

STAT 2223-002

Final Project

Spring 2020

Due on 11:59 PM, May 3, 2020

1 The Project

This project concerns a problem of interest of buying a used car. The calling price of used cars will vary depending on a number of quantitative and qualitative variables: year of production, mileage, brand, color etc.

The purpose of this project is to investigate the predictors affecting the calling price and examine the relationship between the mean calling price $E(y)$ (the price asked for by the owner) of a specific kind of car and the independent variables like:

1. The number of years since production; e.g. If a car is produced on 2010, then $number\ of\ years = 2020 - 2010 = 10$.
2. The Current Mileage of the car(miles).
3. Color
4. Title status
5. Other predictors, you have investigated, that are relevant to calling price of a used car.
6.

Collect your sample data from: <http://charlotte.craigslist.org/>
or <https://www.kaggle.com/orgesleka/used-cars-database>
with sample size $n \geq 40$

2 Format of Your Work

Your work should be clear and easy to understand, follow the following format:

1. **Statement of The Problem:** You need to state your problem here. That is, tell us what your study is about and your purpose of the study.
2. **The Data:** You need to specify how you collect the data and summarize your sample data. The following table must be included.

Obs	X_1	X_2	X_3	...	Y
1					
2					
3					
...
n					

Define the variables.

Y:

X_1 :

X_2 :

X_3 :

...

3. **Variable selection:** Apply the techniques you have learned in chapter 6 and select (3 to 4) variables that are useful to predict calling price of used car, y.
4. **The Model selection:** Propose at least four different models you want to apply and select the best model using partial F test or any other statistical techniques you have learned.
5. **Assumption check:** In this part, you will check the model assumption for the model you have selected in part 4 (apply the concepts learned in chapter 8).
6. **Model Remedy:** Depending on your findings in step 5, do some transformation for y or x to make the model assumptions be satisfied (if necessary) and write down the new model. Use the concepts learned in Chapter 8.
7. **Model adequacy:** Test the adequacy of the final model in part 5 (overall F test).
8. **Conclusion:** Provide adjusted R squared, model standard error and explain is this model good to use.
9. **check:** find new data for a car that are not used in this model building process and predict its price using your final model. Is the predicted price close to the real price?
10. **Technical Appendices R code:** provide the R - codes you have used in this project.

2.1 Note

- If you finish early, you can turn in on earlier date, but late submission will not be accepted.
- Do not forget to write your name.
- Submit a pdf file.