

Mise en œuvre d'un Lakehouse avec Microsoft Fabric (DP-601)

Durée: 1 Jour Réf de cours: M-DP601

Résumé:

Ce cours est conçu pour développer vos compétences fondamentales en ingénierie des données sur Microsoft Fabric, en se concentrant sur le concept Lakehouse.

Ce cours explorera les puissantes capacités d'Apache Spark pour le traitement des données distribuées et les techniques essentielles pour une gestion efficace des données, des versions et de la fiabilité en travaillant avec les tables Delta Lake, ainsi que l'ingestion et l'orchestration des données en utilisant les pipelines Dataflows Gen2 et Data Factory.

Ce cours comprend une combinaison de parties magistrales et d'exercices pratiques qui vous prépareront à travailler avec des Lakehouses dans Microsoft Fabric.

Mis à jour Mars 2024.

Public visé:

Ce cours s'adresse principalement aux professionnels des données qui sont familiers avec la modélisation, l'extraction et l'analyse des données.

Il est conçu pour les professionnels qui souhaitent acquérir des connaissances sur l'architecture Lakehouse, la plate-forme Microsoft Fabric et la façon de mettre en œuvre l'analyse de bout en bout à l'aide de ces technologies.

Objectifs pédagogiques:

- À l'issue de la formation, les participants seront capables de :
- Comprendre l'analyse de bout en bout avec Microsoft Fabric
- Démarrer avec lakehouses dans Microsoft Fabric
- Utiliser Apache Spark dans Microsoft Fabric
- Travailler avec les tables Delta Lake dans Microsoft Fabric
- Ingérer des données avec Dataflows Gen2 dans Microsoft Fabric
- Utiliser les pipelines Data Factory dans Microsoft Fabric

Pré-requis:

Les participants doivent avoir le prérequis suivant :

- Etre familiarisé avec les concepts et la terminologie de base des données

Contenu:

Module 1 : Introduction à l'analyse de bout en bout avec Microsoft Fabric

- Décrire l'analyse de bout en bout dans Microsoft Fabric

Module 2 : Démarrer avec les centres d'hébergement (lakehouses) dans Microsoft Fabric

- Décrire les principales caractéristiques et capacités des centres de données (lakehouses) dans Microsoft Fabric
- Créer un lakehouse
- Intégrer des données dans des fichiers et des tables dans un lakehouse
- Interroger les tables d'un lakehouse avec SQL

Module 3 : Utiliser Apache Spark dans Microsoft Fabric

- Configurer Spark dans un espace de travail Microsoft Fabric
- Identifier les scénarios appropriés pour les notebooks et les jobs Spark
- Utiliser les dataframes Spark pour analyser et transformer les données
- Utiliser Spark SQL pour interroger les données dans les tables et les vues
- Visualiser les données dans un notebook Spark

Module 4 : Travailler avec des tables Delta Lake dans Microsoft Fabric

- Comprendre Delta Lake et les tables delta dans Microsoft Fabric
- Créer et gérer des tables delta à l'aide de Spark
- Utiliser Spark pour interroger et transformer les données dans les tables delta
- Utiliser les tables delta avec Spark structured streaming

Module 5 : Intégrer des données avec Dataflows Gen2 dans Microsoft Fabric

- Décrire les capacités de Dataflow (Gen2) dans Microsoft Fabric
- Créer des solutions de flux de données (Gen2) pour ingérer et transformer des données
- Inclure un Dataflow (Gen2) dans un pipeline

Module 6 : Utiliser les pipelines Data Factory dans Microsoft Fabric

- Décrire les capacités des pipelines dans Microsoft Fabric
- Utiliser l'activité Copy Data dans un pipeline
- Créer des pipelines basés sur des modèles prédéfinis
- Exécuter et surveiller les pipelines

Méthodes pédagogiques :

Les participants réalisent un test d'évaluation des connaissances en amont et en aval de la formation pour valider les connaissances acquises pendant la formation.

Un support de cours officiel sera remis aux participants.

Autres moyens pédagogiques et de suivi:

- Compétence du formateur : Les experts qui animent la formation sont des spécialistes des matières abordées et ont au minimum cinq ans d'expérience d'animation. Nos équipes ont validé à la fois leurs connaissances techniques (certifications le cas échéant) ainsi que leur compétence pédagogique.
- Suivi d'exécution : Une feuille d'émargement par demi-journée de présence est signée par tous les participants et le formateur.
- En fin de formation, le participant est invité à s'auto-évaluer sur l'atteinte des objectifs énoncés, et à répondre à un questionnaire de satisfaction qui sera ensuite étudié par nos équipes pédagogiques en vue de maintenir et d'améliorer la qualité de nos prestations.

Délais d'inscription :

- Vous pouvez vous inscrire sur l'une de nos sessions planifiées en inter-entreprises jusqu'à 5 jours ouvrés avant le début de la formation sous réserve de disponibilité de places et de labs le cas échéant.
- Votre place sera confirmée à la réception d'un devis ou "booking form" signé. Vous recevrez ensuite la convocation et les modalités d'accès en présentiel ou distanciel.
- Attention, si cette formation est éligible au Compte Personnel de Formation, vous devrez respecter un délai minimum et non négociable fixé à 11 jours ouvrés avant le début de la session pour vous inscrire via moncompteformation.gouv.fr.

Accueil des bénéficiaires :

- En cas de handicap : plus d'info sur globalknowledge.fr/handicap
- Le Règlement intérieur est disponible sur globalknowledge.fr/reglement