

Calcul sur des tableaux avec Pandas

Référence : PYCB055

Durée : 1 jour

Certification : **Aucune**

CONNAISSANCES PREALABLES

- Connaissance de base de Python, notion de base de données.

PROFIL DES STAGIAIRES

- Chefs de projet, data-scientists, souhaitant traiter des données avec pandas.

OBJECTIFS

- Comprendre les manipulations de données dans les bibliothèques spécialisées Python. • Savoir traiter des tableaux de données avec Pandas.

CERTIFICATION PREPAREE

Aucune

METHODES PEDAGOGIQUES

- Mise à disposition d'un poste de travail par stagiaire
- Remise d'une documentation pédagogique papier ou numérique pendant le stage
- La formation est constituée d'apports théoriques, d'exercices pratiques, de réflexions et de retours d'expérience
- Le suivi de cette formation donne lieu à la signature d'une feuille d'émargement

FORMATEUR

Consultant-Formateur expert Bigdata

METHODE D'EVALUATION DES ACQUIS

- Auto-évaluation des acquis par le stagiaire via un questionnaire
- Attestation de fin de stage adressée avec la facture

CONTENU DU COURS

Introduction

- Historique
- Pourquoi Pandas ? Comparaison avec les autres bibliothèques
- Tour d'horizon des fonctionnalités
- Utilité Pandas dans l'écosystème Big Data

Manipulations de données

- Syntaxe
- Types de données : Serie, DataFrame, Panel
- Construction d'un DataFrame
- Indexation des données
- Opérations élémentaires : arithmétique, tris, affichage

Calculs avancés

- Algèbre relationnelle
- Modification de la structure d'un DataFrame : ajout de colonne, renommage d'un champ
- Tables pivot, croisement de données

Mise en oeuvre pratique

- Lecture et écriture de fichiers textes
- Traitement des données manquantes
- Travail Pratique : extraction, traitement et analyse de données CSV