

## EAR212 - Web data

### Présentation

#### Prérequis

Pas de pré-requis.

Le cours nécessite d'avoir un ordinateur portable sur lequel Python peut être installé (gratuit).

### Objectifs pédagogiques

Ce cours propose une introduction au langage de programmation Python.

L'objectif est d'utiliser ce langage de programmation pour :

- traiter, fusionner et analyser des bases de données dans différents formats (csv, excel, txt...)
- collecter des données sur Internet via du web scrapping. Le web scraping est une technique d'extraction du contenu de sites Web, via un script ou un programme
- analyser des réseaux sociaux.

L'analyse des réseaux sociaux est une approche basée sur l'étude de la théorie des graphes. La théorie des réseaux sociaux conçoit les relations sociales en termes de nœuds et liens. Les nœuds sont habituellement les acteurs sociaux dans le réseau mais ils peuvent aussi représenter des institutions, et les liens sont les interactions ou des relations entre ces nœuds.

### Compétences

A l'issue du cours l'étudiant devra être capable de :

- travailler avec des bases de données dans des formats différents
- collecter des informations sur des pages web
- analyser des réseaux sociaux: représenter le graphe d'une communauté, définir quels sont les éléments centraux dans un réseau...

### Compétences

A l'issue du cours l'étudiant devra être capable de :

- travailler avec des bases de données dans des formats différents
- collecter des informations sur des pages web
- analyser des réseaux sociaux: représenter le graphe d'une communauté, définir quels sont les éléments centraux dans un réseau...

## Programme

### Contenu

#### Partie 1.

Introduction à Python

#### Partie 2.

Applications (en utilisant Python)

- fusionner des bases de données
- analyser des données
- scraper des pages web (i.e., collecter des informations sur internet)
- collecter des informations à partir de l'API twitter
- représenter et analyser le graphe d'une communauté

### Modalités de validation

- Contrôle continu
- Projet(s)

Valide le 02-07-2022



**Code : EAR212**

Unité d'enseignement de type cours

4 crédits

Volume horaire de référence (+/- 10%) : **40 heures**

**Responsabilité nationale :**

EPN09 - Economie Finance Assurance Banque (EFAB) / 1

**Contact national :**

EPN09 - EFAB

40 rue des jeuneurs

Bureau C 256

75002 Paris

01 40 27 23 66

Virginie Moreau

[virginie.moreau@lecnam.net](mailto:virginie.moreau@lecnam.net)

# Bibliographie

---

Titre	Auteur(s)
Apprenez à programmer en Python	Vincent Le Goff
Web Scraping with Python	Ryan Mitchell