Name: Tushar Muley
Assignment: DSC 680 – Audience Questions and Answers
Date: March 5, 2022

Questions:

1. Why was the data limited to two months?

A: The data actually only has one month worth of data. Originally it appeared as if we had two months of data, based on the volume of posts and posting date. I was not able to use other version of the data due to missing columns.

2. Is there a column to determine dealership listing compared to private party listings?

A: While Craigslist offer this feature on their website this data did not contain that indicator. I was going to attempt to pull data out of the description section. The deadline for this project would not allow time needed to pull an indicator for dealer and private party sales from the description section as the text was freeform.

3. Where able to determine what the final selling prices of the vehicle was?

A: No. This cannot be done as Craigslist is used more to get buyers and sellers together in person compared to being an online market place.

4. Can this data be used to determine where used car prices are headed?

A: Yes. You can use the model to predict future prices of used cars listed on Craigslist. Currently the model is only 72% accurate. I believe that can be increased.

5. Can sell date be determined?

A: Since individuals doing the posting don't always comeback and remove the post a sell date cannot be determined. It was also not available in the dataset.

6. Is the price impacted by location or region?

A: Yes. Location does have an impact on price. The more demand in a location did lead to higher prices. But model, year, odometer reading also impact price. Some types of vehicles are more valued in certain states for their features.

7. Since Craigslist charges $5.00 to list a car how much revenue does Craigslist earn?

A: When I started the analysis I had 426,880 rows of data. That earned Craigslist $2.13 million per month. It is revenue for Craigslist, but the fee is meant to reduce the number of fraudulent ads instead of providing revenue for Craigslist. I am sure they are not complaining.

8. Can you determine what attributes cause a car to sell faster than another car of the same make and model?

A: Not with the current model setup. I plan to add additional feature to this model.

9. What is the limitation of this type of analysis?

A: Limitation is with the data. A lot of information goes into the description field and that is freeform text. Making it hard to pull out feature that add value to the model. Also, there is no way to verify if the vehicle sold faster or for the amount listed. Plus, you have individual listing their vehicles for a dollar to generate views. I would suspect that has to work against them in the viewers eye.

10. What additional insights can be gained from this data?

A: We could have used regional data to pick out where car prices are lower. So, when a car comes up for sale in that area it would help potential buyer get ok deals. You could also perform other analysis on the data by extracting the features of the vehicles from the description section. If sellers remove their listing we could approximate sale dates. Giving an idea of how long it takes to sell a vehicle. You could also use to this to determine if mileage and year of a car determine income of a given area.