Exercise 2: Open Data and Methods

General Information:

- Exercises must be done and handed in individually.
- Submission deadline is **7.11.2018, 23:55 Uhr** through **Moodle**.
- Each exercise in this course contains 50 points. So from all six exercises you can achieve 300 points. One of the requirements to pass this course is that you achieve at least 150 points. Some exercises also contain some bonus points.

Documents for submission on Moodle:

- **PDF file containing the documentation** of how you solved the exercises and the answers to the questions in English or German. Name it according to *"ex02_yourname_matriclenumber.pdf"* e.g. ex02_MelanieMüller_1234556.pdf.

Note: Include your name, matriculation number and GitHub name in the beginning of your documentation file. Your work on GitHub is part of the grading!

In this exercise, you will gain experience in searching for open data and methods. Read the instructions carefully and document your answers in a PDF document that you submit to Moodle.

Some of these exercises require you to work in the command line. In this case, provide all the commands that you used to solve the task and add a brief comment what it does as explained in the seminar. If the question requires you to provide the output of a command that is shown in the command line, copy and paste it from the command line into your documentation file along with the command that was use to create it and a brief comment.

Moodle Forum: If you run into problems (e.g. error messages) while running some of the modules and you cannot figure it out using online research, post it on the Moodle forum to get help. Other students might have similar questions or can help you. **Participation in the form is also part of your participation grade.** You may also post questions about exercises if something is unclear.

Introduction

The data for this exercise is stored in the following GitHub repository:

https://github.com/redfrexx/OpenSourceGIS_exercise02.git

Fork it using your GitHub account and clone your forked repository to your computer using git in the command line or Tortoise git.

Exercise 1 Data methods and gathering 10 points

We are interested in knowing the temperature in nearby cities and in Heidelberg as well as the spatial distribution of those stations. Therefore, we have to first gather the data, then extract, transform and visualize it. The tasks are:

1. Go to https://udo.lubw.baden-wuerttemberg.de/public/ and download the air temperature at the stations "Heidelberg", "Mannheim-Nord" and "Karlsruhe-Nordwest" for the last month. Write down the temperature for each station at 2018-10-01 01:00.

(Go to “Luft” → “Luftmessdaten” → “Komponentenvergleich” and then search and select the stations as well as the time interval)

2. Search the location for each of those stations (under “Luft” → “Luftmessdaten” → “Stationsinformationen”) and write the coordinates in WGS84.
3. What is the original coordinate system?
4. Load the data into QGIS as a delimited text layer.
5. Visualize the position of each station over a map of the “Regierungsbezirk Karlsruhe in Baden-Württemberg/Germany”. Use https://download.geofabrik.de/europe/germany/baden-wuerttemberg/karlsruhe-regbez.html as a base layer. Make a screenshot.
6. What level of Open Data is the LUBW data set and how could it be made to level 4?

Flood risk analysis

In the remaining exercises, a flood risk analysis for a region in South Western Slovenia should be performed. The aim of this analysis is to identify populated places within flood prone areas where many people would be affected by a flood. After doing some research on available methods and data, we will perform the analysis and interpret and verify the results.

Exercise 2: Define the study area 5 points

As a first step of the flood risk analysis, we need to get a data set that delineates the study area.

1. Download the “Municipalities(LAU2)” from https://www.stat.si/gis/Baza.aspx?lang=en.
2. Extract the file and load Obcine.shp into QGIS. If the layer is not displayed properly, open the layer properties and change the coordinate reference system (CRS) to EPSG:3787.
3. Select the municipalities “Izola” and “Piran”. Save the selected features in a new shapefile called “studyarea.shp” with EPSG:3787.
   Hint: use “Save as …” from the layer menu of the shapefile.
4. Upload the new shapefile to your GitHub repository.

Exercise 3: Data and Methods research 15 points

Now we need to find a suitable methodology to perform the flood risk analysis considering the fact that we only have limit time and processing resources available. A short literature search yields a study by Pourali et al (2016). You’ll find it on Moodle. They propose a simple method to derive the flood risk using the Topographic Wetness Index (TWI). Based on the Pourali et al (2016) answer the following questions:

1. Why do they use the TWI instead of a hydrological model to derive flood prone areas?
2. How is the TWI related to flood probability?
3. What are the limitations of using the TWI to map flood prone areas?

We want to test this method, so we have to search for a suitable implementation of the TWI that we can use for our calculation.

4. Search for implementations of the TWI. List two possibilities and answer the following questions for each method. Provide your information sources.
   1. What kind of input data is required and which parameters need to be specified?
   2. How easy is it to run? Is it already part of QGIS or do you need to install something?

   Hint: Search within the QGIS Processing Toolbox, GitHub, R packages, Python packages and other online resource.

5. Which method would you choose and why?
6. After reading the documentation of some of the TWI modules you should have found out that all of them require a Digital Elevation Model (DEM) as input. Do an online search for open (free) DEM data sets that cover our study area. List two potential data sets and answer the following questions for each of them:
    1. What is the resolution in meters?
    2. How accurate is it? Is there information about potential errors in the dataset?
    3. What is the coverage? Is it only locally available or is there a global coverage?

Exercise 4: Calculate the TWI 10 points

One of the implementations of the TWI that you might have found in exercise 3 is the GRASS GIS module *r.topidx.* We will apply this module for the flood risk analysis using the SRTM DEM.

1. Download the SRTM DEM tile covering the study area from https://earthexplorer.usgs.gov (you have to register to download data). Choose the *"SRTM 1 Arc-Second Global"* data set and download it as a raster file in .tif format.
2. The TWI can only be calculated for DEMs with a projected coordinate reference system. Use *gdalinfo* from within the *OSGeo4W shell* to get information about the coordinate reference system of the DEM that you downloaded. Provide the CRS of the raster file as a WKT (well known text) string in the documentation and answer the following questions:
    a. Is it a projected or a geographic coordinate system?
    b. What is the resolution of the data set? How much is it in meters?
3. Load the DEM into QGIS and export it using the coordinate reference system EPSG:3787. Load to the new raster file into QGIS.
4. Open the QGIS processing toolbox. Search for the GRASS GIS module "*r.topidx"*. Apply it to the reprojected DEM. Set the "GRASS GIS 7 region cell size" to 31 meters.
5. Color the resulting TWI layer using a beige to blue color scale where blue indicates high TWI values. Make a screen shot and add it to the documentation.

Exercise 5: Analyse and verify your results 10 points

Now we will interpret our results. If you didn't manage to run the TWI in QGIS in the exercise before, use the TWI provided in the GitHub repository.

1. Load an OSM base map to your QGIS project using the OpenLayers Plugin.
2. Add the TWI raster file to QGIS, if it is not there already. By comparing the TWI and OSM, identify **urban or populated areas that are situated within flood prone areas**. You may also use population data available at https://www.stat.si/gis/Baza.aspx?lang=en for a more detailed analysis. Create a polygon shapefile called "flood_risk_areas.shp". Mark the high risk areas by digitizing them as polygons and adding them to the shapefile.
3. Do an online research about flood events in Slovenia in the last years. Are the areas that were hit by the flood back then amongst the regions that you identified? Do you think the TWI is a suitable method for a flood risk analysis in this case?
4. Add the "flood_risk_areas.shp" shapefile to your GitHub repository.
5. Create a map showing the TWI and the flood risk areas. Add it to your documentation.

Bonus exercise: 5 Points

Write a simple data crawler (e.g. in Python 3) which provides regular data updates for air temperatures from the opensensemap project (https://opensensemap.org/ or https://docs.opensensemap.org/). We propose to use https://scrapy.org/ as a starting point. Submit the code to your GitHub repository.