

Customer Data Analysis Using SQL and Tableau

Objective: The goal of this project is to analyse customer demographics and behaviour to help predict customer retention and support the development of focused customer retention programs. This involves identifying patterns in the dataset, such as customer lifetime value, policy preferences, and retention rates, using SQL for querying the dataset and Tableau for visualizing insights.

Stage 1: Data Understanding and Preparation

1. Dataset Review:

The dataset includes several key columns that provide customer information:

- Demographics: Gender, State, Income, Education, Employment Status, Marital Status, etc.
- Behavioural Data: Customer Lifetime Value, Policy Type, Response, Monthly Premium, Claims, etc.

I chose to focus on columns that offer insights into customer retention, lifetime value, and demographics:

- Customer Lifetime Value to measure long-term profitability.
- Response to assess retention and customer feedback.
- Policy Type to understand product preferences.
- State, Gender, and Income to investigate demographic trends.

2. Data Cleaning and Preparation:

SQL was used to clean the dataset by handling missing values, ensuring consistency in the data format, and performing basic transformations. For example:

- Filtering out invalid or incomplete data entries.
- Formatting categorical data (e.g., state and gender) for easy analysis.

Stage 2: SQL Queries and Analysis

We performed several SQL queries to extract insights from the dataset:

1. Customer Demographics Analysis:

- The query grouped customers by gender and state to identify the distribution of customers.
- **Key Insight:** The majority of female customers are located in California and Oregon, while the majority of male customers are in Oregon and California as well. This insight is useful for targeted regional campaigns.

Results:

- California has the largest number of customers (both male and female).

2. Customer Lifetime Value Analysis:

- I calculated the average customer lifetime value (CLV) grouped by gender.
- **Key Insight:** Female customers have a slightly higher average CLV (\$8,096) than male customers (\$7,909), indicating potential differences in long-term engagement or profitability.

3. Policy Type Distribution:

- The query grouped customers by policy type to understand their distribution across different products.

- **Key Insight:** The majority of customers (around 70%) hold "Personal Auto" policies, indicating this is the most popular product. "Corporate Auto" and "Special Auto" policies have much fewer customers.

4. Retention Analysis:

- We analysed customer responses where the time since the last claim was less than 12 months.
- **Key Insight:** Only 552 out of 3,846 customers responded positively (i.e., indicated a "Yes" for retention), highlighting a challenge in customer retention efforts.

5. Income Distribution:

- The query analysed the distribution of customers by income.
- **Key Insight:** Most customers fall within an income range of \$10,000 to \$11,000, indicating a relatively uniform distribution of income levels.

Stage 3: Tableau Dashboard Creation

Using the SQL insights, I created a dashboard with the following sheets:

1. Customer Demographics:

- **Insight:** Shows the geographical distribution of male and female customers, with California and Oregon having the most customers.

2. Customer Lifetime Value:

- **Insight:** Highlights the slight difference in lifetime value between male and female customers.

3. Policy Type Distribution:

- **Insight:** Emphasizes the popularity of "Personal Auto" policies.

4. Retention Analysis:

- **Insight:** Shows the large gap between customers who responded positively and those who did not.

5. Income Distribution:

- **Insight:** Illustrates the income distribution, which is concentrated between \$10,000 and \$11,000.

Stage 4: Insights from the Dashboard

- **Customer Retention:** The dashboard reveals that customer retention is a challenge, as most customers have not responded positively to retention efforts. This suggests a need for improved retention strategies, particularly for customers who have made recent claims.

- **Policy Preferences:** The majority of customers prefer "Personal Auto" policies, suggesting this is the most lucrative segment for the company to focus on for growth.

- **Demographic Trends:** Female customers, particularly in California and Oregon, have a slightly higher lifetime value than male customers, which can inform targeted retention campaigns.

Conclusion

By leveraging SQL for data preparation and Tableau for visualization, this project provides a clear view of customer demographics, behaviours, and retention patterns. The insights drawn from this analysis can help the company develop targeted strategies to retain high-value customers and expand profitable segments.