CHRIST (Deemed to be University), Bangalore - 560029

Department of Statistics and Data Science

Neekitasingh Rajput

2441339

Class: 3DM

Course: DSC261-3 - Data Visualization

Date: 21/07/2025

Exercise No: Lab Assignment-5

INSURANCE DATASET

About Dataset-

This dataset represents performance metrics of agents working across 10 insurance companies within a regional insurance group. These companies operate in a 17-state area and offer property and casualty insurance, life insurance, and brokerage services. The data captures agent-level activity, product performance, premium details, retention, losses, and demographic indicators. The goal is to analyse agent performance and support segmentation for predictive modelling, ultimately improving business outcomes and operational efficiency.

Key Attributes Used-

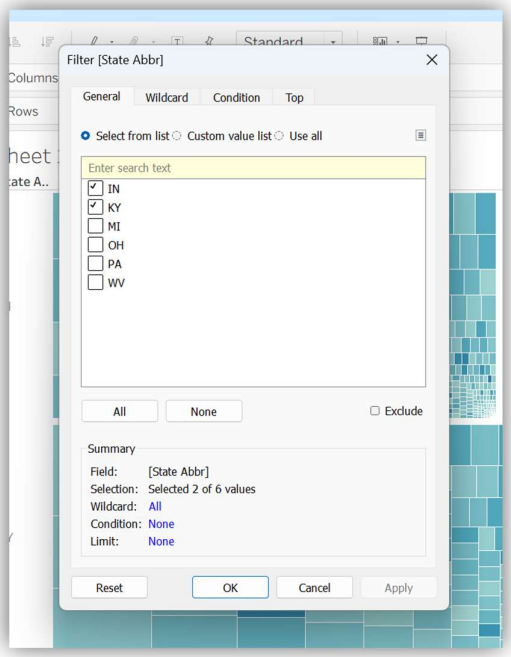
- **AGENCY_ID** – Unique identifier for each insurance agency (Categorical)
- **STATE_ABBR** – State abbreviation where the agency operates (Categorical)
- **PROD_LINE** – Line of insurance product (e.g., Life, Property & Casualty) (Categorical)
- **STAT_PROFILE_DATE_YEAR** – Year of the statistical profile (Date)
- **RETENTION_POLY_QTY** – Number of policies retained by the agency (Numerical)
- **POLY_INFORCE_QTY** – Number of policies currently in force (Numerical)
- **NB_WRTN_PREM_AMT** – New business written premium amount (Numerical)
- **WRTN_PREM_AMT** – Total written premium amount (Numerical)
- **PRD_ERND_PREM_AMT** – Earned premium amount for the period (Numerical)

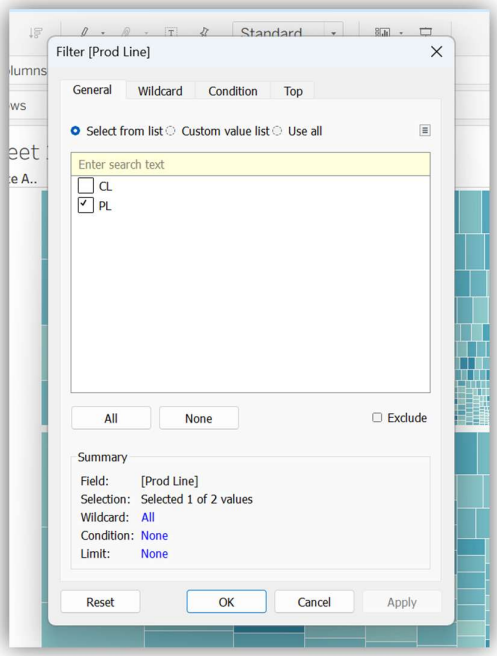
- **PRD_INCRD_LOSSES_AMT** – Incurred losses amount for the period (Numerical)
- **RETENTION_RATIO** – Ratio of retained policies to total policies (Numerical)
- **LOSS_RATIO** – Ratio of incurred losses to earned premiums (Numerical)
- **LOSS_RATIO_3YR** – Three-year average loss ratio (Numerical)
- **GROWTH_RATE_3YR** – Three-year growth rate of premiums (Numerical)
- **ACTIVE_PRODUCERS** – Number of active producers in the agency (Numerical)
- **AGENCY_NAME** – Name of the insurance agency (Categorical)
- **STATE_REGION** – Regional classification of the state (e.g., Midwest, South) (Categorical)
- **CHANNEL_TYPE** – Distribution channel used by the agency (e.g., Direct, Broker) (Categorical)
- **QUOTE_COUNT** – Number of quotes generated by the agency (Numerical)
- **BIND_RATIO** – Ratio of bound policies to quoted policies (Numerical)
- **RENEWAL_PREM_AMT** – Premium amount from renewed policies (Numerical)
- **EFF_DATE** – Effective date of the policy (Date)
- **EXP_DATE** – Expiration date of the policy (Date)
- **POLICY_TERM** – Duration of the policy in months or years (Numerical)
- **CLAIM_COUNT** – Number of claims filed by policyholders (Numerical)
- **AVG_CLAIM_COST** – Average cost per claim (Numerical)
- **ENDORSEMENT_COUNT** – Number of endorsements or changes made to policies (Numerical)

Worksheet 1: Data Source Filters – Regional & Product Focus

This worksheet uses data source filters to focus the analysis on agencies operating in Indiana and Kentucky, specifically within the Personal Lines (PL) product category. The visualization is designed to highlight retention and premium performance across agencies in these states.

Filters Applied:





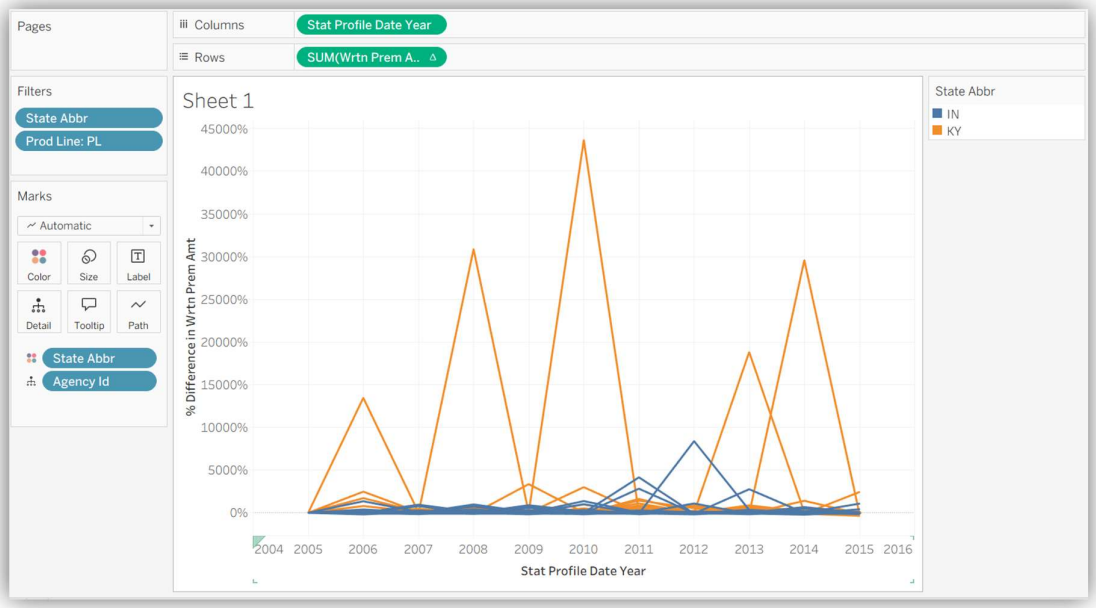
Visualization:

The treemap provides a clear and intuitive view of agency performance across two states. Larger rectangles represent agencies with higher written premium volumes, while darker shades of blue indicate stronger retention ratios. This dual encoding allows for quick identification of high-performing agencies in terms of both business volume and customer retention. By filtering out other states and product lines, the visualization remains focused and relevant to regional strategy and product-specific insights.

Worksheet 2: Quick Table Calculations – Percent Difference

Year-over-Year Premium Change

This worksheet uses Tableau's Quick Table Calculation feature to analyze the year-over-year percentage difference in written premium amounts (WRTN_PREM_AMT) for agencies in Indiana (IN) and Kentucky (KY). The line graph visualizes trends from 2004 to 2015, filtered to Personal Lines (PL) only.



The line chart reveals distinct premium growth patterns between the two states:

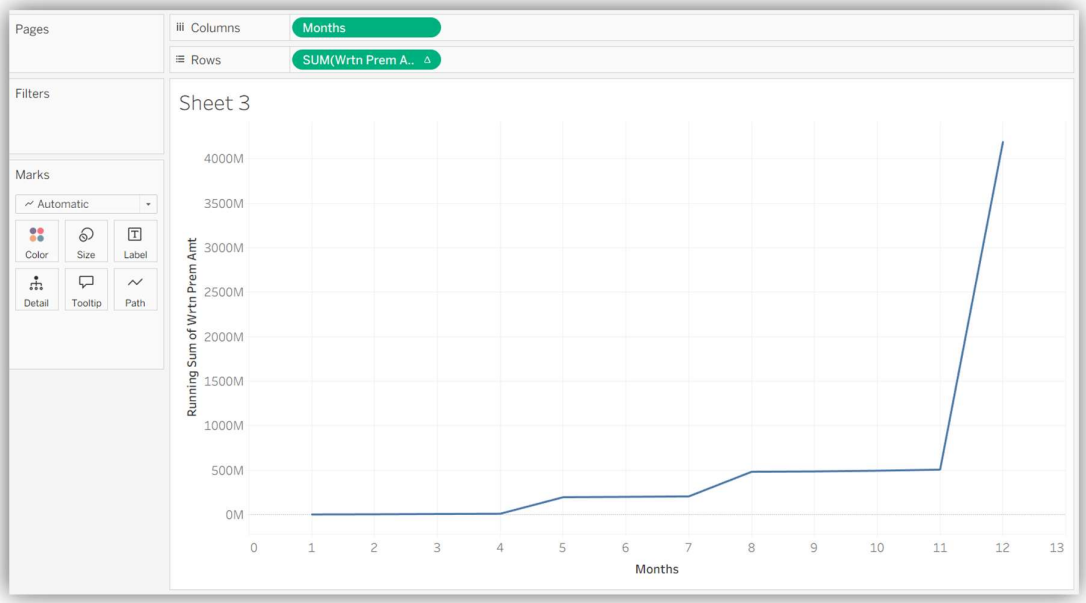
- Kentucky (KY) shows high volatility, with sharp spikes in premium growth during 2006, 2009, 2011, and 2014. These peaks may indicate aggressive expansion, market shifts, or changes in agency performance.
- Indiana (IN) exhibits a more stable trend, with moderate fluctuations and fewer dramatic changes year-over-year.

This analysis helps identify which regions experienced significant growth or instability, offering insights into market dynamics and agency effectiveness. The use of Quick Table Calculations simplifies the process of comparing year-over-year changes without manual computation.

Worksheet 3: Quick Table Calculation – Running Total

Cumulative premium trend

This line chart visualizes the running sum of WRTN_PREM_AMT across a 13-month period. The x-axis represents months (0 to 13), and the y-axis shows the cumulative premium amount in millions (M), ranging from 0M to 4000M.



Using a running total helps reveal cumulative performance trends over time. The sharp rise in Month 12 is a critical insight—it warrants further investigation to understand its cause and implications. This visualization is especially useful for identifying growth patterns, seasonal effects, or reporting lags in insurance premium data.