CH-200

Practical programming in Chemistry

Schwaller Philippe

Cursus	Sem.	Type
Chemistry and chemical engineering	BA4	Obl.
HES - CGC	E	Obl.

Language of teaching	English
Credits	3
Session	Summer
Semester	Spring
Exam	During the semester
Workload	90h
Weeks	14
Hours	**3 weekly**
Courses	1 weekly
Exercises	2 weekly
Number of positions	

Summary

This course offers a comprehensive, practical introduction to computer programming tailored for chemists and chemical engineers. Python is the main language used throughout the course.

Content

This course provides a comprehensive and practical introduction to computer programming with a specific focus on its application for chemists and chemical engineers. The course assumes little programming experience and primarily uses Python as the teaching language. The content covers crucial programming concepts like data types, control structures, string processing, functions, and input/output operations. These concepts are taught within the context of chemical tasks.

The course format consists lessons on diverse topics and hands-on exercise sessions. We will cover the basics of configuring a computer and an introduction to Jupyter(Lab), a popular tool used in data science and cheminformatics. The course will teach basic command line skills, usage of variables, operators, and types, more about operators and conditionals, lists and tuples, iteration, and functions. Lessons cover string methods, dictionaries, packages and modules, file I/O, the use of Git for version control, errors and exception handling, style, comprehensions, and an introduction to Pandas and RDKit for chemical data manipulation. The course will also contain lessons on managing and cleaning data, along with making plots.

In addition to exercises, students are required to complete a project that involves digitizing data from a literature paper, then applying the skills they've learned to analyze, plot and visualize the data. By the end of the course, students will have a solid foundation in programming, enabling them to create basic scripts to operate general and chemistry-specific software packages, analyze data, and generate figures. This course will provide them with the necessary skills to apply programming to their field of work or study, particularly emphasizing how these principles can be applied to common tasks in chemistry. Such skills would be valuable in research and industry, where data analysis and presentation are critical.

Keywords

Python, Cheminformatics, Data Analysis, Data Visualisation

Learning Prerequisites

Required courses

Information, Computation, Communication CS-119(k)

Important concepts to start the course

Basic programming knowledge in Python.

If you want to refresh your Python skills before the start, you can go through the "Introduction to Programming" and "Python" tutorials on https://www.kaggle.com/learn.

Learning Outcomes

By the end of the course, the student must be able to:

- Produce chemical data in a machine-readable format
- Represent molecules and reactions in different chemical representations
- Apply common Python tools for chemical tasks
- Visualize chemical data with Python
- Analyze chemical data with Python

Transversal skills

- Assess progress against the plan, and adapt the plan as appropriate.
- Plan and carry out activities in a way which makes optimal use of available time and other resources.
- Communicate effectively with professionals from other disciplines.
- Demonstrate the capacity for critical thinking
- Demonstrate a capacity for creativity.

Teaching methods

1h lecture
2h hands-on exercises (bring your own laptop)

Assessment methods

The assessment will be project-based. During the project, the students will have the opportunity to apply the concepts and techniques they have learned throughout the course to real-world chemical data.

Resources

Virtual desktop infrastructure (VDI)

No

Bibliography

- [MolSSI Education Resources](http://education.molssi.org/resources.html#programming)
- [Greg Landrum's RDKit blog](https://greglandrum.github.io/rdkit-blog/)
- [Esben Bjerrum's Cheminformania](https://www.cheminformania.com)
- [iwatobipens' blog](https://iwatobipen.wordpress.com)- [Pat Walter's Practical Cheminformatics With Open Source Software](https://github.com/PatWalters/practical_cheminformatics_tutorials)

Moodle Link

- https://go.epfl.ch/CH-200