

## Programmation Python orientée Data science

### Objectifs pédagogiques :

A l'issue de la formation, les participants se seront familiarisés avec les différents paradigmes du Machine Learning et seront capables d'utiliser Python et des bibliothèques pertinentes pour manipuler les algorithmes classiques du domaine pour produire des analyses éclairées de leurs données.

#### Durée :

4 jours (28h)

#### Prérequis :

Posséder des connaissances de base en algorithmique est conseillé, mais aussi en Python ou dans un autre langage de programmation ainsi que des compétences en mathématiques statistiques

#### Publics concernés :

Doctorants, ingénieurs, chercheurs, data analystes, marketing

#### Tarif HT/jour\*

640 €

\*Tarif non applicable si accords-cadres client ou coaching

## Contenu de la formation

### Introduction

- Tour de table de présentation individuelle
- Identification des attentes personnelles de chaque participant
- Présentation des objectifs de la formation

### Rappels des notions de base de la programmation et présentation de Python

- Notions d'algorithmique et langage de programmation
- Où se situe Python dans le domaine de la programmation ?

### La modularité

- Les fonctions définies par le mot clef def
- Les fonctions lambda
- La bibliothèque standard, les modules et le mot clef import
- Les autres bibliothèques et le système de gestion pip
- Sauvegarder, organiser et appeler son code : les fichiers .py

### Les erreurs

- Principes fondamentaux des exceptions
- Gérer les erreurs simplement avec les blocs try et except
- Gérer un type d'erreur particulier
- Aller plus loin avec else, finally et pass
- Les assertions

## Programmation Python orientée Data science

### Objectifs pédagogiques :

A l'issue de la formation, les participants se seront familiarisés avec les différents paradigmes du Machine Learning et seront capables d'utiliser Python et des bibliothèques pertinentes pour manipuler les algorithmes classiques du domaine pour produire des analyses éclairées de leurs données.

#### Durée :

4 jours (28h)

#### Prérequis :

Posséder des connaissances de base en algorithmique est conseillé, mais aussi en Python ou dans un autre langage de programmation ainsi que des compétences en mathématiques statistiques

#### Publics concernés :

Doctorants, ingénieurs, chercheurs, data analystes, marketing

#### Tarif HT/jour\*

640 €

\*Tarif non applicable si accords-cadres client ou coaching

### Python : les objets et les classes

- Les principes du paradigme de la programmation objet
- Travailler avec des séquences d'éléments de base : les listes et les tuples
- Travailler avec des séquences d'objets : les dictionnaires
- Le module os et la gestion des fichiers
- La portée des variables en Python
- Les conventions de nommage et de gestion
- La création d'une classe
- Les attributs
- Les méthodes de classe et les méthodes statiques

### Héritage simple / héritage multiple

- Principe de l'héritage
- Héritage simple et classe mère
- Héritage multiple et classes parentes

### Les méthodes spéciales

- Les méthodes de conteneur
- Les méthodes mathématiques et de comparaison
- Les décorateurs
- Les méthodes de pickle

### Manipulation de données, format de fichier

- Manipuler différents formats de fichiers
- Accéder aux données
- Manipuler des données massives

## Programmation Python orientée Data science

### Objectifs pédagogiques :

A l'issue de la formation, les participants se seront familiarisés avec les différents paradigmes du Machine Learning et seront capables d'utiliser Python et des bibliothèques pertinentes pour manipuler les algorithmes classiques du domaine pour produire des analyses éclairées de leurs données.

#### Durée :

4 jours (28h)

#### Prérequis :

Posséder des connaissances de base en algorithmique est conseillé, mais aussi en Python ou dans un autre langage de programmation ainsi que des compétences en mathématiques statistiques

#### Publics concernés :

Doctorants, ingénieurs, chercheurs, data analystes, marketing

#### Tarif HT/jour\*

640 €

\*Tarif non applicable si accords-cadres client ou coaching

### Les mathématiques et les statistiques en Python

- Les structures de données en NumPy et Pandas
- L'aléatoire en informatique et le module random
- Algèbre linéaire : manipuler des vecteurs, matrices, tenseurs avec NumPy
- Tests statistiques (test Z, test de Student)
- La descente de gradient

### Initiation à l'apprentissage statistique

- Les différents régimes de l'apprentissage statistique : apprentissage supervisé et non supervisé
- Évaluer un modèle (métriques, matrice de confusion, courbes ROC et AUC, jeux de tests, validation croisée)
- Prétraitement des données (encodage, classes déséquilibrées, valeurs manquantes...)
- Les régressions linéaires et logistiques (avec Scikit-learn), les régressions polynomiales, les régressions régularisées
- Le clustering *K-means* et *k-nearest neighbors*
- Le clustering hiérarchique
- Arbres de décisions et de régression
- *Perceptron* et *Support Vector Machines*
- Introduction aux réseaux de neurones

### Visualisation

- Analyses de composantes (principales, indépendantes ...)
- Tracer des graphiques pertinents avec matplotlib et seaborn

### Clôture de la formation

- Récapitulatif
- Conseils, trucs et astuces
- Fiche d'évaluation, synthèse
- Récupération par les participants des fichiers travaillés et des exemples traités

## Programmation Python orientée Data science

### Objectifs pédagogiques :

A l'issue de la formation, les participants se seront familiarisés avec les différents paradigmes du Machine Learning et seront capables d'utiliser Python et des bibliothèques pertinentes pour manipuler les algorithmes classiques du domaine pour produire des analyses éclairées de leurs données.

#### Durée :

4 jours (28h)

#### Prérequis :

Posséder des connaissances de base en algorithmique est conseillé, mais aussi en Python ou dans un autre langage de programmation ainsi que des compétences en mathématiques statistiques

#### Publics concernés :

Doctorants, ingénieurs, chercheurs, data analystes, marketing

#### Tarif HT/jour\*

640 €

\*Tarif non applicable si accords-cadres client ou coaching

### Les méthodes et critères d'évaluation pédagogique

La constitution des groupes homogènes s'établira à partir d'un outil d'évaluation. L'évaluation permettra d'avoir un premier aperçu du niveau de l'apprenant, de ses connaissances et de ses attentes pour la formation appropriée.

Une approche pédagogique sera réalisée par le formateur avant le début de la formation, afin d'adapter le contenu du programme pour répondre aux attentes des apprenants.

Une attestation est fournie à l'apprenant à l'issue de la formation validant les connaissances acquises lors de la formation.

### Les méthodes pédagogiques

Chaque thème du programme sera accompagné d'ateliers pratiques avec suivi et assistance personnalisée.

Les ateliers pourront être réadaptés en fonction des propres modèles des participants.

### Le suivi et les moyens pédagogiques

Un support de formation sera transmis à chacun des participants, reprenant les principaux thèmes de la formation réalisé sous forme de captures d'écran et d'explications de texte.

Les apprenants repartent à l'issue de la formation avec les fichiers travaillés pendant la formation ainsi que les coordonnées du formateur, ce qui leur permettra d'échanger avec ce dernier et de lui poser toute question relative à la formation suivie sans limitation de durée.

Une feuille d'émargement est signée par les stagiaires (matin et après-midi) chaque jour de la formation, afin d'attester de leur présence.

### Les moyens techniques

Salle équipée avec un poste par personne, un tableau blanc, un paperboard, un accès wifi et un vidéo projecteur.

### Les prérequis nécessaires au passage de la certification

Aucun prérequis pour cette certification.