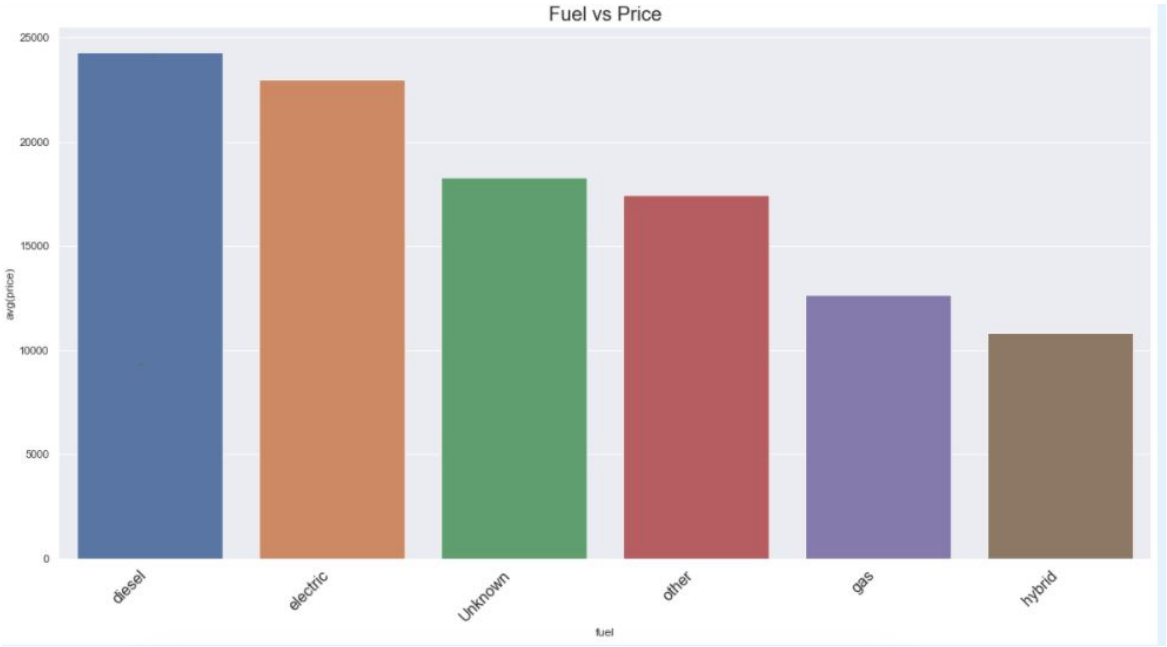
Purpose of the project:

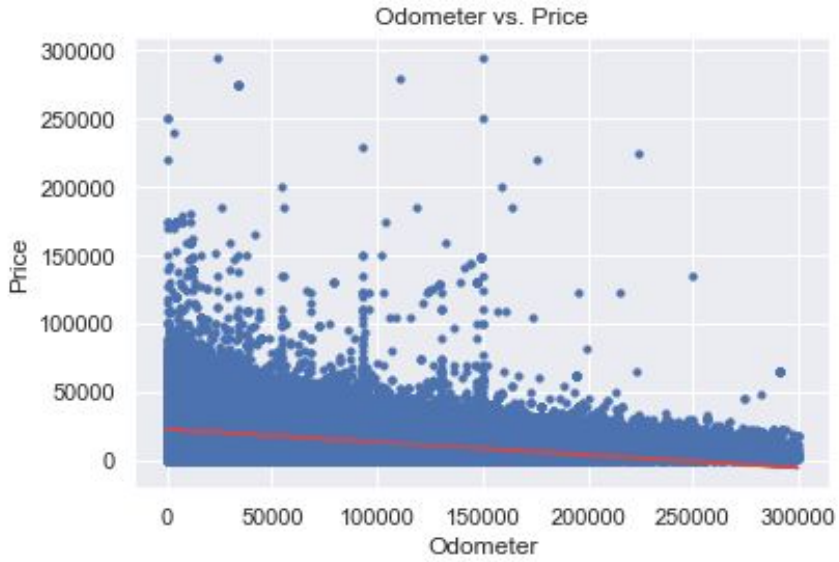
In this project I created a model to predict the price of used cars based on a used cars dataset from Kaggle. I created graphs and to help visualize which features affect the sales of cars and used these features to help build my model. After the data was uploaded, I cleaned and organized the data so that it is useful for other people to analyze and draw their own conclusions. Below I will share a few of my graphs and then explain my final model.

Graphs:

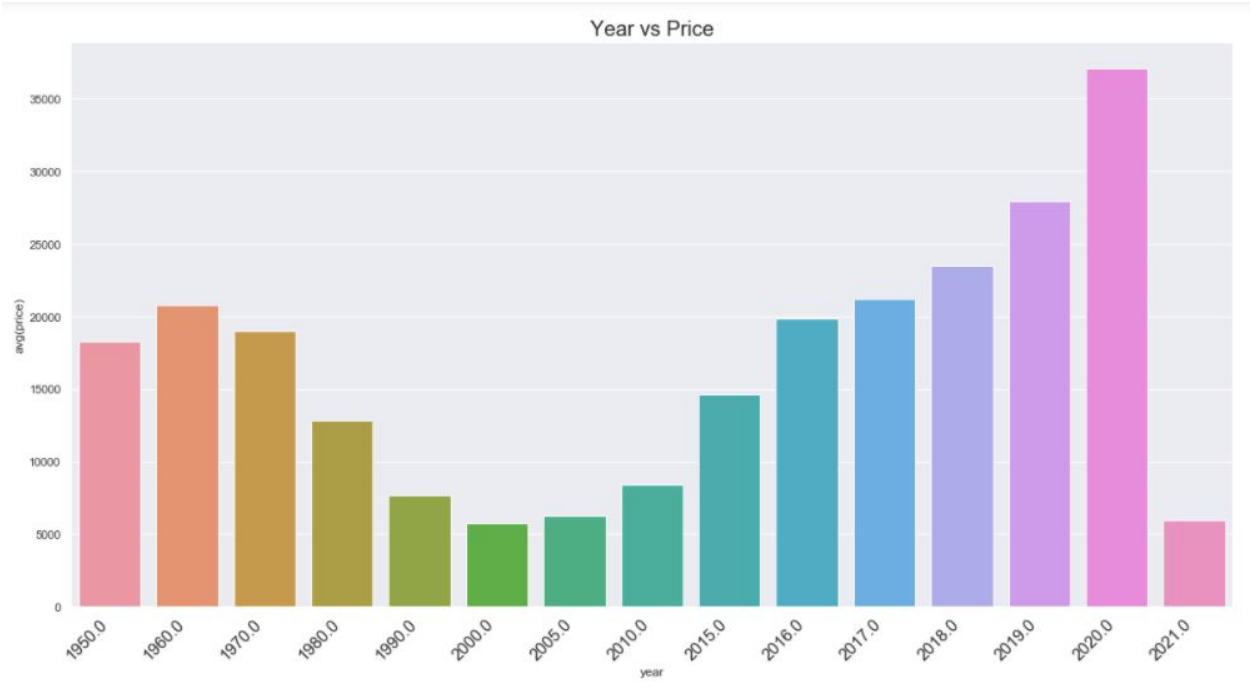
1. I graphed the fuel type against the average price for each fuel type to see which type of used car costs the most. I found that electric cars cost on average $10,000 more than gas cars.



2. I graphed the total miles used by the car (odometer) against the price of the car to see what the effect is. I found that the more miles a car accumulated costs more than a car with fewer miles



3. Lastly, I plotted the age of the car against the average price of the car for that model year. As you can see, there is an inverted bell curve describing the price of cars against the year. As a car approaches the early 2000's the car is relatively cheap, as opposed to a newer car closer to 2020 or a car that is antique from the mid 1900's.



Model:

In my model I included all of these features and many more to predict the price of a car. For example, I included the condition of a car, the manufacturer, the state that the used car is being sold in and many more features to help my model.

Conclusion:

In conclusion, my model was about $7000 off from predicting the true price of a car