EXECUTIVE SUMMARY REPORT

ON

FARMERS INSURANCE

INTRODUCTION

The Pradhan Mantri Fasal Bima Yojana (PMFBY) is a crop insurance scheme in India that aims to provide financial support to farmers facing crop losses due to unforeseen events.We are provided with the dataset containing insurance details of the farmers under PMFBY insurance scheme to generate insights using SQL.

DATA UPLOADING

This dataset is free and publicly available. The data is derived from the National Data and AnalyticsPlatform (NDAP).The data table contains information on farmers' insurance coverage, population demographics, and land area across various states and districts. There are 44 columns in the dataset

The key attributes include:

•Insurance Data: IncludesInsuranceUnits,FarmersPremiumAmount,InsuredLandArea andtotal premiums.

•Farmer Demographics: Percentage coverage of male, female, and social category farmers(e.g., SC, ST, OBC).

•Geographical Data: Information about states, districts, and land area (urban and rural).

•Population Data: Total population and breakdowns by gender, urban/rural areas, andhouseholds

•Year and Regional Metadata: Year, state, district, and country information.

We have created the schema first and then use the table data import wizard in MySQL Workbench.

QUERIES

1. **names of all states**
   There are 25 states whose data is covered in PMFBY.

2. **total number of farmers covered (TotalFarmersCovered) and the sum insured (SumInsured) for each state (srcStateName)**

Madhya Pradesh is the state with highest no of farmers covered followed by Maharashtra then Uttar Pradesh.
Punjab is the state that has covered least no. of farmers.

3. **total number of farmers covered and the sum of premiums for each state where the insured land area (InsuredLandArea) is greater than 5.0 and the Year is 2018**.

Highest is Tamil Nadu and lowest is for Jammu and Kashmir.

4. Average Insured Land area on Yearly Basis.

It was Highest in 2020 and Lowest in 2018
2020>2019>2021>2018

5. total number of farmers covered for each district

District with the highest farmers covered is Bid followed by latur and Nanded.
District with lowest no. of farmers covered is Kheda then Bhatinda.

6. top 5 districts (srcDistrictName) with the highest TotalPopulation in the year 2020

'Pune', '9429408'
'Thane', '8070032'
'Jaipur', '6626178'
'Nashik', '6107187'
'Allahabad', '5954391'

7. ratio of insured farmers to the total population

CHHATTISGARH	2021	1265556	25405378	0.0498
TRIPURA	2020	172048	3673917	0.0468
TRIPURA	2021	170345	3673917	0.0464

8.There are 211 districts that starts from Letter B.

9.There are 14 states that contains the districts ending with string'pur' and 208 such districts are present across 14 states. Uttarpradesh consists of largest no. of such districts like Saharanpur,Balrampur,Fatehpur etc.

10. There are 553 such districts (srcDistrictName) where the TotalFarmersCovered is greater than the average TotalFarmersCovered across all records.

11. Maharshtra is the state where srcStateName where the SumInsured is higher than the SumInsured of the district with the highest FarmersPremiumAmount.

12. Bhiwadi fatehabad and Hisar are top three srcDistrictName where the FarmersPremiumAmount is higher than the average FarmersPremiumAmount of the state that has the highest TotalPopulation.

13. **rank the districts (srcDistrictName) based on the SumInsured (descending) and partition by alphabetical srcStateName.**

Top 3 districts are of

1 Andhra Pradesh

2 Kurnool

3 Srikakulam

14.calculate a cumulative sum of FarmersPremiumAmount for each district (srcDistrictName), ordered ascending by the srcYear, partitioned by srcStateName.

Top 3 districts are(All from Andhra Pradesh)

Anantpur -2018

Anantpur-2019

Chittoor -2018

ANALYSIS

Over the course of this project, We analyzed the **FarmersInsuranceData** table using a wide range of SQL queries. These queries helped uncover valuable insights about how **farmer insurance premiums and coverage** are distributed across **states, districts, and years**.

Here are the key outcomes and observations:

- Using **RANK()** functions, We ranked districts within each state by **SumInsured**, showing which districts had the **highest financial coverage**.
- With the **SUM() window function**, We calculated **cumulative FarmersPremiumAmount** year-by-year for each district. This revealed **trends in premium growth** over time.
- Queries involving **GROUP BY** and **aggregate functions** helped identify **states with the highest number of farmers covered**, and **states where coverage was poor** despite large populations.
- We handled SQL challenges like the **ONLY_FULL_GROUP_BY** error by understanding why **grouping non-aggregated columns** is restricted — helping us write cleaner and more accurate queries.
- We also explored **ratios**, such as **TotalFarmersCovered to TotalPopulation**, to understand **insurance coverage efficiency** per state and year.
- Ordering and filtering techniques like ORDER BY, HAVING, and conditional WHERE clauses allowed us to isolate important records — such as districts with **lowest non-zero premium amounts**.

- This project demonstrates the power of structured data analysis in driving evidence-based decisions for improving agricultural policy and farmer welfare.

CONCLUSION

The analysis of the *FarmersInsuranceData* revealed valuable insights into how effectively farmer insurance schemes are reaching rural populations. By using SQL queries, we identified the top districts by population, the most affordable active insurance regions, and the state-year combinations with the highest farmer coverage ratios. This data-driven approach highlights regional strengths and gaps, supporting smarter decisions to improve agricultural insurance policies.