Formation Pandas Avancé

Objectifs : Exploiter la librairie Pandas du langage Python pour traiter vos problématiques et projets en Data Science

Compétences visées : - Maîtriser la librairie Pandas pour l'analyse de données
- Connaître les subtilités des groupbys
- Savoir manipuler les tables pivots et les tableaux croisés
- Avoir des notions sur l'accélération des calculs avec Pandas
- Connaître les bonnes pratiques en Data Science

Durée : 3 jours (21 heures)

Public : Décideurs, statisticiens, data scientists, chargés d'études en charge de l'exploitation des données de l'entreprise

Pré-requis : Pour suivre ce stage dans de bonnes conditions, il est recommandé d'avoir suivi en amont la formation Python – Bases et introduction aux librairies scientifiques

Méthode pédagogique : Pédagogie active mêlant exposés, exercices et applications pratiques dans le logiciel Python.

Tarifs :
- Présentiel : 1950 € HT
- Distanciel : 1800 € HT
(-10% pour 2 inscrits, -20% dès 3 inscrits)

Option(s) :
- Forfait déjeuners : 75 € HT

Nos prochaines sessions

Distance
du 14 au 16 décembre 2026
du 14 au 16 juin 2027
du 13 au 15 décembre 2027

Lyon
du 14 au 16 avril 2027
du 18 au 20 octobre 2027

Paris
du 10 au 12 mai 2027
du 15 au 17 novembre 2027

Toulouse
du 15 au 17 mars 2027

du 13 au 15 septembre 2027

Programme :

- Tour d'horizon de la librairie Pandas

Vous utilisez Pandas pour vos analyses de données ou vos projets de Data Science, mais vous n'en maîtrisez pas tous les aspects ? Ce tour d'horizon vous permettra d'identifier toutes les possibilités qu'offre cette librairie. Par un exercice guidé, vous pourrez revoir les notions de bases pour l'utilisation de Pandas, des notions de visualisation de données avec Matplotlib, Pandas et Seaborn

- Rappels sur les notions de bases de Pandas
- Lecture de fichiers de données (csv, excel, SQL, parquet)
- Description du jeu de données et analyse statistique simple
- Implémenter des analyses et des visualisations différentes en fonction du type de données
- Gestion des données manquantes
- Manipulation de dates pour les Time Series
- Gestion des chaînes de caractères
- Mise en place des bonnes pratiques en Data Science

- Maîtriser les subtilités des groupbys

Les groupbys peuvent vous permettre de mieux appréhender les modalités dans votre jeu de données.

- Groupby à simple indice avec les fonctions d'agrégations classiques
- Personnalisation des fonctions d'agrégations
- Groupby à multiples indices
- Différence entre les fonctions apply et transform
- Rappels sur les fonctions anonymes

- Tables pivots et tableaux croisés

Vous verrez par le biais de nombreux exemples et exercices, l'intérêt des tables pivots et des tableaux croisés pour une représentation différente des jeux de données.

- Fonctions d'agrégation et tables pivots
- Matrice de contingence
- Tableaux croisés

- Jointure de tables

Pour joindre des tables, il est possible de procéder par indice ou par colonne.

- Notions d'axes
- Concaténation
- Merge selon une ou plusieurs clés
- Jointure par rapport aux indices

- Atelier de mise en pratique sur une journée

Toutes les notions vues précédemment seront mises en pratique lors d'une journée de travaux pratiques sur un problème d'analyse de données complet mettant en œuvre les bonnes pratiques à utiliser en Data Science.

- Accélération du calcul avec Pandas

Vous verrez quelles sont les librairies qui vous permettent de faire du multiprocessing avec Pandas

- Boucler sur les lignes et les colonnes
- Revenir aux basiques avec NumPy
- Exemples avec la librairie Modin
- Exemples avec la librairie Numba

Modalités d'évaluation : Un formulaire d'auto-évaluation proposé en amont de la formation nous permettra d'évaluer votre niveau et de recueillir vos attentes. Ce même formulaire soumis en aval de la formation fournira une appréciation de votre progression.
Des exercices pratiques seront proposés à la fin de chaque séquence pédagogique pour l'évaluation des acquis.
En fin de formation, vous serez amené(e) à renseigner un questionnaire d'évaluation à chaud.
Une attestation de formation vous sera adressée à l'issue de la session.
Trois mois après votre formation, vous recevrez par email un formulaire d'évaluation à froid sur l'utilisation des acquis de la formation.

Qualification du formateur : Les formations DATA VALUE sont animées par des intervenants experts, alliant expérience opérationnelle et compétences pédagogiques reconnues. Ils disposent de plus de 10 ans d'expérience professionnelle et assurent une veille constante sur les pratiques, outils et normes de leur domaine afin de garantir des contenus actualisés et adaptés aux besoins des stagiaires.

Délais d'accès et modalités de formation : Les inscriptions peuvent être effectuées jusqu'à 48 heures avant le début de la formation.
Vous pourrez trouver des informations plus détaillées concernant les modalités de nos formations à cette adresse : https://www.datavalue.fr/modalites-acces-formation.

Accessibilité : Vous souhaitez suivre notre formation Pandas Avancé et êtes en situation de handicap ? Merci de nous contacter afin que nous puissions envisager les adaptations nécessaires et vous garantir de bonnes conditions d'apprentissage

Date de dernière modification : 30 octobre 2026