Homework 2 – 19/09 2023

1. **Create a \*tidy\* spreadsheet/table listing the names of Danish monarchs with their birth- and death-date and start and end of their reign. They should be sortable by year of birth. Suitable source website is for example**[**here**](https://kongehuset.dk/monarkiet-i-danmark/kongerakken)**, but you can also use another source, provided you reference it. (Collaboration is welcome. Remember to attach this spreadsheet to Brightspace submission)**

[**Link to the spreadsheets about the Danish kings**](https://aarhusuniversitet-my.sharepoint.com/:x:/g/personal/au717183_uni_au_dk/Ee7MmvmUvzdBv1lmnssBGlMBl9SCKZuyd0dHpKZWW0aXPg?email=202207246%40post.au.dk&e=bWMLug)

1. **Does OpenRefine alter the raw data during sorting and filtering?**

When you import your spreadsheet into OpenRefine, the sorting of the dataset remains unchanged.

Et billede, der indeholder tekst, skærmbillede, software, nummer/tal

Automatisk genereret beskrivelse

1. **Fix the**[**interviews dataset**](https://ndownloader.figshare.com/files/11502815)**in OpenRefine enough to answer this question: "Which two months are reported as the most water-deprived/driest by the interviewed farmer households?"**

**I forgot to wirte af scribt, but here…..**

Et billede, der indeholder tekst, skærmbillede, Font/skrifttype, nummer/tal

Automatisk genereret beskrivelse

Et billede, der indeholder tekst, skærmbillede, software, display/skærm/fremvisning

Automatisk genereret beskrivelse

1. **OPTIONAL Real-Data-Challenge: What are the 10 most frequent occupations (erhverv) among unmarried men and women in**[**1801 Aarhus**](https://raw.githubusercontent.com/aarhusstadsarkiv/datasets/master/censuses/1801/census-1801-normalized.csv)**? (hint: some expert judgement interpretation is necessary. As an inspiration,**[**check out this chapter Making a living outside marriage from the**](https://academic.oup.com/book/1458/chapter/140849368)**Swedish Gender and Work project of Maria Agren)**

I used OpenRefine – and copied the URLs to open the data.

Et billede, der indeholder tekst, Font/skrifttype, linje/række, skærmbillede

Automatisk genereret beskrivelse

Et billede, der indeholder tekst, skærmbillede

Automatisk genereret beskrivelse

Et billede, der indeholder tekst, skærmbillede

Automatisk genereret beskrivelseThe data looked like this:

Et billede, der indeholder tekst, skærmbillede, nummer/tal, Font/skrifttype

Automatisk genereret beskrivelseI clicked ”include” in the text facet on the ”civilstande” I wanted. In this case ”ugift”.

Then I open a new text facet for “erhverv” and get a lot of mess, which needed to be cleaned. First, I separated people with several different “erhverv” by using: “edit cells” 🡪 “split value cells”.

Et billede, der indeholder tekst, skærmbillede, software, display/skærm/fremvisning

Automatisk genereret beskrivelse

And then I just keep cleaning… with multiple tools. Ex. “Cluster and edit”.

Et billede, der indeholder tekst, software, nummer/tal, Font/skrifttype

Automatisk genereret beskrivelse

I stopped with the cleaning….

Et billede, der indeholder tekst, skærmbillede, Font/skrifttype, software

Automatisk genereret beskrivelse