# Estonian car statistics exploration.

# Project goals and planning.

*GitHub repository:* [*https://github.com/SilverKolde/IDS2020-group-project*](https://github.com/SilverKolde/IDS2020-group-project)

Collaborators:

* Taavi Karelson
* Silver Kolde
* Marten Kuusmann

## Business understanding

**Identifying business goals**

**Background**

The idea of this project started when thinking about recent fuel excise lowering in Estonia. The government’s main goal was to revive the transportation sector by lowering diesel excise. We began to question if it has any effect on private citizens also when buying passenger cars. After some digging into the idea, we discovered a large dataset from Estonian Road Administration that holds all the passenger cars (M1 and M1G category). That spurred our curiosity even further to find out different trends over the years – what features are relevant to Estonians when buying new cars? Have the most important features changed over the years? Does it differ by location (cities, countryside areas)?

So let it be stated here that this project is grown out of curiosity rather than a business need or requirement.

Business goals

One of the main goals of this project is to find out if Estonian’s new car buying preferences change after fuel excise has been altered.

The other one is to discover trends – are there any features that are becoming increasingly important and does it differ by region.

Business success criteria

What we hope to see is that fuel excise has affected the buying habits of Estonians over the years, but we are very aware of the fact that Estonia is small and we can’t make fundamental conclusions without investigating global trends. So the project is considered successful when all members have learned something new solidified basic concepts of data science/data mining.

**Situation assessment**

Inventory of resources

We have a dataset of all cars registered in Estonia (about 800 000 entries). We found some data about fuel excise from tax administration, but it may not be enough. We are still gathering information and waiting for a reply from the tax administration.

Requirements, assumptions, and constraints

The deadline for making the introduction video about the project 12:00, Monday, Dec 14.

Risks and contingencies

It may happen that we can’t get enough useful data (soon enough) about fuel excise and we fail to deliver one of the main goals of the project.

Terminology

At this point, we don’t see anything that needs extra clarification.

*~~Costs and benefits~~ irrelevant*

**Data-mining goals**

Goals

We aim to discover the trends in buying a car (instead of people buying them randomly) and display them in graphs.

Data-mining success criteria

If we discover that fuel excise directly affects people’s buying decision, we consider it a major success. And if we only manage to find some features that have become increasingly important over the years, we consider it good enough.