

Data is key for the training a Machine Learning model.

Injection of data is the first piece of the construction of a data lake. Let's play with some APIs to retrieve and transform data.

Resources

Read or watch:

Requests package

Learning Objectives

At the end of this project, you are expected to be able to explain to anyone, without the help of Google:

General

- How to use the Python package requests
- How to make HTTP GET request
- How to handle rate limit
- How to handle pagination
- How to fetch JSON resources
- How to manipulate data from an external service

Requirements

General

- Allowed editors: vi, vim, emacs
- All your files will be interpreted/compiled on Ubuntu 16.04 LTS using python3 (version 3.5)
- All your files should end with a new line
- The first line of all your files should be exactly `#!/usr/bin/env python3`
- A README.md file, at the root of the folder of the project, is mandatory
- Your code should use the pycodestyle style (version 2.4)
- All your modules should have documentation (python3 -c 'print(__import__("my_module").__doc__)')
- All your classes should have documentation (python3 -c 'print(__import__("my_module").MyClass.__doc__)')
- All your functions (inside and outside a class) should have documentation (python3 -c 'print(__import__("my_module").my_function.__doc__)' and python3 -c 'print(__import__("my_module").MyClass.my_function.__doc__)')
- All your files must be executable