About Data Science Lab
Activities in 2018

More info: https://datalab.science.ku.dk
Aim and history

Overall aim: Enhance the quality of scientific data analyses in research carried out at the Faculty of Science.

- Merge of faculties with different traditions, 2013
- Laboratory for Applied Statistics 2013–2017
- Data Science Lab 2018–
Organization

Collaboration between Department of Mathematics (MATH) and Department of Computer Science (DIKU).

Funded by Faculty of Science 2018–20 with 2.4 mio kr per year, 50:50 between MATH and DIKU.

- MATH: 4 full/associate professors + 1 postdoc, all with reduced teaching obligations due to work in the Lab
- DIKU: 1 associate professor + 1 postdoc, both without teaching obligations outside the Lab
- Other colleagues whenever appropriate

Organizational meetings 1–2 times per month.
Overview of activities

• Research collaboration
• Consultancy meetings with students and researchers
• Four PhD courses per year
• Short courses about Python, R, and Data Science
• Supervision of BSc and MSc theses in collaboration with other departments or private companies
• For payment: Criminal court cases, counseling for public offices and private companies
Research collaboration and publications

- Open for counseling of all researchers from SCIENCE (PhD students and up)
- Range from a single meeting or email to co-authorship and development of new stat/cs methods
- In 2018 we had ≈ 70 requests, of which ≈ 18 have potential for collaboration, joint papers and/or grant submission
- 20 cross-disciplinary journal/conference publications in 2018
Ongoing research collaboration outside SCIENCE

- Universitetshospitalernes Center for Sundhedsfaglig Forskning (partly financed by collaborator, Novo Nordisk Fonden)
- Greenland Institute of Natural Resources (partly financed)
- Landbrugets Veterinære Konsulenttjeneste (industrial PhD)
- Departments/centres at UCPH: Economics, Psychology, Veterinary and Animal Sciences, Neuroscience, CBMR
- Osteoarthritis Biomarkers Consortium, Foundations of the National Institute of Health, US
PhD courses

- **Statistical Methods for the Biosciences I**
  4.5 ECTS, 40 participants in 2018/19 (27 from SCIENCE).

- **Statistical Methods for the Biosciences II**
  Project course, 3 ECTS, 9 in 2017/18.

- **Machine Learning and Imaging Methods**
  3 ECTS, first time from April 2019

- **Machine Learning and Imaging Projects** (previously *Big Data in the Biosciences*)
  Project course, 4.5 ECTS, 11 participants in 2018
Short courses on software for data science

- *Introduction to R.*
  1.5 days, only for PhD students and up, almost exclusively people from SCIENCE, 43 participants in 2018.

- *Introduction to Python.*
  2 days, for students and employees at SCIENCE, 34 participants in 2018 (80 participants in January 2019)

- *Data Science With RStudio*
  5 days, Copenhagen Summer University, 18 participants in 2018
Consultancy for students

Consultancy meetings with students two times per week, each time with 3 bookable time slots followed by „open door“.

Students at all levels are welcome, and meetings last ≈ 20 min.

In 2018 we had 398 such meetings.

- All departments represented in 2018, but NEXS, PLEN, FOOD and BIO account for ≈ 75%.
- Most students write their MSc thesis (285), but PhD students (66), BSc students (21), and others (26) also show up.
BSc and MSc supervision

- MSc theses in collaboration with Novo Nordisk, Lundbeck, Mærsk, deCode, Greenland Institute of Natural Resources, Biology at UCPH, Biostatistics at UCPH
- BSc theses and projects with analyses of data from Data Science Lab collaborators
Cases

*Plant growth quantification and phenotyping*
Automated image analysis of sub-soil multi-spectral root pictures

*Diving behaviour and sound production of narwhals*
Long time series of tagging data.

*Establishing metabolic phenotypes in varying environments*
Profiles from Chromatography–mass spectrometry, high-dim. data, generalized linear mixed models, NMDS scaling.

*Improvement of sales*
Design of experiment and analyses for supermarket.