IJNRD.ORG ISSN : 2456-4184
INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG
An International Open Access, Peer-reviewed, Refereed Journal

Visualization of Vehicle Insurance Customer Data

Ms. Jacquelin Anushya P, Agathiyan V, Alan John, Antony Sunny, Baskarasethupathy S,
Assistant Professor, Department of Computer Science and Engineering,
"IV'th year, Department of Computer Science and Engineering, SNS College Of Technology, Coimbatore, Tamil Nadu, India

Abstract: Due to the intense competition in the auto insurance market, it is crucial for insurance providers to continuously assess and comprehend their clientele. In this project, we demonstrate the visualization of a vehicle insurance customer dataset using Tableau for dashboard development and Python for exploratory data analysis (EDA) and data cleaning. Our data sheds light on the traits, habits, and preferences of people who purchase auto insurance. We may study many parts of the data, such as customer demographics, car attributes, insurance coverage, and claims history, using the visualization dashboard's various charts and graphs, including bar charts, line charts, scatter plots, and heat maps. Our work highlights the value of data analysis and visualization for comprehending customer data *for automobile insurance, and it can assist insurance providers in better understanding their clients and creating more successful marketing campaigns.*

Keywords: Tableau, EDA, Data Cleaning, Dashboard, Insurance

1. INTRODUCTION

Because the auto insurance market is so dynamic and fiercely competitive, insurance providers must continually assess and comprehend their clientele in order to remain competitive. With the help of Tableau for creating visualisation dashboards and Python for exploratory data analysis (EDA) and data cleaning, we want to give a thorough study and visualisation of a dataset of automobile insurance customers in this project.

The dataset was cleaned using Python as the initial stage in our study. Customer demographics, car attributes, insurance coverage, and claims history were only a few of the elements in the dataset. We started by finding outliers and missing values in the dataset, which were then either deleted or imputed depending on how much data was missing. In order to establish additional variables that would aid in our understanding of the data, we also transformed the data and engineered features.

Following the dataset's cleaning and transformation, we utilised Tableau to provide a visualisation dashboard. Bar charts, line charts, scatter plots, and heat maps were just a few of the graphs and charts that were displayed on the dashboard. We were able to investigate several parts of the data, such as client demographics, car characteristics, insurance coverage, and claims history, thanks to these charts and graphs.

Our investigation revealed some important information about customer demographics. The bulk of the patrons were men, and their average age was 40, according to our research. We also discovered that the majority of the clients were married and most had completed high school or college. We discovered that the bulk of the clients held professional or technical jobs in terms of occupation.

Our study also yielded important information on vehicle attributes. Sedans were the most common type of vehicle owned by consumers, followed by SUVs and trucks. Also, we discovered that the majority of the automobiles were under $20,000 in value and that the average age of the vehicles was around five years.

The majority of the clients, according to our analysis of insurance policies, had comprehensive coverage, which was followed by collision and liability coverage. The bulk of the consumers had no recent claims, according to our analysis of their claims history.

Overall, our investigation gave us useful information about the traits, habits, and preferences of the people who buy auto insurance. The visualisation dashboard gave us a better way to showcase the findings and allowed us to examine many facets of the data. We can assist insurance firms in better understanding their clients and creating more successful marketing campaigns by utilising data analysis and visualisation approaches.

Due to the intense competition in the auto insurance market, it is crucial for insurance providers to continuously assess and comprehend their clientele. In this project, we demonstrate the visualization of a vehicle insurance customer dataset using Tableau for dashboard development and Python for exploratory data analysis (EDA) and data cleaning. Our data sheds light on the traits, habits, and

preferences of people who purchase auto insurance. We may study many parts of the data, such as customer demographics, car attributes, insurance coverage, and claims history.Our work highlights the value of data analysis and visualisation for comprehending client data for automobile insurance and can offer insurance businesses insightful information to remain competitive in a continuously shifting market. Our study can be used to find trends and patterns in the data that may be helpful for creating new goods or services, in addition to offering insights into client demographics, car attributes, insurance coverage, and claims history.

In conclusion, this project offers a helpful illustration of how data analysis and visualisation methods can be utilised to uncover insights into client data for vehicle insurance. We were able to provide a thorough analysis of the dataset using Python for EDA and Tableau for creating visualisation dashboards. This analysis can assist insurance companies in better understanding their clients and creating more successful marketing campaigns. We hope that our research can serve as a template for future studies of consumer data from automobile insurance and that it can aid in the creation of more creative and useful insurance products and services.

2. METHODOLOGY

2.1 Data Cleaning

All data analysis effort must include data cleansing since it makes sure the data is correct, full, and consistent. For the data cleansing process in our project, which included numerous processes, we employed Python. The dataset was first imported using the Pandas library in Python. After that, we evaluated the data's quality by looking for outliers, mismatched data types, and missing values. We used the info() method to look for any missing values and the describe() function to generate an overview of the dataset.

We choose to impute the missing values using the mean, median, or mode of the relevant column after discovering the missing values. This choice was made based on the characteristics of the missing values, and we made sure that imputed values wouldn't have a substantial effect on the analysis as a whole.

In order to construct additional variables that would aid in our understanding of the data, we also did feature engineering and data transformations. For instance, by deducting the year of manufacturing from the present year, we were able to generate a new variable for the age of the vehicle.

We used domain expertise to validate the data and make sure it was logically sound in order to guarantee data consistency. For instance, we looked to see if the customer's age matched the age of the car they owned.

Next, in order to avoid bias in our study, we eliminated any redundant data entries. To get rid of any duplicate rows, we utilised the drop duplicates() function.

Overall, the data cleaning procedure was crucial to assuring the objectivity and accuracy of our study. We were able to automate the data cleaning process and apply sophisticated data transformations thanks to the Python programming language and its libraries. We were able to produce insightful visualisation dashboards using Tableau and derive relevant conclusions from our investigation by making sure the data was accurate and consistent.

2.2 Data Visualization

We utilised Tableau for data visualisation after using Python to clean up the dataset. Tableau is an effective application for data visualisation that enables us to build interactive dashboards and visuals that aid in understanding the data.

Finding the important factors that would aid in understanding the data and providing answers to our research questions was the first stage in the data visualisation process. Age, gender, the type of vehicle, and the location were some of the factors we listed.

Afterwards, we produced a number of visualisations to aid in our data exploration and insight-gathering. To illustrate the data, we employed a variety of charts, including bar charts, line charts, scatterplots, and heat maps. For instance, we produced a scatterplot showing the correlation between age and the insured vehicle's worth, a bar chart showing the distribution of clients by gender, and a heat map displaying the areas with the most customers.

Also, to help us emphasise crucial insights and improve the visualisations' readability, we applied colour coding and labelling. To differentiate between male and female clients, for instance, we utilised various colours, and we named each section on the map.

We used filters and settings that let people examine the data and alter the visuals to make them more interactive. For instance, we developed a filter that let users choose a certain automobile brand and view how it was spread across various locations.

Also, in order to link several visualisations and make it simpler for users to browse the data, we implemented dashboard actions. For instance, we developed a dashboard activity that let users select a location on the map and view the breakdown of clients by age and gender.

Ultimately, we effectively shared our views and findings using narrative strategies. We developed a narrative to guide readers through the data and emphasise significant findings, such as the fact that male consumers were more likely to possess expensive vehicles or that the West area had the greatest number of customers.

We used colour coding and labelling to highlight significant insights, interactive visualisations with filters and parameters, dashboard actions to link various visualisations, and storytelling techniques to communicate our findings. To summarise, our data visualisation methodology involved identifying key variables, creating visualisations that helped us explore the data and gain insights, using colour coding and labelling to highlight important insights, and using interactive visualisations with filters and parameters. As a result, we were able to develop a dashboard that was both insightful and interesting and offered information on the customer data for auto insurance.

3. INSIGHTS AND DISCUSSION

Many significant findings from our examination of the customer data for automobile insurance can assist insurance businesses in better understanding their clients and streamlining processes.

The correlation between the insured vehicle's worth and the customer's age was one of the most important findings we made. According to our data, younger consumers were more likely to insure automobiles with lower values, whereas older customers were more likely to insure vehicles with greater values. This discovery may enable insurance providers to more effectively target various age groups with their marketing campaigns and pricing plans.

Also, we discovered that compared to female consumers, male customers were more inclined to insure higher-value vehicles. This knowledge can aid insurance businesses in better comprehending the wants and preferences of their consumers that are specific to gender and modifying their services accordingly.

The distribution of customers across different geographies was another intriguing finding. We discovered that the West region had the most customers, followed by the South and Northeast. By focusing their marketing efforts in areas with the greatest client bases, insurance companies can more efficiently manage their resources.

In addition, we found that Ford, Toyota, and Honda were the most popular car brands with our clients. Insurance businesses can use this information to better understand the requirements and preferences of their consumers and change the products they offer as a result.

A association between the customer's age and the kind of coverage they had was also found by our investigation. Younger customers had a higher propensity to have liability-only coverage, whereas older customers had a higher propensity to have comprehensive coverage, according to our research. This knowledge will enable insurance companies to provide more specialised coverage alternatives and age-specific product offerings.

Furthermore, our data revealed that the average premium was lowest for compact automobiles and sedans and greatest for sports cars and luxury vehicles. Based on the sort of car that their clients own, insurance companies might modify their pricing tactics with the aid of this information.

Overall, our research of the customer data for vehicle insurance offers some significant insights that might aid insurance businesses in better understanding their clients and modifying their marketing approaches accordingly. This data can be used by insurance businesses to enhance their product lines, marketing plans, and pricing ideas, which will ultimately improve customer satisfaction and boost profitability.

. CONCLUSION

Our project demonstrates the usefulness of data analysis and visualization in understanding vehicle insurance customer data. By using Python and Tableau, we were able to explore and clean the dataset, and create a visualization dashboard that provides insights into the characteristics and behavior of the customers. We believe that our work can help insurance companies to better understand their customers and develop more effective marketing strategies. Future work could include the use of machine learning algorithms to predict customer behavior and preferences based on the available data.