





## Programme de la formation

## Statistiques et optimisation des données

-  En groupe sur site
-  16 heures
-  Salarié en poste · De 3 à 6 participants
-  Sur devis

## Description courte

L'objectif de cette formation est de **mieux comprendre** les **données** manipulées au **quotidien**. Pour atteindre cet objectif, plusieurs thèmes sont abordés durant la formation : **analyse, visualisation et nettoyage** des données avec le langage **Python** et transition d'un tableur (Excel) vers le langage Python.

## Profil des participants

Personnes manipulant des données avec des **tableurs** (type Excel), ingénieurs

## Prérequis

Connaissances de **base en programmation en Python**, utilisation des fonctionnalités de base des **Jupyter Notebooks**.

## Objectifs

Cette action a pour objectif de :

- Présenter les méthodes statistiques d'**analyse de données**
- Présenter différentes méthodes de **visualisations de données**
- Présenter l'optimisation de la traçabilité du processus de **nettoyage de données**
- Présenter l'optimisation de la **manipulation de données** dépassant les capacités d'un tableur

Tous les exemples seront illustrés à l'aide du langage **Python** et des bibliothèques de **data sciences** associées (**Numpy, Pandas, Matplotlib, Seaborn, Plotly, Scikit-learn...**)

## Aptitudes et compétences visées / attestées

Les **compétences professionnelles** visées à l'issue de cette formation sont de :

- Amélioration des informations extraites des données (**analyse et visualisation**)
- Amélioration du processus de **traitement des données** atteignant les **limites des capacités des tableurs**

## Contenu

- **Qualité et traçabilité de l'analyse de données**

Méthodes de nettoyage des données

Traçabilité des méthodes de nettoyage de données

Bonnes pratiques d'écriture et de construction du code Python

- **Statistiques pour l'analyse de données**

Analyse univariée : centralité et dispersion des données

Analyse bivariée : corrélation et covariance

Analyse multivariée : analyse en composantes principales

- **Visualisation et présentation des données**

Visualisations de données et leurs cas d'utilisation

Formats de sortie pour les visualisations de données

- **Automatisation des processus d'ingestion et de manipulation de données**

Remplacement des manipulations de données dans un tableur par du code

Étude de la collecte de données par du code

## Pédagogie et organisation

Pédagogie **active** avec les données manipulées au quotidien (à préciser en amont de la formation)

## Positionnement et Modalités d'évaluation des acquis

Quiz et **activités** de mise en application sur les données

## Référent handicap

**PERSONNES EN SITUATION DE HANDICAP** : (*mise en place de compensation...*)

**Contact** : Référent handicap 02 51 84 95 55 / 06 28 70 45 28 / [nadiahadjeri@cadresenmission.com](mailto:nadiahadjeri@cadresenmission.com)

## Intervenant

Spécialités du consultant formateur : Pierre-Loïc Bayart

- Formateur école de l'intelligence artificielle Microsoft – Simplon Lyon
- **Intelligence Artificielle** : Certification Azure AI100/102 – Azure AI Engineer Associate
- **Science des données** : Certification Azure DP100 – Azure Data Scientist Associate
- **Pédagogie** : Microsoft Certified Trainer – 2021/2022/2023