Quantitative Data Analysis – Exercises

(Week 05)

In these exercises, you will learn:

- basic GIS functionalities of Python
- how to perform geocoding of addresses
- how to perform a nearest neighbor analysis
- how to perform point-in-polygon analysis
- how to perform catchment area analysis
- how to work with Swiss buildings data set
- how to work with raster data from Swisstopo
- how to extract, analyze and visualize spatial entities from Tripadvisor data

In the data analytics process model, these exercises cover part of the steps "Preparing & storing data" and "Exploratory Data Analysis (EDA)" (see figure 1). Results of the exercises must be uploaded as separate files (**no .zip files**!) by each student on Moodle. Details on how to submit the results can be found in the tasks below.

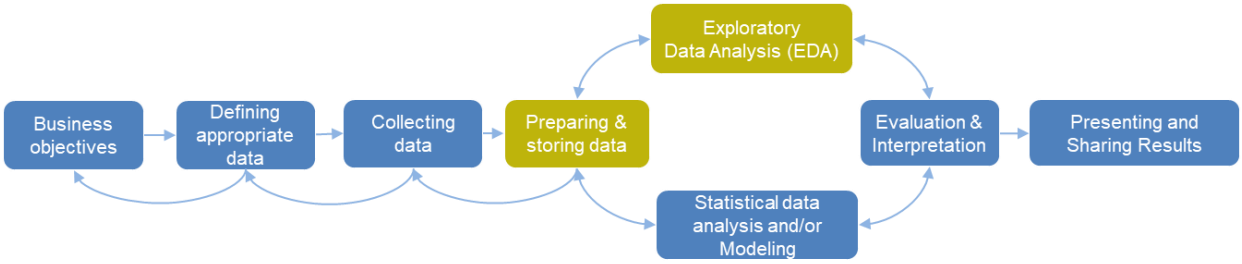

Figure 1: Data analytics process model (see slides of week 01)

Task 1

In this exercise, you will learn the fundamental Geographic Information System (GIS) functionalities available in Python. Python is a powerful language for handling geospatial data due to its rich ecosystem of libraries like geopandas, rasterio, and others. This exercise will guide you through performing common GIS tasks such as loading, manipulating, analyzing, and visualizing spatial data using Python. The tasks are:

a) Create a GitHub Codepaces environment based on the GitHub repository: https://github.com/mario-gellrich-zhaw/spatial_data_analysis.
b) Run the Jupter notebook 'basic_GIS_functionalities_Python.ipynb' in the folder '01_Python_Basic_GIS_Functionalities', run step by step.
c) Create a subset of the data and a map showing the municipality of 'Wädenswil'.

To be submitted on Moodle:

d) A screenshot of the map as 'waedenswil_map.png'.
e) A screenshot the attribute table as 'waedenswil_attribute_data.png'.

Task 2

In this exercise, you will learn how to geocode addresses using Python and the GeoAdmin Web API. Geocoding is essential for integrating location-based data into mapping applications, spatial analysis, or any project that involves geospatial data. The tasks are:

a) Go to the website https://tools.retorte.ch/map, search for a single address and view the coordinates of this address on the left-hand side of the website. You can find both the Swiss coordinates and the World Geodetic System 1984 (WGS84) coordinates of this address.
b) Run the Jupter notebook 'geocoding_addresses.ipynb' in the folder '02_Python_Geocoding_Addresses' step by step.
c) In the 'Geocoding a single address' section of the Jupyter notebook, change the address "8400 Winterthur, Theaterstrasse 17" to an address of your choice and then geocode the address in the 'Server request & response' section of the Jupyter notebook. Compare the coordinates with those on https://tools.retorte.ch/map.

To be submitted on Moodle:

- A screenshot of your own geocoded address as 'address_geocoded.png' analog to the example from the Jupyter notebook below:

	attrs
featureId	2323240_0
label	Theaterstrasse 17 8400 Winterthur
lat	47.503517
lon	8.727852
x	262215.90625
y	697129.5

Task 3

In this exercise, you will learn to perform a nearest neighbor analysis using apartment and supermarket data. This type of analysis is useful for understanding accessibility, identifying patterns in urban infrastructure, and making decisions related to urban planning, real estate development, and service accessibility. The tasks are:

a) Run the Jupter notebook 'nearest_neighbor_analysis.ipynb' in the folder '03_Python_NearestNeighbor_Analysis' step by step.
b) Create subsets containing only 'ALDI' and 'VOLG' supermarkets.
c) Perform a nearest neighbor analysis based on this subset of supermarkets.

To be submitted on Moodle:

- A screenshot of the map with apartments and nearest supermarkets as 'nearest_neighbor_analysis_aldi_volg.png'.

Task 4

In this exercise, you will learn to perform a catchment area analysis for supermarkets. Catchment area analysis is a geospatial technique used to identify the geographic area from which a business, service, or facility draws its customers. This type of analysis is crucial for understanding market reach, accessibility, and customer behavior, and can be used for planning new locations, optimizing service coverage, or understanding competition in a specific region. The tasks are:

a) Obtain an ors token as described in the file ORS_TOKEN_README.md in the folder '04_Python_CatchmentArea_Analysis'. Save the token in the file 'ors_token.txt'.
a) Run the Jupter notebook 'catchment_area_analysis_supermarkets.ipynb' in this folder step by step.
b) Look at the output. What is the number of residents in the isochrone area? What is the estimated purchasing power?
c) In the 'Create isochrone for a single supermarket' section, change the parameters to municip = 'Wädenswil', transport = 'cycling-regular' and traveltime = 10. Run the Jupyter notebook with the changed parameter values.
d) Now look at the output. What is the number of inhabitants in the isochrone area? What is the estimated purchasing power?

To be submitted on Moodle:

- A screenshot of the isochrone area as 'isochrone_map_waedenswil_bicycle.png'.
- A screenshot with the number of residents and purchasing power as 'residents_pp_waedenswil_bicycle.png''.

Task 5

In this exercise, you will learn to work with the Swiss Buildings dataset, known in German as the Gebäude und Wohnungsregister (GWR). This dataset contains detailed information about buildings and residences across Switzerland, including attributes such as location, building type, year of construction, and housing characteristics. The dataset is maintained by the Federal Office for Statistics and is a valuable resource for a wide range of analyses related to urban planning, housing, infrastructure, and socio-demographic studies. The tasks are:

a) Run the Jupter notebook 'gwr_import.ipynb' in the folder '05_Python_GWR_Data' step by step.
b) Create a subset of all buildings in the municipality of Wädenswil and plot the buildings on a map.

c) Select and count the number of existing buildings with one apartment (GKLAS 1110) in the city of Zürich.

To be submitted on Moodle:

- A screenshot of the map as 'buildings_map_waedenswil.png'.
- A screenshot of the selection as 'bldg_selection_zuerich.png'.

Task 6

In this exercise, you will learn to work with raster data, specifically aerial photos from Swisstopo. Working with aerial imagery is particularly useful for geospatial analysis, urban planning, environmental monitoring, and many other fields where spatial patterns and features need to be studied from above. The tasks are:

a) Run the Jupter notebook 'raster_analysis.ipynb' in the folder '06_Python_Raster_Data' step by step.
b) Change the Swisstopo image in the Jupyter Notebook (choose your own) from: https://www.swisstopo.admin.ch/de/orthobilder-swissimage-10-cm.
c) In the section 'Masking a raster using a shapefile' change the coordinates (x1, x2) and the side length in meters according to your own Swisstopo image.
d) Create a raster based on your changed coordinates and side length.

To be submitted on Moodle:

- A screenshot of the new raster as 'selected_raster_area.png'.

Task 7

In this exercise, you will learn how to extract, analyze, and visualize spatial units from Tripadvisor data. This is particularly helpful in cases where you want to explore the geographical distribution of reviews, locations, or venues across various cities, regions, or countries. Analyzing spatial data can offer valuable insights into patterns of user preferences, tourist hotspots, or the influence of location on ratings and reviews. The tasks are:

a) Look at the information about day trips available on Tripadvisor and the data in 'tripadvisor.csv' in the folder '07_Python_Tripadvisor'. Which data is available?
b) Run the Jupter notebook 'tripadvisor_spatial.ipynb' step by step.
c) Look at the results of this data analysis and briefly summarize the most important results. What can you learn from the results?

To be submitted on Moodle:

- Nothing 😊 but try to understand what the Python code does!