

From Data to Insights with Google Cloud Platform

Des données aux informations avec Google Cloud Platform

Cours Pratique de 3 jours

Réf : GCE - Prix 2022 : 2 750€ HT

Avec cette formation, vous apprendrez à extraire des insights de données à grande échelle à l'aide de BigQuery, l'entrepôt de données cloud sans serveur, hautement évolutif et rentable de Google Cloud. Vous découvrirez les principes fondamentaux de BigQuery, notamment comment créer un pipeline de transformation de données, créer un tableau de bord BI, ingérer de nouveaux ensembles de données et concevoir des schémas à grande échelle.

OBJECTIFS PÉDAGOGIQUES

À l'issue de la formation l'apprenant sera en mesure de :

Obtenir des insights à partir des données et à l'aide des outils d'analyse et de visualisation sur Google Cloud Platform

Charger, nettoyer et transformer les données à grande échelle avec Google Cloud Dataprep

Explorer et visualiser les données à l'aide de Google Data Studio

Dépanner, optimiser et écrire des requêtes hautes performances

S'entraîner avec des API de ML prédéfinies pour la compréhension des images et du texte

Former la classification et la prévision des modèles de ML à l'aide de SQL avec BQML

MÉTHODES PÉDAGOGIQUES

Animation de la formation en français.
Support de cours officiel en anglais.

CERTIFICATION

Cours officiel sans certification.

PARTICIPANTS

Data analysts, business analysts, professionnels de la business intelligence.

PRÉREQUIS

Maîtrise de base de ANSI SQL.

COMPÉTENCES DU FORMATEUR

Les experts qui animent la formation sont des spécialistes des matières abordées. Ils sont agréés par l'éditeur et sont certifiés sur le cours. Ils ont aussi été validés par nos équipes pédagogiques tant sur le plan des connaissances métiers que sur celui de la pédagogie, et ce pour chaque cours qu'ils enseignent. Ils ont au minimum trois à dix années d'expérience dans leur domaine et occupent ou ont occupé des postes à responsabilité en entreprise.

MODALITÉS D'ÉVALUATION

Évaluation des compétences visées en amont de la formation.

Évaluation par le participant, à l'issue de la formation, des compétences acquises durant la formation.

Validation par le formateur des acquis du participant en précisant les outils utilisés : QCM, mises en situation...

À l'issue de chaque stage, ITTCERT fournit aux participants un questionnaire d'évaluation du cours qui est ensuite analysé par nos équipes pédagogiques. Les participants réalisent aussi une évaluation officielle de l'éditeur. Une feuille d'émargement par demi-journée de présence est fournie en fin de formation ainsi qu'une attestation de fin de formation si le participant a bien assisté à la totalité de la session.

MOYENS PÉDAGOGIQUES ET TECHNIQUES

Les ressources pédagogiques utilisées sont les supports et les travaux pratiques officiels de l'éditeur.

MODALITÉS ET DÉLAIS D'ACCÈS

L'inscription doit être finalisée 24 heures avant le début de la formation.

ACCESSIBILITÉ AUX PERSONNES HANDICAPÉES

Vous avez un besoin spécifique d'accessibilité ? Contactez Mme FOSSE, référente handicap, à l'adresse suivante psh-accueil@orsys.fr pour étudier au mieux votre demande et sa faisabilité.

LE PROGRAMME

dernière mise à jour : 09/2021

1) Introduction aux données sur Google Cloud Platform

- Mettre en évidence les défis analytiques rencontrés par les analystes de données.
- Comparer le big data on-premises versus 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.

Travaux pratiques : Premiers pas avec Google Cloud Platform

2) Présentation des outils de big data

- Présentation des tâches, des défis et des outils de données de Google Cloud Platform.
- Démonstration : analyser 10 milliards d'enregistrements avec Google BigQuery.
- Découvrir les neuf fonctionnalités fondamentales de Google BigQuery.
- Comparer les outils GCP pour les analystes, les scientifiques de données et les ingénieurs de données.

Travaux pratiques : Explorer des ensembles de données avec Google BigQuery.

3) Exploration de vos données avec SQL

- Comparer les techniques courantes d'exploration de données.
- Apprendre à coder du SQL standard de haute qualité.
- Explorer les ensembles de données publics Google BigQuery.
- Aperçu de la visualisation : Google Data Studio.

Travaux pratiques : Résoudre les erreurs SQL courantes.

4) Tarification de Google BigQuery

- Présentation d'une tâche BigQuery.
- Calculer les tarifs de BigQuery : coûts de stockage, d'interrogation et de streaming.
- Optimiser les requêtes pour le coût.

Travaux pratiques : Calculer les tarifs de Google BigQuery.

5) Nettoyage et transformation de vos données

- Examiner les cinq 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.

Travaux pratiques : Explorer et mettre en forme des données avec Cloud Dataprep

6) Stockage et exportation de données

- Comparer les tables permanentes et temporaires.
- Enregistrer et exporter les résultats de requête.
- Performance Preview : Query Cache.

Travaux pratiques : Créer de nouvelles tables permanentes.

7) Ingestion de nouveaux ensembles de données dans Google BigQuery

- Créer des requêtes à partir de sources de données externes.
- Éviter les pièges de l'ingestion de données.
- Ingérer de nouvelles données dans des tables permanentes.
- Discuter des inserts en streaming.

Travaux pratiques : Ingérer et interroger de nouveaux ensembles de données.

8) Visualisation des données

- Présentation des principes de visualisation des données.
- Approches d'analyses exploratoires versus explicatives.
- Démonstration : Google Data Studio UI.
- Connexion de Google Data Studio à Google BigQuery.

Travaux pratiques : Explorer un ensemble de données dans Google Data Studio

9) Union et fusion des ensembles de données

- Fusionner des tables de données historiques avec UNION.
- Introduire des caractères génériques de table pour des fusions faciles.
- Passez en revue les schémas de données : liaison de données entre plusieurs tables.
- Procédure pas à pas JOIN Exemples et pièges.

Travaux pratiques : Joindre et regrouper des données à partir de plusieurs tables.

10) Fonctions et clauses avancées

- Passer en revue les instructions de cas SQL.
- Présentation des fonctions de fenêtre analytique.
- Protection des données avec le cryptage de champs unidirectionnel.
- Discussion autour de la conception efficace des sous-requêtes et des CTE.
- Comparaison entre les UDF SQL et JavaScript.

Travaux pratiques : Dériver des informations avec des fonctions SQL avancées.

11) Conception de schémas et structures de données imbriquées

- Comparer l'architecture de données Google BigQuery à l'architecture de données SGBDR traditionnelle.
- Normalisation versus dénormalisation : compromis sur les performances.
- Revue de schéma : le bon, le mauvais et le laid.
- Tableaux et données imbriquées dans Google BigQuery.

Travaux pratiques : Interroger des données imbriquées et répétées.

12) Plus de visualisation avec Google Data Studio

- Créer des instructions Case et des champs calculés.
- Éviter les pièges de performance avec des considérations de cache.
- Partager les tableaux de bord et discuter de l'accès aux données.

13) Optimisation de la performance

- Éviter les pièges de performance de Google BigQuery.
- Empêcher les points d'accès dans ses données.
- Diagnostiquer les problèmes de performances avec la carte d'explication des requêtes.

Optimisation et dépannage des performances des requêtes.

14) Accès aux données

- Comparer les rôles d'ensemble de données IAM et BigQuery.
- Éviter les pièges d'accès.
- Passer en revue les membres, les rôles, les organisations, l'administration des comptes et les comptes de service.

15) Notebooks in the cloud

- Cloud Datalab.
- Compute Engine et Cloud Storage.
- Analyse de données avec BigQuery.

Travaux pratiques : Rent-a-VM pour traiter les données sur les tremblements de terre.

16) Google et le machine learning

- Introduction au machine learning pour les analystes.
- Entraînement avec les API de ML pré-entraînées pour la compréhension des images et du texte.

Travaux pratiques : Pretrained ML APIs.

17) Application de l'apprentissage automatique aux ensembles de données BQML

- Création d'ensembles de données machine learning et analyse des fonctionnalités.
- Création de modèles de classification et de prévision avec BQML.

Travaux pratiques : Prédire les achats des visiteurs avec un modèle de classification dans BQML. Prédire les tarifs des taxis à l'aide d'un modèle de prévision BigQuery ML.

LES DATES

CLASSE A DISTANCE

2022 : 21 nov.