
From Data to Insights with Google Cloud

Durée: 2 Jours **Réf de cours: GO6589** **Méthodes d'apprentissage: Virtual Learning**

Résumé:

Apprenez à interroger et traiter des pétaoctets de données en quelques secondes. Approfondissez vos connaissances sur l'analyse des données qui évolue automatiquement à mesure que vos données augmentent.

Dans le cours Data Insights, les participants tirent des enseignements grâce à l'analyse et à la visualisation des données à l'aide de Google Cloud Platform. Le cours présente des scénariis interactifs et des laboratoires pratiques où les participants explorent, extraient, chargent, visualisent et extraient des informations à partir de divers ensembles de données Google BigQuery.

Seront également abordés le chargement des données, l'interrogation, la modélisation de schéma, l'optimisation des performances, la tarification des requêtes et la visualisation des données.

Mise à jour : 11.05.2023

Public visé:

Cette formation s'adresse aux Data Analysts, Business Analysts, Business Intelligence, et Cloud Data Engineers qui s'associeront à des Data Analysts pour créer des solutions de données évolutives sur Google Cloud Platform

Objectifs pédagogiques:

- A l'issue de la formation, les participants seront capables de :
 - Explorer et visualiser les données à l'aide de Google Data Studio
 - Obtenir des informations à partir de données à l'aide des outils d'analyse et de visualisation de Google Cloud Platform
 - Optimiser et écrire des requêtes hautes performances et résoudre les problèmes associés
 - Analyser des données à l'échelle avec BigQuery
 - Dériver des informations à partir de données à l'aide du Machine Learning
 - Charger, nettoyer et transformer les données à grande échelle à l'aide de Google Cloud Dataprep
-

Pré-requis:

- Maîtriser ANSI SQL
-

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

Acquerrez une vue plus large de Google Cloud Platform à l'aide du Big Data et des principes de base du ML

Contenu:

Introduction à Google Cloud Platform

- Mettre en évidence les défis analytiques rencontrés par les analystes de données
- Comparer le Big Data On-Premises vs dans le Cloud
- Apprendre des cas d'utilisation réels d'entreprises transformées grâce à l'analyse dans le cloud
- Parcourir les bases d'un projet Google Cloud Platform

Analyse de grands ensembles de données avec BigQuery

- Parcourir les tâches et défis d'un analyste de données et introduction aux outils de données sur Google Cloud Platform
- Démo: Analyser 10 milliards d'enregistrements avec Google BigQuery
- Explorer 9 fonctionnalités fondamentales de Google BigQuery
- Comparer les outils GCP pour les analystes, les scientifiques et les ingénieurs de données
- Lab : Principes de base de BigQuery

Exploration de l'ensemble de données public avec SQL

- Comparer les techniques courantes d'exploration de données
- Apprendre à coder du SQL standard de haute qualité
- Explorer les jeux de données publics Google BigQuery
- Aperçu de la visualisation: Google Data Studio
- Lab : Explorer votre ensemble de données de commerce électronique avec SQL dans Google BigQuery

Nettoyage et transformation des données avec Cloud Dataprep

- Examiner les 5 principes de l'intégrité des ensembles de données
- Caractériser la forme et le biais d'un jeu de données
- Nettoyer et transformer des données à l'aide de SQL
- Nettoyer et transformer les données à l'aide d'une nouvelle interface utilisateur: présentation de Cloud Dataprep
- Lab : Création d'un pipeline de transformation de données avec Cloud Dataprep

Visualisation des informations et création de requêtes planifiées

- Présentation des principes de visualisation des données

Stockage et ingestion de nouveaux ensembles de données

- Comparer les tables permanentes et temporaires
- Enregistrer et exporter les résultats de requête
- Aperçu des performances: cache de requête
- Lab : Ingestion de nouveaux jeux de données dans BigQuery

Enrichissement de l'entrepôt de données avec JOINS

- Fusionner les tables de données historiques avec UNION
- Présentation des wildcards de tables pour des fusions faciles
- Examen des schémas de données: liaison de données entre plusieurs tables
- Exemples de jointures et pièges liés aux jointures
- Lab : Dépannage et résolution des problèmes liés aux jointures

Partitionnement des requêtes et tables pour des informations avancées

- Examen des instructions SQL Case
- Présentation des fonctions de fenêtre analytique
- Protéger vos données avec le cryptage de champs unidirectionnel
- Discussion autour de la conception efficace de sous-requêtes et CTE
- Comparer les UDF SQL et Javascript
- Lab : Création de tables partitionnées par date dans BigQuery

Conception de schémas à l'échelle : tableaux et structures dans BigQuery

- Comparer Google BigQuery à l'architecture des bases de données RDBMS traditionnelle
- Normalisation vs dénormalisation: compromis sur les performances
- Revue de schéma: le bon, le mauvais et le laid
- Tableaux et données imbriquées dans Google BigQuery
- Lab : Interrogation de données imbriquées et répétées
- Lab : Conception de schéma pour la performance: tableaux et structures dans BigQuery

Optimisation des requêtes pour la performance

- Exploration d'un job BigQuery
- Calculer les tarifs BigQuery: coûts de

Contrôle de l'accès avec les meilleures pratiques de sécurité des données

- Meilleures pratiques de sécurité des données
- Contrôle des accès avec les vues autorisées

Prédiction des achats de retour des visiteurs avec BigQuery ML

- Introduction au ML
- Sélection de caractéristiques
- Types de modèles
- Machine learning dans BigQuery
- Lab : Prédire les achats des visiteurs avec un modèle de classification BigQuery ML

Dérivation des informations à partir de données non structurées à l'aide du Machine Learning

- ML structuré vs non structuré
- Modèles ML préconstruits
- Atelier: Extraire, analyser et traduire du texte à partir d'images avec les API Cloud ML
- Lab : Entraînement avec des modèles ML prédéfinis à l'aide de l'API Cloud Vision et d'AutoML

- Approches exploratoires vs approches explicatives
- Démonstration: interface utilisateur de Google Data Studio
- Connecter Google Data Studio à Google BigQuery
- Lab : Comment créer un tableau de bord BI à l'aide de Google Data Studio et BigQuery

- stockage, de requête et de streaming
- Optimiser les requêtes pour le coût

Méthodes pédagogiques :

Support de cours officiel Google Cloud 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.
- Modalités d'évaluation : le participant est invité à s'auto-évaluer par rapport aux objectifs énoncés.
- Chaque participant, à l'issue de la formation, répond à un questionnaire de satisfaction qui est 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 vous utilisez votre Compte Personnel de Formation pour financer votre inscription, vous devrez respecter un délai minimum et non négociable fixé à 11 jours ouvrés.