

DEPARTMENT OF GLOBAL PUBLIC HEALTH

K9F5316, Fundamentals of Using Python in Health Related Research, 1.5 credits (hec)

Grunderna i att använda Python för hälsorelaterad forskning, 1,5 högskolepoäng

Third-cycle level / Forskarnivå

Approval

This syllabus was approved by the The Committee for Doctoral Education on 2023-12-08, and was last revised on 2025-02-14. The revised course syllabus is valid from autumn semester 2025.

Responsible department

Department of Global Public Health, Faculty of Medicine

Prerequisite courses, or equivalent

Epidemiology I: Introduction to epidemiology; Biostatistics I: Introduction for epidemiologists and Biostatistics II: Logistic regression for epidemiologists, or equivalent courses.

Purpose & Intended learning outcomes

Purpose

This course aims at introducing students to the fundamental elements of the Python programming language. Motivating examples arising from health-related research will be used to demonstrate how to use the programming language to answer a variety of relevant questions. Learning activities will give students the possibility to learn Python the hard yet easier way – that is – problem, code, and run.

Intended learning outcomes

After successfully completing this course you as a student should be able to:

- import and describe different types of data
- produce high quality figures of statistics
- estimate multivariable regression models (linear, logistic) including spline analysis
- conduct statistical inference based on the statistical model
- simulate plausible data generating mechanisms

• automatize code using looping and comprehension

Course content

The course is a full-time hands-on practice of Python language answering relevant health related questions based on either empirical or simulated data. The participant will learn how to import a dataset, create visualizations of distributions and statistics, estimation using popular regression models (linear, logistic), inference (likelihood based statistical tests, pointwise confidence intervals) on predicted responses or changes in predicted responses, draw pseudo-random values from theoretical probability distributions, Monte-Carlo simulations of common data generating mechanisms (interaction, non-linearity), and basic elements of programming such as creating new functions and avoid looping using comprehensions. The packages pandas/polars, numpy, scipy, statsmodels, scikitlearn, matplotlib will be used.

Forms of teaching and learning

Lectures, group work, exercises, and individual coding workout using Python.

Language of instruction

The course is given in English

Grading scale

Pass (G) /Fail (U)

Compulsory components & forms of assessment

Compulsory components

The individual examination (summative assessment) is compulsory.

Forms of assessment

Individual written examination. Students who do not obtain a passing grade in the first examination will be offered a second chance to resubmit the examination within two months of the final day of the course. Students who do not obtain a passing grade at the first two examinations will be given top priority for admission the next time the course is offered.

Course literature

Useful link: https://www.python.org/about/gettingstarted/