://en.wikipedia.org/wiki/Main_Page)
- Google  
<https://www.google.co.in/>
- YouTube  
[https://www.youtube.com/watch?v=COI0BUMNHT8&list=PLyqSpQzTE6M\\_JcleDbrVyPnEOPixKs2JE](https://www.youtube.com/watch?v=COI0BUMNHT8&list=PLyqSpQzTE6M_JcleDbrVyPnEOPixKs2JE)  
[https://www.youtube.com/watch?v=VPZD\\_aij8H0&list=PLUI4u3cNGP60uVBMaoNERc6knT\\_MgPKSO](https://www.youtube.com/watch?v=VPZD_aij8H0&list=PLUI4u3cNGP60uVBMaoNERc6knT_MgPKSO)  
<https://www.youtube.com/watch?v=hnDU1G9hWqU>

- Link to dataset that we used in this analysis:

[https://drive.google.com/file/d/1KHSTIOtJ7QiRgOINaExtj0v5mdGPDQV /view?usp=sharing](https://drive.google.com/file/d/1KHSTIOtJ7QiRgOINaExtj0v5mdGPDQV/view?usp=sharing)

- Link to complete analysis:

[https://colab.research.google.com/drive/15dnbdwltXjt5q6XkPXSuylqLixU- SWQu?usp=sharing](https://colab.research.google.com/drive/15dnbdwltXjt5q6XkPXSuylqLixU-SWQu?usp=sharing)

(You may need to send an access request to view it)

THANK

you