



ORGANISME DE FORMATION AUX TECHNOLOGIES ET METIERS DE L'INFORMATIQUE

Formation Python avancé pour data-scientists

N° ACTIVITÉ : 11 92 18558 92

TÉLÉPHONE : 01 85 77 07 07

E-MAIL : inscription@hubformation.com

Objectifs

- | utiliser les principaux outils de traitement et d'analyse des données
- | extraire des données d'un fichier et les manipuler
- | mettre en place un modèle d'apprentissage simple

Public

- | développeurs en Python, Développeurs de logiciels, programmeurs, Data analysts, Data scientists

Prérequis

- | Bonne connaissance de la programmation Python.

Programme de la formation

Positionnement Python dans l'analyse de données

- | Besoins des data-scientists : calculs, analyse d'images, machine learning, interface avec les bases de données
- | Apports de python : grande variété d'outils, expertise dans le domaine du calcul scientifique
- | Tour d'horizon des outils:
- | pandas, pyarrow, agate, bokeh, scikit-learn, pybrain, tensorflow, keras, mxnet, caffe, Pytorch

Calculs et graphiques

- | NumPy : Base du calcul sur des tableaux
- | SciPy : Scientific Tools for Python, couche scientifique
- | Manipulation de tableaux, fonctions mathématiques.
- | Représentation graphique avec basemap et matplotlib.
- | Atelier : Mise en oeuvre de SciPy/NumPy : manipulation d'images, détection de contours

Être capable d'extraire des données d'un fichier

- | Pandas : manipulation de tables de données. Notion de dataframe.
- | Manipulation de données relationnelles
- | Tableaux avec Pandas: indexation, opérations, algèbre relationnelle
- | Stockage dans des fichiers: CSV, JSON
- | Comparaison et performances Pandas / pyarrow / NumPy
- | Atelier : construction d'ETL de base entre json et csvkaggle.com,

Comprendre les mécanismes d'interconnexion aux bases de données

- | Définitions : pilotes, connexions, curseurs, CRUD, transactions
- | Les pilotes : postgresql, mysql, mariadb, ... Présentation de sql-alchemy
- | Opérations : gestion du curseur, chargement de données, insertion et modification d'enregistrements
- | Atelier : mise en oeuvre avec postgresql. Construction d'ETL SQL/json

Comprendre les principaux outils de traitement et d'analyse de données pour Python

- | Présentation des outils d'apprentissage Python : scikit-learn, pybrain,

Référence	DSL010
Durée	4 jours (28h)
Tarif	3 230 €HT
Repas	92 €HT(en option)

SESSIONS PROGRAMMÉES

A DISTANCE (FRA)

- du 30 juil. au 2 août 2025
- du 29 sept. au 2 octobre 2025

PARIS

- du 30 juil. au 2 août 2025
- du 29 sept. au 2 octobre 2025

[VOIR TOUTES LES DATES](#)

TensorFlow/keras, mxnet, caffe
| Atelier : mise en oeuvre de scikit-learn

Créer des sélections et des classements dans de grands volumes de données pour dégager des tendances

| Présentation de pyspark
| Machine learning et deep learning : les solutions Python,
| TensorFlow : principe de fonctionnement, plateformes supportées, distribution

Sites de références data-sciences

| Ressources d'apprentissage, datasets, modèles de données pré-entraînés, etc ..
| Présentation de : kaggle.com, data-puzzles.com, huggingface.co

Optimisation des développements

| Tour d'horizon des outils actuels et futurs:
| Jupyter notebook, Aide à la vérification de code, respect des recommandations PEP8 :
| exemples avec pydecoestyle, Pylint, Black
| Analyse et production de code informatique avec une IA.
| Génération de code avec OpenAI : démonstrations ChatGPT, apports, bonnes pratiques.
| Atelier : utilisation de la génération de code et de snippets Python avec ChatGPT

Méthode pédagogique

Chaque participant travaille sur un poste informatique qui lui est dédié. Un support de cours lui est remis soit en début soit en fin de cours. La théorie est complétée par des cas pratiques ou exercices corrigés et discutés avec le formateur. Le formateur projette une présentation pour animer la formation et reste disponible pour répondre à toutes les questions.

Méthode d'évaluation

Tout au long de la formation, les exercices et mises en situation permettent de valider et contrôler les acquis du stagiaire. En fin de formation, le stagiaire complète un QCM d'auto-évaluation.

Suivre cette formation à distance

Voici les prérequis techniques pour pouvoir suivre le cours à distance :

- | Un ordinateur avec webcam, micro, haut-parleur et un navigateur (de préférence Chrome ou Firefox). Un casque n'est pas nécessaire suivant l'environnement.
- | Une connexion Internet de type ADSL ou supérieure. Attention, une connexion Internet ne permettant pas, par exemple, de recevoir la télévision par Internet, ne sera pas suffisante, cela engendrera des déconnexions intempestives du stagiaire et dérangera toute la classe.
- | Privilégier une connexion filaire plutôt que le Wifi.
- | Avoir accès au poste depuis lequel vous suivrez le cours à distance au moins 2 jours avant la formation pour effectuer les tests de connexion préalables.
- | Votre numéro de téléphone portable (pour l'envoi du mot de passe d'accès aux supports de cours et pour une messagerie instantanée autre que celle intégrée à la classe virtuelle).
- | Selon la formation, une configuration spécifique de votre machine peut être attendue, merci de nous contacter.
- | Pour les formations incluant le passage d'une certification la dernière journée, un voucher vous est fourni pour passer l'examen en ligne.
- | Pour les formations logiciel (Adobe, Microsoft Office...), il est nécessaire d'avoir le logiciel installé sur votre machine, nous ne fournissons pas de licence ou de version test.
- | Horaires identiques au présentiel.

Accessibilité



Les sessions de formation se déroulent sur des sites différents selon les villes ou les dates, merci de nous contacter pour vérifier l'accessibilité aux personnes à mobilité réduite.
Pour tout besoin spécifique (vue, audition...), veuillez nous contacter au 01 85 77 07 07.