

Sessions garanties disponibles

Formation Python pour la Data Science

Réf. SF33186

Maîtrisez la programmation Python pour l'analyse de données, la visualisation et le Machine Learning !

3 j Durée	2 400 € HT Inter / participant	4 950 € HT Intra (≤12 pers.)	43 Sessions à venir
--------------	-----------------------------------	---------------------------------	------------------------

Objectifs pédagogiques

- ✓ Comprendre les principes fondamentaux de la programmation Python et son application à la manipulation et à l'analyse des données
- ✓ Appliquer des techniques de nettoyage et de prétraitement des données pour préparer les ensembles de données à l'analyse
- ✓ Créer et interpréter des visualisations de données à l'aide de bibliothèques Python
- ✓ Développer des modèles de Machine Learning basiques pour l'analyse prédictive
- ✓ Évaluer les performances des modèles d'apprentissage automatique

Public visé & prérequis

Public visé Data analysts, Data scientists, Développeurs, Ingénieurs Big Data

Prérequis Compréhension de base des concepts de programmation. Familiarité avec les concepts mathématiques fondamentaux (statistiques et algèbre).

Programme détaillé

Introduction à Python pour la science des données

Les bases de la programmation Python
Introduction à la syntaxe de Python, aux types de données (chaînes, entiers, flottants), aux structures de contrôle (instructions if, boucles), aux fonctions et aux modules.
Vue d'ensemble des environnements de développement Python (Jupyter Notebooks, PyCharm).
Configuration de l'environnement Python
Installation de Python et des bibliothèques essentielles : Pandas, NumPy, Matplotlib, Seaborn, Scikit-learn.
Environnements virtuels et gestion des paquets avec pip.

Activités pratiques :

Mise en place d'un environnement de développement Python.
Écrire des scripts Python de base pour effectuer des manipulations de données simples.

Manipulation de données avec Pandas

Comprendre Pandas pour l'analyse de données
Exploration des objets Series et DataFrame.
Indexation, sélection, filtrage et tri des données.
Fusionner, joindre et concaténer des données.
Techniques de nettoyage des données
Gestion des valeurs manquantes, des données dupliquées et des types de données incorrects.
Traitement des données textuelles et encodage des données catégorielles.

Activités pratiques :

Nettoyage d'un ensemble de données réelles à l'aide de Pandas.
Fusion de plusieurs ensembles de données en un seul DataFrame pour l'analyse.

Visualisation des données

Introduction à Matplotlib et Seaborn

Création de graphiques de base : graphiques linéaires, diagrammes à barres et histogrammes.

Visualisations avancées : diagrammes de dispersion, diagrammes en boîte et cartes thermiques.

Analyse visuelle des données

Utiliser les visualisations pour comprendre les distributions et les relations entre les données.

Personnaliser les graphiques avec des styles, des couleurs et des annotations.

Activités pratiques :

Visualisation des distributions de données et des relations dans un ensemble de données.

Conception de visualisations personnalisées pour communiquer des informations sur les données.

Les bases de l'apprentissage automatique (machine learning)

Apprentissage automatique avec Scikit-learn

Vue d'ensemble de l'apprentissage supervisé et non supervisé.

Modèles de régression et de classification : Régression linéaire, régression logistique, k-voisins les plus proches.

Métriques de régression et classification et techniques d'évaluation : R²-score, erreur moyenne absolue, précision, rappel, validation croisée...

Feature engineering et sélection des modèles

Prétraitement des données pour l'apprentissage automatique.

Sélection et mise au point des modèles de machine learning.

Activités pratiques :

Construction et évaluation d'un modèle de régression linéaire pour prédire des valeurs numériques.

Développement d'un modèle de classification pour catégoriser les points de données.

Machine Learning avancé et meilleures pratiques

Techniques avancées d'apprentissage automatique

Arbres de décision et forêts aléatoires.

Surajustement, sous-ajustement et régularisation du modèle.

Introduction aux réseaux neuronaux et aux concepts d'apprentissage profond.

Gestion de projet en science des données

Meilleures pratiques pour la collaboration dans les projets de science des données.

Considérations éthiques en science des données

Aborder la confidentialité des données, la sécurité et l'utilisation éthique des données.

Comprendre l'impact des données et des modèles biaisés.

Activités pratiques :

Mise en œuvre d'un modèle d'arbre de décision pour améliorer les prédictions.

Prochaines sessions

DATE	LIEU	STATUT
27/10/2026	à distance	Sur confirmation
27/10/2026	Paris	Sur confirmation
27/10/2026	Nantes	Sur confirmation
27/10/2026	Lille	Sur confirmation
25/11/2026	Lyon	Sur confirmation
25/11/2026	à distance	Session garantie

Pourquoi choisir Sparks Formation ?

20 ans d'expertise IT, des formateurs experts et des sessions garanties contre l'annulation.

20 ans
d'expertise

1 068+
formations

11 178
apprenants formés/an

+67
Net Promoter Score

Avis des participants

4,9/5 — 8 avis

Anthony G. — Formation très bien organisée, très claire Alison explique très clairement et est très à l'écoute, avec toujours une approche pratique essentielle. Je n'avais jamais fait de Python et...

Pierre B. — Formation très intéressante, j'ai appris beaucoup de chose. Un grand merci au formateur.

Votre formateur

Camilo R. — Formateur Machine Learning & Google Cloud

Camilo intervient sur des formations portant sur la Data Science et le Machine Learning. Parmi les formations qui lui sont confiées figurent notamment « MLOps » et « PyTorch »...

Ils nous font confiance

Cette année notre filière Data Science & Machine Learning a accompagné plus de 28 entreprises dont :

CEGEDIM LBC FRANCE SNCF - E.SNCF SOLUTIONS CSTB GERS SA ADENTIS ALSTOM BECTON DICKINSON FRANCE SAS CEGEDIM ACTIV CEGEDIM CLOUD +18 autres

Financement & certifications

Cette formation est éligible aux financements OPCO et FNE-Formation. Nos équipes vous accompagnent dans le montage de votre dossier.



Réservez votre place dès maintenant

20 ans d'expertise · 1 000+ formations · NPS +67 · réponse sous 24h ouvrées

0 805 950 800

demande@sparks-formation.com