CODE
7007

TECHNOLOGIES DU NUMÉRIQUE ET SYSTÈME D'INFORMATION / BASES DE DONNÉES - BIG DATA - BLOCKCHAIN

NOUVEAU **POSSIBLE EN INTRA**

ANALYSE DE DONNÉES AVEC PYTHON

L'exploitation des données constitue aujourd'hui un enjeu majeur pour de nombreux métiers, de l'analyse décisionnelle à la Data Science. Python et ses principales bibliothèques offrent un environnement complet pour préparer, explorer, analyser et visualiser des données.

1800 € HT | **3 JOURS** (21 H.) | **VILLEURBANNE**
DU 15/06/2027 AU 17/06/2027

COMPÉTENCE PRINCIPALE VISÉE

Analyser, nettoyer et visualiser des données avec Python pour produire des indicateurs décisionnels fiables.



LES + DE LA FORMATION

60 % du temps est consacré à des travaux pratiques, permettant aux apprenants de mettre immédiatement en œuvre les notions abordées.

PUBLIC

- Analystes de données, data analysts et data scientists
- Ingénieurs et professionnels métiers
- Développeurs Python

PRÉREQUIS

- Connaissances de base en Python : variables, fonctions, listes et dictionnaires
- Notions de manipulation de données et d'Excel

OBJECTIFS PÉDAGOGIQUES

- Préparer et nettoyer des jeux de données avec NumPy et Pandas
- Explorer et analyser des données à l'aide de méthodes statistiques et d'outils d'analyse exploratoire
- Visualiser et interpréter les résultats afin d'identifier des tendances, corrélations et informations pertinentes.

OBJECTIFS PÉDAGOGIQUES

- Préparer et nettoyer des jeux de données avec NumPy et Pandas
- Explorer et analyser des données à l'aide de méthodes statistiques et d'outils d'analyse exploratoire
- Visualiser et interpréter les résultats afin d'identifier des tendances, corrélations et informations pertinentes.

97,2%
de clients satisfaits*
* enquête réalisée auprès de nos clients en septembre 2025

CONTENU

PARTIE 1 - Rappels essentiels et préparation de l'environnement

- Mise en place de l'environnement de travail
  - Installation et configuration : Anaconda, Miniconda, uv / Jupyter Notebook, JupyterLab
  - VS Code (optionnel)
  - Gestion des environnements virtuels
  - Introduction à pip et conda
- Rappels Python pour la data
  - Types de données utiles : listes, dictionnaires, tuples
  - Fonctions et compréhension de listes
  - Modules essentiels : math, datetime, os
- Introduction aux bibliothèques data
  - NumPy : tableaux et calcul vectorisé
  - Pandas : Series et DataFrame
  - Notions clés : Indexation, Sélection, Types de données

TRAVAUX PRATIQUES – Prise en main de l'environnement Python

- Création et gestion d'un environnement Python
- Manipulation de tableaux avec NumPy
- Création et manipulation de DataFrames avec Pandas

PARTIE 2 - Nettoyage et préparation des données

- Importation de données
  - CSV, Excel, JSON
  - Lecture de fichiers avec Pandas
  - Gestion des encodages
- Exploration des données (EDA basique)
  - .head(), .info(), .describe()
  - Détection des valeurs manquantes
  - Analyse des distributions
- Nettoyage des données
  - Traitement des valeurs manquantes : suppression, imputation (moyenne, médiane, mode)
  - Gestion des doublons
  - Correction des types de données
- Transformation des données
  - Création de nouvelles colonnes
  - Application de fonctions (apply, lambda)
  - Encodage de variables catégorielles

TRAVAUX PRATIQUES – Nettoyer et préparer un jeu de données

- Importation d'un jeu de données
- Exploration et identification des anomalies
- Traitement des valeurs manquantes et des doublons
- Correction des types et transformation des données

PARTIE 3 - Analyse exploratoire et visualisation des données

- Analyse exploratoire (EDA avancée)
    - Analyse univariée et bivariée
    - GroupBy et agrégations
    - Tableaux croisés dynamiques (pivot tables)
- Visualisation des données
    - Matplotlib : histogrammes, courbes, scatter plots
    - Seaborn : boxplots, heatmaps, pairplots
- Analyse des tendances
    - Corrélations entre variables
    - Détection de patterns
    - Analyse temporelle (time series basique)

TRAVAUX PRATIQUES – Explorer et visualiser les données

- Réalisation d'une analyse exploratoire
- Production de graphiques avec Matplotlib et Seaborn
- Analyse des corrélations et des tendances
- Interprétation des résultats obtenus

PARTIE 4 - Analyse avancée et mini-projet

- Statistiques appliquées à la data
    - Moyenne, variance, écart-type
    - Corrélation vs causalité
    - Introduction aux tests statistiques simples
- Introduction au machine learning
    - égression linéaire
      Classification simple (KNN, Logistic Regression)
      Utilisation de scikit-learn
- Construction d'un pipeline d'analyse
    - Structuration d'un projet data
    - Nettoyage ⁄ analyse ⁄ visualisation ⁄ interprétation
    - Reproductibilité (notebooks propres)

TRAVAUX PRATIQUES – Mini-projet d'analyse de données

- Mise en œuvre d'un pipeline complet d'analyse
- Nettoyage et préparation des données
- Analyse statistique et exploration des données
- Visualisation et interprétation des résultats
- Introduction à un modèle simple de machine learning
- Structuration et documentation du notebook final

ÉQUIPE PÉDAGOGIQUE

L'équipe enseignante est composée de plusieurs enseignants-chercheurs de l'INSA Lyon, membres du laboratoire LIRIS, avec une activité de recherche reconnue en sciences des données. Cette équipe dispose d'une forte expertise technico-scientifique, obtenue via des nombreux projets de recherche partenariale, avec des grands groupes français mais également avec des start-up, PME ou ETI.

INSA INSTITUT NATIONAL DES SCIENCES APPLIQUÉES LYON

MOYENS ET MÉTHODES PÉDAGOGIQUES

Alternance d'exposés, illustrations, études de cas, travaux pratiques et échanges avec des experts du domaine Un support de cours sera remis à chacun des participants

PROCHAINE SESSION

VILLEURBANNE : DU 15/06/2027 AU 17/06/2027

Frais pédagogiques individuels : 1 800 € H.T. (* Repas inclus)

L'ouverture de la session est conditionnée par un nombre minimum de participants. Nous consulter pour d'autres dates.

ÉVALUATION ET RÉSULTATS

Évaluation des acquis de la formation

Evaluation des acquis des apprenants par auto-examen. 94.3% des apprenants ont acquis la compétence principale visée. (sur 177 apprenants évalués sur cette thématique depuis 2020)

Évaluation de la satisfaction des participants en fin de formation (Niveau 1 KIRKPATRICK)

4.3 par les participants. (sur 199 participants ayant suivi une formation dans la thématique depuis 2020)

RENSEIGNEMENTS ET INSCRIPTION
Tel : +33 (0)4 72 43 83 93
Fax : +33 (0)4 72 44 34 24
mail : formation@insavalor.fr
Préinscription sur formation.insavalor.fr

Accueil des personnes en situation de handicap nécessitant un besoin spécifique d'accompagnement : nous contacter à l'inscription. Nos locaux sont accessibles aux personnes à mobilité réduite.

Actualisée le 05/10/2026