Business Strategy

One piece of additional information that Salamander could share is the date and time of the car purchased by the users belonging to the control group and recorded date and time for when the online ad was viewed by the experimental group.

The difference between these times would give us a time window and depending on how big or small this window is, we can analyze the impact of the graph. If the window is too big eg; 3 years, it wouldnt be logical for the user to make the purchase base on a 3 yr old ad.

This being said, It would have been easier to carry out analysis if the rate at which ad views got converted to purchases would have been provided, this conversion would become a point of reference for Salamander and it would be easy to identify the factors that affected this conversion.

Similarly keeping track of the number of purchases made before the ad campaign and number of purchases made after the ad campaign can also help to check if there is an increase in purchases.
Since our lists have repeated users it would also be beneficial to track how many times did the user click on the ad. This could help predict if the user would make the purchase or not.

The most prominent KPI that AppleCart should consider is the correlation between edges and the one_degree_target users who clicked and bought the car. In our observations, they are neighbors and the people who are colocated.

Also fetching additional data about the location of the users could also be beneficial as every location is observed to be culturally bounded for example people living in Cleveland would have a different set of requirements when it comes to cars and insurances as compared to people residing in Cincinnati.

Reason for purchasing an item depends on various factors, these factors need to be same between different populations. Different factors lead to different patterns of behavior which may not give us conclusive results.

For example, finding out what cars people buy based on their incomes would show different trends as opposed to comparing cars bought by people who are co-workers.
Or another example would be if one population has people residing in Malibu and one showing people residing in Cupertino, both these data sets would have different results.

One way to create control groups could be to have a process or a pipeline where datasets that need to be compared are provided with some degree of similar demographics before handing

out to the client. This process would help convincing the client that our data is reliable and for us, it would give way better conclusive results.

To conclude, I would like to thank you for giving me the opportunity to work on this project and introducing me to company based platforms such as databricks.

-Ravina Gaikwad