Hello, so today we are going to talk about some results we got from our project answering to our problematic which is “How do the buildings' characteristics impact their yearly energy consumption and can we cluster some different consumption profiles”

As a quick reminder, we will go trough some of the concepts I introduced during the last seminar:

The first program I talked about was about the categorical and continuous energy consumption averages we can retrieve using the given properties of the buildings. The categorical features such intended uses or postal codes or clearly separated because of their different names as the continuous features, such as the gross area or the volume have a large range of possible values.

The algorithm consists of creating a list of characteristic categories or ranges. Then we browse this list and retrieve the average energy consumption of a certain number of buildings per category or range by merging the metadata and the consumption records using the property id.

At the same time, we are checking if we can actually retrieve the consumption, it happens that the records are partly missing or invalid.

Finally, we group these means by category or range to have a wider viewpoint of energy consumption.

And we can display these results with graphs to visualize them.

So, these are some interesting graphs I collected using this program.

First, on these 3 graphs, I noticed that the size of a building linearly impacts the general energy consumption as we can see on these three graphs: the more building’s gross area, volume or floor count is, the more energy it consumes. But surprisingly, we can that the electricity consumption average decreases as the floor count exceed 4. Unfortunately, this is probably due to the lack of recordings for buildings having more than 4 floors.

I also realized that some intended uses have more impact, for example, here ??? clearly show a difference with the other buildings. Contrarily as kindergartens or storage buildings.

And finally, although the year of construction doesn’t seem to have a clear impact on buildings’ energy consumption, the year of renovation shows some interesting results: buildings which were renovated between 1998 and 2009 are more consumptive than the buildings renovated before and after this time window.

So with these findings, we can try to focus on some particular buildings’ properties relationship towards the energy consumption.

For example, I started to focus on the most recorded intended uses because the more data we have, the more reliable it is. And so, the 5 most recorded intended uses are Educational buildings, Kindergartens, Health Centers, Sports buildings and Retirement homes.

If we look at the size of these intended uses using volume and gross area attributes, we have this graph, showing that Educational buildings, Retirement homes and Health centers are the biggest building by average.

The educational high average can easily be explained by the large campuses in the region such as Linnanmaa and Kontinkangas. However, since health centers use a lot of energy consumptive infrastructures (scanners, heating for every bedrooms or bathrooms), we can wonder if their energy consumption will exceeds the educational buildings.

This why I created this graph by using some parts of my first feature extraction program. As you can see, I am using the continuous program by using the volume as the x axis and the energy consumption average as the y axis. And finally, instead of grouping the data by the energy consumption type, I choose heat (or electricity) and I group the data by the 3 most recorded intended uses (Educational buildings, Retirement homes and Health centers).

Here, we can see that ???