Dataiku : concevoir et automatiser ses pipelines de données

Construire son premier pipeline de bout en bout

Stage pratique
Culture Data

Code :
AIKU1

Durée :
1 jour(s) (7,00 heures)

Exposés : **20 %**
Cas pratiques : **70 %**
Echanges d'expérience : **10 %**

Inter-entreprises :
Prochaines sessions disponibles sur notre site web.
Tarif : 1 920,00 € HT / participant

Intra-entreprise :
Tarifs et dates sur demande.

DESCRIPTION

Dans un contexte où la maîtrise des données devient un levier stratégique, les organisations doivent être capables de structurer, analyser et industrialiser leurs projets data & ia de manière efficace. Dataiku s'impose comme une plateforme collaborative permettant de couvrir l'ensemble du cycle de vie des données, de leur préparation à leur mise en production.

Cette formation propose une montée en compétence progressive en couvrant les modules essentiels des parcours Core Designer et Advanced Designer. Les participants acquièrent les fondamentaux pour explorer leurs données, concevoir des workflows visuels et collaborer efficacement, puis approfondissent leur maîtrise en optimisant et automatisant des pipelines data complets.

Grâce à une approche orientée pratique, elle permet de gagner rapidement en autonomie, de structurer des projets end-to-end et de s'approprier les concepts fondamentaux requis pour le passage des certifications (Core Designer et Advanced Designer), tout en développant des compétences directement applicables en contexte métier.

En sortie de formation, il est recommandé aux stagiaires de compléter les apports de celle-ci par un travail personnel d'approfondissement avant de passer les examens de certification.

OBJECTIFS PEDAGOGIQUES

- Expliquer les concepts fondamentaux de Dataiku et identifier leur rôle dans un projet data
- Concevoir et mettre en œuvre un projet data complet dans Dataiku
- Construire, automatiser et industrialiser des workflows de traitement de données
- Sélectionner et utiliser de manière autonome les fonctionnalités clés correspondant aux compétences des parcours Core Designer et Advanced Designer
- Résoudre des problèmes et réaliser des exercices pratiques mobilisant les compétences requises pour les certifications Dataiku Core Designer et Advanced Designer

PUBLIC CIBLE

- Data Analyst

- Data Engineer
- Data Scientist
- Business Analyst
- Business User (statisticien, actuaire...)

PRE-REQUIS
Une première expérience en analyse, traitement ou exploitation de données est recommandée. Des connaissances de base en SQL et/ou en manipulation de données constituent un plus.

METHODE PEDAGOGIQUE
Formation avec apports théoriques, échanges sur les contextes des participants et retours d'expérience pratique des formateurs, complétés de travaux pratiques et de mises en situation.

PROFIL DES INTERVENANTS

Cette formation est dispensée par un·e ou plusieurs consultant·es d'OCTO Technology ou de son réseau de partenaires, expert·es reconnus des sujets traités.

Le processus de sélection de nos formateurs et formatrices est exigeant et repose sur une évaluation rigoureuse leurs capacités techniques, de leur expérience professionnelle et de leurs compétences pédagogiques.

MODALITÉS D'ÉVALUATION ET FORMALISATION À L'ISSUE DE LA FORMATION
L'évaluation des acquis se fait tout au long de la session au travers des ateliers et des mises en pratique. Afin de valider les compétences acquises lors de la formation, un formulaire d'auto-positionnement est envoyé en amont et en aval de celle-ci. Une évaluation à chaud est également effectuée en fin de session pour mesurer la satisfaction des stagiaires et un certificat de réalisation leur est adressé individuellement.

PROGRAMME PEDAGOGIQUE DETAILLE

Matin : Construire un pipeline de données fiable

OUVERTURE DE SESSION

- Accueil des participants et tour de table des attentes
- Présentation du déroulé détaillé de la jourée

CONNECTER ET EXPLORER

- Prendre en main Dataiku : projet, Flow, jeux de données et recettes visuelles
- Se connecter à une base de données cloud et juger la qualité de

la donnée en quelques clics (statistiques, valeurs invalides, mise en forme conditionnelle)

CONSOLIDER

- Réunir des données réparties sur plusieurs tables de même structure (Stack)
- Croiser plusieurs sources en une seule opération pour obtenir une vue unifiée (Join)
- Comprendre comment Dataiku choisit automatiquement le moteur de calcul le plus efficace

PRÉPARER ET ENRICHIR

- Nettoyer, harmoniser et rendre la donnée lisible (Prepare)
- Exploiter la dimension géographique des données (géopoints, géocodage inversé)
- Traduire des règles métier en nouvelles colonnes, sans coder (Switch, If/Then/Else)

PARTAGER ET VISUALISER

- Documenter et publier un jeu de données de référence pour toute l'équipe (Data Collections, génération assistée de métadonnées)
- Faire parler la donnée : graphiques avec exploration par niveau de détail, cartes, tableaux croisés
- Assembler un tableau de bord et en gérer les droits d'accès

Après-midi - Industrialiser et partager avec le métier : les premiers réflexes

ANALYSER DANS LE TEMPS

- Agréger la donnée par période et par entité (Group)
- Calculer des indicateurs glissants pour détecter tendances et pics, sans une ligne de code (Window)
- Enrichir avec une source externe et faire ressortir les cas prioritaires (Join, Top N)

OPTIMISER : LES BONS REFLEXES

- Lire son Flow par connexions et par moteurs de calcul pour repérer ce qui ralentit
- Faire exécuter les traitements directement en base plutôt que sur le serveur Dataiku (Sync, calcul in-database)
- Choisir le bon mode de reconstruction selon la situation

AUTOMATISER : PREMIERS SCÉNARIOS

- Piloter tout un pipeline depuis un paramètre unique (variables de projet)
- Planifier un rafraîchissement régulier, sans intervention manuelle (Scenarios)

PACKAGER

- Transformer un projet en application simple, utilisable par des profils non techniques (Application Designer)
- Tester, instancier et versionner l'application pour la faire vivre dans le temps

CLÔTURE DE SESSION

- Revue des concepts clés et plan d'action personnel
- Temps d'échange autour des questions et réponses additionnelles

Accessibilité

L'inclusion est sujet important pour OCTO Academy.
Nos référent·es sont à votre disposition pour faciliter l'adaptation de votre formation à vos besoins spécifiques.
Pour les contacter : academy.accessibilite@octo.com