Team Name: statsUP

Introduction:

The motivation of our study is based on the interesting observation that many customers opted to abandon their initial choice of car class in their final decision makings. Therefore, we set out to investigate the factors that affect this change. Understanding of these factors have multiple benefits both to the customer and business side.

Customers: The potential benefits of our study to the customers are mainly reducing the time and energy involved in their car purchasing process, in other words, making buying cars for them easier.

Business owner: Business owners can use the results to better direct customers to make their decision. For example, business can tailor their search results, recommendation and advertisement based on the potential needs of the customers. Business can better strategize their promotion plan to increase their return rate, for example, matching specific promotion based on the time of the year.

Methods:

Data: We classified the car make as luxury and standard and created a variable called "shift" to denote the change of car make in final purchase from lead submission. "-1" denotes shift from luxury to standard, "0" denotes no change and "1" denotes upgrade. We also created a predictor variable "income", using IRS data to assign income values for each individual by postal code.

Statistical Analysis:

- 1) In initial analysis, we grouped the shift variable into 0(no change) and 1(change). We used Logistic regression to probe factors that have an impact on shift.
- 2) Based on their brand selection (luxury/standard) in the leads, we fit the logistic regression for the two subgroups.
- 3) In final step, we used the shift variable with three levels (-1, 0, 1). We used proportional odds model to discover the factors affecting the propensity to upgrade in final decision making.

Results: In initial logistic regression analysis, we identified eight statistically significant predictors that impact whether customer make brand class change: "Price_bought", "make_count" "quarter" and "tot_dwell_time" have a positive impact, whereas "page_views", "dealer_distance" , "income" and "mileage" have a negative one. Subgroup group analysis shows some discrepancies between the luxury and standard lead group. Proportional odds model finds "mileage" and "tot_dwell_time", two previously significant covariates, non-significant.

Discussion: In the future study, we could measure the shift more accurately by factoring in the models, country, body type and other information of the car into the design of our response variable. In conclusion, it is advisable to perform more studies to detect factors affecting the change in customers' decision. Business owners can make policies to improve customers' car shopping experience accordingly.