

# EXPLORATION DE DONNÉES EN PYTHON - PY2

## Objectifs de la formation

- Apprendre à manipuler les données en Python avec les principales bibliothèques dédiées (pandas, numpy)
- Apprendre les bases de la visualisation de données :
  - Savoir faire des graphiques simples intégrables aux rapports
  - Savoir faire des visualisations interactives
- S'initier à l'utilisation des outils Big Data (PySpark)

## Publics - Prérequis

- Toute personne ayant un **profil technique** ou ayant validé la formation « Introduction Intensive à Python – PY1 » et souhaitant s'initier à la manipulation et à l'exploration de données en Python.

## Moyens pédagogiques

- Exposé théorique sur le langage et exercices pratiques
- Fil rouge autour d'un projet concret
- Utilisation des cahiers (notebooks) Jupyter pour rassembler support et exercices pratiques dans un même document.

## Programme de la formation

- Lecture / écriture de données dans les formats de données courants (csv, excel, etc.)
- Sélections et transformations de colonnes
- Analyses univariées (descriptions statistiques et visualisations avec matplotlib, pandas, seaborn, plotly)
- Groupements, agrégations, jointures, tableaux croisés
- Analyses multivariées (mesures d'association des variables, visualisation, projections)
- Cartographie (visualisation de données géographiques)
- Visualisation interactive dans le notebook (utilisation de ipywidgets)
- Création de Dashboards (Streamlit)
- Premiers pas avec PySpark

## Formateur

Alexis est docteur en informatique spécialisé en vision par ordinateur et apprentissage automatique. Il travaille depuis 2014 chez Probayes en tant qu'ingénieur recherche et développement et chef de projet sur des sujets d'apprentissage automatique. Il utilise Python dans le cadre de son travail et de ses projets personnels depuis plus de 15 ans.

## Contact

Probayes - 53 Avenue Kuntzmann 38330 Montbonnot Saint Martin  
04.76.42.64.13 - [formation@probayes.com](mailto:formation@probayes.com) - [www.probayes.com](http://www.probayes.com)

N° SIRET : 45053883000073 / Code NAF : 6202A / TVA : FR72 450 538 830

Déclaration d'activité enregistrée sous le numéro 84 38 06761 38 auprès du préfet de la région Auvergne-Rhône-Alpes. Cet enregistrement ne vaut pas agrément de l'Etat.



Exposé : 70%

Exercices : 30%



3 jours / 21 heures



6 à 15 personnes



Accessible à toutes les personnes porteuses d'un handicap



INTER : 2990€ HT/stagiaire

INTRA : sur devis

## Parcours possible

Machine Learning - Tronc  
Commun 1 - MLTC1

Machine Learning - Tronc  
Commun 2 - MLTC2

TAL/NLP en Python -  
TAL1

TAL/NLP pour Chatbot -  
TAL2

 probayes