



## IDF – Python orienté Data Science

<b>Objectifs</b>	A l'issue de la formation, les participants se seront familiarisés avec les différents paradigmes du Machine Learning et seront capables d'utiliser Python et des librairies pertinentes pour manipuler les algorithmes classiques du domaine pour produire des analyses éclairées de leurs données
<b>Public</b>	Chercheurs, ingénieurs, techniciens, doctorants, chef de projet, data analystes
<b>Pré-requis</b>	Posséder des connaissances de base en algorithmique est fortement conseillé, mais aussi en Python ou dans un autre langage de programmation ainsi que des compétences en mathématiques statistiques
<b>Programme</b>	<p><b>Rappels des notions de base de la programmation et présentation de Python</b></p> <ul style="list-style-type: none"><li>- Notions d'algorithmique et langage de programmation</li><li>- Où se situe Python dans le domaine de la programmation ?</li></ul> <p><b>La modularité</b></p> <ul style="list-style-type: none"><li>- Les fonctions définies par le mot clef def</li><li>- Les fonctions lambda</li><li>- La bibliothèque standard, les modules et le mot clef import</li><li>- Les autres bibliothèques et le système de gestion pip</li><li>- Sauvegarder, organiser et appeler son code : les fichiers .py</li></ul> <p><b>Les erreurs</b></p> <ul style="list-style-type: none"><li>- Principes fondamentaux des exceptions</li><li>- Gérer les erreurs simplement avec les blocs try et except</li><li>- Gérer un type d'erreur particulier</li><li>- Aller plus loin avec else, finally et pass</li><li>- Les assertions</li></ul> <p><b>Python : les objets et les classes</b></p> <ul style="list-style-type: none"><li>- Les principes du paradigme de la programmation objet</li><li>- Travailler avec des séquences d'éléments de base : les listes et les tuples</li><li>- Travailler avec des séquences d'objets : les dictionnaires</li><li>- Le module os et la gestion des fichiers</li><li>- La portée des variables en Python</li><li>- Les conventions de nommage et de gestion</li><li>- La création d'une classe</li><li>- Les attributs</li><li>- Les méthodes de classe et les méthodes statiques</li></ul> <p><b>Héritage simple / héritage multiple</b></p> <ul style="list-style-type: none"><li>- Principe de l'héritage</li><li>- Héritage simple et classe mère</li><li>- Héritage multiple et classes parentes</li></ul> <p><b>Les méthodes spéciales</b></p> <ul style="list-style-type: none"><li>- Les méthodes de conteneur</li><li>- Les méthodes mathématiques et de comparaison</li><li>- Les décorateurs</li><li>- Les méthodes de pickle</li></ul>



	<p><b>Manipulation de données, format de fichier</b></p> <ul style="list-style-type: none"><li>- Manipuler différents formats de fichiers</li><li>- Accéder aux données</li><li>- Manipuler des données massives</li></ul> <p><b>Les mathématiques et les statistiques en Python</b></p> <ul style="list-style-type: none"><li>- Les structures de données en NumPy et Pandas</li><li>- L'aléatoire en informatique et le module random</li><li>- Algèbre linéaire : manipuler des vecteurs, matrices, tenseurs avec NumPy</li><li>- Tests statistiques (test Z, test de Student)</li><li>- La descente de gradient</li></ul> <p><b>Initiation à l'apprentissage statistique</b></p> <ul style="list-style-type: none"><li>- Les différents régimes de l'apprentissage statistique : apprentissage supervisé et non supervisé</li><li>- Évaluer un modèle (métriques, matrice de confusion, courbes ROC et AUC, jeux de tests, validation croisée)</li><li>- Prétraitement des données (encodage, classes déséquilibrées, valeurs manquantes...)</li><li>- Les régressions linéaires et logistiques (avec Scikit-learn), les régressions polynomiales, les régressions régularisées</li><li>- Le clustering K-means et k-nearest neighbors</li><li>- Le clustering hiérarchique</li><li>- Arbres de décisions et de régression</li><li>- Perceptron et Support Vector Machines</li><li>- Introduction aux réseaux de neurones</li></ul> <p><b>Visualisation</b></p> <ul style="list-style-type: none"><li>- Analyses de composantes (principales, indépendantes ...)</li><li>- Tracer des graphiques pertinents avec matplotlib et seaborn</li></ul>
<b>Dates et lieu</b>	<p><b>Dates :</b></p> <ul style="list-style-type: none"><li>• <b>les 07, 08, 09 et 10 avril 2026</b></li></ul> <p><b>Clôture des inscriptions :</b> le 16 février 2026 sur <a href="https://www.sirene.inserm.fr/">https://www.sirene.inserm.fr/</a></p> <p>Formation en présentiel</p>
<b>Contact</b>	<p><b>INSERM DR Paris IDF Sud</b> <b>Formation Continue</b></p> <p>48-50 rue Albert 75013 Paris</p> <p>@ : <a href="mailto:formation.paris_11@inserm.fr">formation.paris_11@inserm.fr</a> ① : 01.85.55.38.44</p>