

Research Project Proposal

Abstract:

There are many reasons an auto dealer needs to know the market price of their inventory. This often requires lots of experience and attention to the alternative channels available to customers. If an auto dealer was able to rapidly determine the distribution of the advertised price for comparable vehicles they could market vehicles more effectively.

Using python, it is possible to rapidly retrieve the prices of comparable vehicles from multiple different channels. This information can then be transformed into an “advertised price” distribution. Not only would this allow the mean price to be determined, but the distribution could also be used to decide on upper limits for the advertised price of cars in high demand and the lower limit of cars that weren’t selling. This could potentially minimize revenue lost to excessive discounting or pricing.

Methods:

A python web scraper would take user inputs corresponding to the vehicle to be advertised. This would establish the search parameters and consist of year, make, model, a range of mileage, and a range of dates. Python would then query a range of sites and record the distribution of pricing. From here, the matplotlib and atkinter would calculate the mean price, the variance, and also return the percentage of values that would be below a proposed sale price.

Data sources:

1. Craigslist-Private seller
2. Craigslist-Dealer seller
3. AutoTrader-Dealers and private
4. Individual Dealership webpages

Use Cases (tentative):

1. As a price checking tool for dealers, allowing them to estimate the market value of vehicles prior to purchasing
2. As a research tool to allow valuation of inventory

Additional:

With the given data it may be possible to estimate the expected price difference between two separate populations of the same make and model.

$$P_1 = \text{mean price of cars sold by dealers}$$
$$P_2 = \text{mean price of cars sold by independent owners}$$

Although this might be susceptible in general, it could give a dealer a better idea of how much premium he could expect in general. It would be interesting to learn if this amount was generally constant (ex. 1000) or a proportion of the overall price.