Universität Potsdam

Coding-Bootcamp: Introduction to Webscraping with Python

CEPA Course by Gerome Wolf
Location: Campus Griebnitzsee or Online
Date: 2021, April 14-16 | 10-16.30 pm
Local Organizer: Philipp Schrauth

Contact Information

Philipp Schrauth
Chair of Public Economics – University of Potsdam
August-Bebel-Str. 89 (House 7)
14482 Potsdam
schrauth@uni-potsdam.de

CEPA
Center for Economic
Policy Analysis

PCQR
Potsdam Center for
Quantitative Research

About the Course

For any empirical researcher it is a key determinant of research success to have a broad, complementary toolkit of methods and domain knowledge. Moreover, it becomes ever more important to present scientific results visually and appealingly to address a wider audience.

Data-Scraping, the programmatic automation of data retrieval (e.g. from Websites), is a powerful tool that facilitates the construction of novel datasets at relatively low costs. Python is a powerful programming language, which experiences growing popularity in the worlds of data science and academia.

In this 3-day Coding-Bootcamp participants will learn the basics and best-practices of data projects and software development in Python with a focus on implementation and application within the Web. In the workshop, participants will furthermore apply their knowledge to spatial data and learn how to visualize it. The course is self-contained and participation from beginning to end is highly encouraged as subsequent concepts and tools build on those introduced before.

The Lecturer

Gerome Wolf is an economist (MSc UCL, BSc + MSc HU Berlin) with a balanced background in empirical and theoretical work. Having worked across commercial as well as central banks and technology start-ups at the intersection of intelligence and technology he aims to support capacity building and informed decision-making based on scientific methods.

Registration

The content is tailored towards PhD candidates and early career researchers with no or little previous programming experience. We kindly ask you to register via **schrauth@ uni-potsdam.de** by **April 1st**. There will be no participation fees, but space availability is limited. It is required to bring your own Laptop! *(Instructions on software required to be installed will be given before the start of the workshop)*

Program

Wednesday	April 14	
10:00 – 13:00	Welcome and Introduction, Up and Running with • GitHub* • Python/Jupyter
13:00 – 14:00	Break
14:00 – 16:30	HTML Basics, Writing your first Webscraper
Thursday | April 15	
10:00 – 13:00	Introduction to Web APIs (Twitter**, GENESIS**, Spotify**)
13:00 – 14:00	Break
14:00 – 16:30	Continuation of Web APIs, Visualization techniques in Plotly**
Friday | April 16	
10:00 – 13:00	Scraping dynamic content (Selenium), Spatial data and analysis (QGIS, Tableau Public**)
13:00 – 14:00	Break
14:00 – 16:30	Web-Application development in Dash + Wrapping it up

*Developer login credentials required: please register and obtain them beforehand. Precise instructions will be sent ahead of the course to the participants.