Summary :

* What tests did you use? What were the results?

We did T – Test and Qi square tests , not Anova Test.

Chi-Square Test:

We use Chi- square test for the categorical variables . we have six categorical variables here which are : Gender, driving License, Region Code, Previously Insured, Vehicle Age, Vehicle damage and Response to calculate the p- value for them. The type of chi square test which we use is showing the independency of variables. If the P-value is less than 0.05 which is our alpha, we can understand from it that our variables are not independent, we fail to reject the null hypothesis and it is statistically significant for our model.

For example: the first variable is Gender which are female and male. The p-value which came out from the test is less than 0.05 which shows that is not independent. We fail to reject the null hypothesis . and it is statistically significant for our model.

The p-value from the Pearsson’s test between Region Code vs response is almost zero and we realize that the region code is not independent.(It is dependent)

The p-value from the Pearsson’s test between Driving License vs response is 6e-09 and we realize that the Driving License is not independent.(It is dependent)

The p-value from the Pearsson’s test between Previously Insured vs response is almost zero and we realize that the Driving License is not independent.(It is dependent)

The p-value from the Pearsson’s test between vehicle age vs response is almost zero and we realize that the vehicle age is not independent.(It is dependent)

The p-value from the Pearsson’s test between vehicle damage vs response is almost zero and we realize that the vehicle damage is not independent.(It is dependent)

We don’t use Anova test here because our target variable which is response is categorical (0 and 1)

* Why did your team choose this topic?

At first we chose insurance topic to start our project with because we were familiar with it , and after that we saw that the cross sell is the popular topic around and related with insurance .

* What prior research and analysis have been done on this topic?

We were trying to find some effective ways to understand the cross sells.

The link is :

https://www.yieldify.com/blog/cross-selling/

<https://www.podium.com/article/cross-selling/>

<https://www.business.com/articles/how-to-boost-sales-with-cross-selling-and-cross-promotion/>

* Your SMART questions, and how did they come up?

After analyzing the data set we realized them.

* After the EDA, did your questions change? If so, how?

The main smart question did not changed but the sub questions changed (we rearranged them) after analyzing data .

* Draw conclusions. How do these answer the SMART questions?

From the smart questions we were trying to realized how many variables impact the response.

All the variables had impact except the vintage.