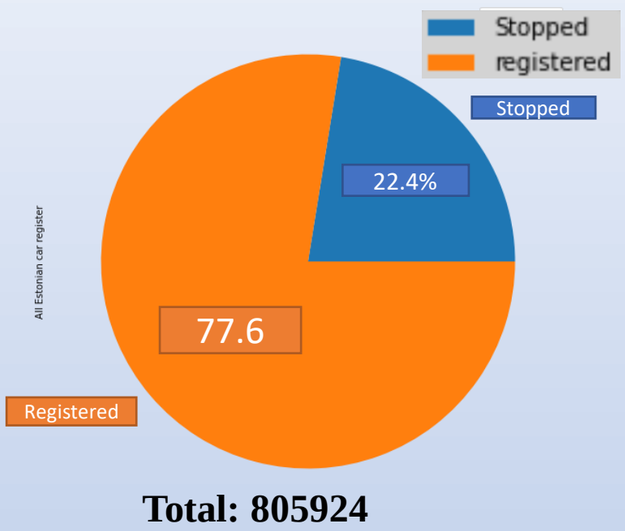
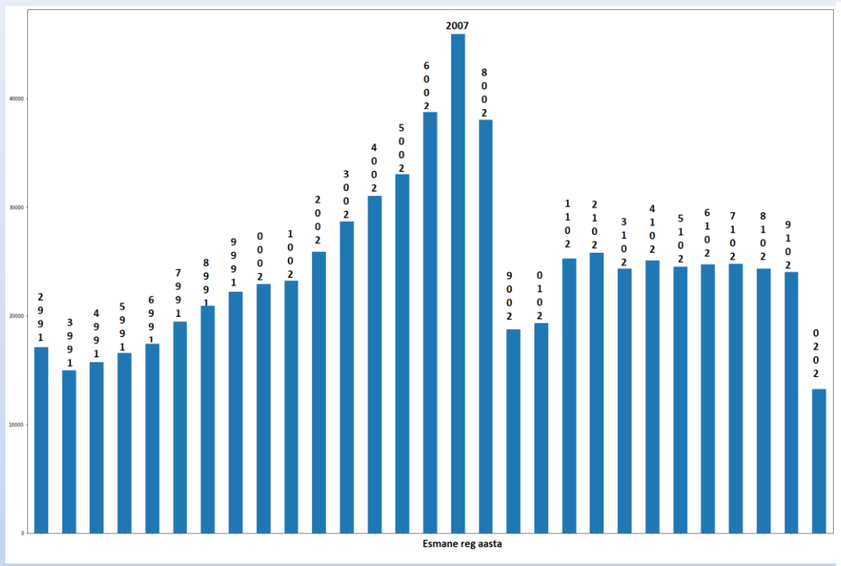
Estonian car statistics exploration

Marten Kuusmann, Taavi Karelson, Silver Kolde

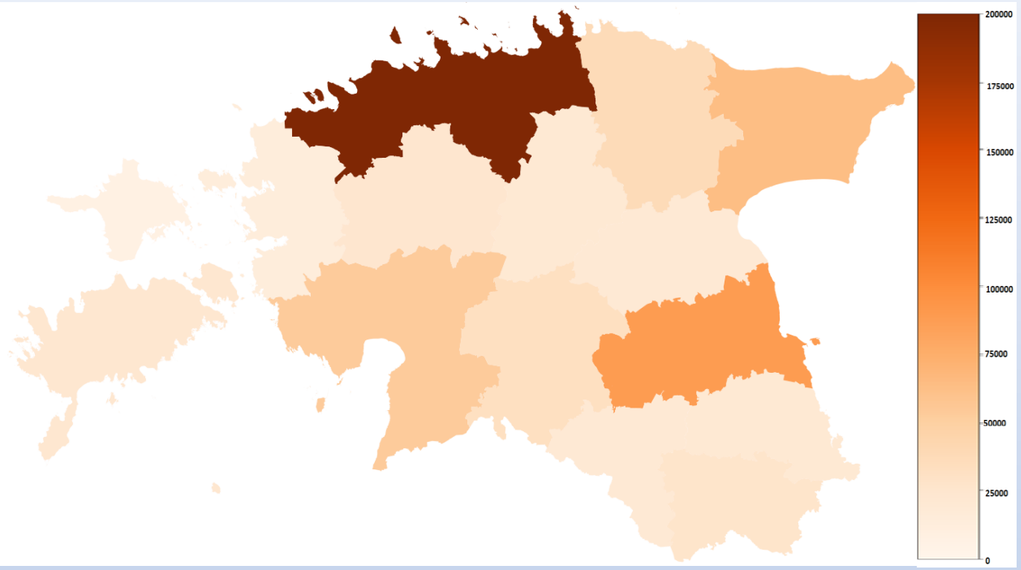
Overview registered cars in Estonia



Total: 805924

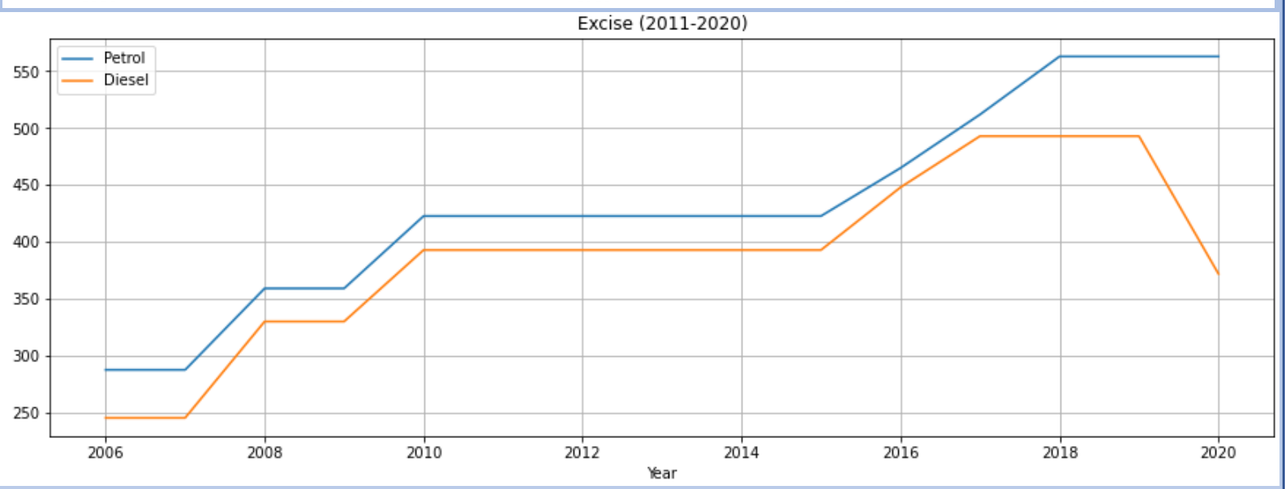


On this diagram we see data of registered cars in re independent Estonia by yearly. In range 2007 to 2011 we can see drastic decline which is caused by economic downturn.
(Estonia regained independence in 1991 August)



Distribution of car registrations by counties

Excise



2011-2016 (Diesel excise is 7% less than petrol excise)

From the "Engine types ratio (2011-2020)" chart and excise chart we can see that, in the period of 2011-2016 (when petrol excise was approx. 4-7% higher) both fuel types did not change more than +-5% against average. What is more, the two fuel types had been mirroring each other against the 0.5 point middle line. So if petrol lost market percentage the diesel gained approximately the same amount of market

2017-2020

From 2017-2020 (first 9 months), the petrol engine cars lost almost 0.15 of its numbers. At the same time the diesel part has remained the same. The drop in petrol cars can be explained by an increase of 0.07 point for petrol hybrid cars and CNG cars by 0.04 points.

2020 excise (Diesel excise is 34% less than petrol excise)

Then after the difference by 34% in excise, the pure petrol engine cars have dropped 0.08 points, at the same time diesel cars have risen by a small margin 0.02 points for the first time since 2013. Similar 0.02 growth can be seen for CNG and electric cars. Biggest upsurge has been for petron hybrid cars cars by 0.04 points.

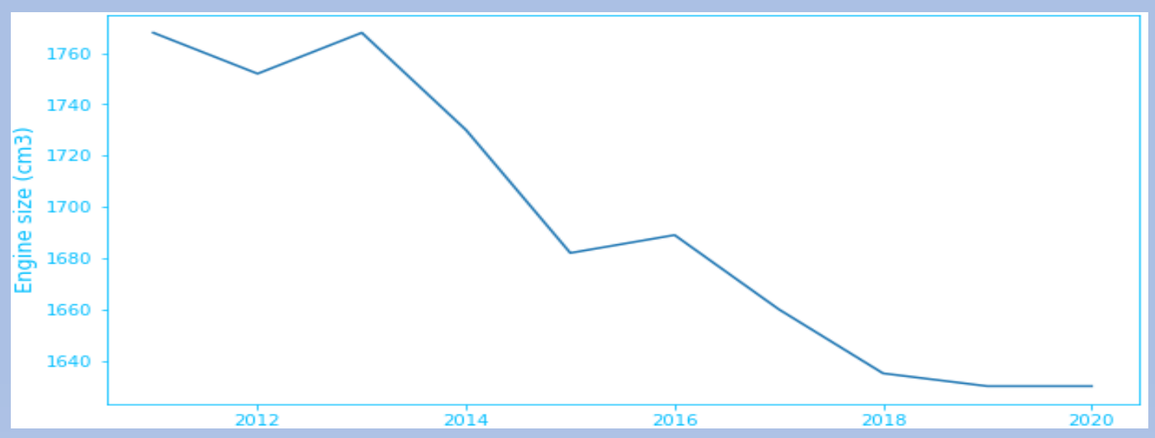
Electric cars

The electric vehicle numbers have been higher than 0 twice. As seen on the chart in 2012-2014 and 2020, and this can be easily identified as the two periods when the Estonian government has had aid for electric car buyers. In 2011-2014 there was a 50% (or up to 18000€) discount for electric car buyers and in 2020 the governmental discount has been 5000€. Note that both discounts were for a limited number of customers and were capped in a few hours.

Observations

Weight

Analyzing the data set of first registrations from 2011 to 2020 we observed that the average weight of cars has increased. We assumed it to be lower because as technology improves, materials usually become lighter, electronics more compact, and lighter cars should be less fuel-hungry, hence more environment friendly. On the contrary, it turns out that cars sold this year weigh about 60kg more than 10 years ago. The most probable reason is the economic recovery from the last crises - people tend to buy cars that have more accessories if they have more money in the pocket.



Engine

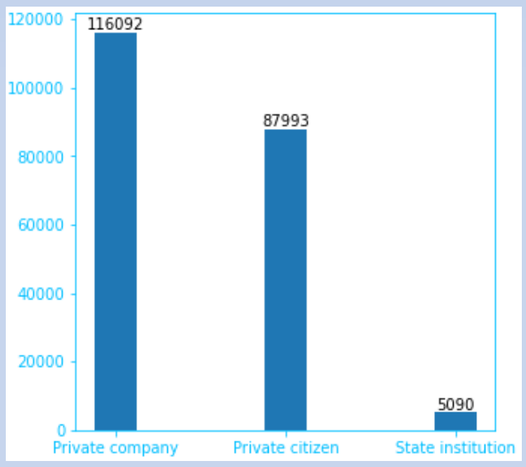
We observed that since 2014 engine sizes have shrunk about 0.15L, it's about a 7-9% drop. It would be interesting to have data about fuel economy, engine power, and CO2 levels also. In that case, we could see if smaller engines produce the same amount of power, take less fuel, pollute less, etc. Unfortunately, we only have that data since summer 2016.

	Actual	Predicted
76959	UNIVERSAAL	UNIVERSAAL
103352	UNIVERSAAL	UNIVERSAAL
80643	LUUKPÄRA	LUUKPÄRA
104467	LUUKPÄRA	LUUKPÄRA
136339	UNIVERSAAL	UNIVERSAAL
...	...	...
124040	UNIVERSAAL	UNIVERSAAL
119468	LUUKPÄRA	LUUKPÄRA
86781	MAHTUNIVERSAAL	MAHTUNIVERSAAL
17768	SEDAAN	SEDAAN
41662	MAHTUNIVERSAAL	MAHTUNIVERSAAL

Predicting the body type based on other features.

Our goal was to predict the right body type for a car based on other features. Using KNN and basically all the features from our dataset, we managed to predict the right body type with 93% accuracy. We set k equal to 1. We tried different values (1,3, 5, 7, 9, 11), but the result did not change.

We tried to remove features one by one to find the features that are most relevant. Although we could remove some columns without any significant loss in accuracy, we were still in the dark after a few hours gave that up.



Ownership

We can see that most of the cars registered in Estonia belong to companies (55.5%) and only 2.4% are owned by the state. It came as a surprise that most of the cars driving around belong to companies. Of course, it would be interesting for us (and for the Tax office also) to know how many of the corporate cars are actually in use for executive management's private purposes, but that information is hard to gather.

GitHub: https://github.com/SilverKolde/IDS2020-group-project