Team 1-10: Yuchen Zhao, Tanner Reichard, Yuwei Yao, Jiaxin Lu,

Executive Summary

**Introduction:**

This dataset contains over 8000 records of customers at an auto-insurance company. Each row represents an existing customer with their relevant information. Our goal is to find insights and improvement from insurance dataset by using tableau and R studio.

**Major findings:**

1. Roughly ¼ of all customers of the insurance company have filed a claim for an accident or collision.
2. Customers that have been in an accident in the past year, on average, have a lower yearly income (~$50K) than drivers that have not filed for a claim(~$65K).
3. By proportion, those who are under the age of 21 file more claims. In fact, more than half of our youngest insured customers filed claims in this past year. Further, they have a higher average The age group of under 21 gets more tickets(Motor Vehicle Record Points) than other age groups. The median of the age group under 21 is 4. This means of the crashed incidents are more likely to happen among young people who get more tickets during other time.
4. Given that a customer has previously filed a claim before(40%), they are more likely than their accident-free counterparts(18%) to file a new claim.
5. The age range with the largest number of claims is 33-42, yet customers 43-52 represent the largest proportion of our market base.
6. The amount of claims by proportion of age group start to decline at age 33, with those 43-52 filing the fewest amount of claims by proportion (though not by volume). For example, while customers aged 43-52 filed 703 claims, their entire age range represents 3,404 of our customers. This means that about 21% of them filed claims.
7. There are a large number of null values reported in the variables “Car\_age” , “Years on Job”, and “Income”. This may be due to confidentiality issues in which customers did not share this information.
8. The average time traveled by customers who claimed an insurance case is slightly longer than people with no claims. Simply put: more time on the road translates to a higher chance for an accident.
9. Minivan and SUV have more number of claims among all types of cars.
10. The most 2 popular types of cars are Minivans and SUVs. This likely signals that many of our customers have families or other dependents that they need to transport.

**Interesting Insights:**

1. Out of customers that have filed for a claim, 28.6% of them had red cars. Out of customers that had not filed a claim, 29.3% had red cars. This signals no correlation between red cars and claims.

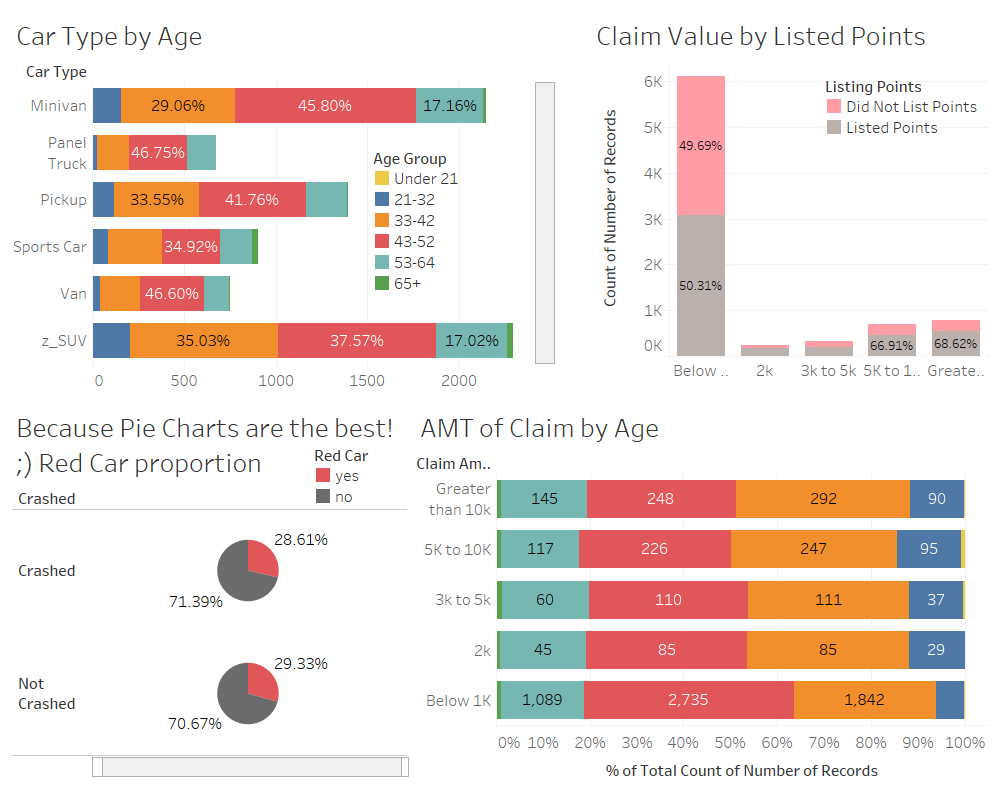
**Conclusion:**

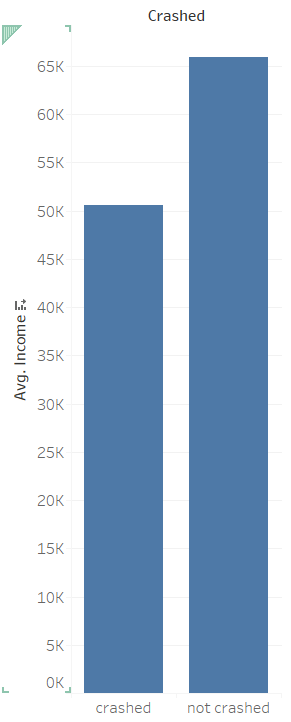
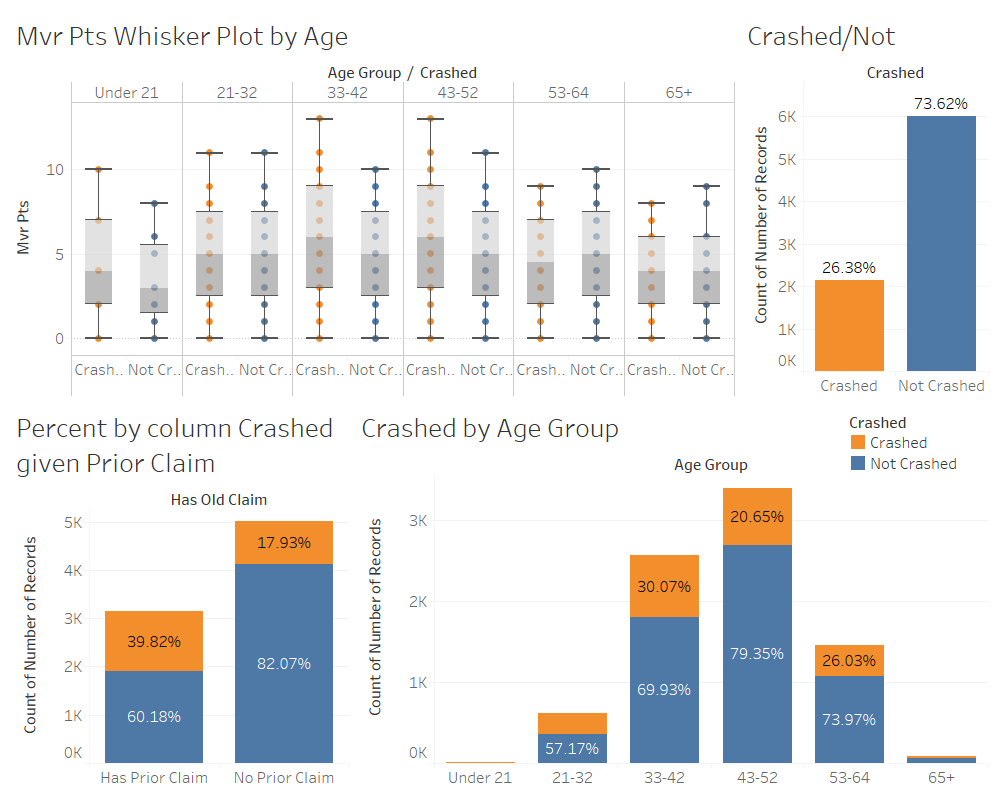
Compared to customers that have not filed a claim, those that have filed for an accident tend to have lower income, are typically younger, are more likely to drive SUVs than other vehicle types, have more Motor Vehicle Record Points, and have previously filed claims with us. For customers with these attributes, we would suggest charging a reasonable premium. Further, providing incentives to attract more customers in the 43-52 age range, such as sign-up bonuses, would ultimately provide a more stable base of income with fewer payouts.

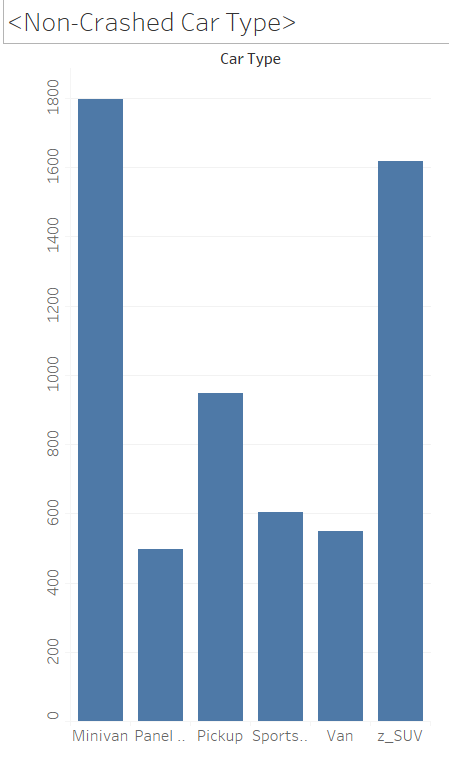
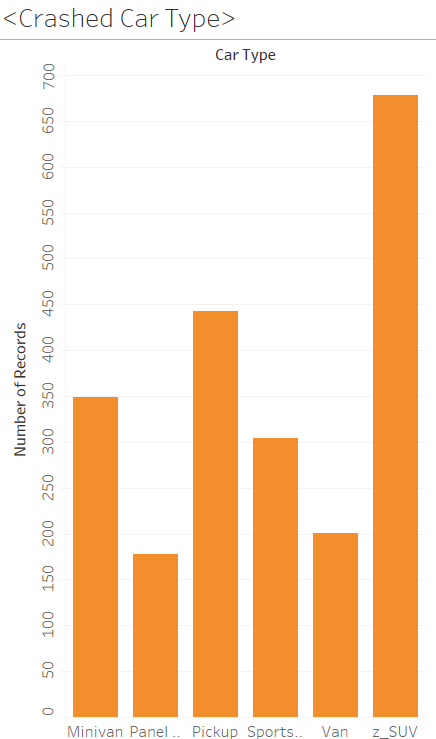
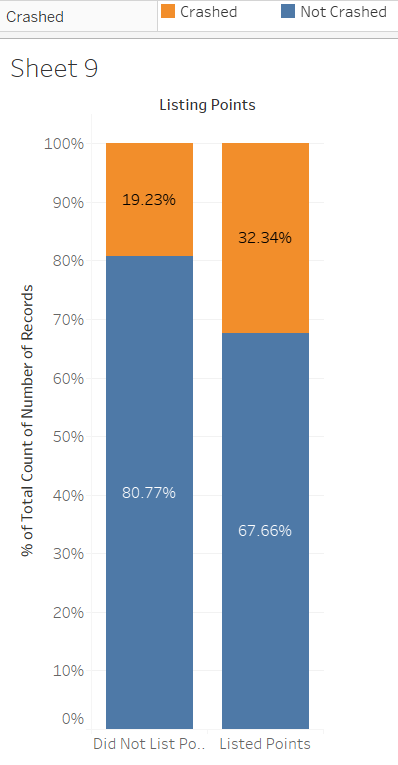
Those who did not list points are no more likely to have a large claim than those who listed points. In fact, those who did not list point typically had a lower proportion of large claims than those who did list their points. This implies that those who did not list points likely had 0 points on their record, and were not trying to hide it.

Further, if a driver did not list points, they were less likely to file for a crash than those who did. This further signals good driving habits (especially in relation to our earlier claim that customers that have been in a prior accident are more likely to get in another one), and leads us more towards the hypothesis that the average customer who does not write down their motor vehicle points likely has 0 points.

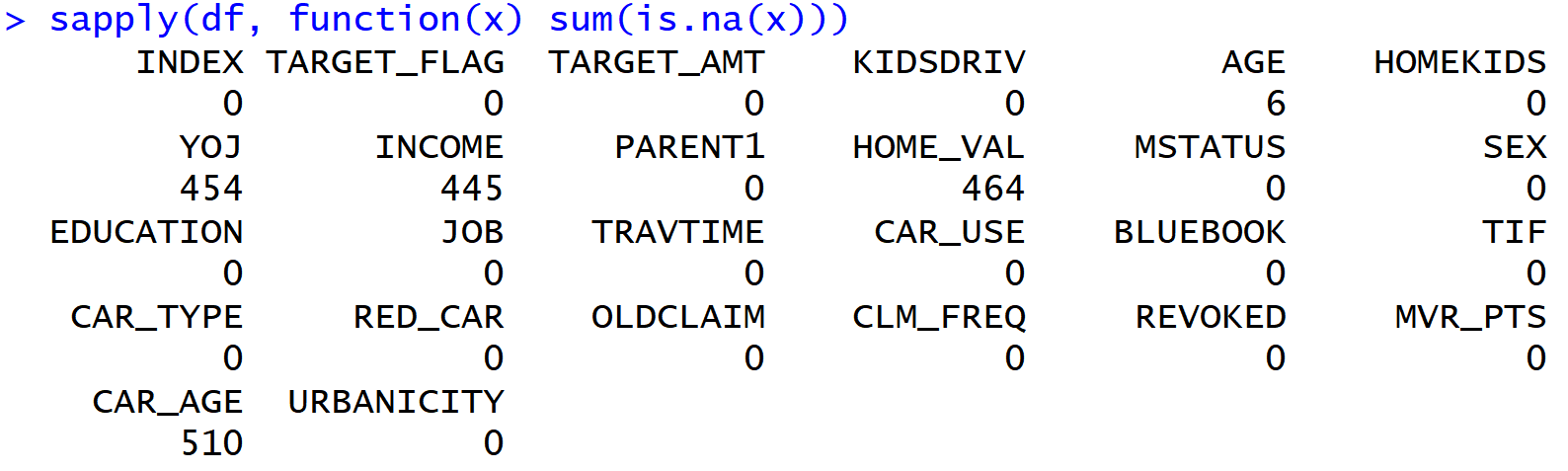
**Supporting Files**



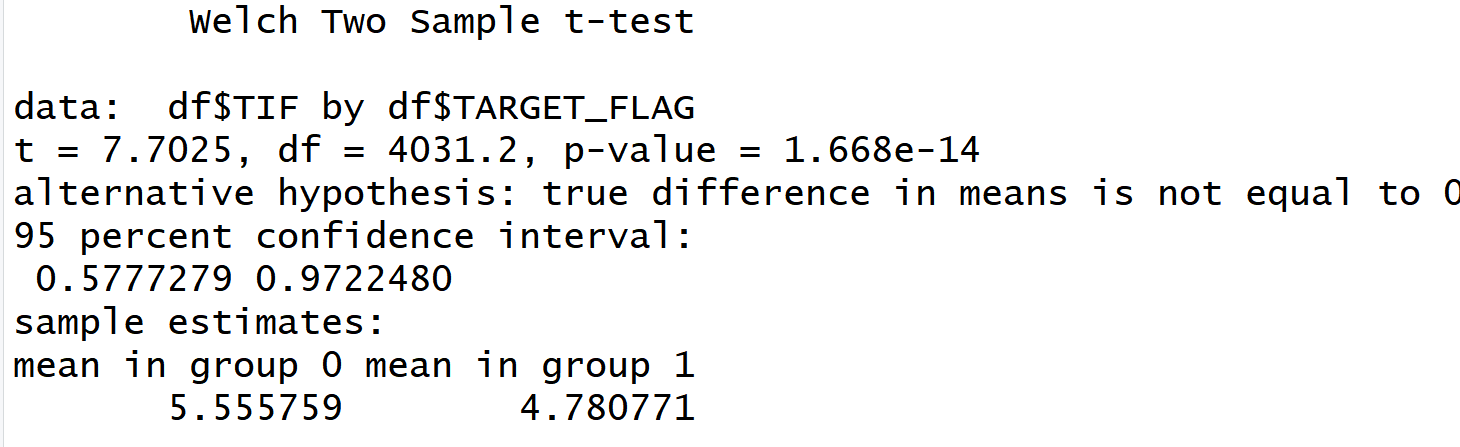


****

* + Missing values



* + Time in Force



* + Distance to work

