**CarMax Report**

Group Member: Yixuan Long(Leader), Zitao Zeng, Zhaohan Xing

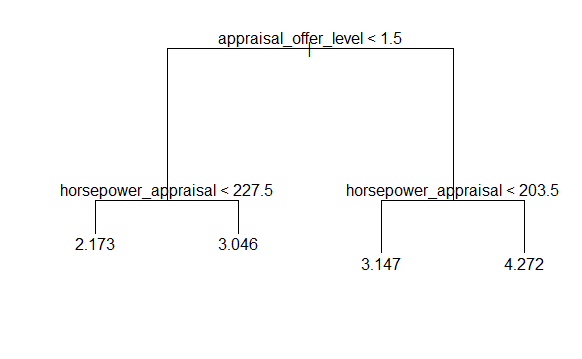
**Problem definition**

Using the database provided by CarMax, we can gain insight into customer preferences and needs, and identify trends in the marketplace. This information can be used to improve business operations, such as identifying information about the vehicles that customers care about most and providing a more personalized shopping experience for customers. In addition, CarMax can use this information to better market their inventory and services to the right customer segments, ultimately increasing sales and customer satisfaction.

**Methodology**

Firstly, we do data mining to remove all the observations that have missing data. Then, we notice that there are some string variables that are distributed in some intervals, such as price, appraisal offer, mileage, etc. Consequently, we convert the variables to some digit levels, so that we can use these variables to make our prediction. In addition, we choose some variables that can be used to predict the price level. After that, we firstly use linear regression to find out the relationship between variables and the price level of the car, and we find out that the linear regression model is significant. To improve our model, we choose the tree-based model to classify how we should triage the customers of various different models. Finally, use random forest to improve further to provide customers with the corresponding new car price level information. Comparing these models by determining the MSE of the model, we then summarize the most important information for customers to purchase a new car.

**Result**



**Conclusion**

We have deduced the range where the user's possible new car is located based on the data of all the customer's previous cars, and developed a new training model to successfully reduce the accuracy of the predicted range from level 6 to about level 2, which will bring more effective help and detailed service to the user.

In the future, the introduction of more accurate, mathematical models with more data and better fitting in used car transactions would revolutionize our recommendation system. We believe that today's experiment is just the beginning, and the continuous progress will eventually make the CarMax platform give users more detailed, thorough and rapid recommendations and services.