Predicting Car Listing Prices in the UK

Capstone | GA DSI 523 Giovanni Ceretti



Just like in the US, Car Prices in the UK have been crazy over the last 2 years



Value of UK's most popular used cars rose dramatically over past year

CarGurus says its most popular 20 models are now worth 26.3 percent more on average.

Cars are often family's 2nd largest assets after their homes

- Car related spending represents a meaningful part of people's income, between 10%-15% of post-tax income
- The immense financial commitment required to buy a car means families look at the purchase of a vehicle as an investment
- With car price volatility and predatory industry practices, it can be very difficult for buyers and sellers to price vehicles correctly



The Project

To create an application to provide car buyers and sellers insight into the British auto reselling market and predict vehicle listing prices

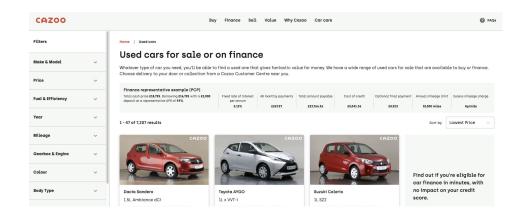
The Data

Cazoo is one of the UK's most popular and largest online car retailers

- 7000+ car listing on the platform
- The most data-rich platform in the UK
- The highest customer satisfaction in score the industry
- Comparable to Carvana in the US

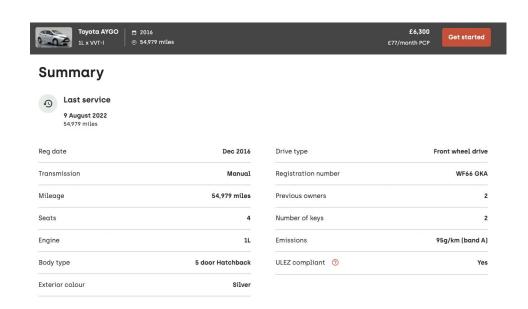






Key Statistics being Scraped from Cazoo

- Car Brand
- Car Model
- Listing Price
- Mileage
- Registration Date
- Trim Specs
- Engine Specs
- Car Body Type



The Model

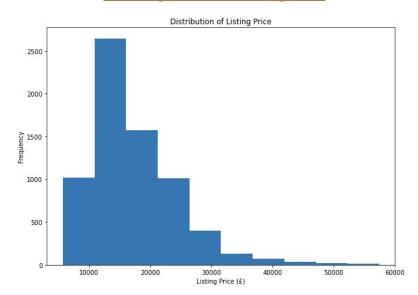
Linear Regression

Listing Price target variable shows a strong right skew

Listing Price Statistics

| mean | 17345 |
|------|-------|
| std | 6968 |
| min | 5700 |
| 25% | 12350 |
| 50% | 15550 |
| 75% | 21050 |
| max | 57500 |
| | |

Listing Price Histogram



The Multi Linear Regression Model showed very strong predictive capabilities

Predictive Model Metrics

Model MSE: £1,853,225

Model RMSE: £1,361

Model MAE: £867

R2 Statistics

Test R2 Statistic: 0.960

Training R2 Statistic 0.986

Null Model Metrics

Null Model Mean Squared Error: £48,556,014

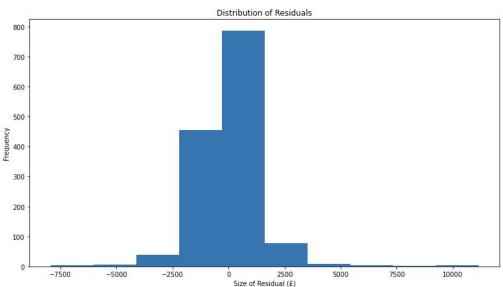
Null Model Root Mean Squared Error: £6,968

Null Model Mean Absolute Error: £5,368

The null model assigns all predictions to be the mean listing price of the sample of cars

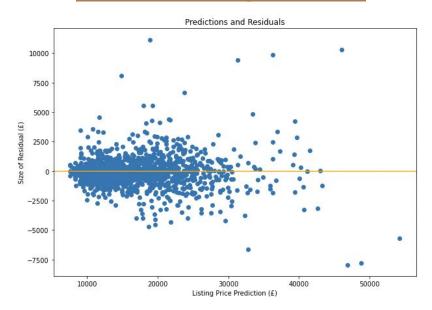
The Model tends to predict slightly higher car listing prices

Histogram of Residuals

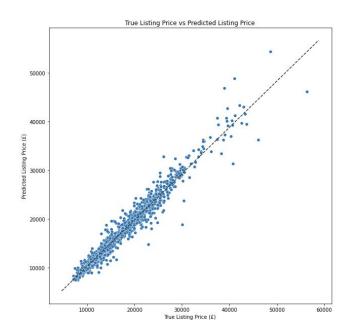


The Model excels at predicting listing prices of < £30,000

Plot of Residuals vs Listing Price Predictions



Predicted vs True Listing Price Plot



The Application

Conclusion & Next Steps

The Project is a strong predictor of car listing prices in the UK

The project is able to predict car prices with a Mean Absolute Error of 1/8th of a
 Standard Deviation

Next Steps/Future Improvements

- Collect more data on cars with values of £30,000+, improving predictions
- Collect more data to increase the number of car models in the app
- Identify a source for actual transactional data rather than only listing price data

Thank You! Questions?