







Formation : Introduction à la Data Science

-  **Dates** : 6–7 octobre 2025
-  **Horaires** : 9h00–12h30 / 13h30–17h00
-  **Lieu** : À distance (visioconférence Zoom + plateforme d'accompagnement)
-  **Durée** : 2 jours – 14 heures de formation
-  **Participants** : 20 maximum
-  **Prix** : 1 200 € HT

Objectifs pédagogiques

Cette formation a pour but de permettre aux participants de :

- Comprendre les fondements et les enjeux de la Data Science.
- Prendre en main l'écosystème Python (Anaconda, Jupyter Notebook).
- Manipuler et analyser des jeux de données à l'aide de **Pandas**.
- Visualiser des données et produire des graphiques utiles.
- Appliquer les bases du **Machine Learning** sur un mini-projet.

Public concerné

- Débutants en Data Science ou Python souhaitant se professionnaliser.
- Analystes, chefs de projet, ingénieurs, profils métier souhaitant acquérir une compétence technique en data.
- Enseignants ou chercheurs souhaitant intégrer la Data Science à leur pratique.

Prérequis

- Bases en programmation (Python de préférence).
- Goût pour les chiffres, la logique et l'analyse.

Matériel requis

- Ordinateur personnel avec webcam et micro.
- Connexion Internet stable.
- Installation de Python (Anaconda) et Jupyter. Un guide d'installation est fourni à l'inscription.

Méthodologie pédagogique

- Formation en **classe virtuelle Zoom interactive**.
- Alternance entre **cours en direct, démonstrations et exercices pratiques**.
- Mini-projet final en binôme ou individuel.
- Plateforme en ligne avec supports, notebooks et ressources complémentaires.

Programme détaillé avec répartition horaire

Horaire	
Jour 1 – Fondamentaux de la Data Science & Python	
09h00 – 10h00	Module 1 : Panorama de la Data Science (définitions, enjeux, cas d'usage)
10h00 – 11h00	Module 2 : Environnement de travail (Anaconda, Jupyter, bibliothèques)
11h00 – 12h30	Module 3 : Python pour la Data Science (structures, listes, dictionnaires, fonctions)
13h30 – 15h30	Module 4 : Manipulation de données avec Pandas (import, nettoyage, transformations)
15h30 – 17h00	Module 5 : Analyse exploratoire des données (statistiques, visualisation avec Matplotlib/Seaborn)
Horaire	
Jour 2 – Statistiques, Machine Learning & Projet	
09h00 – 10h30	Module 6 : Statistiques appliquées (corrélations, tests d'hypothèses)
10h30 – 12h30	Module 7 : Introduction au Machine Learning (régression, k-means, modèles de base)
13h30 – 16h00	Module 8 : Mini-projet (analyse d'un dataset réel, clustering, interprétation)
16h00 – 17h00	Module 9 : Conclusion & Perspectives (restitution, ressources avancées, recommandations)

En résumé

Élément	Détail
Durée	14 heures (2 jours)
Format	À distance (Zoom + supports partagés)
Niveau	Initiation → Intermédiaire → Avancé
Supports fournis	Notebooks Python, datasets, tutoriels
Accompagnement	Formateur en direct + messagerie post-formation