

Mis à jour le 12/04/2024

S'inscrire

Formation Google Cloud Dataform

2 jours (14 heures)

Présentation

Depuis son émergence, Dataform s'est imposé comme un pilier dans le domaine de la gestion des données dans le cloud. Cet outil, qui se positionne comme une alternative novatrice, offre aux équipes analytiques la possibilité de créer des pipelines de données efficaces en s'inspirant des meilleures pratiques de l'ingénierie logicielle.

Au cours de cette formation Google Cloud Dataform, vous apprendrez à configurer un environnement de développement spécifique à Dataform, ce qui vous permettra de développer, tester et déployer vos premiers modèles de données.

Vous découvrirez également comment [Dataform](#) s'intègre parfaitement dans les architectures de données actuelles et comment il répond aux défis les plus pressants auxquels sont confrontées les entreprises en matière de gestion des données.

À la fin de cette formation, vous maîtriserez les fonctionnalités principales de Dataform. Nous vous présenterons également les dernières avancées de Dataform avec la [version Core v2.8](#).

Objectifs

- Comprendre l'importance de la gestion des données dans les entreprises
- Configurer un projet Dataform et établir des connexions à des sources de données variées
- Développer des modèles de données avec SQL
- Automatiser les workflows de données
- Collaborer efficacement sur des projets de données

Public visé

- Data Scientists
- Data Analysts
- Chefs de projet

- Ingénieurs de données
- Business Analysts

PRÉ-REQUIS

- Avoir un compte Google Cloud Platform (GCP) actif
- Maîtriser les concepts de base de la gestion des données
- Disposer d'un compte GitHub lié au compte Gmail

PROGRAMME DE NOTRE FORMATION DATAFORM

INTRODUCTION À DATAFORM ET GESTION DES DONNÉES

- Présentation de Dataform et de son rôle dans la gestion des données modernes
- Importance de la gestion des données dans les entreprises
- Concepts clés : données, bases de données, entrepôts de données
- Présentation de l'interface utilisateur de Dataform et de ses composants principaux
- Création d'un compte et configuration initiale de l'environnement de travail

CONFIGURATION DU PROJET DATAFORM

- Création et configuration d'un fichier de projet dataform.json
- Établissement de la connexion à des sources de données variées
- Définition des dépendances entre les tables et les fichiers de modélisation
- Gestion des environnements de développement, de test et de production
- Bonnes pratiques pour la structuration de projets de données

DÉVELOPPEMENT DE MODÈLES DE DONNÉES AVEC SQL

- Introduction à la modélisation des données relationnelles : entités, relations, attributs
- Création de tables, vues et fonctions avec Dataform et SQL
- Techniques de transformation, de filtrage et d'agrégation des données
- Exercices pratiques sur des cas d'usage réels
- Conseils pour l'écriture de code SQL propre et maintenable

AUTOMATISATION DES WORKFLOWS DE DONNÉES

- Présentation des tâches Dataform et de leur configuration
- Planification et exécution des tâches : déclencheurs et planificateurs
- Gestion des dépendances et orchestration des workflows
- Surveillance en temps réel des exécutions de tâches et débogage
- Exemples pratiques d'automatisation des workflows de données

COLLABORATION ET PARTAGE DANS LES PROJETS DE DONNÉES

- Techniques de collaboration au sein des équipes de données
- Utilisation des fonctionnalités de révision et de partage de code de Dataform
- Gestion des accès et des permissions pour sécuriser les projets
- Mise en place de notifications et de commentaires pour améliorer la communication
- Stratégies de résolution de conflits et de révisions de code

SURVEILLANCE, OPTIMISATION ET DÉPANNAGE

- Introduction aux principes de surveillance des pipelines de données
- Techniques d'optimisation des performances des requêtes SQL
- Utilisation des index, des partitions et des agrégations dans Dataform
- Stratégies pour détecter et résoudre les problèmes de performance
- Étude de cas : dépannage d'un workflow de données avec Dataform

Sociétés concernées

Cette formation s'adresse à la fois aux particuliers ainsi qu'aux entreprises, petites ou grandes, souhaitant former ses équipes à une nouvelle technologie informatique avancée ou bien à acquérir des connaissances métiers spécifiques ou des méthodes modernes.

Positionnement à l'entrée en formation

Le positionnement à l'entrée en formation respecte les critères qualité Qualiopi. Dès son inscription définitive, l'apprenant reçoit un questionnaire d'auto-évaluation nous permettant d'apprécier son niveau estimé sur différents types de technologies, ses attentes et objectifs personnels quant à la formation à venir, dans les limites imposées par le format sélectionné. Ce questionnaire nous permet également d'anticiper certaines difficultés de connexion ou de sécurité interne en entreprise (intraentreprise ou classe virtuelle) qui pourraient être problématiques pour le suivi et le bon déroulement de la session de formation.

Méthodes pédagogiques

Stage Pratique : 60% Pratique, 40% Théorie. Support de la formation distribué au format numérique à tous les participants.

Organisation

Le cours alterne les apports théoriques du formateur soutenus par des exemples et des séances de réflexions, et de travail en groupe.

Validation

À la fin de la session, un questionnaire à choix multiples permet de vérifier l'acquisition correcte des compétences.

Sanction

Une attestation sera remise à chaque stagiaire qui aura suivi la totalité de la formation.

