PROJECT ON

DATA VISUALIZATION TECHNIQUES

Name: Reena Roy

Date: 18-04-2025

Contents List:

Problem Statement:

Business Context

We are all aware that accidents are prone everywhere due to negligent driving or climatic conditions. An insurance company always needs to be prepared to estimate the number of accidents and the claims that they can receive at a given point time. Also understanding the pattern of claims would help the companies to frame different types of policies for the users providing better benefits and at the same time increasing the premium to the company.

Problem Statement:

Consider that you are a Lead Data Analyst at an Insurance Claims company that has provided you with the Car Insurance Claims dataset. You have been given a task to explore the data, create different plots and interpret useful insights/findings. Your end goal here will be to create a storyboard that you have to present to the Senior Management and the story has to have an end objective and should follow a logical flow to display that you are heading towards achieving the end objective. This will help the Senior Management in taking some decisive actions on the current claims system in place. This storyboard will be an open-ended story for you to explore various different features in the data and try to showcase different plots. Make sure to have minimum clutter in the plots, follow a consistent color scheme across all the plots, and use proper colours to highlight a specific insight. Moreover, your plots on all the dashboards should be interactive and responsive. *There should be 1 dashboard that should cover the summary of the story as well as your recommendations.*

Summary:

We have been given the dataset of car insurance claim. The dataset has 7648 rows and 21 columns. Based on the dataset different plots and graphs are designed to show the correlation between variables and derive insights. The dataset is first extracted on the Tableau public version. The variables are segregated into dimensions and measures. We have two main measure values ---- Old Claim Amt and Claim Amount which is compared against other variables.

Data Dictionary:

Assumption	Car Owner and Driver are same Amounts are in Dollars ($)
ID	Identification Variable
KIDSDRIV	Number of teenagers among the car owner's children who can drive a car.
BIRTH	Date of birth of the driver
HOMEKIDS	No of childern the car owner has
YOJ	Years on Job. How many years has the owner of the car been working?
INCOME	Income of the driver
PARENT1	Is the car owner a Single Parent
HOME_VAL	Value of the house owned by the car owner
MSTATUS	Marital status of the car owner
GENDER	Gender of the driver

EDUCATION	Maximum Education level of the driver
OCCUPATION	Occupation of the driver
TRAVTIME	Time taken to get to work on an average
CAR_USE	Purpose of using the car
BLUEBOOK	What is the worth of the car. Value of the Vehicle(in dollars)
CAR_TYPE	Car type
OLDCLAIM	Total claim (in past 5 years - in dollars)
CLM_FREQ	Number of claims (in past 5 years)
CLM_AMT	If car was in a crash, what is the currently claimed amount(in dollars)
CAR_AGE	Age of car
URBANICITY	Where the car is being driven primarily

Worksheet 1- Bar Graph of Old Claim against Car Type and Car Use. The graph is coupled with filter for Urbancity and Gender.

https://public.tableau.com/app/profile/reena.roy/viz/DVT_Project_16294429484690/OldClaim-BarChart?publish=yes

Insights:

1. The Old Claim amount for commercial use is more than private use ---18,615,157> 11,165,486. Amongst the Private use, car type, SUV, Minivan and Sports car is max claim. For Commercial use, car type- Pickup, SUV and Minivan have max claim.
2. For Highly Rural area the car type claim for Private and commercial sue have similar trend. Private use claim amt is more than commercial use. 11 : 8
3. For Highly Urban area the trend for car type and car sue is similar. Claim amount ration for Private against commercial sue is 173: 103
4. For Female gender, Highly Rural the claim amt ratio is 8: 4 for private to commercial sue. The trend is similar for car type and car use. For Urban it is 117: 43
5. For Male, Highly Rural ratio claim amount for commercial use is more than private--- 5.5 : 3. The car type is also different with max claim for Minivan and Pick up for private type whereas for commercial type the car type is pickup, Minivan and panel truck.
6. For Male, Highly Urban---the claim amount is more for private than commercial 61:59. The car type is Minivan, Pickup and van for private and Pickup, Panel and van for commercial.

Worksheet 2- Histogram of old claim amt against count of data.

https://public.tableau.com/app/profile/reena.roy/viz/DVT_Project_16294429484690/Histogram-oldclm?publish=yes

1. The graph shows the relation between old claim amount which is split into bins and data count. There are 4810 claims between 0 and 1K claims amount. Towards the tail end of graph we see that the no of claims decreases with increased value.
2. Most of the claims are between 0 K and 15K claim amount. There are lessor no of claims for high claim value.

Worksheet 3- **Boxplot of Old Claim amount against Car Type**

https://public.tableau.com/app/profile/reena.roy/viz/DVT_Project_16294429484690/Boxplot-oldclm?publish=yes

1. The median claim amount is at 5 M. The lowest whisker is at 1.7M and lowest hinge value at 2.4M
2. The Upper hinge value is at 7.4M and upper whisker at 9 M

Worksheet 4- Bubble Chart-Old Claim Vs Occupation and Urbancity

https://public.tableau.com/app/profile/reena.roy/viz/DVT_Project_16294429484690/Bubble-Oldclm?publish=yes

1. Blue Collar has highest claim value at 7.7 M. Next is clerical at 5.4M, Professional at 4.4M and least for doctor at 0.7M
2. For Highly Urban the trend for Occupation wise Claim value is similar with Blue Collar, Clerical and Professional claiming high value
3. For Highly Rural area the Clerical claim highest value followed by Student and Blue Collar

Worksheet 5- Line and Area Plot of Old Claim against Car Age

https://public.tableau.com/app/profile/reena.roy/viz/DVT_Project_16294429484690/Trend-oldclm?publish=yes

1. The Claim amount is maximum at car age of 1 Year. The Line Graph is plotted alongside the area plot for moving average
2. The trend for line graph and moving average is same with peak average at car age of 1 and between 7 & 10
3. The graph shows that most of claims are claimed at earlier stages of car age and mean average for most claims are done between 5 to 15 years of car age, after which the claims come down. The max car age is 28 years

Worksheet 6- Pareto Chart-Old Claim against Car Type

https://public.tableau.com/app/profile/reena.roy/viz/DVT_Project_16294429484690/Pareto-oldclm?publish=yes

1. Shows cumulative percentage of total of Old Claim amount for different Car types

2. More than 80% of the claims are taken by car types SUV, Minivan, Pickup and Sports car
3. The Panel Truck and van claim only 12% to 14% of total claim

Worksheet 7- Funnel Chart- Claim amount. The Claim amount is the amount claimed currently, if the car was in a crash.

https://public.tableau.com/app/profile/reena.roy/viz/DVT_Project_16294429484690/Clm-Amt-Funnel?publish=yes

1. The Funnel shows the claim amount for different cars for highest at the top to lowest at the bottom.
2. The highest amount claimed is for the SUV cars at 3.6 M which also means that the maximum car crashes occurred for SUV cars
3. The second claimed amount is for Pickup cars at 2.0 M and the least is for Panel Truck at 0.9 M

Worksheet 8- Pareto chart for Claim Amount against Car Type

https://public.tableau.com/app/profile/reena.roy/viz/DVT_Project_16294429484690/pareto-Clmamt?publish=yes

1. More than 80% of claimed amount is from the cars type- SUV, Minivan, Pickup and Sports car
2. This also shows that most of car accidents have occurred with these cars and will need further analysis on the type of use and the profile of drivers

Worksheet 9- Wordcloud for Claim amount for different occupation types of the owners and the count of claims respectively

https://public.tableau.com/app/profile/reena.roy/viz/DVT_Project_16294429484690/Wordcloud?publish=yes

1. People who are into blue collared jobs have maximum claims of 1853 and 3.8 M value.
2. Second is people with clerical jobs at 1287 claims of value ~2 M
3. Third is Professional occupation with 1097 claims and value around 1.7M
4. Though people with occupation role of Manager and lawyer have higher no of claims at 999 and 823 claims, the students even with lessor no of claims count of 688 have higher claim value of 1.2 M which means more severe accidents and casualty caused by students driving the cars

Worksheet 10- Histogram of Claim Amount

https://public.tableau.com/app/profile/reena.roy/viz/DVT_Project_16294429484690/Histogram-Clmamt?publish=yes

1. 5691 claims are claimed between claim amount of 0K and 1K
2. 233 claims are claimed between 1k and 2 K
3. As the claim value increases the no of claims done decreases significantly which means the most of the accidents are less severe and do not cause much casualty

Worksheet 11- Boxplot of Claim Amount

https://public.tableau.com/app/profile/reena.roy/viz/DVT_Project_16294429484690/Boxplot-ClmAmt?publish=yes

1. The boxplot shows the claim amount against the car type. We see there is an outlier present in the boxplot
2. The lower whisker is at 0.9 M whereas lower hinge is at 1.0 M
3. The Median of Claimed amount is at 1.8 M
4. The upper hinge is at 2.0 M

Worksheet 12- Bubble Chart-Claim Amount

https://public.tableau.com/app/profile/reena.roy/viz/DVT_Project_16294429484690/Bubble-Clmamt?publish=yes

1. The Bubble Chart shows the relation between claimed amount and the occupation and unbancity
2. Clerical Occupation is highest claim 0.167 M with 442 records for High Rural Area
3. Next is Students with 0.142 M and 231 records
4. Blue collar has 460 no of counts but the claim amount is less than clerical and students at 0.132 M
5. For Highly Urban region the Blue Collar has highest no and highest value claims—1393 count and 3.7 M
6. Next is Clerical with 845 records and 1.8 M followed by Professional Occupation

Worksheet 13- Claim Category

https://public.tableau.com/app/profile/reena.roy/viz/DVT_Project_16294429484690/ClmCategory?publish=yes

1. We have used If Then else condition to create calculated field called Claim Category for different categories of claim amount. This helps to have a glance to look at the count and value for different categories of claim like High Claim Value, Low and medium claim values. Also we can see that how many zero claim counts are present.

Worksheet 14- Line Graph for Old Claim and Claim Amount

https://public.tableau.com/app/profile/reena.roy/viz/DVT_Project_16294429484690/LineGraph-OldClmClmAmt?publish=yes

1. Both Old Claim and Claim Amount are converted to percentage of total and plotted car age
2. % Total for Old claim amount is more that recent claim amount but the trend is majorly similar with highest claim values at earlier car age between 1 and 2 years and again between 6 & 13 years post which the claim value drops significantly.

Worksheet- Bar Chart for Old Claim and Claim Amount

https://public.tableau.com/app/profile/reena.roy/viz/DVT_Project_16294429484690/CompbwoldclmClmamt?publish=yes

1. A Bar Chart is plotted for old claim and claim amount against car type and measure names. We also apply filter for Gender and Urbancity
2. SUV car has highest claim followed by Minivan
3. For Rural Region and for female the claim is highest SUV and Sports car at 0.8M and 0.35M whereas the Male have highest claim of 0.39 M for Minivan and 0.2 M for Pickup car
4. For Urban Region Female claim highest claim of 8M for SUV and 4M for Sports car. Male claim highest 4M for Minivan and 3M for Pickup car

Worksheet – Stack bar chart for Old Claim and Claim Amount

https://public.tableau.com/app/profile/reena.roy/viz/DVT_Project_16294429484690/InsuranceMix-Both?publish=yes

1. People with Blue Collar Occupation and High School Education have highest claim value followed by Bachelors and Less than High School
2. Doctors with PHD degree have least claim
3. Amongst Male and single parent with kids the blue collared jobs have highest claims for last five years. In these high school and less than high school have higher claims. Current claims have high school, bachelors and less than high school have higher claims
4. Amongst female single parent with kids, old claim has highest claim by clerical jobs with high school, bachelors and less than high school degree. Next highest claims is by lawyers and doctors
5. Current Claim has highest claims by professionals with Bachelors degree

Worksheet- Dashboard 1 shows the trend chart for Old Claims made in the last five years.

https://public.tableau.com/app/profile/reena.roy/viz/DVT_Project_16294429484690/scatterplot-both?publish=yes

1. The Pareto Chart which shows the Cumulative % of Total Claim Amount for different car types shows that more than 80% of the claims are made by SUV, Pickup, Minivan and Sports.
2. The trendline for Old Claim amount against car age shows the correlation between Claims amount and the Car Age wherein we see that maximum claims are made between 5 and 15 years of car age

Worksheet –Dashboard 2- Four Graphs are placed in the four quadrant to see different relation between attributes and old claims made in the last five years. Graph

1. Bar Chart showing correlation between old claim and car type and car use with filters for urbancity and gender
2. Histogram to show claim count and claim amount
3. Boxplot to show the claim details. Median Claim amount is around 5 M
4. Bubble chart to show claim amount occupation and urban city wise

Worksheet –Dashboard 3- Four Graphs to show the different aspect of single variable Claim Amount for current car crash cases

1. Boxplot with median claim value at 1.7M and presence of outlier
2. Histogram showing claim amount and claim count for the current claims for car accidents
3. Pareto Chart for cumulative % of claim amount
4. Funnel Chart to show the claim amount in descending order for different car types

Worksheet- Dashboard 4- Three Graphs are placed in this dashboard to show the Relation for Recent Claim Amount against the occupation for people who have claimed maximum amount and the different categories of claim

1. Wordcloud to show the occupation wise relation for the claimed amount
2. Bubble Chart to show occupation wise relation with urbancity
3. Claim Category through calculated field

Worksheet- Dashboard 5- Shows Comparative picture for the old claims and the recent claim with four graph placed in four quadrant.

1. Stacked Bar Chart to show comparative relation against occupation, education, gender, whether single parent with kids etc.
2. Bar Chart to show comparison against car type and car use
3. Line Graph to show comparative picture against car age
4. Scatter Plot shows the social status like income and home value alongwith occupation and car type for the people who have claimed insurance

Dashboard 6- Summary of Insights

1. The social status of people who have claimed maximum claim is in the income range of 0 M to 5M with occupation into blue collar jobs, clerical, professionals and students
2. Females have claimed more than males and are driving luxury brand cars like SUV's and Sports car
3. Claim Insurance is more in Urban city for private type use
4. High end claims value are less which means most the insurance is claimed for minor accidents
5. Maximum Claims are made by SUV, PickUp car, Minivan and Sports Car with car age lying between 5 years and 15 years

Storyboard- Key Insights

1. Histogram for Old Claim- The old claims for the car insurance in the past five years have been claimed between 0$ and 10,000$ value. Higher value claims are less
2. Pareto Chart for Old Claim- More than 85% of the claims were made by SUV, Minivan, Pickup and Sports Car
3. Bar Chart for Old Claim- Further we get more insights from gender wise, car use type and Urban city wise that higher no of claims are made by females for Private sue type who majorly drive SUV's and Sports Car. Even rural area have females who are driving SUV's and Sports claiming higher nos.
4. Bubble Chart for old claim- occupation wise, people with blue collar jobs have made highest claims. Next are people with clerical and Professional jobs. Very highly educated with occupation like doctors and homemakers have claimed least. Important point to note is that even students with high school and graduation have claimed significantly
5. Histogram for Recent Claims- Claims made recently majorly for car crashes shows that most claims are made between 0$ and 8000$ and very less claims are made for higher value. this goes on to say that most car crashes were not very fatal with any major damages
6. Pareto Chart for Recent Claims- Recent Claims for car insurance have ~83% claims made by SUV, Pick Up, Minivan and Sports Car
7. Category wise chart for Recent Claim- In case of comparison of category wise claims we see that a significant no of people ~5598 out of 7647 have not claimed any insurance. Only 27% people have claimed insurance
8. Word cloud- Further we see that Blue Collar jobs have higher claims followed by Clerical and Professional jobs. Even Students have significant claims at 1.2M $
9. Stack bar Chart for Both Claims- The Stack bar chart in comparison mode for both types of claims shows Professionals with bachelor's degree who are female, single parent with kids have claimed highest for recent claims whereas for old claims female, single parent with kids who are into clerical with education of high school, bachelors and less than high school have highest claims
10. Line Graph-Both- Trend Line for old claims as well as recent claims is similar when compared against car age. Most claims are between 5 years and 15 years of the car age

Link for the DVT Project

https://public.tableau.com/views/DVT_Project_16294429484690/Story2?:language=en-US&publish=yes&:display_count=n&:origin=viz_share_link

https://public.tableau.com/app/profile/reena.roy/viz/DVT_Project_16294429484690/Story2?publish=yes

https://public.tableau.com/app/profile/reena.roy/viz/DVT_Project_16294429484690/Dashboard1?publish=yes

https://public.tableau.com/app/profile/reena.roy/viz/DVT_Project_16294429484690/Dashboard2?publish=yes

https://public.tableau.com/app/profile/reena.roy/viz/DVT_Project_16294429484690/Dashboard3?publish=yes

https://public.tableau.com/app/profile/reena.roy/viz/DVT_Project_16294429484690/Dashboard4?publish=yes
https://public.tableau.com/app/profile/reena.roy/viz/DVT_Project_16294429484690/Dashboard5?publish=yes
https://public.tableau.com/app/profile/reena.roy/viz/DVT_Project_16294429484690/Dashboard6?publish=yes

---

THE END