



ORGANISME DE FORMATION AUX TECHNOLOGIES ET METIERS DE L'INFORMATIQUE

Formation Data Integration with Cloud Data Fusion

N° ACTIVITÉ : 11 92 18558 92

TÉLÉPHONE : 01 85 77 07 07

E-MAIL : inscription@hubformation.com

Ce cours de deux jours présente aux participants la capacité d'intégration de données de Google Cloud à l'aide de Cloud Data Fusion.

Dans ce cours, nous discutons des défis liés à l'intégration de données et de la nécessité d'une plate-forme d'intégration de données (middleware).

Nous discutons ensuite de la manière dont Cloud Data Fusion peut aider à intégrer efficacement des données provenant de diverses sources et 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 de diffusion en temps réel avec une conception de pipeline visuel, un suivi approfondi des métadonnées et de la lignée des données, et comment déployer des pipelines de données sur divers moteurs d'exécution.

Objectifs

- | Identifier le besoin d'intégration de données
- | Identifier les fonctionnalités fournies par Cloud Data Fusion en tant que plate-forme d'intégration de données
- | Identifier les cas d'utilisation pour une éventuelle mise en oeuvre avec Cloud Data Fusion
- | Répertorier les composants principaux de Cloud Data Fusion
- | Concevoir et exécuter des pipelines de traitement de données par lots et en temps réel
- | Travailler avec Wrangler pour créer des transformations de données
- | Utiliser des connecteurs pour intégrer des données de différentes sources et formats
- | Configurer l'environnement d'exécution ; Surveiller et dépanner l'exécution du pipeline
- | Identifier la relation entre les métadonnées et la lignée des données

Prérequis

| Avoir suivi le cours « Google Cloud Platform - Les fondamentaux du Big Data et du Machine Learning » ou avoir des connaissances équivalentes

Public

- | Data Engineer
- | Data Analysts

Programme de la formation

Introduction à l'intégration de données et Cloud Data Fusion

- | Intégration de données : quoi, pourquoi, défis
- | Outils d'intégration de données utilisés dans l'industrie
- | Personas utilisateur
- | Introduction à la fusion de données cloud
- | Capacités critiques d'intégration de données
- | Composants de l'interface utilisateur Cloud Data Fusion

Référence	GCP200CDF
Durée	2 jours (14h)
Tarif	1 400 €HT

SESSIONS PROGRAMMÉES

A DISTANCE (FRA)

- du 9 au 10 mai 2024
- du 10 au 11 juillet 2024
- du 25 au 26 septembre 2024
- du 20 au 21 novembre 2024
- du 18 au 19 décembre 2024

[VOIR TOUTES LES DATES](#)

Construire des pipelines

- | Architecture de Cloud Data Fusion
- | Concepts de base
- | Pipelines de données et graphes acycliques dirigés (DAG)
- | Cycle de vie des pipelines
- | Conception de pipelines dans Pipeline Studio

Construire des pipelines complexes

- | Branchement, fusion et jointure
- | Actions et notifications
- | Gestion des erreurs et macros
- | Configurations de pipeline, planification, importation et exportation

Environnement d'exécution du pipeline

- | Horaires et déclencheurs
- | Environnement d'exécution : profil de calcul et provisionneurs
- | Surveillance des pipelines

Construire des transformations et préparer des données avec Wrangler

- | Wrangler
- | Directives
- | Directives définies par l'utilisateur

Connecteurs et pipelines de streaming

- | Comprendre l'architecture d'intégration de données.
- | Lister les différents connecteurs.
- | Utilisez l'API Cloud Data Loss Prevention (DLP).
- | Comprendre l'architecture de référence des pipelines de streaming.
- | Construire et exécuter un pipeline de streaming.

??Métadonnées et lignage des données

- | Métadonnées
- | Lignage des données

Méthode pédagogique

Chaque participant travaille sur un poste informatique qui lui est dédié. Un support de cours lui est remis soit en début soit en fin de cours. La théorie est complétée par des cas pratiques ou exercices corrigés et discutés avec le formateur. Le formateur projette une présentation pour animer la formation et reste disponible pour répondre à toutes les questions.

Méthode d'évaluation

Tout au long de la formation, les exercices et mises en situation permettent de valider et contrôler les acquis du stagiaire. En fin de formation, le stagiaire complète un QCM d'auto-évaluation.

Suivre cette formation à distance

Voici les prérequis techniques pour pouvoir suivre le cours à distance :

- | Un ordinateur avec webcam, micro, haut-parleur et un navigateur (de préférence Chrome ou Firefox). Un casque n'est pas nécessaire suivant l'environnement.
- | Une connexion Internet de type ADSL ou supérieure. Attention, une connexion Internet ne permettant pas, par exemple, de recevoir la télévision par Internet, ne sera pas suffisante, cela engendrera des déconnexions intempestives du stagiaire et dérangera toute la classe.
- | Privilégier une connexion filaire plutôt que le Wifi.
- | Avoir accès au poste depuis lequel vous suivrez le cours à distance au moins 2 jours avant la formation pour effectuer les tests de connexion préalables.
- | Votre numéro de téléphone portable (pour l'envoi du mot de passe d'accès aux supports de cours et pour une messagerie instantanée autre que celle intégrée à la classe virtuelle).
- | Selon la formation, une configuration spécifique de votre machine peut être attendue, merci de nous contacter.
- | Pour les formations incluant le passage d'une certification la dernière journée, un voucher vous est fourni pour passer l'examen en

ligne.
| Pour les formations logiciel (Adobe, Microsoft Office...), il est nécessaire d'avoir le logiciel installé sur votre machine, nous ne fournissons pas de licence ou de version test.
| Horaires identiques au présentiel.

Accessibilité

Les sessions de formation se déroulent sur des sites différents selon les villes ou les dates, merci de nous contacter pour vérifier l'accessibilité aux personnes à mobilité réduite.
Pour tout besoin spécifique (vue, audition...), veuillez nous contacter au 01 85 77 07 07.