

DATA2001: Data Science, Big Data, and Data Diversity

Semester 1, 2021

Tutorial Week 7: Exercises on Web APIs

This week's exercises are about the topic of the week: Retrieving data from Web Services using a programmable API. If you haven't done so, you also should get organised for the upcoming practical group assignment.

Exercise 1. Accessing Web Service APIs to gather data

In this exercise, we learn how to use a Web Service API (application programming interface) such as from OpenStreetMap to retrieve data or augment existing data with further information (e.g. for geo-coding some data). The data in these cases comes in form of a semi-structured textual format, either JSON or XML. Please have a brief recap of this week's lecture on semi-structured data (JSON and XML).

In this tutorial, we are using Python via Jupyter notebooks to access web service APIs and to extract data from there. Please download the corresponding notebook from our Canvas website, and then upload it to your personal Jupyter notebook space.

The exercise has several sup-parts. Open the Jupyter notebook for Week 7, and start working on Exercises 1.

Exercise 2. Assignment Hints and Group Forming

Finally, you should also prepare for the upcoming practical assignment in DATA2001/2901. The assignment is groupwork for teams of typically 2 students (max 3), and besides some data analysis it will involve some programming, data integration, data cleansing and even database work. You should form teams now and sign up with your chosen team members in the group pages on Canvas (under 'People' - 'Groups'). Please follow the corresponding instructions at the end of this week's Jupyter notebook.