

Mis à jour le 11/07/2025

S'inscrire

Formation Dataiku : La plateforme DSS

2 jours (14 heures)

Présentation

Dataiku est une plateforme data science permettant la création de modèles tout en industrialisant en toile de fond l'ensemble des chaînes de traitement : collecte, préparation des données...

Grâce à cette formation Dataiku, votre équipe pourra en apprendre davantage sur les fonctionnalités de Dataiku. Elle saura s'interfacer sur des [environnements Cloud](#) et améliorer votre organisation en mettant en valeur la donnée au cœur de votre stratégie.

Cet outil collaboratif va permettre à votre entreprise d'avoir à disposition des fonctionnalités autour du Machine Learning et de la Data Science.

Mais aussi de convertir les données en prédiction, utiliser pour explorer, développer vos propres produits data. Vous pourrez préparer votre plateforme afin d'automatiser le workflow et déployer la production.

Pour cette formation, nous utilisons la dernière version : [DSS 11](#).

Objectifs

- Identifier les concepts fondamentaux de Dataiku DSS et son interface (Flow, Flowzones, recettes, moteurs, projets)
- Importer, préparer et enrichir des jeux de données à l'aide des recettes visuelles
- Structurer un projet complet avec une logique de Flow claire et export des résultats
- Créer des visualisations interactives dans un dashboard
- Automatiser l'exécution de traitements via les scénarios
- Mettre en œuvre un projet complet de bout en bout dans Dataiku

Public visé

- Data analysts
- Data scientists
- Data engineers
- Analyst SAS

Pré-requis

- Connaissance de base en SQL
- Connaissances de base en Data Science

Pré-requis logiciel

- Au minimum, avoir un accès actif sur la [version trial de 14 jours de Dataiku](#)
- Une instance Dataiku DSS installée et accessible aux participants
- Un accès aux bases de données (SQL, NoSQL) ou des fichiers de données pour les exercices
- Connexion aux API et services cloud si nécessaire

Programme de notre formation Dataiku

Introduction à Dataiku DSS

- Présentation de la plateforme DSS : interface, logique projet/Flow
- Comprendre les Flowzones pour organiser les traitements
- Types de stockage et moteurs d'exécution
- Importation de fichiers (CSV, Excel)
- Cas pratique : création d'un projet DSS et premier import de données

PRÉPARATION ET NETTOYAGE DES DONNÉES

- Exploration des données importées : types de colonnes, formats, qualité
- Nettoyage des données avec la recette "Prepare" :
 - Normalisation des formats
 - Traitement des valeurs manquantes
 - Renommage et filtrage de colonnes
- Cas pratique : préparer un jeu de données clients

ENRICHISSEMENT ET TRANSFORMATION DES DONNÉES

- Enrichissement par jointure de plusieurs datasets (recette Join)
- Regroupement et agrégation des données (recette Group)
- Structuration d'un Flow de traitement avec zones dédiées
- Cas pratique : enrichir un jeu de données avec des commandes ou des informations produit

STRUCTURATION ET EXPORT DU FLOW

- Création de Flowzones et organisation logique du projet
- Documentation des recettes et gestion des dépendances
- Export de données (CSV, Excel, base de données)
- Cas pratique : finaliser un Flow complet, prêt à l'export

DATA VISUALISATION ET DASHBOARDS

- Bonnes pratiques en data visualisation
- Création de graphiques à partir d'un dataset
- Mise en page et filtrage dans un Dashboard interactif
- Cas pratique : construire un dashboard des ventes ou clients

PERSONNALISÉ : RECETTES ET OUTILS ADAPTÉS AUX BESOINS MÉTIERS

- Écoute et identification des cas d'usage spécifiques des participants
- Présentation de recettes ou plug-ins pertinents pour leurs contextes métier
- Démonstration ciblée sur des données adaptées
- Accompagnement à l'exploration libre ou approfondissement au cas par cas

AUTOMATISATION DES TRAITEMENTS

- Introduction aux scénarios dans Dataiku
- Création d'un scénario simple : nettoyage, enrichissement, export
- Ajout d'étapes et déclencheurs (triggers)
- Cas pratique : automatiser l'exécution d'un projet

ÉTUDE DE CAS FINALE : PROJET COMPLET

- Réalisation d'un projet de A à Z :
 - Import, nettoyage, enrichissement
 - Structuration du Flow
 - Visualisation via dashboard
 - Automatisation via scénario
- Objectif : démontrer la maîtrise de toutes les étapes vues pendant la formation
- Feedback final et retour individuel sur le projet

Module complémentaire (+1/2 journée par module)

Pré-formation SQL (1/2 journée)

- Logique d'une base de données relationnelle
- Identifier les rôles des clés primaires et étrangères
- Ecrire et de lire une jointure en SQL
- Fonctions d'agrégation de base (COUNT, SUM, etc.)
- Utiliser GROUP BY pour synthétiser les données

- Parallèle entre les concepts SQL et les recettes visuelles de Dataiku DSS
- HERE ? recette Prepare
- JOIN ? recette Join
- GROUP BY ? recette Group
- Logique de traitement de données structurée

Utilisation avancée de SQL dans DATAIKU

- Comprendre les espaces SQL dans Dataiku DSS
- Utiliser la recette Prepare avec moteur SQL
- Explorer ses données dans un SQL Notebook
- Créer une recette SQL personnalisée
- Choisir entre recette visuelle et SQL (pro et cons de SQL)

Sociétés concernées

Cette formation s'adresse à la fois aux particuliers ainsi qu'aux entreprises, petites ou grandes, souhaitant former ses équipes à une nouvelle technologie informatique avancée ou bien à acquérir des connaissances métiers spécifiques ou des méthodes modernes.

Positionnement à l'entrée en formation

Le positionnement à l'entrée en formation respecte les critères qualité Qualiopi. Dès son inscription définitive, l'apprenant reçoit un questionnaire d'auto-évaluation nous permettant d'apprécier son niveau estimé sur différents types de technologies, ses attentes et objectifs personnels quant à la formation à venir, dans les limites imposées par le format sélectionné. Ce questionnaire nous permet également d'anticiper certaines difficultés de connexion ou de sécurité interne en entreprise (intraentreprise ou classe virtuelle) qui pourraient être problématiques pour le suivi et le bon déroulement de la session de formation.

Méthodes pédagogiques

Stage Pratique : 60% Pratique, 40% Théorie. Support de la formation distribué au format numérique à tous les participants.

Organisation

Le cours alterne les apports théoriques du formateur soutenus par des exemples et des séances de réflexions, et de travail en groupe.

Validation

À la fin de la session, un questionnaire à choix multiples permet de vérifier l'acquisition correcte des compétences.

Sanction

~~Une attestation sera remise à chaque stagiaire qui aura suivi la totalité de la formation.~~

[Page Web du Programme de Formation](#) - Annexe 1 - Fiche formation

Organisme de formation enregistré sous le numéro 11 75 54743 75. Cet enregistrement ne vaut pas agrément de l'État.

© Ambient IT 2015-2025. Tous droits réservés. Paris, France - Suisse - Belgique - Luxembourg