



## Course information

<b>Language</b>	English
<b>Credit</b>	7,5 ECTS
<b>Level</b>	Full Degree Master Bachelor
<b>Duration</b>	1 semester
<b>Placement</b>	Summer And Autumn
<b>Schedule</b>	Summer 2016: August 8 to 26. Teaching: August 8 to 12 and 15 to 19, 9 AM - 4 PM. Writing assignment August 19 to 26  Autumn: The course will not be offered in Autumn 2016.

## Continuing and further education

**Price** 320 DKK per ECTS

**Study board** Department of Economics, Study Council

## Contracting department

- Department of Economics

## Course responsables

- Sebastian Barfort (sebastian.barfort@econ.ku.dk)
- David Dreyer Lassen (david.dreyer.lassen@econ.ku.dk)

Saved on the 18-03-2016

## Education

Recommended elective from 3. year at BSc in Economics

Elective at MSc in Economics

## Content

The objective of this course is to learn how to analyze, gather and work with modern quantitative social science data. Increasingly, social data--data that capture how people behave and interact with each other--is available online in new, challenging forms and formats. This opens up the possibility of gathering large amounts of interesting data, to investigate existing theories and new phenomena, provided that the analyst has sufficient computer literacy while at the same time being aware of the promises and pitfalls of working with various types of data.

## Learning Outcome

1. We will introduce students to the state of the art social science literature using computational methods and social data.
2. We will present students with an overview of key benefits and challenges of working with different kinds of social data. We will show how various kinds of data (survey, web-based, experimental, administrative, etc.) can be used to answer different questions within the social sciences. Furthermore, we will discuss ethical challenges related to the use of different types of data.
3. We will introduce students to statistical techniques for predicting and classification, known as statistical learning, and we will discuss how these methods relate to existing empirical tools within economics such as causal inference and regression.
4. We will present modern data science methods needed for working with computational social science and social data *in practice*. Being an effective economist and data scientist means spending large fractions of our time writing and debugging code. In this section you will learn how to write code to clean, transform, scrape, merge, visualize and analyze social data.

In addition to core computational concepts, the class exercises will focus on the following topics

- 1. Generating new data:** We will learn how to collect and scrape data from websites as well as working with [APIs](#).
- 2. Data manipulation tools:** Participants will learn how to import, transform, munge and merge data from various sources.
- 3. Visualization tools:** We will learn best practices for visualizing data in different steps of a data analysis. Participants will learn how to visualize raw data as well as effective tools for communicating results from statistical models for broader audiences.
- 4. Reproducibility tools:** Participants will learn how to use version control and social coding using [Github](#) and how to effectively communicate the insights of an analysis using markdown.
- 6. Prediction tools:** We will cover key implementations of statistical learning algorithms and participants will learn how to apply and interpret these models in practice.

1.

## After the course the student should:

- Have strong knowledge of the state of the art social science literature using computational methods and social data.
- Have strong knowledge of advantages and challenges in using different kinds of data to answer various questions in the social sciences
- Strong practical data science skills such as the ability to scrape web pages, import and export data from numerous sources, basic knowledge of functional programming and effective data visualization skills.
- Have knowledge of widely used statistical prediction algorithms as well as the ability to estimate these models in practice.

- Strong working knowledge of the R programming language for statistical computing.

## Literature

A comprehensive reading list as well as detailed information about the course is available on the course website at

[http://sebastianbarfort.github.io/sds\\_summer/](http://sebastianbarfort.github.io/sds_summer/)

## Teaching and learning methods

The course will consist of 3 hours of lectures and 2 hours of exercises and problem solving per day. The lectures will focus on the broad topics covered in the course (part 1-3 listed above). In the exercise classes we will get our hands dirty and present data science methods needed for collecting and analyzing real-world data.

2 hours of exercises a day is not a large amount of time for learning how to code. We will use some of this time like development meetings: going over assignments, having detailed code reviews of various forms, and discussing blocking issues and potential solutions.

Schedule (tentative):

3 hours of lectures: August 8 to 12, 15, 16 and 19.th, 9.00-12.00 hrs

2 hours of exercise classes: August 8 to 12, 15 and 16.th 13.00-15.00 hrs and 17.th and 18.th 9-12 hrs.

For enrolled students please find more information of courses, schedule, rules etc at

[https://intranet.ku.dk/economics\\_ma/courses/Pages/default.aspx](https://intranet.ku.dk/economics_ma/courses/Pages/default.aspx)

Timetable and classroom:

For time and classroom please press the link under "Se skema" (See schedule) at the right side of this page.

## Academic qualifications

Because the course builds on a wide range of techniques, we do not have any hard requirements, but students are expected to have an interest in some subset of: statistics, econometrics, linear algebra, and a scripting language (we will use R in this course).

## Sign up

Self Service at KUNet

## Exam (Written take-home)

<i>Credit</i>	7,5 ECTS
<i>Type of assessment</i>	Written assignment, 7 days The exam is an 7-days assignment written in groups, but where the assignment must be handed in individually before deadline.
<i>Exam registration requirements</i>	Students are expected to complete at least 2 out of 3 mandatory assignments.
	Full participation at the summerschool is mandatory and the student must actively participate in all activities.
<i>Aid</i>	All aids allowed
<i>Marking scale</i>	7-point grading scale
<i>Censorship form</i>	External censorship 100% censorship
<i>Exam period</i>	Summer 2016:  The assignment will be given the 19.th of August and has to be handed in not later than 26 August at noon.  For enrolled students more information about examination, exam/re-sit, rules etc. is available at the <a href="#">student intranet for Examination (English)</a> , <a href="#">student intranet for Examination (KA-Danish)</a> and <a href="#">student intranet for Examination (BA-Danish)</a> .
<i>Re-exam</i>	Same as the ordinary exam.

## Criteria for exam assesment

The student must in a satisfactory way demonstrate that he/she has mastered the learning outcome of the course.

## Workload

Category	Hours
Lectures	42
Preparation	124
Project work	40
Total	206

Saved on the 18-03-2016