Faculty of Social Sciences

SIMP59, Data Selection and Visualisation, 7.5 credits

Urval och visualisering av data, 7,5 högskolepoäng

Second Cycle / Avancerad nivå

Details of approval

The syllabus was approved by Graduate School Board on 2024-06-03. The syllabus comes into effect 2024-06-03 and is valid from the spring semester 2025.

General information

The course is offered as an interdisciplinary single subject course in Social Sciences at the second-cycle level, and as a mandatory course in the second semester within the Master Programme in Social Scientific Data Analysis.

Language of instruction: English

Main field of study	*Specialisation*
Sociology of Law	A1N, Second cycle, has only first-cycle course/s as entry requirements
Human Geography	A1N, Second cycle, has only first-cycle course/s as entry requirements
Development Studies	A1N, Second cycle, has only first-cycle course/s as entry requirements
Social Work	A1N, Second cycle, has only first-cycle course/s as entry requirements
Political Science	A1N, Second cycle, has only first-cycle course/s as entry requirements
Gender Studies	A1N, Second cycle, has only first-cycle course/s as entry requirements
Social Anthropology	A1N, Second cycle, has only first-cycle course/s as entry requirements
Sociology	A1N, Second cycle, has only first-cycle course/s as entry requirements

This is a translation of the syllabus which was established in Swedish.

Learning outcomes

Upon completion of the course, the student shall be able to:

Knowledge and understanding

- demonstrate knowledge and understanding about the variety of possible data sources;
- demonstrate knowledge and understanding about the variety of data visualisation methods.

Competence and skills

- discuss the strengths and weaknesses of various data sources in relation to a given research question;
- discuss the strengths and weaknesses of various data sampling strategies in relation to a given research question;
- argue for the selection of particular data sources as well as sampling strategies in relation to their appropriateness for specific research questions;
- create and adapt data visualisations relevant for specified scientific research questions using a modern programming language;
- clearly communicate scientific insights through reference to data visualisations in relation to scientific research questions;
- clean and transform a dataset in preparation for analysis and/or visualisation using a modern programming language.

Judgement and approach

- evaluate the implications of using a given data source for research analysis and results;
- plan necessary data wrangling, or the cleaning, transformation, and potential integration of data sources appropriate for given data sources and research aims;
- evaluate the implications of using a given data visualisation in relation to a given research question.

Course content

The student will learn how to select data relevant for their research questions, how to clean, transform, and potentially integrate (i.e. "data wrangling") those data for analysis, and how to effectively visualise and communicate patterns in their data for a given scientific aim.

There are two parts in the course: i.) questions related to data selection and data wrangling and ii.) data visualisation and communication. Both parts will be discussed not merely as technical exercises but rather as choices to be grounded in arguments in relation to particular research questions as well as practical constraints. The selected exercises will be implemented where relevant with a modern programming language.

The part of the course dealing with data selection and preparation will focus primarily on data that can be analysed through quantitative and computational methods. Data sources as well as their sampling will be discussed in relation to building arguments that link them to particular scientific aims/research questions. The student will also learn how to 'wrangle' their data in order to ready it for analysis /visualisation. This may include cleaning, transforming (e.g. long vs. short formats, network vs. non-network data), and/or integrating data.

In the part dealing with data visualisation, the student will learn about a wide variety of data visualisations and exploratory data analysis. The student will be instructed and practice to creatively evaluate, implement, and adapt them in relation to the communication of scientific insights. This will be implemented using a programming language.

Course design

The teaching consists of lectures, labs, seminars, and workshops.

Unless there are valid reasons to the contrary, compulsory participation is required in labs and seminars. Students who have been unable to participate due to circumstances such as accidents or sudden illness will be offered the opportunity to compensate for or re- take compulsory components. This also applies to students who have been absent because of duties as an elected student representative.

Assessment

The assessment of the course is based on two lab reports where each count for 50 % of the final grade:

- one lab report that will evaluate the arguments put forward for selection of a given data source and sample for a given research question (50 %)
- one lab report will evaluate how well the student communicates scientific insights in relation to a research question through a chosen implementation of data visualisation (50 %)

The course includes opportunities for assessment at a first examination, a re-sit close to the first examination and a second re-sit for courses that have ended during that school year. Two further re-examinations on the same course content are offered within a year of the end of the course. After this, further reexamination opportunities are offered but in accordance with the current course syllabus.

The examiner, in consultation with Disability Support Services, may deviate from the regular form of examination in order to provide a permanently disabled student with a form of examination equivalent to that of a student without a disability.

Grades

Grading scale includes the grades: U=Fail, E=Sufficient, D=Satisfactory, C=Good, B=Very Good, A=Excellent
The grade for a non-passing result is Fail. The student’s performance is assessed with reference to the learning outcomes of the course. For the grade of E the student must show acceptable results. For the grade of D the student must show satisfactory results. For the grade of C the student must show good results. For the grade of B the student must show very good results. For the grade of A the student must show excellent results. For the grade of Fail the student must have shown unacceptable results.

The grade for the entire course is based on the two lab reports, each of which are weighted equally (50/50). In cases where the average falls between two consecutive letter grades (e.g. B & C), the overall grade is rounded up. For a grade of Pass on the entire course, the student must have been awarded at least E on all assessments for which the grading scale A-E+Fail applies. The student must also have participated in all compulsory components.

At the start of the course, students are informed about the learning outcomes stated in the syllabus and about the grading scale and how it is applied on the course.

Entry requirements

To be eligible for the course students must have a Bachelor's degree with a major (i.e. at least 90 ECTS credits) in Development Studies, Gender Studies, Human Geography, Political Science, Social Anthropology, Social Work, Sociology or Sociology of Law, or the equivalent.

Oral and written proficiency in English equivalent to English 6/B (advanced) from Swedish upper secondary school is a requirement. International qualifications will be assessed in accordance with national guidelines.

Further information

The course cannot be included in a degree together with *SIMP56: Using Social Theory, 15 credits.*