

Case Study

1. Import claims_data.csv and cust_demographics.csv and combine the two datasets appropriately to create a 360-degree view of the data. Use the same for the subsequent questions.
2. Calculate the age of the customers in years. Based on the age, categorize the customers according to the below criteria:
 - a. Children < 18
 - b. Youth 18 – 30
 - c. Adult 30 – 60
 - d. Senior > 60
3. What is the average amount claimed by the customers from the various segments?
4. What is the total Claim amount based on incident cause for all the claims that have been done at least 20 days prior to 1st October,2018 ?
5. How many adults from TX, DE and AK claimed insurance for driver related issues and causes?
6. Draw a Pie chart between the aggregated value of claim amount based on gender and segment. Represent the claim amount as a percentage on the pie chart.
7. Among males and females, which gender had claimed the most for any type of driver related issues? E.g. Bar chart
8. Which age had the maximum fraudulent policy claims? Visualize it on a bar chart
9. Visualize the monthly trend of the total amount that has been claimed by the customers. Ensure that on the “month” axis, the month is in a chronological order and not alphabetical order.
10. What is the average claim amount for gender and age categories and suitable represent the above using a faceted bar chart, one facet that represents fraudulent claims and the other for non-fraudulent claims.

Based on the conclusions from the exploratory analysis as well as suitable statistical tests, answer the below questions. Please include a detailed write-up on the parameters taken into consideration, the Hypothesis testing steps, conclusion from the p-values and the business implications of the statements.

11. Is there any similarity in the amount claimed by males and females?
12. Is there any relationship between age category and segment?
13. The current year has shown a significant rise in claim amounts as compared to 2016-17 fiscal average which was \$10,000.

- a. Is there any difference between age groups and insurance claims?
- b. Is there any relationship between total number of policy claims and the claimed amount?