

A hand is shown in the upper right corner, reaching down to assemble a structure using various colored LEGO bricks. The bricks are scattered on a white surface, with some already partially assembled into a structure. The colors include yellow, green, red, and black. The background is slightly blurred, showing more bricks and a hand.

Programming for DA

Joseba Carricas
jcarricasga@external.unav.es

Working with data files

Where can I find good data?

How can I work with raw data?

"Studying the museum activity around the world"

Group 1: US qualify market & revenue

<https://www.kaggle.com/imls/museum-directory>

Group 2: EU qualify tipology.

Group 3: EU qualify financing, revenue and employees.

Group 4: EU qualify attendance evolution by time and distribution by country.

<https://www.egmus.eu/es/statistics/>

"Studying the museum activity around the world"

Group 5: SP qualify attendants:

<https://estadisticas.mecd.gob.es/CulturaDynPx/culturabase/index.htm?type=pcaxis&path=/t11/p11/a2018/c07/&file=pcaxis>

Group 6: SP qualify tipology:

<https://estadisticas.mecd.gob.es/CulturaDynPx/culturabase/index.htm?type=pcaxis&path=/t11/p11/a2018/c02/&file=pcaxis>

Group 7: SP qualify services:

<https://estadisticas.mecd.gob.es/CulturaDynPx/culturabase/index.htm?type=pcaxis&path=/t11/p11/a2018/c04/&file=pcaxis>

"Studying the museum activity around the world"

Group 8: SP qualify activities:

<https://estadisticas.mecd.gob.es/CulturaDynPx/culturabase/index.htm?type=pcaxis&path=/t11/p11/a2018/c06/&file=pcaxis>

Group 9: SP qualify financing and employees:

<https://estadisticas.mecd.gob.es/CulturaDynPx/culturabase/index.htm?type=pcaxis&path=/t11/p11/a2018/c09/&file=pcaxis>

<https://estadisticas.mecd.gob.es/CulturaDynPx/culturabase/index.htm?type=pcaxis&path=/t11/p11/a2018/c08/&file=pcaxis>

Data origins

- Open Data initiatives:
 - <https://www.kaggle.com/>
 - <https://datos.gob.es/>
 - <https://data.europa.eu/>
 - ... (lots of public and private initiatives)
- Google Cloud Datasets.
- System exports.

Data formats

- Concept of structured data.
- Common formats: XML, JSON, CSV, XLS
- Let's play a bit with data formats
 - Download US_museum_directory_short.
 - Open each file and see how the data is structured.

... any problem? 

Working with CSV files

Get control of your files

- How to download and save a CSV file.
- How to open a CSV file.
Excel vs Notepad++ (<https://macromates.com/> o
<https://brackets.io/>)
- Where is my file?

Working with CSV files

Understand data structure

- Comma vs Semicolon.
- Quotation for text delimiter.
- Number format.

Character encoding (charset)

- UTF-8.
- ISO-8859-1, OEM850.

Install a good text editor

Windows:

<https://notepad-plus-plus.org/downloads/>

Mac:

<https://brackets.io/>

or

<https://macromates.com/>

Exercise 1

Goal: first touch with CSV files.

1. Open “US_museum_directory_short.csv” with the text editor.

Have a look at its structure.

2. Read about CSV files:

<https://www.spreadsheetmadeeasy.com/understanding-csv-files-in-excel/>

<https://www.fundrecs.com/blog/working-with-csv-files>

Go through some videos

Automate the Boring Stuff with Python:

https://www.youtube.com/watch?v=&list=PL0-84-yl1fUnRuXGFe_F7qSH1LEnn9LkW

Lesson 2: expressions, data types, variables

Lesson 3: your first program

A hand is shown in the upper right corner, reaching towards a collection of LEGO bricks. The bricks are primarily yellow, with some green, red, and black ones scattered around. Some bricks are already assembled into small structures, while others are loose. The background is a plain white surface.

Programming for DA

Joseba Carricas
jcarricasga@external.unav.es

Install Python

<https://www.python.org/downloads/>

Get in touch with the working environment...

... and type your first code ;-)

Exercise 2

Goal: first touch with Python and working environment.

1. Create a program that prints “Hello world!”.