

Formation Python scientifique

Mise à jour janvier 2025

Inter 3000€ HT/participant

Intra 7500€ HT* groupe de 6 participants

***hors frais de déplacement et personnalisation de programme sur-mesure**

Python est devenu un langage de choix pour les scientifiques, en raison de sa simplicité de mise en œuvre et de la richesse de son écosystème, notamment grâce à ses nombreuses et performantes bibliothèques de calculs numériques bien souvent développées par les scientifiques eux-mêmes.

Python est probablement le seul langage à offrir aux informaticiens un environnement open-source complet dédié à la recherche scientifique, aux sciences de l'ingénieur et aux mathématiques.

Durée: 35.00 heures (5.00 jours)

À QUI S'ADRESSE CETTE FORMATION ?

Profil du participant

- Toute personne travaillant dans des secteurs ayant de forts besoins en calculs et simulations mathématiques
- Informaticiens, scientifiques, chercheurs et ingénieurs travaillant dans les domaines comme les sciences physiques, la chimie, les sciences des matériaux, la génétique, la mécanique des fluides, l'aérospatiale, les statistiques, le monde de la finance, etc...

Prérequis

- Connaître l'algorithmie
- Avoir une première expérience du langage Python
- Connaître quelques rudiments mathématiques, trigonométrie (sinus, cosinus) et algèbre linéaire (calcul matriciel)

OBJECTIFS PÉDAGOGIQUES

1. Se repérer dans l'écosystème Python Scientifique
2. Créer et utiliser un notebook Jupyter
3. Réaliser des opérations sur des tableaux numériques
4. Manipuler des données issues de fichiers tabulaires
5. Calculer des statistiques à partir de séries temporelles
6. Prédire une valeur par interpolation
7. Créer des graphiques 2D statiques
8. Créer des graphiques 2D interactifs
9. Créer des cartes à partir de données géographiques
10. Manipuler des images
11. Classifier automatiquement un jeu de données
12. Mettre en place une méthode d'apprentissage supervisé
13. Paralléliser une boucle for
14. Mesurer les performances d'un programme

CONTENU (PROGRESSION PÉDAGOGIQUE)

Poser les fondations

- Vu d'ensemble de l'écosystème scientifique
- Présentation des différentes bibliothèques
- Limites du calcul numérique :
 - Les problèmes liés au stockage en mémoire
 - Les problèmes d'erreurs numériques
- Anaconda
- Notebooks Jupyter
- Conda

Les tableaux

- Limites des pistes Python
- Numpy (permet de manipuler des vecteurs et matrices à plusieurs dimensions. Elle offre des types de données avancés pouvant être gérés très simplement. Implémentée en langage C, elle mixe le meilleur des 2 mondes: la souplesse du langage Python et les performances du C)
- Pandas (permet de manipuler des matrices dont chaque colonne peut être d'un type différent et offrant de nombreuses et puissantes fonctionnalités de manipulation, réduction, recherche et statistiques, permet d'analyser avec efficacité des structures de données complexes)
- Xarray (permet de travailler avec des tableaux multidimensionnels et multi-types. Elle propose des fonctions d'analyse avancées qui s'appuient sur et étendent les forces de base de NumPy et Pandas.)

Gestion des données temporelles

- Fonctionnalités de la bibliothèque Pandas (manipulation de données temporelles)
- Calcul de statistiques temporelles
- Interpolation de données manquantes
- Découverte des fonctionnalités Spicy (ensemble d'algorithmes mathématiques et de fonctions utilitaires qui s'appuient sur la bibliothèque Numpy)

Visualisation des données

- Présentation du panorama des bibliothèques actuelles et représentation graphique en Python
- Graphiques statiques ou interactifs, univariés ou multivariés

Traitement d'images

- Traitements classiques d'analyse d'images avec la bibliothèque Scikit-image :
 - La binarisation
 - L'application de filtres de convolution
 - Segmentation d'images

Optimisation et calculs parallèles

- Parallélisation en Python (multi-threading et multi-processing)
- Bibliothèque Dask (permet d'assurer le passage à l'échelle des analyses de jeux de données trop volumineux pour être chargés d'un bloc en mémoire par Numpy ou Pandas et simplifie la parallélisation et la répartition sur des clusters des calculs scientifiques.)

Machine learning

- Concept de base du machine learning
- Bibliothèque Scikit-Learn

En option

- Dans le cadre d'une formation intra-entreprise, un cas pratique ou jeu de données représentatif peut être fourni par l'entreprise afin que le formateur l'intègre dans le cursus.

ORGANISATION

Formateur

Formation assurée par un expert-formateur Python

Moyens pédagogiques et techniques

- Accueil des stagiaires dans une salle dédiée à la formation
- Documents supports de formation partagés
- Exposés théoriques
- Étude de cas concrets
- Quiz en ligne
- Mise à disposition en ligne de documents supports à la suite de la formation

Dispositif de suivi de l'exécution de l'évaluation des résultats de la formation

- Feuilles de présence dématérialisées
- Questions orales ou écrites (QCM)
- Mises en situation
- Formulaire d'évaluation de la formation
- Durant la formation : essentiellement axé sur des exercices pratiques et de mise en situation sous forme de travaux pratiques corrigés avec l'expert-formateur.

Délais d'accès

La convocation est envoyée 7 jours avant le début de la formation après réception du bon de commande signé.

Coordonnées de l'équipe pédagogique

- Responsable formation, handicap et votre formateur : Cécile Chardonneau formation@makina-corpus.com
- Suivi facturation : Nathalie Carles Salmon administration@makina-corpus.com