

Building fires in Estonia

Karolin Kivilaan, Geitrud Pank, Kati Ilus

Informaatika bakalaureuse 2. kursuse, arvutiteaduse instituut

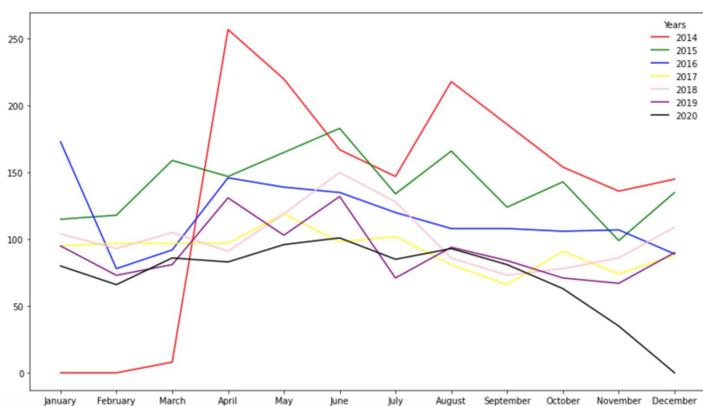
Github link: https://github.com/Katillus/IDC_Project_Fires_in_Estonia

Introduction

Our goal with choosing the topic of the project was to find one that would be relevant in real life. We browsed through Estonia's open government data portal and found a dataset about building fires. We chose the project, because we think it is an important topic and we couldn't find any reports about building fires, that analysed more than a year's worth of data at a time. Our dataset hasn't been analysed before and we think that the results could be useful in real life.

We had two main and one additional goal. First, we wanted to analyse how the amount of building fires in Estonia varies by county and month, which would tell us which counties have had the most fires and why. We also wanted to see if there are some kind of patterns, for example when we also took months into account, would we see if seasons have an impact on the number of fires. Our second goal was to cluster fires (originally by cause, but we realised that by county and year is more reasonable) to make prevention work more effective. Our additional goal was to predict next year's fires by year, month and cause.

Analysing building fires throughout 2014 - 2020



We tried to find out whether some months or counties have distinguishably more or less fires than others.

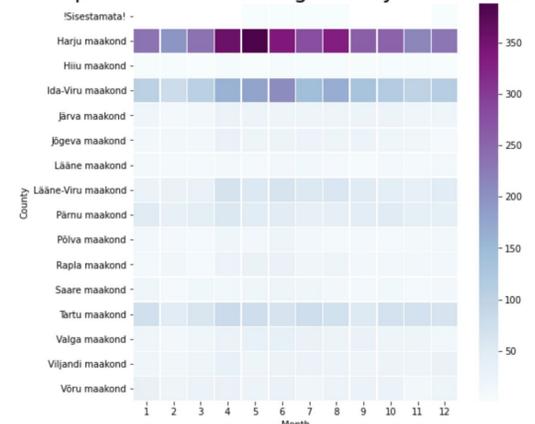
← The left map shows the number of fires taken place by month and year with a line graph.

□ It's interesting to notice how the overall amount of fires has decreased over the years (generally speaking the lines are lower by year).

→ The graph on the right shows fires taken place by month and county without taking the year into account.

□ As you can see on both plots, the largest amount of building fires happen in April, May and June, which was really surprising as we thought that with candles, heating etc, the largest amount would happen in the winter. On the right plot you can see the correlation of building fires between county and month. The most fires take place in Harju, Ida-Viru and Tartu counties.

Heatmap of count of building fires by month and county

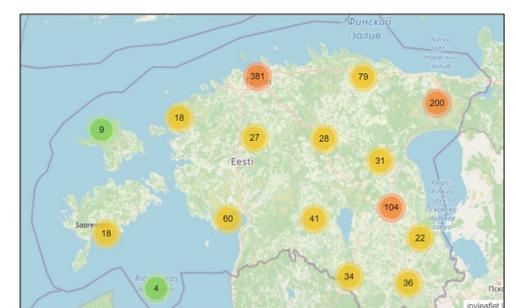
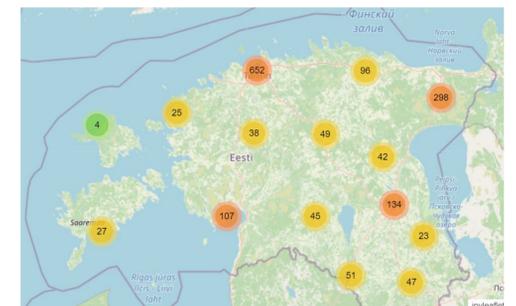
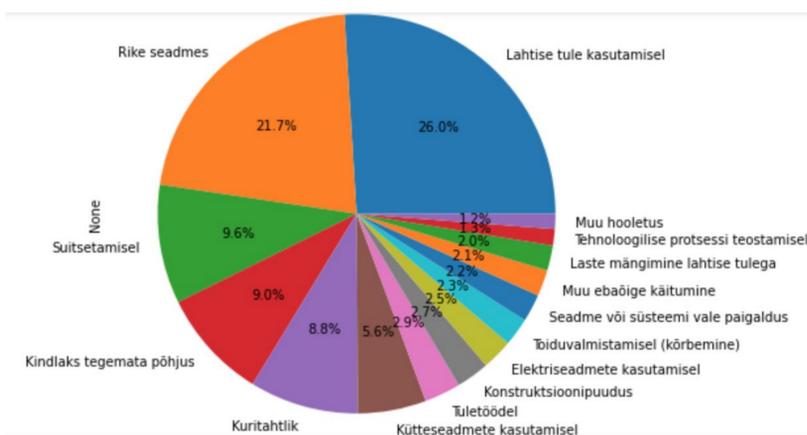


Common causes and clustering

As our second goal was to make prevention work more effective, we found out what have been the most common causes of fires overall in years 2014-2020. We clustered similar causes (for example "Rike elektriseadmes" and "Rike kütteseadmes" are now "Rike seadmes".

Most common turned out to be "using open fire" and "fault in device". We also brought out fires by counties by years separately to visualize how the amount of fires taken place in county has changed over the years. Here we brought out the results of 2014 and 2019.

For visualization we clustered those fires by county and year on the map of Estonia.



Additional facts

- Our third goal was to try to predict fires by month, county and cause. We tried to divide the data into train and test and then use Random Forest and Decision tree classifiers, but the results were unusable, because we can't really predict fires as they happen quite randomly.
- We found out that in the county of Jõgeva the most common cause of building fires is malicious arson. This is quite interesting as that means there's more fires caused on purpose than by accident.
- The largest amount of fires happened in 2015. The difference between 2014 and 2015 is really small and as the dataset begins with March of 2014, it is quite possible that there were actually more fires in 2014.
- If you look at the line plot, you can see that there is a big spike in the amount of fires in April of 2014. We couldn't figure out any reasoning behind it as there aren't any major holidays in April (for example there's a spike in June and December and these are probably caused by Midsummer day and Christmas).

Conclusion

The analysis revealed a lot of well-known facts: for example the most common causes of fires are using open fire and fault in either electric or some kind of heating device. It's also quite predictable that the most fires take place in Harjumaa county where the number of people living there is the highest and the least fires took place in Hiiumaa where the number of people living there is also smaller.

But also some quite interesting facts came into our attention: over the years 2014-2020 the number of fires per year has decreased. The most common cause of fires over the years in Jõgeva county was malicious arson. Over 250 fires took place in 2014 April. Over the years 2014 to 2020 there were always less than 10 fires in Hiiumaa and more than 100 fires in Harjumaa and Ida-Virumaa.

In conclusion it's possible to say that due to the decrease in number of fires, we are moving in the right direction in fire prevention but since the most fires took place due to 2 distinguishable reasons, people should pay more attention to using open fire more safely and the inspection of electric and heating devices could use some improvement. Also the number of malicious arsons in Jõgeva could indicate the rise of criminal activity in that region and might be something the police could look into.