**E-mail for Nicklas/Andreas?**

Topic: The remote labour market past covid-19

1. What is your research question?
2. How are you going to find out/what methods are you planning to use?

First, we want to web-scrape jobindex/jobnet/careergate for historic and recent job listings to see how the possibility of working remotely has evolved, especially since the covid-19 pandemic has hit Denmark. The web-scraped data will be analysed by creating a list of words/phrases that indicate remote working, that we can search for in the job postings, and then visualize and describe the patterns in the possibility to work remotely.

Second, with a starting point in careergate.com’s filter “Fjernarbejde”, that indicates whether the job listing is a remote job posting or not, we would compare our method from the web-scraping to the filter to see the difference.

Third, we could try to evaluate our method for filtering remote job postings

1. Why is this going to be interesting to read?

Other thoughts: if working remotely will become more normal and available, then commute times and distances may not be as important to workers in the future. Could we model this (maybe from DST API-data?) and try to predict this?

**Exam project**

At the end of the course your group must hand in a independent exam project.

The content of the exam project is something that you choose. You and your group must find a subject, data, choose methods etc.

**Grading**

The grade for this course is exclusively determined by the project handed in. The project will be judged on a number of dimensions, these include:

* how the data was obtained (setting up new data collection);
* how the data was processed;
* how machine learning methods are applied and which methods are used;
* how results are explained (writing, figures, tables with model output etc.);
* the research question and its originality as well as how it is answered.

Some advice about the grading. It is essential that spend time on motivating your project and conveying your results. In addition, it is important that you spend time on calibrating and validating the models you work with rather than using as many models as possible. We emphasize that using machine learning is NOT necessary to make a great project, many of the best projects gain insights from the data without modelling.

**Requirements for project**

The exam projects have a number of requirements that must be met, these are: requirement

* Research question (you should discuss with TAs)
* Groups with two to four members
* Project formalia
  + Project must consist of a report (.pdf file) and a documentation as Jupyter Notebook (.ipynb file).
  + The report should be written like a brief research article (short literature review, references to methods, results etc.). The report is limited to the following maximum number of pages (normalsider).
    - 2 members, 12 pages;
    - 3 members, 15 pages;
    - 4 members, 18 pages.
    - Note that 1 page (normalside) corresponds 2,400 and does not count figures, abstract, list of reference, frontpage, appendix.
    - The report should contain your exam numbers and possibly your names (optional). The exam numbers (or names) MUST show who contributed with writing which parts of the report. At most 20 pct. of the report can be written shared. If you fail to provide this the submission of your project may get rejected!
  + Grading will be based on the report but process but data collection, computations etc. should be well documented in the supporting Jupyter Notebook.

**Possible data sources**

Students in previous years of Social Data Science have used a large variety of data sources including:

* news on [DR](https://dr.dk/) (Danish Broadcasting Company) and the Danish newspaper Information
* price of cars for sales on [bilbasen](https://bilbasen.dk/)
* analyzing linguistic content on Twitter
* Airbnb pricing in Copenhagen
* Prediction of bitcoin prices from Reddit data.

If you are interested in working with one or more of these datasets or see the projects by the students who made them please contact us and we will put you in touch.

subject,

data,

methods

The report should be written like a brief research article (short literature review, references to methods, results etc.). The report is limited to the following maximum number of pages (normalsider).