



## EN BREF

**Référence** : PYDS

**Durée** : 4 jours

**Niveau** : intermédiaire

**Lieu** : Paris, Orléans, sur site ou à distance

**Dates** : voir site internet ou nous contacter

**Tarif** (inter-entreprises) : 2500 € HT

**Mise à jour** : 19/11/2024.

## NOS FORMATEUR-RICE-S

- ✓ Expert·e·s dans leur domaine, ils et elles ont l'expérience du développement logiciel et de la réalisation de projets informatiques.
- ✓ Ils et elles ont dispensé, avec succès, de nombreuses heures de formation vis-à-vis d'un public exigeant d'enseignant·e·s, de technicien·ne·s, d'ingénieur·e·s et de scientifiques.

## INSCRIPTION / RENSEIGNEMENTS



**02 38 21 30 57**  
(du lundi au vendredi)



**contact@numgrade.com**



## LA BOÎTE À OUTILS POUR LA DATA SCIENCE EN PYTHON

Ce stage intensif vous amènera à maîtriser progressivement les bibliothèques de Data Science du monde Python pour traiter, analyser et visualiser vos données !

## L'ENVIRONNEMENT

**Python** est un langage clair, intuitif, massivement répandu, et devenu incontournable dans le traitement d'une multitude de tâches informatiques.

## LA FORMATION

Cette formation vous permettra de traiter, analyser et visualiser vos données avec les bibliothèques scientifiques les plus populaires du langage Python : NumPy, pandas, Matplotlib, Plotly, seaborn.

Les notions liées à l'utilisation des bibliothèques sont présentées de manière progressive. De nombreux exercices pratiques sont proposés tout au long de la formation.

Cette formation vous prépare à la certification API Society référencée auprès de France Compétences dans le répertoire spécifique sous le numéro [RS6763](#). Toutes les modalités pratiques pour le passage de la certification vous seront communiquées.

## FINANCEMENT

**La formation est finançable avec votre CPF.** Le tarif inclus la formation et le passage de la certification.

## OBJECTIFS PÉDAGOGIQUES

À l'issue de la formation, le ou la stagiaire sera capable de :

- Créer et paramétrer son environnement de travail Python Data Science
- Utiliser efficacement les bibliothèques Python de Data Science
- Analyser ses données avec NumPy et pandas
- Visualiser ses données avec Matplotlib, pandas, seaborn et Plotly

## PRÉREQUIS

Connaître les bases du langage de programmation Python ou avoir suivi la formation « *Python débutant* ». Un questionnaire de pré-formation est proposé pour valider l'adéquation entre le besoin et la formation.





## NOS MOYENS PÉDAGOGIQUES

- ✓ Un kit de cours complet en couleur
- ✓ De nombreux exercices pratiques
- ✓ Des programmes régulièrement révisés
- ✓ Un rythme adaptable à chacun·e
- ✓ Un suivi pour toutes vos questions post-formation

## INSCRIPTION / RENSEIGNEMENTS



02 38 21 30 57

(du lundi au vendredi)



contact@numgrade.fr

## ÉVALUATION DE LA FORMATION

À l'issue de chaque demi-journée de formation, une **feuille d'émargement** est signée, à la fois par le·la stagiaire et le·la formateur·rice.

Des **QCMs** en début et fin de formation permettent de vérifier l'**acquisition des compétences** durant la formation.

Un **certificat de formation**, ainsi qu'une **attestation de compétences** sont remis aux stagiaires en fin de formation, leur permettant de **faire valoir l'acquisition de leurs nouvelles connaissances**.

## PUBLIC CONCERNÉ

Technicien·ne·s, ingénieur·e·s, développeur·se·s, scientifiques, enseignant·e·s ou tout public souhaitant découvrir la Data Science avec Python.

## SITUATION DE HANDICAP

NUMGRADE s'engage à accueillir les personnes en situation de handicap. N'hésitez pas à nous contacter pour que le suivi de la formation se déroule dans les meilleures conditions possibles.

## PROGRAMME DE LA FORMATION

### L'écosystème Python pour la science des données

- Tour d'horizon de packages Python de Data Science
- Installation dans un environnement virtuel des bibliothèques scientifiques
- Les environnements de développement

### La bibliothèque NumPy

- Intérêts de travailler avec les tableaux
- Création de tableaux et de matrices
- Initialisation avec des données aléatoires
- Les types de données et changer le type de données
- Les attributs ndim, shape, size, dtype, itemsize, bytes
- Manipulation de tableaux : indexation, slicing, indexation avancée
- Broadcasting pour l'indexation et les opérations sur les données
- Copie et vue d'un tableau
- Changer les dimensions d'un tableau et le transposer
- Concaténer et découper des tableaux
- Pourquoi éviter les boucles for ?
- Fonctions classiques (addition, soustraction...), fonctions trigonométriques, exposants et logarithmes, fonctions statistiques...
- Apprendre à utiliser l'option axis
- Fonctions de comparaisons
- Extraire des informations avec des masques booléens
- Charger et sauvegarder les tableaux





## NOS PRESTATIONS

- ✓ Formations en INTER ou INTRA
- ✓ Développement logiciel
- ✓ Conseil et assistance technique

## INSCRIPTION / RENSEIGNEMENTS



02 38 21 30 57

(du lundi au vendredi)



contact@numgrade.fr

## FORMATIONS CONNEXES

- ✓ [Python Interfaces Graphiques](#)
- ✓ [Python Perfectionnement](#)
- ✓ [Python Big Data](#)

## NOS CLIENT·E·S TÉMOIGNENT

« Très bonne formation expliquée de manière pédagogique »

« Très bon formateur, très pédagogue et sympa ainsi que compétent »

« I have got what I expected from the course. It was clear, interesting and with normal speed of explanation »

« La formation était très bien organisée et le fait qu'elle soit à distance n'était d'aucune façon limitante.

Le cours a complètement répondu à mes attentes, le formateur, à l'écoute et bienveillant, maîtrisait parfaitement le sujet »

## La bibliothèque pandas

- Créer une série avec la classe Series
- Créer un tableau 2D (DataFrame) avec la classe DataFrame
- Extraire les indices de ligne et de colonnes
- Lire et exporter des données dans différents formats (csv, excel...)
- Indexation, slicing et sélection avancée implicite et explicite
- Sélectionner des données avec des masques
- Insérer, modifier et concaténer des données
- Copier des données : copie superficielle ou profonde
- Traiter les données manquantes
- Transposer des données
- Fusion et jointure de données
- Mettre une colonne en indice
- Trier les indices et les valeurs
- Supprimer des données
- Agréger des données
- Grouper et analyser des données avec la fonction groupby()
- Analyser les données avec les fonctions aggregate(), filter(), apply() et transform()
- Créer des tableaux croisés dynamiques
- Segmenter les données avec les fonctions qcut() et cut()
- Calculer des moyennes glissantes
- Multi-indices :
  - Créer des multi-indices
  - Ajouter des dimensions aux Series et DataFrames avec l'objet MultiIndex
  - Indexer les Series et DataFrames multi-indexés
- Chaînes de caractères :
  - Indexation et slicing
  - Traiter les données avec les méthodes de chaînes de caractères implémentées dans pandas
  - Utiliser des expressions régulières avec pandas
- Les données temporelles
  - Créer des dates, des durées et des périodes
  - Indexation et slicing des données temporelles
  - Échantillonnage avec les fonctions asfreq() et resample()





## MOYENS PÉDAGOGIQUES

- Un kit de cours en couleur, reprenant l'intégralité du contenu présenté
- Des quiz et de nombreux exercices pratiques en cohérence avec les objectifs de la formation
- Des programmes révisés régulièrement afin de répondre à l'état de l'art
- Un rythme adaptable à tous
- Un suivi pour toutes questions post-formation

- **En visio-conférence :**

Des outils de partage de code et des commentaires en temps réels sur les exercices pratiques pour aider les stagiaires à progresser.

- **En présentiel :**

Des salles agréables et fonctionnelles.

- **En entreprise :**

Nous nous déplaçons partout en France et à l'étranger pour vous former sur vos propres sites.

## La bibliothèque Matplotlib

- Afficher les graphiques depuis un script Python ou un notebook
- Afficher les graphiques en utilisant le style MATLAB ou le style orienté objet
- Modifier le style du graphique
- Les objets Figure et Axes
- Tracer des courbes avec la méthode plot() :
  - Modifier la couleur, le style du tracé et la largeur du tracé,
  - Ajuster la longueur des axes, les nommer et changer les graduations
  - Ajouter un titre et une légende
- Afficher des nuages de points
- Afficher des barres d'erreurs
- Remplir la surface entre 2 lignes avec la méthode fill\_between()
- Tracer un histogramme
- Tracer plusieurs graphiques :
  - Placer un graphique à l'intérieur d'un autre
  - Créer une grille avec la fonction subplots()
- Tracer des graphiques en 3 dimensions
- Interagir avec les graphiques dans le Jupyter notebook avec le widget interact
- Utiliser pandas plot pour réaliser rapidement des tracés depuis un objet Series ou DataFrame : tour d'horizon des méthodes plot(), bar(), barh(), hist(), box(), scatter(), pie()

## La bibliothèque seaborn

- Fonctionnement de l'API seaborn : Figure-level et Axes-level
- Les "relational plots"
- Tracer des distributions
- Tracer des données qualitatives (categorical data)
- Tracer des cartes thermiques
- Tracer des modèles de régression linéaire
- Changer le rendu de la figure :
  - Ajouter un titre
  - Changer les couleurs
  - Les fonctions set\_theme(), set\_style(), set\_context() et despine()

### INSCRIPTION / RENSEIGNEMENTS



**02 38 21 30 57**  
(du lundi au vendredi)



**contact@numgrade.fr**





### QUELQUES CHIFFRES

- ✓ **4.7/5** de moyenne sur l'ensemble des questionnaires de **satisfaction**
- ✓ **+ 850** personnes formées
- ✓ **+ 19000** heures de formations stagiaires
- ✓ **98.85 %** des stagiaires recommandent nos formations !

### INSCRIPTION / RENSEIGNEMENTS



**02 38 21 30 57**  
(du lundi au vendredi)



**contact@numgrade.fr**

### EXIGEZ LA DIFFÉRENCE NUMGRADE

- ✓ Des consultant·e·s expert·e·s et passionné·e·s
- ✓ Concentré·e·s sur la réussite de leurs clients et clientes
- ✓ En quête permanente d'excellence technique

### La bibliothèque Plotly

- Présentation de la librairie Plotly et de Kaleido
- Exploration de plotly.express
- Tracer des courbes avec la fonction `line()` et explorer les options pour :
  - Modifier la figure
  - Tracer plusieurs courbes
  - Ajouter des informations
  - Tracer plusieurs graphiques
  - Ajouter des barres d'erreurs
  - Modifier le style de la figure
- Changer le thème des figures par défaut
- Sauvegarder la figure
- Tracer des graphiques en aires :
  - Ajouter des motifs avec l'option `pattern_shape`
- Afficher des nuages de points et explorer les options pour :
  - Définir la taille des points
  - Générer de la transparence
  - Ajouter des symboles
  - Modifier la barre de couleur la méthode
- Modifier la position de la légende et de la barre de couleurs
- Réaliser des graphiques en 3D avec `scatter_3d()` et `line_3d()`
- Tracer des diagrammes en barres
- Tracer des histogrammes
- Tracer des cartes :
  - Les fonctions `line_map()` et `scatter_map()` avec les options `zoom` et `center`
  - Les fonctions `line_geo()` et `scatter_geo()` avec les options `scope` et `projection`,
  - La fonction `choropleth()`

