DATA VISUALIZATION (DATA 230) – SEC21

POWER BI DASHBOARD on VEHICLE INSURANCE DATASET

Assignment 10 (Final Submission)

Sakshi Mukkirwar
sakshimanish.mukkirwar@sjsu.edu

Dataset Used: Vehicle Insurance EDA and Boosting Models
Source: Kaggle (link to the dataset)

OBJECTIVE:

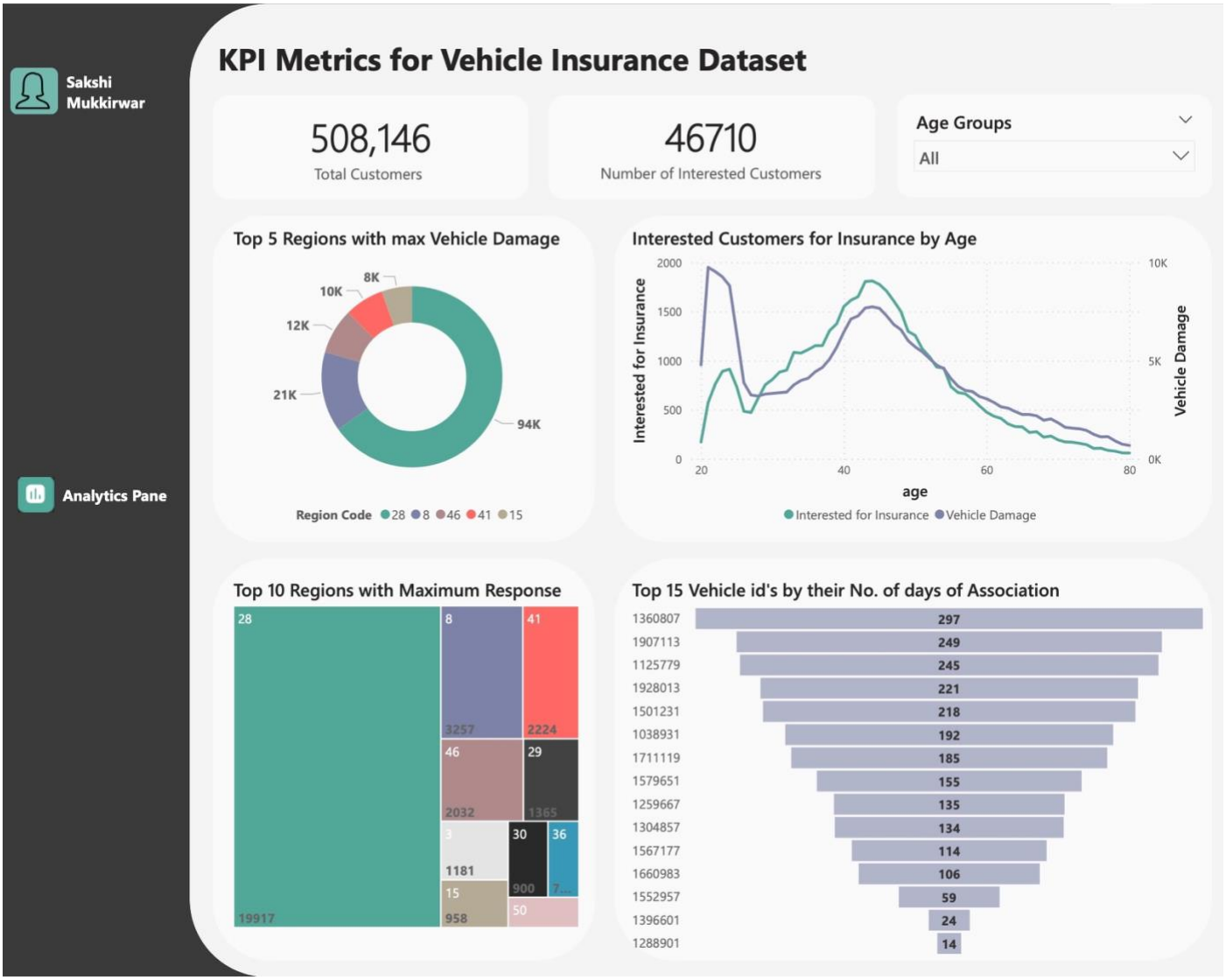
The goal of developing a Power BI dashboard using the "Vehicle Insurance EDA and Boosting Models" dataset is to get understanding of the variables impacting the purchase of vehicle insurance and to identify trends that can be used to improve the marketing strategy and boost cross-selling possibilities. The dataset is intended for exploratory data analysis (EDA) and building boosting models for vehicle insurance prediction. It likely contains information about customers and their vehicle insurance policies.
The aim is to identify significant trends, patterns, and links between customer variables (such as age, gender, area, and vehicle information) and the likelihood to buy insurance for vehicles by visualizing and evaluating the dataset with Power BI.
The Power BI dashboard will give a complete picture of the dataset and offer informative graphics and precise statistics.

STEPS TO DEVELOP:

Steps followed to create the dashboard:

1. **"Get Data":** Connecting to our dataset using the get data option in Power BI Desktop.
2. **Data Preparation**: I checked for null/missing values, checked for duplicate data and then normalized the data.
3. **Planning the Dashboard Layout**: I planned the layout and structure of the dashboard, including the placement of visualizations, filters, and other interactive elements. I even planned the visualizations and the kind of charts that I can display based on my dataset.
4. **Create the Visualizations**: Built individual visualizations using Power BI's visualizations pane and then configured each visualization by selecting appropriate data fields, adjusting properties (such as colors, labels, and axes), and applying filters or slicers as needed.
5. **Interactivity and Slicers:** Added interactive features to the dashboard, such as slicers and turned on the interaction of every visual with another.
6. **Testing:** Test the functionality and usability of the dashboard by interacting with the visualizations and validating the accuracy of the displayed information.

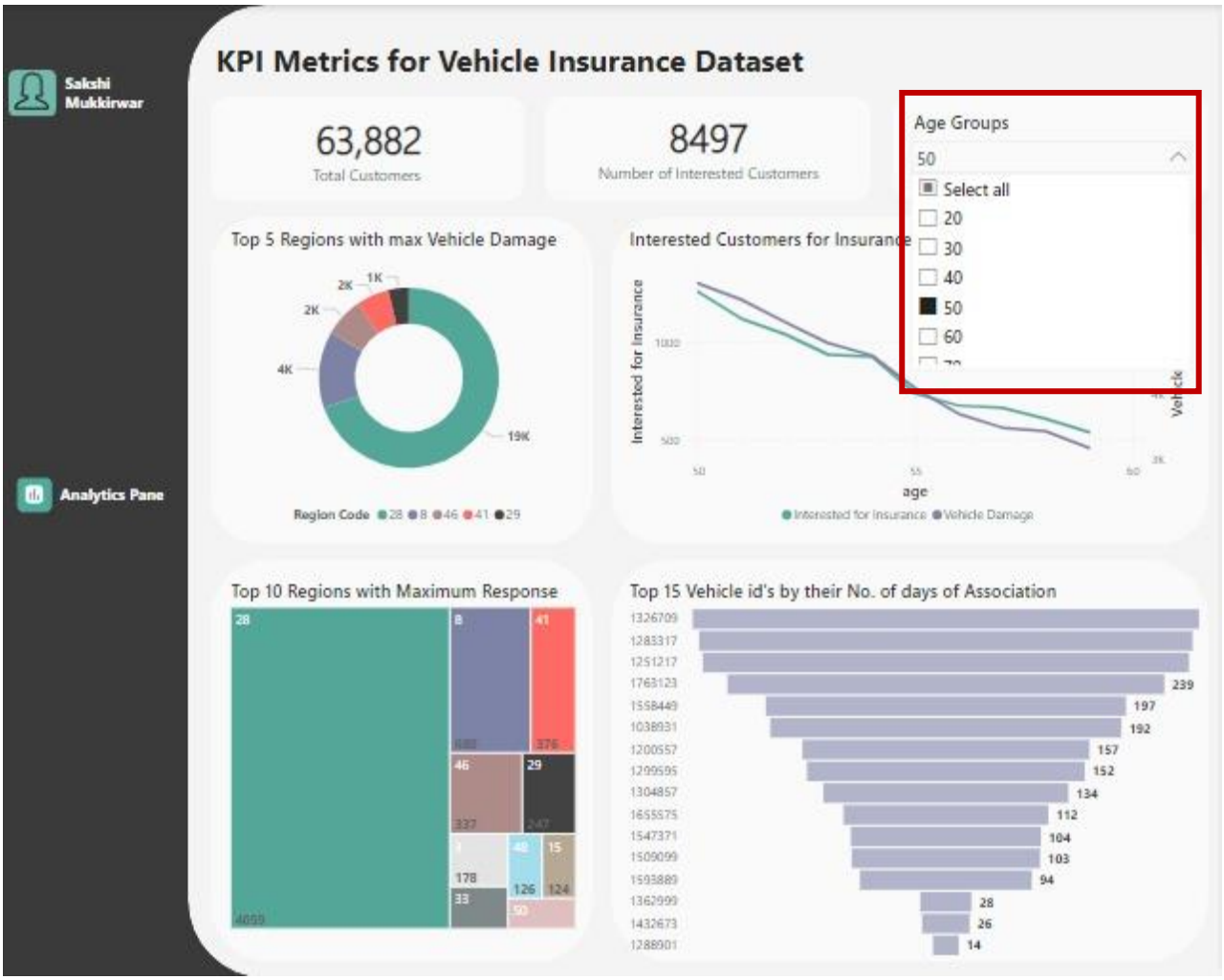
Dashboard:



a) The top two cards show us most important metric that is the total number of customers and the number of interested customers in the vehicle insurance.
b) The first visual is a donut chart that helps us to immediately find the top 2 regions that has maximum number of vehicle hazards.
c) The second visual is a line chart that shows us the trendlines for the customers interested in the insurance and vehicle damage by age. We can easily say that for the age 20 to 80 almost all the customers who have faced damage are interested in the vehicle insurance.
d) The treemap shows us the top 10 regions from where the maximum number of customers are interested to buy the insurance.
e) The fourth visual is a funnel chart that gives us an idea of the existing model. It shows the businesspeople that current people who are associated with the insurance company. Top 15 vehicle ids are mentioned here based on their number of days associated with the company.

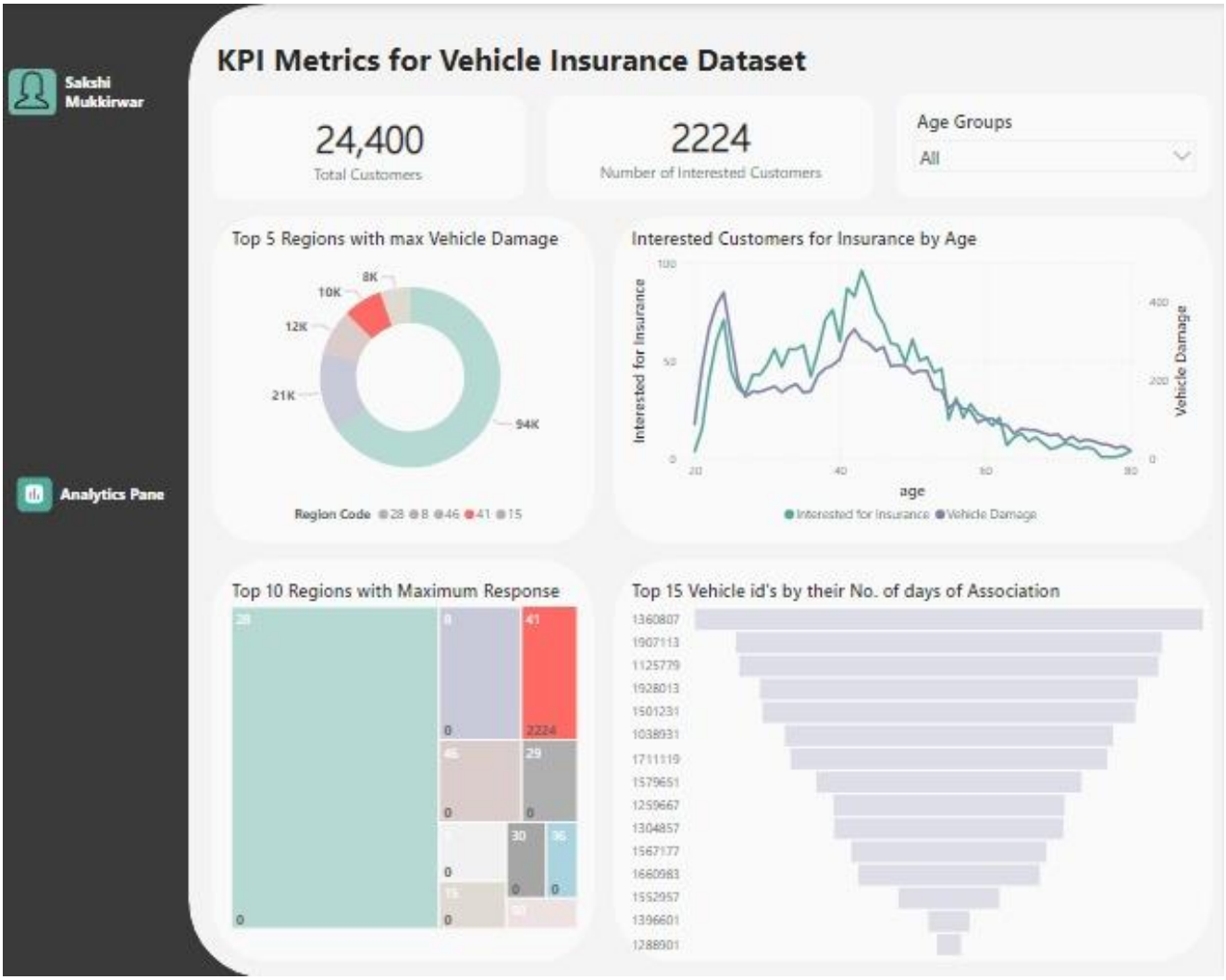
KEY PRINCIPLES EMPLOYED:

1. **Slicers:** Slicers are interactive filters that allow users to dynamically slice and dice the data displayed in the visualizations. I have used an Age Group slicer that will allow the end user to slice through the data based on various age groups.



The slicer has filtered the data for all the visuals for the age group “50”.

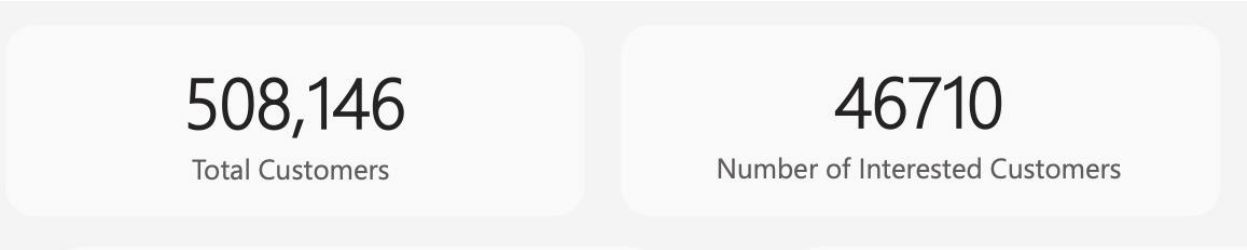
2. **Interactive Visuals**: Interactive visuals are visualizations that respond to user interactions. They enable users to drill down into the data, highlight specific data points, or toggle between different views.



As we can see, on selecting the region code 41, the data for region code 41 has been highlighted in all the visuals. This is how all the visuals have been made interactive.

3. **Color Theme Used:** I have used a consistent color theme throughout the dashboard, that has toned down shades of all the colors and maintain a perfect contrast. It helps to maintain visual consistency and ensure that different visualizations work together harmoniously.
4. **Gestalt Principles**: Gestalt principles are design principles that describe how humans perceive and organize visual information. I have incorporated the **principle of enclosure** (all the visuals are enclosed inside a confined shape) and **principle of continuity** (the data labels are differentiated using the distinct colors throughout the dashboard).
5. **Fonts and Typography**: Fonts and typography choices play a significant role in the dashboard's readability and aesthetics. I have selected appropriate fonts that are easy to read, considering factors like font size, line spacing, and font styles.

6. **Use of Icons where necessary:** I have used icons at two places (at profile picture and the dashboard pane). I thought here the icons could represent actions in a concise and visually engaging way.
7. **Major KPI Described:** Key Performance Indicators (KPIs) are essential metrics that provide insights into the performance of a business or process. The two cards at the top of the dashboard are most important KPIs for this dataset. It shows a direct relation between the total number of customers and the number of interested customers in the insurance.



8. **Types of Visual:** The choice of visualizations is crucial for effectively presenting the data and conveying insights. I have used four different types of visualizations that are very intuitive and convey the information in the best possible way.