

PYTHON POUR LA DATA SCIENCE - 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)

Programme de la formation

- Lecture / écriture de données dans les formats de données courantes (csv, excel, etc.)
- Sélection et transformation 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, projection)
- Cartographie (visualisation de données géographiques)
- Visualisation interactive dans le notebook (utilisation de ipywidgets)
- Création de Dashboards (Streamlit)
- Premiers pas avec PySpark

Public - Prérequis

- Toute personne ayant un profil technique ou ayant validée la formation « Le langage Python – PY1 » et souhaitant s'initier à la manipulation et à l'exploration de données en Python

Formateurs

David est Data Scientist à Probayes depuis 2004. Chef de projets sur de nombreux sujets R&D ou d'industrialisation, il utilise quotidiennement Python pour l'analyse de données, l'apprentissage automatique ainsi que la mise en production des solutions développées.

Alexis est docteur en informatique spécialisé en vision par ordinateur et apprentissage automatique. Il travaille depuis 2014 chez Probayes en tant qu'ingénieur R&D et chef de projets 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.

Moyens pédagogiques

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



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

Le Langage Python
-
PY1

Exploration de
données en Python
-PY2

Machine Learning
-Tronc Commun 1 -
MLTC1

Machine Learning
-Tronc Commun 2-
MLTC2

TAL/NLP en Python
-
TAL1

TAL/NLP pour Chatbot
-
TAL2