Scraped data of used cars listings. 100,000 listings, which have been separated into files corresponding to each car manufacturer. I collected the data to make a tool to predict how much my friend should sell his old car for compared to other stuff on the market, and then just extended the data set. Then made a more general car value regression model.

previous version

Picked two fairly common cars on the British market for analysis (Ford Focus and Mercedes C Class). The hope is to find info such as: when is the ideal time to sell certain cars (i.e. at what age and mileage are there significant drops in resale value). Also can make comparisons between the two, and make a classifier for a ford or Mercedes car. Can easily add more makes and models, so comment for any request e.g. if you want a big data set of all Mercedes makes and models.

Content

The cleaned data set contains information of price, transmission, mileage, fuel type, road tax, miles per gallon (mpg), and engine size. I've removed duplicate listings and cleaned the columns, but have included a notebook showing the process and the original data for anyone who wants to check/improve my work.

task

make a regression model that could inform whether the car he wanted to buy was good value in relation to the market in general. I've extended this data set to include lots of other makes, so it would be cool to extend the car value prediction model.

Expected Submission

Some sort of value regression model.

Evaluation

Low error, and easy to understand output. Should be able to put in custom car details and get a value estimate.