Python et Data Science : Maîtrise des Fondamentaux et Bonnes Pratiques

Lien : https://innov-systems.com/formation/python-et-data-science-matrise-des-fondamentaux-et-bonnes-pratiques

DURÉE	RÉFÉRENCE	CATÉGORIE
3 jours (21h)	**BSI436**	**Big Data - Langages**

OBJECTIFS DE LA FORMATION

À l'issue de cette formation, vous serez capable de :

- Comprendre et utiliser efficacement Python pour la manipulation et l’analyse de données
- Maîtriser les outils incontournables : Jupyter, Numpy, Pandas, Matplotlib, Seaborn
- Savoir préparer, transformer et visualiser des données de manière optimale
- Appliquer les bonnes pratiques pour préparer des données en vue d’analyses avancées ou de Machine Learning
- Passer efficacement du développement exploratoire en notebook à un environnement de production

POUR QUI ?

- Data Scientists, Data Engineers, Data Analysts
- Développeurs ou Ingénieurs souhaitant se spécialiser en traitement et analyse de données
- Professionnels de la BI ou du Machine Learning débutant avec Python

Innov Systems

Programme détaillé

1 / Introduction et environnement de travail

- Comprendre les bases de Python et ses spécificités pour la data
- Installation et configuration de l'environnement (Anaconda, Python, bibliothèques)
- Prise en main de Jupyter Notebook et JupyterLab : raccourcis, cellules, extensions utiles
- Structurer un projet de data science : organisation et bonnes pratiques

2 / Manipulation de données avec Numpy et Pandas

- Les bases de Numpy : tableaux, indexation, opérations vectorisées
- Introduction à Pandas : Series et DataFrames
- Chargement et export de données (CSV, Excel, SQL, JSON, API)
- Sélection, filtrage et modification de données
- Gestion des valeurs manquantes et nettoyage initial

3 / Accès aux données et intégration avec SQL

- Connexion à des bases de données SQL depuis Python
- Requêtes SQL simples et avancées dans Pandas
- Jointures, agrégations et regroupements de données
- Optimiser les performances des requêtes et lectures de données

4 / Manipulations avancées avec Pandas

- Transformations complexes : pivot, melt, merge et concat
- Gestion des données temporelles : datetime, resampling, time series
- Création de colonnes calculées et utilisation des fonctions apply, map, lambda
- Bonnes pratiques pour travailler sur de gros volumes de données

Inov Systems

5 / Visualisation des données avec Matplotlib et Seaborn

- Principes de base de la visualisation de données
- Graphiques essentiels : histogrammes, barres, lignes, scatter plots
- Visualisations avancées avec Seaborn : boxplots, heatmaps, pairplots
- Choisir la bonne visualisation selon le type de données et l'objectif
- Personnalisation des graphiques et intégration dans un rapport

6 / Préparation des données pour le Machine Learning

- Normalisation et standardisation
- Encodage des variables catégorielles (One-Hot Encoding, Label Encoding)
- Gestion des données déséquilibrées
- Séparation jeu d'entraînement / test et validation croisée
- Introduction aux pipelines de transformation

7 / Du Notebook à la mise en production

- Structurer le code pour le rendre réutilisable (scripts, modules)
- Documentation et commentaires efficaces
- Utilisation de Git et GitHub pour le versioning des notebooks
- Introduction à Streamlit et Dash pour créer des interfaces simples
- Export des notebooks en scripts ou rapports automatisés

8 / Atelier pratique de synthèse

- Analyse complète d'un jeu de données (chargement, nettoyage, visualisation, préparation)
- Création d'un mini-projet avec visualisations et préparation pour ML
- Présentation et discussion des résultats en groupe

Innov Systems

Approche pédagogique

- Support Ecrit et Projection
- Exposés Interactifs, Podcasts et Vidéos
- Brainstorming et Jeux de Rôle
- Cas Pratiques et Labs inclus pour leur impact opérationnel
- Test de Validation des Acquis des Connaissances

Prochaines dates programmées

13 au 15 Jan. 2027	Présentiel - Casablanca
10 au 12 Mars 2027	Présentiel - Casablanca
05 au 07 Mai 2027	Présentiel - Casablanca
30 Juin au 02 Juil. 2027	Présentiel - Casablanca
01 au 03 Sep. 2027	Présentiel - Casablanca

Autres dates possibles sur demande. Contactez-nous pour organiser une session intra-entreprise.

Réservation & Renseignements

Téléphone : +212 522 247 210
Email : contact@innov-systems.com
Web : https//www.innov-systems.com

Innov Systems — Tous droits réservés