Exercises gdms0200

Version V002, 2019-10-08, rb

The task is to select in QGIS the DWD stations in NRW with <u>annual</u> time series from before 1951 until the latest date (2018-12-31) and to export the list as a CSV file. Afterwards you have to import the list with "long time series" by means of Pandas to a Jupyter Notebook and to write a script which automatically downloads the respective annual time series data (zip archive).

1 Download and convert DWD station descriptions for different time resolutions.

1.1 Download

Download from ftp://opendata.dwd.de the files:

- KL_Jahreswerte_Beschreibung_Stationen.txt
- KL_Tageswerte_Beschreibung_Stationen.txt
- TU_Stundenwerte_Beschreibung_Stationen.txt

You can do it manually or use the Jupyter Notebook gnb0201_DWD_station_compare_V001.ipynb.

1.2 Convert to CSV

Use the Jupyter Notebook to convert the fixed width format files to CSV with English column names and ";" as separator. You can use **gnb0201_DWD_station_compare_V001.ipynb** or your own Python scripts.

2 Import to QGIS and compare the point vector layers.

2.1 NRW boundary in the background

Import the NRW boundary to QGIS. Set the project coordinate reference system to **EPSG:25832** if it has not been set automatically.

After conversion of the text files describing the DWD stations you should have generated the CSV files:

- KL_Jahreswerte_Beschreibung_Stationen.csv
- KL_Tageswerte_Beschreibung_Stationen.csv
- TU_Stundenwerte_Beschreibung_Stationen.csv

2.2 Import the CSV files as point layers.

Import the CSV to QGIS. Maybe you have to change the order of the layers and move the NRW layer to the bottom (back) so that the other point layers are displayed on top of it.

2.3 Compare.

Compare the point clouds just by showing and hiding the layers. What do you observe?

3 Crop the layer *KL_Jahreswerte_Beschreibung_Stationen* to NRW

Use the crop tool to crop (cut) only the points of the layer which are inside the NRW polygon.

Save the result as a shape file named "DWD_stations_annual_NRW". Add this layer to your QGIS project.

4 Select stations with "long time series".

Open the attribute table of the new *DWD_stations_annual_NRW* layer. Click on the Greek letter ε to open the *Select by Expression* dialog. The middle window pane comprises all kinds of classes, e.g. *Fields and Values* or *Operators*. Try to create a query to select all stations in NRW with time series from before 1951 and until the latest date of all the time series.

You should have 12 stations selected. Check on the map where they are located.

5 Export the station list providing "long time series" as CSV.

In the Layer panel right click on the layer with selected features and select
Export -> Save Selected Features as …

Select the right format (CSV), right delimiter (";"), folder (data/generated/) and give it a meaningful filename. In the subsequent Jupyter Notebook it is assumed that the CSV file is:

data\generated\DWD_stations_NRW_long_TS.csv

6 Read the station list with Pandas and download the respective time series via FTP.

Use pd.read_csv() to read the selected station list with "long time series" in NRW.

Search for the FTP directory with the annual climate data and store the string in the variable *ftpdir*.

Iterate to the Pandas station_id column and generate the appropriate filenames from it. Store (append) them to a list named *filenames*.

Use ftplib together with the ftp directory and filenames to download the zip archives into the local directory data/DWD.