

# Data analysis and machine learning – Hjemmeopgave 2

Martin Nørgaard

Assistant Professor, DIKU

KØBENHAVNS UNIVERSITET



## Hvad går det ud på?

- formuler en problem stilling med det givne data I har fundet og så brug machine learning på at løse det
  - 1 problemstilling der anvender supervised learning
  - 1 problemstilling der anvender unsupervised learning
- Visualiser dataene vha. værktøjer som matplotlib og seaborn
- Gør Jeres kode tilgængelig online via e.g. github eller colab, så andre kan reproducere Jeres resultater (open and reproducible science)

## Aflevering

- præsentere det som **en word, pdf eller en jupyter notebook (ipynb?)**
- Sende dokumenter til Martin som gruppe **DEADLINE 5. april kl. 22**
- give feedback til de andre grupper **DEADLINE 11. april kl. 22**
- samlet feedback tager vi så op ved næste mødegang

## Start med det nu!

- Brug lidt tid nu på at diskutere ideer til Jeres opgave. Hvilket datasæt vil I bruge? Overvej gerne om disse data muligvis kan bruges i en undervisnings sammenhæng bagefter.
- Sæt en dato til det næste gruppemøde nu eller afklar hvordan I kan arbejde på opgaven sammen online (Google docs?)

## Praktiske ting

- Næste mødegang er 12. april 2024 kl. 9
- Lokale 4-0-10, Københavns Biocenter, Københavns Universitet, Ole Maaløes Vej 5, 2200 København N <https://www.biocenter.ku.dk/>
- Kaffe og frokost er inkluderet

# Spørgsmål ?

# Tak for I dag!