..er / 800 Learnings
Learnings

Learnings regarding the Masterthesis

- The first analysis shows, that heatpump is the major energy consuming device, followed by electric vehicles. The energy consumption of other devices such as dishwashers and washing machine largely varies accross households and depends on households preferences and behaviour. The energy consumption of other devices such as freezers, refrigerators and circulation pumps are rather stable accross households.
- The second analysis shows, that pv generated energy can cover a large share of households overall energy consumption. Where as these installation produces the by far largest part of energy during middays in summer, they show a surprisingly substantial energy production also during middays in winter.
- To estimate households energy consumption in my Masterthesis, I might need several datasets containing energy consumption as well as the properties of households and buildings such as household size or appartments area.

Learnings regarding Spark

Far more important than the results, is to get to know the handling of PySpark and Zeppelin.

Challenges

- I tried to preprocess the data with PySpark. It was quite challenging and tedious, thus I switched to the local machine and used Pandas.
- Then, after a few days working with Zepellin and Spark, my Zeppelin crashed. Fortunately, I got backup files.

Learnings

- In future projects, I will keep the things simple. I mainly work on the local machine with an IDE like PyCharm, that I am quite comfortable. However, if I ran into a big data problem, I got with Spark a solution in my backhand.

- Further, I will do more local backups especially when working on virtual machines.
- Thanks to the two big data lab courses, I got quite familiar with the (Linux) Terminal and will keep practicing.

...er / 800 Learnings