

# Big Data - Python pour l'analyse de données

Le langage Python dispose d'un écosystème scientifique, permettant entre autres, les traitements statistiques : de la construction de modèles d'analyse, à leur évaluation jusqu'à leur représentation. Ce cours vous permet d'analyser des données d'horizon divers avec les bibliothèques Python.

**Formateur·rice(s) :**

## Objectifs visés

- Comprendre le principe de la modélisation statistique
- Savoir utiliser les principaux outils de traitement et d'analyse de données pour Python
- Savoir appliquer les pratiques optimales en matière de nettoyage et de préparation des données avant l'analyse
- Choisir entre la régression et la classification en fonction du type de données
- Apprendre à mettre en place un modèle d'apprentissage simple
- Être capable d'extraire des données d'un fichier

## Méthodes d'évaluation

- Cas pratique
- Quiz

## Déroulé pédagogique

### Journée 1

- Introduction, rappels Python
- Pause
- Socle scientifique Python

### Journée 2

- Introduction Modélisation/ Modèles supervisés
- Pause Midi
- Modèles supervisés

### Journée 3

- Procédure d'évaluation des modèles
- Pause Midi
- Modèles non supervisés

## Informations complémentaires

**Durée** : 22 heures de formation sur 3 jours

**Pré-requis** :

- Maîtrise de la programmation Python. Connaissances de base en statistiques ou avoir suivi le stage "Statistiques, maîtriser les fondamentaux"

**Public concerné** : Développeurs en Python, responsables infocentre, développeurs de logiciels, programmeurs, data analysts, data scientists.

**Tarif** 2 190,00 € HT (Exonération de TVA - Art.261.4.4° a du CGI)