



Big Data Analytics avec Python

 DURÉE
5 jours (35h)

 RÉFÉRENCE
BSI96

 CATÉGORIE
**Big Data - Techniques
D'analyse**

OBJECTIFS DE LA FORMATION

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

- ✓ Comprendre le principe de la modélisation statistique
- ✓ Choisir entre la régression et la classification en fonction du type de données
- ✓ Évaluer les performances prédictives d'un algorithme
- ✓ Créer des sélections et des classements dans de grands volumes de données pour dégager des tendances

POUR QUI ?

- ✓ Responsables Infocentre
- ✓ Utilisateurs
- ✓ Gestionnaires métiers de bases de données



☰ Programme détaillé

1/ Introduction au langage Python

- Les caractéristiques du langage Python
- Pourquoi choisir Python pour l'analyse de données ?
- Installation et configuration
- Philosophie de Python (indentation, objet, etc.)
- Les types de données
- Appels de fonctions et méthodes
- Structures de contrôles (boucle, test, exceptions)
- Structures de données et séquences (tuple, liste, primitives, dict)

2/ Introduction à la modélisation

- Introduction au logiciel Jupiter Notebook
- Les étapes de construction d'un modèle
- Les algorithmes supervisés et non supervisés
- Le choix entre la régression et la classification

3/ Procédures d'évaluation de modèles

- Les techniques de ré-échantillonnage en jeu d'apprentissage, de validation et de test
- Test de représentativité des données d'apprentissage
- Mesures de performance des modèles prédictifs
- Matrice de confusion, de coût et la courbe ROC et AUC

4/ Les algorithmes supervisés

- Le principe de régression linéaire univariée
- La régression multivariée

- La régression polynomiale
- La régression régularisée
- Le Naive Bayes
- La régression logistique

5/ Les algorithmes non supervisés

- Le clustering hiérarchique
- Le clustering non hiérarchique
- Les approches mixtes

6/ Analyse en composantes

- Analyse en composantes principales
- Analyse factorielle des correspondances
- Analyse des correspondances multiples
- Analyse factorielle pour données mixtes
- Classification hiérarchique sur composantes principales

7/ Analyse de données textuelles

- Collecte et prétraitement des données textuelles
- Extraction d'entités primaires, d'entités nommées et résolution référentielle
- Étiquetage grammatical, analyse syntaxique, analyse sémantique
- Lemmatisation
- Représentation vectorielle des textes
- Pondération TF-IDF
- Word2Vec

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

📅 20 au 24 Juil. 2026 📍 Casablanca - Maroc

📅 14 au 18 Sep. 2026 📍 Casablanca - Maroc

📅 09 au 13 Nov. 2026 📍 Casablanca - Maroc

📅 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>

▼
Scannez pour accéder
à la fiche en ligne