

Formation : Microsoft Fabric data engineer

Formation pratique - 3j - 21h00 - Réf. FDN
Prix : 2360 CHF H.T.

NEW

En trois jours, cette formation vous plongera au cœur de Microsoft Fabric pour tirer parti de tout son potentiel en ingénierie des données. De l'ingestion batch au temps réel, vous apprendrez à concevoir des pipelines fiables, automatiser vos traitements et optimiser la performance. Sécurité, gouvernance et pratiques DevOps compléteront cet apprentissage, afin d'industrialiser vos projets et transformer vos données en leviers stratégiques à grande échelle.

Objectifs pédagogiques

À l'issue de la formation, le participant sera en mesure de :

- ✓ Savoir configurer un environnement de travail adapté aux besoins d'un projet analytique
- ✓ Comprendre l'impact des paramètres (Spark, Domain, OneLake...) sur la performance
- ✓ Adapter les modèles de chargement aux besoins du projet
- ✓ Comprendre les différentes méthodes d'ingestion batch et savoir transformer et nettoyer les données
- ✓ Maîtriser les outils d'orchestration et automatiser selon les événements
- ✓ Comprendre les concepts de traitement de données en temps réel
- ✓ Assurer un suivi opérationnel des traitements
- ✓ Diagnostiquer rapidement les erreurs
- ✓ Intégrer les pratiques DevOps dans un environnement data
- ✓ Garantir la sécurité des données

Public concerné

Data engineers, data architects, développeurs ou consultants data.

Prérequis

Connaissances des bases de données relationnelles, de Python et de SQL, connaissance de la programmation data et connaissances générales en ingénierie de donnée.

Vérifiez que vous avez les prérequis nécessaires pour profiter pleinement de cette formation en faisant [ce test](#).

PARTICIPANTS

Data engineers, data architects, développeurs ou consultants data.

PRÉREQUIS

Connaissances des bases de données relationnelles, de Python et de SQL, connaissance de la programmation data et connaissances générales en ingénierie de donnée.

COMPÉTENCES DU FORMATEUR

Les experts qui animent la formation sont des spécialistes des matières abordées. Ils ont été validés par nos équipes pédagogiques tant sur le plan des connaissances métiers que sur celui de la pédagogie, et ce pour chaque cours qu'ils enseignent. Ils ont au minimum cinq à dix années d'expérience dans leur domaine et occupent ou ont occupé des postes à responsabilité en entreprise.

MODALITÉS D'ÉVALUATION

Le formateur évalue la progression pédagogique du participant tout au long de la formation au moyen de QCM, mises en situation, travaux pratiques...

Le participant complète également un test de positionnement en amont et en aval pour valider les compétences acquises.

Méthodes et moyens pédagogiques

Exercice

Échanges, travaux pratiques.

Méthodes pédagogiques

Active.

Modalités d'évaluation

Le formateur évalue la progression pédagogique du participant tout au long de la formation au moyen de QCM, mises en situation, travaux pratiques...

Le participant complète également un test de positionnement en amont et en aval pour valider les compétences acquises.

Programme de la formation

1 Présentation de la plateforme Fabric

- Comprendre l'écosystème Microsoft Fabric et ses composants clés
- Identifier les cas d'usage typiques dans l'ingénierie des données
- Situer Fabric dans une architecture de données moderne

2 Configurer les paramètres d'un workspace Microsoft Fabric

- Créer et configurer un workspace pour les besoins d'un projet data

Travaux pratiques

Créer et configurer les paramètres d'un workspace. Justifier ses choix de configuration selon le contexte.

3 Concevoir et mettre en œuvre des modèles de chargement

- Compréhension des enjeux des chargements complets et incrémentaux
- Préparation des données pour un modèle dimensionnel
- Définition d'un modèle de chargement pour le streaming
- Lakehouse, warehouse et eventhouse

Travaux pratiques

Concevoir le bon modèle selon le besoin. Gérer efficacement les flux entrants.

4 Ingérer et transformer des données batch

- Choix du type de stockage
- Utilisation de dataflows, notebooks, KQL et T-SQL
- Gestion des raccourcis, duplications miroir
- Ingestion avec pipelines
- Transformation avec PySpark, SQL
- Dénormalisation, agrégation
- Gestion des doublons, données manquantes ou décalées

Travaux pratiques

Définir une stratégie d'ingestion batch opérationnelle. Appliquer les bonnes transformations.

MOYENS PÉDAGOGIQUES ET TECHNIQUES

- Les moyens pédagogiques et les méthodes d'enseignement utilisés sont principalement : aides audiovisuelles, documentation et support de cours, exercices pratiques d'application et corrigés des exercices pour les formations pratiques, études de cas ou présentation de cas réels pour les séminaires de formation.
- À l'issue de chaque formation ou séminaire, ORSYS fournit aux participants un questionnaire d'évaluation du cours qui est ensuite analysé par nos équipes pédagogiques.
- Une feuille d'émargement par demi-journée de présence est fournie en fin de formation ainsi qu'une attestation de fin de formation si le participant a bien assisté à la totalité de la session.

MODALITÉS ET DÉLAIS D'ACCÈS

L'inscription doit être finalisée 24 heures avant le début de la formation.

ACCESSIBILITÉ AUX PERSONNES HANDICAPÉES

Pour toute question ou besoin relatif à l'accessibilité, vous pouvez joindre notre équipe PSH par e-mail à l'adresse psh-accueil@orsys.fr.

5 Orchestrer les processus

- Pipeline ou notebook
- Création des planifications et déclencheurs événementiels
- Intégration des paramètres et expressions dynamiques dans les notebooks/pipelines

Travaux pratiques

Identifier la bonne méthode d'orchestration. Mettre en œuvre un pipeline avec automatisation basée sur un événement.

6 Présentation de l'ingestion en temps réel

- Utilisation du stockage natif ou des raccourcis
- Traitement avec Eventstreams, Spark structured streaming, KQL

Travaux pratiques

Maîtriser les concepts et outils liés à l'ingestion en temps réel.

7 Monitorer les ressources Fabric

- Suivi de l'ingestion, la transformation, l'actualisation du modèle
- Configuration des alertes

Travaux pratiques

Identifier les bons KPI. Réagir à une alerte.

8 Identifier et résoudre les erreurs

- Identification des erreurs dans les pipelines, flux, notebooks, eventhouses, eventstreams, T-SQL

Travaux pratiques

Détecter, analyser et corriger une erreur.

9 Mettre en œuvre la gestion du cycle de vie dans Fabric

- Présentation de l'intégration avec Git
- Mise en œuvre des projets de base de données
- Création et configuration des pipelines de déploiement

Travaux pratiques

Créer un pipeline de déploiement opérationnel. Identifier les impacts d'une nouvelle version.

10 Configurer la sécurité et la gouvernance

- Mise en œuvre des contrôles d'accès au niveau du workspace
- Configuration des contrôles d'accès par ressource, ligne, colonne, objet
- Mise en place du masquage dynamique
- Activation et exploitation du logging du workspace

Travaux pratiques

Appliquer des stratégies d'accès fines. Démontrer le masquage dynamique. Analyser les logs de manière pertinente.

