Google Intégration des données avec Cloud Data Fusion

Durée: 2 Jours Réf de cours: GO8334

Résumé:

Découvrez la capacité d'intégration de données de Google Cloud à l'aide de Cloud Data Fusion.

Dans ce cours de 2 jours, nous abordons les défis liés à l'intégration de données et de la nécessité d'une plateforme d'intégration de données (middleware).
Nous discutons ensuite de la manière dont Cloud Data Fusion peut aider à intégrer efficacement les données provenant d'une variété de sources et de formats et à générer des informations.
Nous examinons les principaux composants de Cloud Data Fusion et leur fonctionnement, comment traiter les données par lots et les données en streaming en temps réel avec une conception de pipeline visuelle, un suivi riche des métadonnées et de la lignée des données, et comment déployer des pipelines de données sur divers moteurs d'exécution.

Updated 20/05/2026

Public visé:

- Ingénieur de données
- Analystes de données

Objectifs pédagogiques:

- A l'issue de ce cours, les participants seront capables de :
- Identifier le besoin d'intégration de données
- Identifier des cas d'utilisation pour une mise en œuvre possible avec Cloud Data Fusion
- Concevoir et exécuter des pipelines de traitement de données par lots et en temps réel
- Utiliser des connecteurs pour intégrer des données de diverses sources et formats
- Comprendre la relation entre les métadonnées et la lignée des données
- Comprendre les capacités que Cloud Data Fusion fournit en tant que plateforme d'intégration de données
- Lister les composants principaux de Cloud Data Fusion
- Travailler avec Wrangler pour construire des transformations de données
- Configurer l'environnement d'exécution ; surveiller et dépanner l'exécution du pipeline

Pré-requis:

- Avoir terminé la formation "Introduction à l'ingénierie des données"

Test et certification

Pas de Test associé à ce cours

Après cette formation, nous vous conseillons le(s) module(s) suivant(s):

Aucune recommandation

Contenu:

01 Introduction

02 Introduction à l'intégration de données et à Cloud Data Fusion

- Comprendre le besoin d'intégration de données
- Lister les situations/cas où l'intégration de données peut aider les entreprises
- Lister les plateformes et outils d'intégration de données disponibles
- Identifier les défis de l'intégration de données
- Comprendre l'utilisation de Cloud Data Fusion comme plateforme d'intégration de données
- Créer une instance Cloud Data Fusion
- Se familiariser avec le framework de base et les principaux composants de Cloud Data Fusion

03 Construction de pipelines

- Comprendre l'architecture de Cloud Data Fusion
- Définir ce qu'est un pipeline de données
- Comprendre la représentation DAG d'un pipeline de données
- Apprendre à utiliser Pipeline Studio et ses composants
- Concevoir un pipeline simple à l'aide de Pipeline Studio
- Déployer et exécuter un pipeline

04 Conception de pipelines complexes

- Effectuer des opérations de branchement, de fusion et de jonction
- Exécuter un pipeline avec des arguments d'exécution en utilisant des macros
- Travailler avec des gestionnaires d'erreurs
- Exécuter des actions pré et post-pipeline à l'aide d'actions et de notifications
- Planifier l'exécution des pipelines
- Importer et exporter des pipelines existants

05 Environnement d'exécution de pipeline

- Comprendre la composition d'un environnement d'exécution
- Configurer l'environnement d'exécution, la journalisation et les métriques de votre pipeline. Comprendre des concepts comme le profil de calcul et le provisionneur
- Créer un profil de calcul
- Créer des alertes de pipeline
- Surveiller le pipeline en cours d'exécution

06 Construire des transformations et préparer des données avec Wrangler

- Comprendre l'utilisation de Wrangler et de ses principaux composants
- Transformer des données à l'aide de l'interface utilisateur de Wrangler
- Transformer des données à l'aide de directives/méthodes CLI
- Créer et utiliser des directives définies par l'utilisateur

07 Connecteurs et pipelines de streaming

- Comprendre l'architecture d'intégration de données
- Lister divers connecteurs
- Utiliser l'API de prévention de la perte de données (DLP) du Cloud
- Comprendre l'architecture de référence des pipelines de streaming
- Construire et exécuter un pipeline de streaming

08 Métadonnées et lignée des données

- Lister les types de métadonnées
- Différencier les métadonnées métier, techniques et opérationnelles
- Comprendre ce qu'est la lignée des données
- Comprendre l'importance de maintenir la lignée des données
- Différencier les métadonnées et la lignée des données

09 Résumé

- Passer en revue les objectifs et les concepts du cours

Méthodes pédagogiques :

Un support de cours officiel sera fourni 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