Assignment 1: Introduction to GIS and Python

1 Introduction

This assessment can be worked on individually or in a group.
This assignment is not marked, however it is an essential exercise to familiarise yourself with GIS and spatial data, which will be used in the course.

2 Aim

This practical exercise introduces you to Geographic Information Systems (GIS). GIS are expert tools to manage, explore, query, analyse and visualise spatial data. They come in various shapes (desktop-based, web-based, mobile) and purposes (mapping, analytics, dedicated services).

GIS are standard tools in the spatial information profession and beyond, similar to CAD in construction, Photoshop in photography/advertising, or SAP in accounting. You will desire to get deep skills in using this tool, and this practical exercise will form your first step towards this goal. It is based on self-study (which can be done in groups, of course) and self-motivated exploration. It attempts to make you curious, and invites you to go beyond beaten paths.

This practical exercise is essential for all other assignments in your course and for your professional work in the future. In this regard you have to do it; full stop. Failing to answer the following questions (in time) will deduct points from your final subject marks. But answering the questions covers only a small portion of the learning experience (say, it is equivalent to a "Pass" mark). Any further effort will pay off later, indirectly, so by all means: challenge yourself. This is your chance.

3 Assignment

Follow the assignment worksheet through the three tasks and answer the following questions in a word document submitted to LMS.

4 Questions

4.1 Task 1

1. The shape file data type is still in use 20 years after its creation. During this time, computers have changed a lot. Why do you think we still use this file type, and can you name one disadvantage to using it?

2. Aside from the database and geometry files included in the shapefile package. There are two types of index files, one is to assign an index to a feature, but there is also a spatial index, what might this be used for?

3. Furthermore, there is a file for projection. This information is important for QGIS to position the data on a 2D interface such as QGIS. The projection of the project is located at the bottom right of the window and is set to the the projection of the first layer that is imported. We can see that it is EPSG:28355. What might this mean, and what is the name of this projection?

4. Shape files are not the only method of importing data to GIS. Are you familiar with any other file types or methods that store spatial information?

4.2 Task 2

5. Which parish has the most number of cows in 1995?
6. What is the name of the most westerly parish in Gippsland?
7. Take a look at the column names in the metstations layer. Take note, for example at the first column. Why do you think the columns are named this way?

4.3 Task 3

8. Why do the values in metstations differ?
9. Now you have the coordinates in latitude and longitude, you should be able to put these into, for example, Google Maps to see where they are. Where is the data actually located? (You can also try installing the OpenLayers plugin and import maps from many different sources right into QGIS. Remember to turn on experimental plugins in the plugins settings menu, first).
10. What do you think has happened to the projection of the dataset?
11. What does this tell you about the rest of the data, for example area and perimeter, in the parish shapefile?
12. Is latitude the X or the Y axis?
13. When using the Python console, what is the variable f corresponding to?

4.4 Task 4

14. A GeoPackage is an open data format published in 2014 that allows many vector and raster layers to be contained into one file, and overcomes limitations found in legacy fomats such as ESRI Shapefiles. Which database format is used as its method of containerisation? Create a GeoPackage in QGIS and export the four layers used in this exercise into it. Include it with your assessment submission.
15. In the GEOM90042's assessment 1 git folder, there is a python file rdp.py. After downloading it it can be executed from a command line by running:

    python rdp.py

    (a) The file outputs a series of tuples to the standard output. Open the file with a text editor and familiarise the program and its syntax. Briefly explain the functions *distance()*, *point_line_distance()* and *rdp()*.

    (b) This program is implementing an algorithm. What is the name of it and its intended purpose?

    (c) Looking at the statement *if __name__ == '__main__'* what is its purpose?

5 Marking scheme

The assignment's completion will be a consideration in the marking of future assignments.

6 Tips

This task is an introduction to GIS and spatial data, it is encouraged to follow the links at the bottom of the worksheet for more information, and practice the Python QGIS tutorials.

7 Submission

Submit a single **.zip** file (where **studentno** is your student number) containing both the PDF with your answers to all the questions along with the geopackage from Q14:

studentno_A1.zip

You submit your file by uploading the file to Canvas. Only the last upload before the submission deadline will be marked.