LAB2-PYTHON

OPERATING SYSTEMS-DR. LÉONARD JANER

2021-2022

February 2022

TecnoCampus
Escola Superior Politècnica

Centre adscrit a la

upf. Universitat Pompeu Fabra *Barcelona*

LAB2-PYTHON

Objectives

Part 1: The Transparency Portal

Part 2: Selected Data Set

Part 3: Process Selected Data Set

Background / Scenario

In this lab, you review some files formats to manage information. Then you will write a Python script to extract information from a RDF file on represent visually the information

Required Resources

- 1 PC with operating system of your choice
- Virtual Box or VMWare
- DEVASC Virtual Machine

Instructions

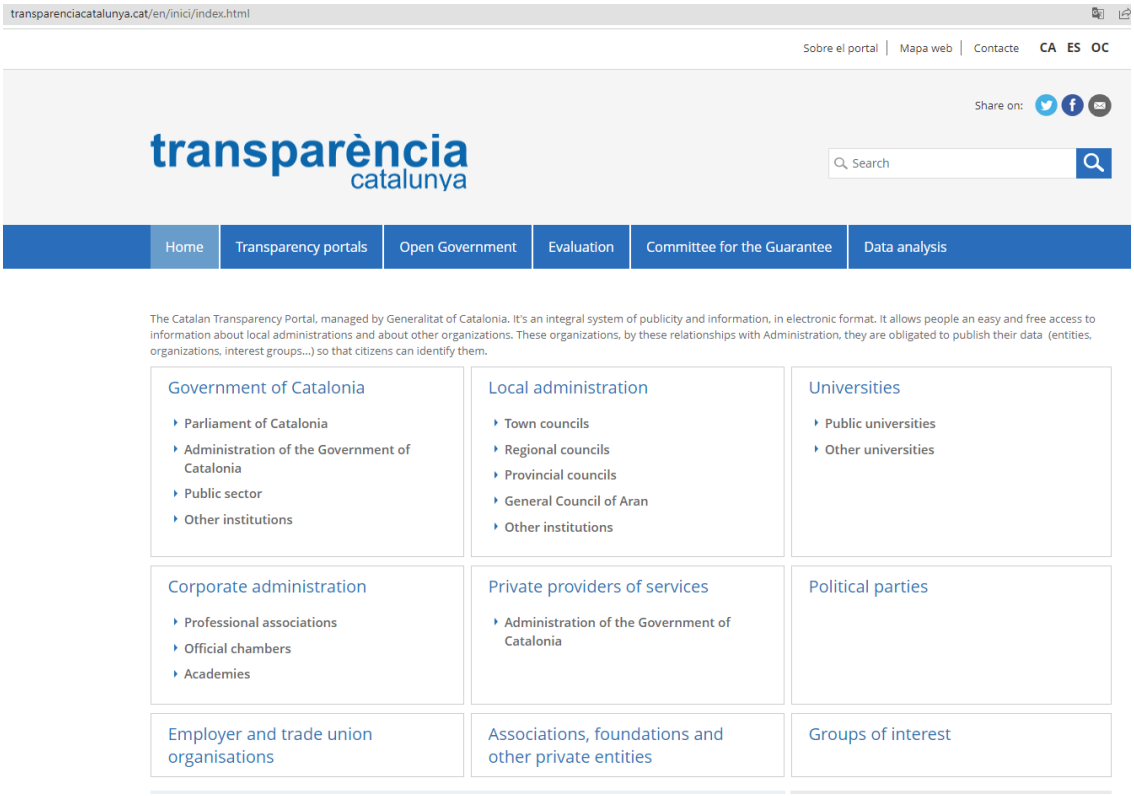
Part 1: **The Transparency Portal**

The objective of this python scripting challenge is to have the ability to plot information on a graph, based on some geolocalization.

The information to be used on the challenge is information form Open Data Repository from Generalitat de Catalunya.

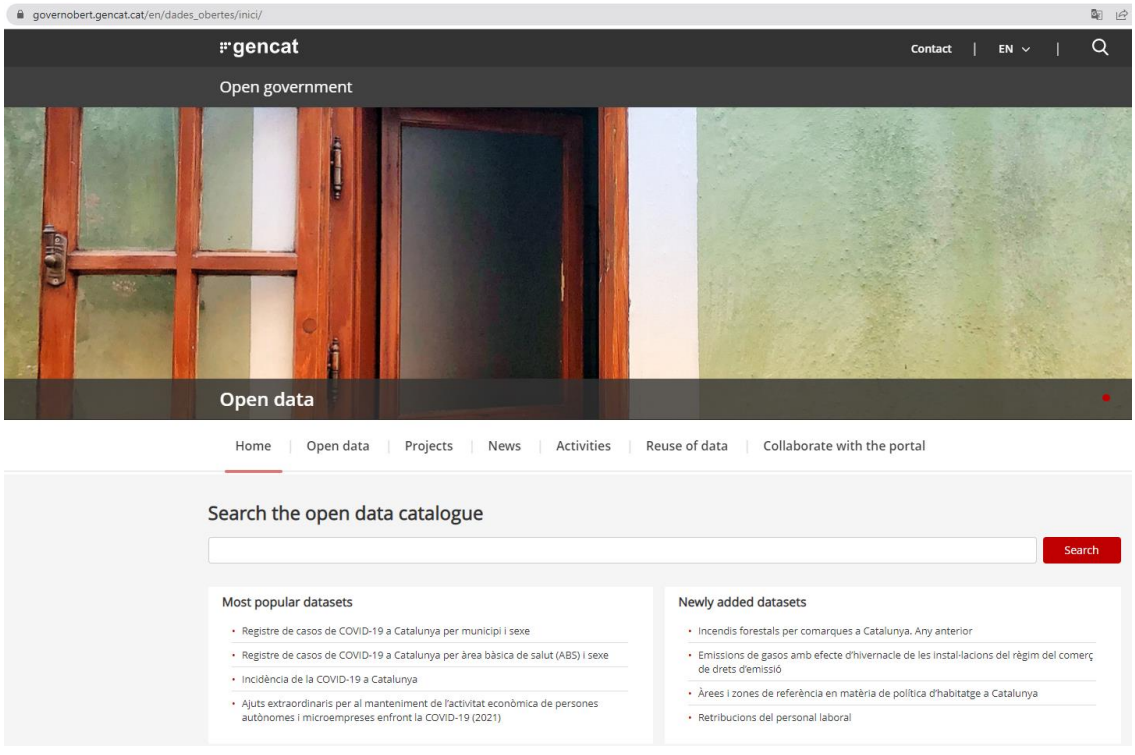
There is the TRANSPARENCY PORTAL where different OPENDATA SETS are published to be used:

http://www.transparenciacatalunya.cat/en/inici/index.html



There is the OPEN GOVERNMENT PORTAL, where you can find a huge amount of public data from so different sectors:

(http://governobert.gencat.cat/en/dades_obertes/index.html)



Have a look to the portal to have a basic knowledge about the kind of information you can have (find)

Part 2: **Selected Data Set**

The dataset to work with is:

Registre d'Entitats Esportives

The link to the dataset is:

https://analisi.transparenciacatalunya.cat/en/Esport/Registre-d-Entitats-Esportives/qrgc-u7pk

Write on the report information about:

1.- What is the purpose of the data set

2.- Which is the information you have on each row of the data set

3.- How can the dataset be exported. List of the formats and explained all the formats the dataset can be exported to. How the information is represented in the exported file format. Which is the number of records of each exporting option

Part 3: **Process selected Data Set**

Write a Python script, to process the data set, from an exported **CSV file**. The file would be downloaded in the local directory where the script will be executed from.

The script will print a menu with a list of options (to be executed):

1.- Print a table with the name of **PROVINCIES** (and for them the number of items on the file. Number of **ENTITATS**. Not ROWS).

2.- Print a table with the MODALITATS and for each one the number of entities, order TOP DOWN

3.- Print a list of the **PROVINCIES** on the file, then the user will select one of them. The script will print the **COMARQUES** there. The user will select one of them. And then print the list of kind (MODALITAT) of entitats and the number for each type.

4.- Ask for a **COMARCA**, and then print, all the municipalities (**MUNICIPIS**) for it. The user will select one municipality… then the list of entities there will be printed, the user

will select one, and that one will be shown/plotted on a MAP (right resolution to be able to identify the location). With the right zoom in.

5.- Ask for a **POSTAL CODE …** then all the entities with that postal code will be shown/plotted on a MAP (right resolution to be able to identify the location). With the right zoom in.

6.- Plot/draw/represent on a map (**MAPA DE COMARQUES DE CATALUNYA**) per each Comarca the number of entities on a colour map per Comarca. Print information only for the TOP10 values (the other one not represented)

7.- Plot/draw/represent on a map (**MAPA DE COMARQUES DE CATALUNYA**) per each Comarca the number of entities on a colour map per Comarca. Print information only for the DOWN10 values (the other one not represented)

8.- Plot/draw/represent on a map (**MAPA DE COMARQUES DE CATALUNYA**) per each Comarca the number of entities on a colour map per Comarca. Do not print information for BARCELONES (all the comarques except that one)

9.- Plot/draw/represent on a map (**MAPA DE COMARQUES DE CATALUNYA**) per each Comarca the number of entities on a colour map per Comarca. Print information for all the comarques

10.- Plot/draw/represent on an histogram the number of entities per Comarca

11.- ask if the user wants to execute a new option or exit

The task report must include:

- Information about how the environment has been saved by the student.
- Information about how the environment must be imported into the instructor system.
- Information about all the packages that must be used on the lab (with a brief description of the package purpose for the lab)
- Information about the structure of the program, and description of the most relevant parts of the code
- (COLOURMAP REPRESENTATION must be explained with deep details)
- (GEOLOCALIZATION COMPUTE AND REPRESENTATION must be explained with deep details)
- (HISTOGRAM REPRESENTATION must be explained with deep details)
- Screenshots with the execution of the program.