Part 1 – General Pricing Model

XYZ corp provides pricing products to its clients to help them understand the value of their inventory and to set reasonable sale prices for their vehicles. The value of a vehicle is derived from numerous factors including vehicle specific properties like year, make, and model to market conditions like the time and region where a vehicle is sold. The ask for the students is to **develop a pricing model** that takes into account the **vehicle attributes**, **seasonality**, and **regional differences**. The success of this pricing model will be determined by comparing the results to real world sales and XYZ corp’s existing proprietary algorithm.

### Part 2 – Reduce Auction Reruns

XYZ corp incurs operational costs every time a vehicle does not sell in an auction. It takes labor to register/stage a vehicle in a new sale. The majority of revenue for an auction vehicle comes from the buy/sell fee that is only realized when a sale is finalized. Because of this XYZ corp has an incentive for clients to sell vehicles quickly and minimize the number of reruns. A key reason a vehicle might not sell is it is priced inappropriately. If a client has an unrealistic expectation about the value of their vehicle, it may need to run in several auctions before it sells. We would like to explore the relationship between the final sale price of a vehicle and the number of times it runs in an auction. If possible, we would like to provide sellers with an estimated probability that their vehicle will sell at different pricing levels.

We are also interested in this solution because of the revenue that we forego in a no-sale incident could be significant. If a seller runs his/her car through the sale, and then decides not to sell it, XYZ corp charges the seller $100. If the seller decides to sell the car, XYZ corp charges the seller $400. Therefore since XYZ corp receives more when the car sells, we are interested in helping the seller sell his/her car for a fair price. The obvious approach to this problem would be to create a logistic regression model to calculate the probability the car will sell on the first run. However, we don’t have the last bid price for the cars that did not sell in the auction. Therefore, there are a few different things that you might want to consider for this problem:

1. From Part 1, how much variation is there in the sales prices of cars of a particular make and model?
2. Are there any anomalies between the prices of cars that sell the first time and the cars that don’t?
3. Create a way to calculate the confidence and prediction intervals for cars that do not sell on the first run. Are there any trends? Are the mean sales prices significantly different for the cars that ran through the sale more than once versus the cars that ran just once?
4. Based on the findings in parts 1-3 above, can you generate possible last bid price before the car was “No-saled”?
5. Create a logit model to calculate the probability that a car sells using your generated bid price in part 4.

Other suggestions from the students to reduce reruns based on the data are welcomed.