

Dataiku DSS

INFORMATIONS GÉNÉRALES

Type de formation : Formation continue

Éligible au CPF : Non

Domaine : IA, Big Data et Bases de données

Action collective : Non

Filière : Big Data

Rubrique : ElasticSearch,...

Code de formation : DB034

€ Tarifs

Prix public : 1510 €

Tarif & financement :

Nous vous accompagnons pour trouver la meilleure solution de financement parmi les suivantes :

- **Le plan de développement des compétences de votre entreprise** (rapprochez-vous de votre service RH).
- **Le dispositif FNE-Formation.**
- **L'OPCO** (opérateurs de compétences) de votre entreprise.
- **Pôle Emploi** sous réserve de l'acceptation de votre dossier par votre conseiller Pôle Emploi.
- **CPF -MonCompteFormation.**

Contactez nous pour plus d'informations.

PRÉSENTATION

Objectifs & compétences

Savoir installer, configurer, Dataiku DSS, l'utiliser depuis l'interface web ou des API.
Présentation, concepts DSS
Connexion aux données
Préparation des données
Graphiques et statistiques
Machine learning
Flow/Recipes
Interfaces de programmation

Public visé

Chefs de projet, développeurs, data-scientists, utilisateurs de Dataiku DSS.

Pré-requis

Connaissance des principes de la manipulation de données et du machine learning.

📍 Lieux & Horaires

Durée : 14 heures

Délai d'accès :

Jusqu'à 8 jours avant le début de la formation.

PROGRAMME

Présentation, concepts DSS

Fonctionnalités :
Mise à disposition des méthodes et outils de data-sciences à partir d'une interface graphique ou de langages de requêtage ou de programmation (Python, SQL, R)
Notions de projet, data, dataset, recipes
Premiers pas avec Dataiku DSS : exemples

Connexion aux données

Connecteurs disponibles, formats de fichiers, formats spécifiques hadoop/Spark
Chargement des fichiers, emplacement des fichiers,
Système de fichiers, HDFS, Amazon S3, stockage Google Cloud, etc ...
FTP, connexion ssh, connexion aux bases de données SQL
Autres stockages : Cassandra, MongoDB, Elasticsearch, ...

Préparation des données

Scripts de préparation, échantillonnage, exécution,
les différents moteurs d'exécution: DSS, Spark, base de données SQL, , Spark

Graphiques et statistiques

Présentation des types de graphiques disponibles et configuration
Echantillonnage et exécution,
graphiques de base, tables, nuages de points,
visualisation cartographique, etc ...
Réalisation d'exercices pratiques.
Fonctionnalités statistiques disponibles,
démonstrations

Machine learning

Fonctionnalités disponibles:
apprentissage supervisé et non supervisé,
algorithmes utilisés, scoring, deep learning, optimisation de modèles, ...

📅 Prochaines sessions

Consultez-nous pour les prochaines sessions.

Flow/Recipes

Interface graphique de gestion des cycles de traitement, des datasets, des traitements, paramétrage de l'exécution en parallèle, des données géographiques, export en format pdf ou images, etc ...
Travaux pratiques

Interfaces de programmation

Présentation des différentes interfaces : SQL, Python, R
Exemples d'interrogations SQL et Python
Interfaçage avec Spark

MODALITÉS**Modalités**

L'inscription est ferme à partir de la signature de la convention ou de la réception d'un bon de commande ou de tout document écrit mentionnant nom, prénom, intitulé de la formation, dates, prix sauf s'il existe un accord cadre.

Méthode

Fin de formation : entretien individuel

Satisfaction des participants : questionnaire de satisfaction réalisé en fin de formation

Assiduité : certificat de réalisation (validation des acquis)