

Analysis of vehicles driving in Estonia

Goals

Get an overview of which cars are driving on Estonian roads, broken down by counties. Investigate whether there is a correlation between county demographic indicators and car characteristics.

County classification based on technical data of cars.

Datasets

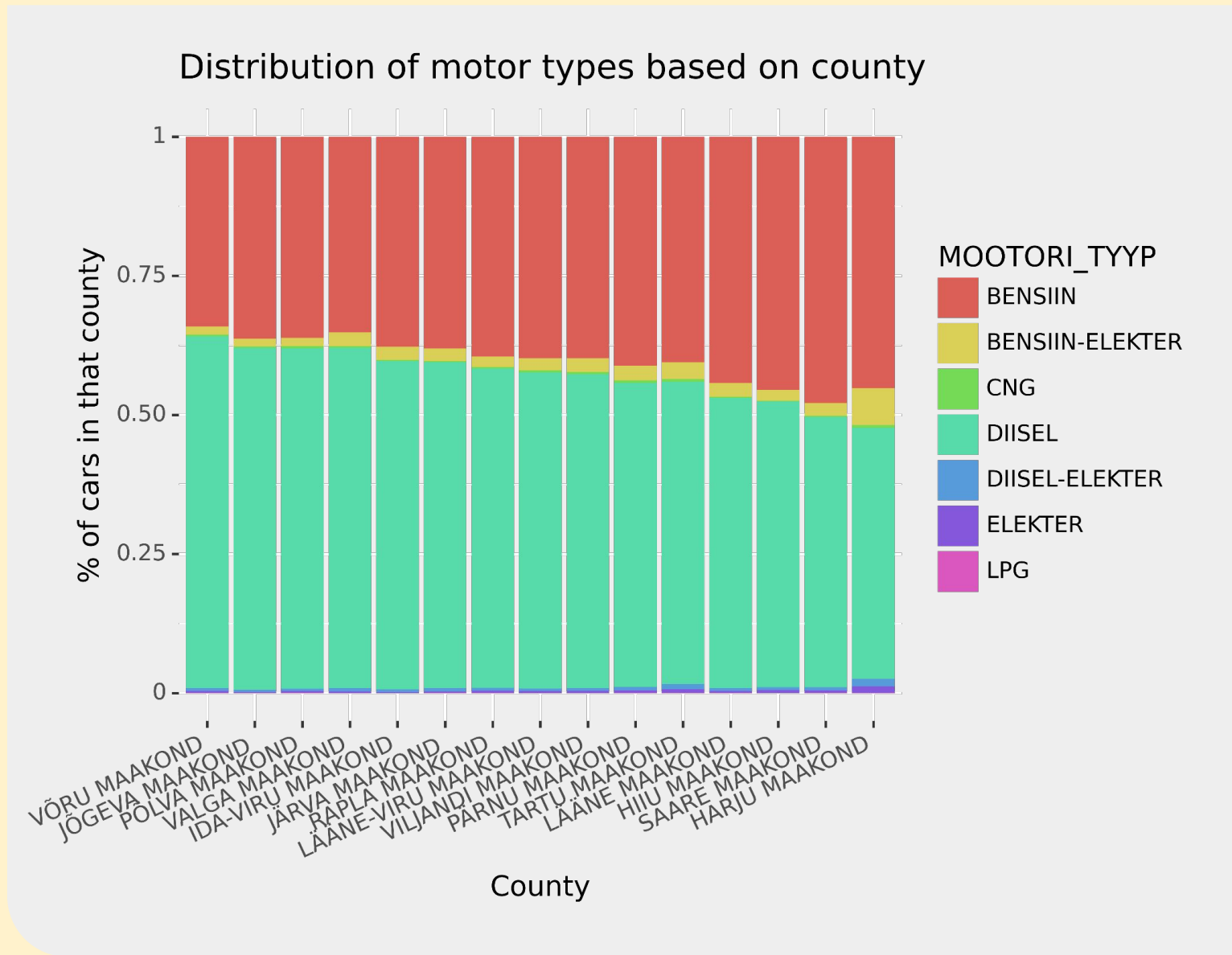
We used two datasets from Estonian open data portal, Status of vehicles in Estonia and Roadworthiness tests of land vehicles in Estonia. We ignored data that was not about B-category cars (classes M1, M1G, N1, N1G), because we wanted to make this about regular road vehicles. This left us data about roughly 734 000 cars.

Appearance

We assumed that there would be some difference in appearance based on county, but any differences seem to be negligable. Most attributes seemed to have roughly the same distribution across counties, like car make, type, and colour.

Project E5:
Vehicle Analysis

Team:
- Markus Karl Mägi
- Gregor Suurvarik



Using nearby cars to determine your location

We developed a machine learning model that tries to guess what county you are in based by the cars around you. This did not prove particularly successful due to the lack of significant regional differences in cars, but we still managed to achieve a positive identification rate of 13% when provided the information of 25 random cars that exist in a county.

Whilst this is not particularly impressive or useful, it is still significantly better than guessing randomly, which would net a positive identification rate of 9%. We also had a model that reached 44% accuracy by guessing ‘Harjumaa’ every single time before we balanced the dataset, but we concluded that this was not a good example of a machine learning model.

Car technical details

Our analysis of cars' technical details revealed no significant variations in distribution by county. However, it was interesting to observe a near-equal split between cars with manual and automatic transmissions. We also observed a higher prevalence of hybrid/electric cars in areas with larger cities.

Surprisingly, we found that most engine types offer both manual and other transmission options, though newer engine types tend to have fewer manual transmission cars.

The most common configurations we identified were diesel, gasoline, and gasoline-electric

Cars with demographic data

Our research indicated that regions with larger populations also have a higher concentration of cars per square kilometer, yet a lower ratio of cars per individual.

Notably, in Hiiu county, there is an average of 0.87 cars per person, which is quite significant given the county's relatively small size.

Roadworthiness tests

This year, up to December 1st, 2023, there were 475,000 inspections conducted. The most frequent critical issues were associated with the braking system, with around 900 cases. The most common minor fault was the absence of a fire extinguisher, noted in approximately 44,000 instances. This particular issue, while not an immediate cause for test failure, required prompt correction.

The most prevalent faults overall were related to the braking system, followed closely by issues concerning emissions. Specific figures include 22,000 instances of insufficient braking performance, 7,960 cases where exhaust gases exceeded prescribed limit values, and 6,713 instances of other environmental defects, such as hazardous liquid leaks, excluding water.

