# Formation Réaliser ses analyses statistiques avec Python

**Objectifs :** Apprendre à utiliser le langage Python pour analyser des données. Mettre en œuvre avec Python les méthodes de statistique descriptive, décisionnelle, analyse de la variance, régression linéaire et analyse de données multidimensionnelles.

**Compétences visées :** - Maîtriser les bases de la manipulation de données avec Pandas
- Réaliser des analyses descriptives et des tests statistiques avec SciPy et Statsmodels
- Produire des visualisations avancées avec Matplotlib et Seaborn
- Implémenter des modèles de régression linéaire et logistique
- Automatiser les workflows d'analyse statistique
- Découvrir les principales méthodes d'analyse multidimensionnelle des données

**Durée :** 3 jours (21 heures)

**Public :** Analystes de données, chercheurs ou ingénieurs souhaitant maîtriser Python pour l'analyse statistique, plus généralement toute personne souhaitant analyser des données avec Python

**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 : 1650 € HT
- Distanciel : 1500 € HT
(-10% pour 2 inscrits, -20% dès 3 inscrits)

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

**Nos prochaines sessions**

**Distance**
du 26 au 28 avril 2027
du 28 au 30 juin 2027
du 25 au 27 octobre 2027

**Lyon**
du 7 au 9 décembre 2026
du 12 au 14 avril 2027
du 6 au 8 décembre 2027

**Paris**
du 30 nov. au 2 déc. 2026
du 8 au 10 mars 2027
du 14 au 16 juin 2027
du 29 nov. au 1 déc. 2027

**Toulouse**
du 22 au 24 mars 2027
du 27 au 29 septembre 2027

**Programme :**

## - Introduction et mise en place de l'environnement

- Installation et configuration des environnements (Anaconda, Jupyter Notebook)
- Présentation des bibliothèques Python clés pour les statistiques (Pandas, Matplotlib, Scipy, Statsmodels)

## - Manipulation et nettoyage des données

- Chargement et exploration des données avec Pandas
    - Lecture/écriture de fichiers CSV, Excel, JSON
    - Traitement des données manquantes
    - Gestion des types de données et transformations

## - Analyses exploratoires des données (EDA)

- Calculs de statistiques descriptives
    - Moyennes, médianes, quartiles, variance, écart-type
    - Création de résumés statistiques groupés
- Visualisations avec Matplotlib et Seaborn
    - Histogrammes, boxplots, nuages de points

## - Introduction aux tests statistiques

- Bases des tests d'hypothèses
    - Concepts : p-value, seuils de significativité, hypothèses nulle et alternative
- Tests courants avec SciPy
    - Test t (indépendant et pairé)

- Tests de normalité (Shapiro-Wilk)
- Tests du chi-carré pour les tables de contingence
- Interprétation des résultats

## - Régression linéaire et introduction à la modélisation

- Introduction aux modèles linéaires
  - Régression linéaire simple avec Statsmodels
  - Analyse des coefficients et des résidus
- Régression multiple
  - Construction d'un modèle avec plusieurs variables explicatives
  - Validation du modèle (tests de significativité, $R^2$)
- Introduction à la régression logistique
  - Concepts de base et mise en œuvre pour les données catégoriques

## - Découvrir les principales méthodes d'analyse multidimensionnelle des données à l'aide de la bibliothèque Prince

**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 Formation par ville 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*