

# Formation Python pour l'analyse de données

## Description

Vous êtes data analyste ou développeur souhaitant découvrir ou approfondir Python pour le traitement de données? Cette formation Python est idéale pour acquérir les bases et pratiques avancées en analyse de données. Trois jours intensifs pour acquérir les bases de l'utilisation du langage python pour des applications en analyse de données. Après une introduction au langage python (ou un rappel si vous avez déjà utilisé python), vous apprendrez à maîtriser les bibliothèques spécialisées en analyse de données et data visualisation. Basée sur de nombreux exercices pratiques, cette formation vous permet d'être immédiatement opérationnel pour le traitement de vos données grâce à python.

## Durée

3.00 jours (21.00 heures)

## Profils des stagiaires et prérequis

### Profils :

- Data analyst désirant découvrir python en analyse de données
- Manager désirant comprendre les spécificité de python
- Développeur spécialisé dans d'autres langages

### Prérequis :

- Connaissances de base en traitement de données (statistique et tables de données)
- Pas de prérequis en python

## Objectifs pédagogiques

- Maîtriser les bases de python
- Etre capable d'analyser des données avec python
- Etre capable d'automatiser le traitement de données avec python

## Programme détaillé

- Python aujourd'hui : bases et bonnes pratiques
  - Introduction au langage python
  - Pour qui? Pour quoi faire? Comment?
  - Les consoles et le débogage en python
  - Le langage (la syntaxe, les structures (string, listes, dictionnaires...), les opérateurs et les fonctions)
  - Construction d'un outil de calcul de statistiques simples
  - Les classes, les objets et la création de bibliothèques
  - Python pour le traitement des données
  - Le processus de développement d'un outil en python pour la data science
- Manipulation et préparation des données avec pandas
  - Importation de données CSV, Excel
  - Fichiers texte

Stat4decision

37-39 avenue Ledru Rollin Paris 75012

Tel. +33172254082 | E-mail : [info@stat4decision.com](mailto:info@stat4decision.com) | [www.stat4decision.com](http://www.stat4decision.com)

Numéro SIRET: 81048985600015 | Numéro de déclaration d'activité: 11755352275 (auprès du préfet de région de: Ile de France)

Organisme de formation certifié Qualiopi pour ses actions de formation

- Bases de données (SQL – principes, exemples)
- Données issues du web (API, scraping léger)
- Manipulation de données avec pandas
- DataFrame : concepts fondamentaux
- Sélection, filtrage, tri
- Création et transformation de variables
- Jointures et concaténations
- Gestion des données manquantes
- Préparation des données pour l'analyse
- Nettoyage et contrôles de qualité
- Typage des variables
- Variables catégorielles
- Travail sur les dates et les données textuelles
- Agrégations (groupby, pivots)
- Automatisation des traitements
- Fonctions appliquées aux DataFrames
- Chaînage des transformations
- Atelier fil rouge : préparation complète d'un jeu de données
- Visualisation et analyse exploratoire
  - Principes de la data visualisation
  - Pourquoi visualiser?
  - Erreurs classiques
  - Choisir le bon graphique pour la bonne question
  - Visualisation avec matplotlib & seaborn
  - Histogrammes, boxplots, scatter plots
  - Comparaison de groupes
  - Visualisations multivariées
  - Personnalisation minimale mais lisible
  - Analyse exploratoire (EDA)
  - Lire une distribution
  - Déetecter des anomalies
  - Comprendre des relations entre variables
  - Interpréter sans sur-interpréter
  - Visualisation interactive (ou dashboards légers)
  - Introduction aux visualisations interactives
  - Plotly
  - Streamlit
  - Atelier : exploration visuelle complète d'un jeu de données

## Organisation de la formation

### Moyens pédagogiques et techniques

- JupyterHub dédié à la formation

### Dispositif de suivi de l'exécution et d'évaluation des objectifs de la formation

- Quiz et QCM à la fin de chaque module pour valider l'acquisition des connaissances théoriques.
- Mini-projets guidés (manipulation de DataFrames, visualisation)

Stat4decision

37-39 avenue Ledru Rollin Paris 75012

Tel. +33172254082 | E-mail : [info@stat4decision.com](mailto:info@stat4decision.com) | [www.stat4decision.com](http://www.stat4decision.com)

Numéro SIRET: 81048985600015 | Numéro de déclaration d'activité: 11755352275 (auprès du préfet de région de: Ile de France)

*Organisme de formation certifié Qualiopi pour ses actions de formation*

- Jupyter notebooks
- Auto-évaluation sur la compréhension et la mise en application des concepts.
- Feedback individuel : chaque participant se voit proposer un bilan de ses points forts et axes d'amélioration.

#### Stat4decision

37-39 avenue Ledru Rollin Paris 75012

Tel. +33172254082 | E-mail : [info@stat4decision.com](mailto:info@stat4decision.com) | [www.stat4decision.com](http://www.stat4decision.com)

Numéro SIRET: 81048985600015 | Numéro de déclaration d'activité: 11755352275 (auprès du préfet de région de: Ile de France)

*Organisme de formation certifié Qualiopi pour ses actions de formation*