

Hands on Projects

There are several data sets to choose from

- European classified car ads
- Home Mortgage Disclosure Act Data
- Flight Delays and Cancellations

European classified car ads

There are roughly 3.5 Million rows and the following columns:

- maker - normalized all lowercase
- model - normalized all lowercase
- mileage - in KM
- manufacture_year
- engine_displacement - in ccm
- engine_power - in kW
- body_type - almost never present, but I scraped only personal cars, no motorcycles or utility vehicles
- color_slug - also almost never present
- stk_year - year of the last emission control
- transmission - automatic or manual
- door_count
- seat_count
- fuel_type - gasoline, diesel, cng, lpg, electric
- date_created - when the ad was scraped
- date_last_seen - when the ad was last seen. Our policy was to remove all ads older than 60 days
- price_eur - list price converted to EUR

European classified car ads since 2015

Which factors determine the price of a car?

<https://www.kaggle.com/mirosval/personal-cars-classifieds>

Home Mortgage Disclosure Act Data

- Can you predict mortgage decisions based on the criteria provided here?

2015 Flight Delays and Cancellations