Dataiku DSS Pour La Data Science Collaborative

Lien : https://innov-systems.com/formation/dataiku-dss-pour-la-data-science-collaborative

DURÉE
4 jours (28h)

RÉFÉRENCE
BSI532

CATÉGORIE
Data Science et Data Engineering

OBJECTIFS DE LA FORMATION

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

- Savoir installer, configurer et administrer Dataiku DSS, depuis l’interface web ou via API
- Comprendre l’architecture complète (Design, Automation, API, Govern) et son intégration dans un SI
- Maîtriser la préparation, l’analyse et la visualisation des données dans un cadre collaboratif
- Construire, entraîner et déployer des modèles de Machine Learning de bout en bout
- Développer des extensions via Python, R, SQL et Spark pour des cas avancés
- Acquérir les bonnes pratiques de gouvernance, sécurité et cycle de vie des projets Dataiku DSS

POUR QUI ?

- Chefs de projet Data
- Développeurs et ingénieurs logiciels
- Data Scientists et Data Analysts
- Administrateurs et utilisateurs avancés de Dataiku DSS

Innov Systems

Programme détaillé

1 / INTRODUCTION ET ARCHITECTURE DSS

- Présentation de l'outil : principes de Self-Service Analytics et collaboration
- Concepts clés : projets, datasets, recipes, flows
- Architecture des clusters DSS : Design, Automation, API et Govern Nodes
- Gestion des utilisateurs, rôles et permissions
- Exemples concrets et premiers pas dans l'interface

2 / ACCÈS ET GESTION DES DONNÉES

- Connexion aux sources : bases SQL, fichiers plats, API, cloud (AWS S3, GCP, Azure)
- Support Big Data : HDFS, Spark, Hadoop
- Connexion aux bases NoSQL : Cassandra, MongoDB, Elasticsearch
- Organisation des données : catalogues, stockage et structuration
- Atelier pratique : ingestion et exploration de données hétérogènes

3 / PRÉPARATION ET TRANSFORMATION DES DONNÉES

- Data Preparation recipes : nettoyage, enrichissement, normalisation
- Échantillonnage et optimisation des traitements
- Utilisation des différents moteurs : DSS, SQL, Spark
- Gestion des données géographiques et temporelles
- Atelier pratique : pipeline complet de préparation de données

4 / ANALYSE STATISTIQUE ET VISUALISATION

- Types de graphiques disponibles et configurations
- Visualisations avancées : cartes, nuages de points, comparaisons interactives
- Fonctionnalités statistiques intégrées : corrélations, distributions, tests

Innov Systems

- Atelier pratique : création d’un dashboard interactif

5 / MACHINE LEARNING ET IA

- Présentation du laboratoire ML intégré
- Apprentissage supervisé et non supervisé : cas d’usage
- Gestion des algorithmes, hyperparamètres et scoring
- Deep Learning : intégration TensorFlow et Keras
- Interprétation, optimisation et comparaison de modèles
- Déploiement des modèles et gestion du cycle de vie
- Atelier pratique : création et déploiement d’un modèle prédictif

6 / AUTOMATISATION ET RECIPES AVANCÉES

- Gestion des flows complexes et parallélisation des traitements
- Recipes avancées : SQL, Python, R, Spark
- Export des résultats (PDF, images, API, systèmes externes)
- Atelier pratique : automatiser un projet complet avec des recipes hybrides

7 / EXTENSIONS ET API

- Utilisation des APIs Dataiku : Python, R, REST
- Développement de plugins et scénarios personnalisés
- Intégration avec Spark et notebooks externes
- Cas pratique : automatisation via API Python

8 / GOUVERNANCE, SÉCURITÉ ET BONNES PRATIQUES

- Gestion des rôles, droits et sécurité des données
- Suivi des projets : tags, documentation intégrée, wiki collaboratif
- Bonnes pratiques de cycle de vie : versionning, monitoring et gouvernance
- Cas d’entreprise : mise en place d’une organisation collaborative DSS

Inov 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

05 au 08 Jan. 2027	Présentiel - Casablanca
02 au 05 Mars 2027	Présentiel - Casablanca
27 au 30 Avr. 2027	Présentiel - Casablanca
22 au 25 Juin 2027	Présentiel - Casablanca
24 au 27 Août 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