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 Committee for Doctoral Education on 2023-12-08, and was last revised on 2026-09-04. The revised course syllabus is valid from spring semester 2027.

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

The course aims to introduce students to the fundamentals of Python programming and its applications in data science and research, including data management, analysis, and presentation. Examples from health-related research will be used to demonstrate how Python can support common research tasks. This is not a statistics course; however, it will demonstrate how Python can be used to implement statistical methods introduced in the prerequisite courses. The course focuses on equipping students with the tools needed to use Python in their research while providing a solid understanding of core programming concepts.

Intended learning outcomes

After successfully completing the course, the student should be able to:

- apply fundamental programming concepts in Python to research-related data-science tasks
- write clear and reproducible Python code following established coding best practices
- use complementary tools, such as Git, Just, and Quarto, in a data-science workflow

- import, explore, transform, join, group, and summarise data
- create data visualisations for data exploration and presentation
- apply and interpret linear and logistic regression models in Python, building on previously acquired knowledge of regression methods.

Course content

This course is a full-time, in-person crash course in Python programming designed to provide students with practical skills for working with health-related research data. Participants will learn how to import and explore datasets, manage and summarise data, create visualisations, and apply regression methods introduced in the prerequisite courses. The course will also introduce coding best practices for data-related tasks, including working with functions and using vectorised operations instead of loops where appropriate. Python packages used in the course include Polars, Matplotlib, and statsmodels.

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.

Individual written examination. Students who do not obtain a passing grade in the first examination will be offered a second examination occasion within two months of the final day of the course. Students may be granted additional examination sessions, up to six in total, subject to discussion with and approval by the examiner.

Course literature

Useful link: https://pythondatascience.dev/