

BigQuery for Data Analysts

La formation BigQuery pour apprendre à ingérer, transformer et interroger vos données avec BigQuery

2 jours / 14h

Présentation du cours

Cette formation est conçue pour les analystes de données souhaitant apprendre à utiliser BigQuery pour leurs besoins en analyse de données.

À travers une combinaison de présentations, de travaux pratiques et de démonstrations, nous couvrons divers sujets portant sur l'ingestion, la transformation et l'interrogation de vos données dans BigQuery afin d'en extraire des insights utiles à la prise de décisions métier.

Méthodes mobilisées : Ce cours alterne parties théoriques sous forme de lectures (slides), démos et parties pratiques sous forme de labs dirigés.

Objectifs pédagogiques

- Comprendre l'objectif et la valeur de BigQuery, l'entrepôt de données d'entreprise de Google Cloud, et découvrir ses fonctionnalités d'analyse de données.
- Analyser de grands ensembles de données dans BigQuery à l'aide de SQL.
- Nettoyer et transformer vos données dans BigQuery avec SQL.
- Importer de nouveaux jeux de données dans BigQuery et explorer les options pour les sources de données externes.
- Revoir les principes de visualisation et utiliser Connected Sheets et Looker Studio pour représenter visuellement les insights issus de BigQuery.

- Utiliser Dataform pour développer des pipelines de transformation de données à grande échelle dans BigQuery.
- Profiter des nouvelles intégrations et des capacités d'assistance introduites avec BigQuery Studio.

Modalités d'évaluation : Les objectifs pédagogiques sont évalués à travers la réalisation des parties pratiques (labs dirigés) sous la supervision du formateur délivrant la session de formation.

Public cible

- Analystes de données souhaitant apprendre à utiliser BigQuery pour leurs besoins en analyse de données.

Prérequis

Avoir suivi le cours [Introduction to Data Analytics on Google Cloud](#) ou avoir des connaissances équivalentes

Programme

Module 01: BigQuery pour analyste de données :

- Aperçu
- Analyse de données sur Google Cloud
- Des données aux insights avec BigQuery
- Cas d'usage réels d'entreprises transformées grâce à l'analyse sur Google Cloud

Objectifs :

- Identifier les défis analytiques des data analysts et comparer le big data sur site et dans le cloud
- Comprendre l'objectif et la valeur de BigQuery, l'entrepôt de données d'entreprise de Google Cloud, et découvrir ses fonctionnalités d'analyse de données

Module 02: Explorer et préparer vos données avec BigQuery

Sujets :

- Aperçu
- Techniques courantes d'exploration de données
- Analyse de grands ensembles de données avec BigQuery
- Principes de base des requêtes
- Utilisation des fonctions
- Enrichissement de vos requêtes avec UNION et JOIN

Objectifs :

- Lister les techniques courantes d'exploration de données
- Réviser les bases des requêtes SQL
- Enrichir les requêtes avec des fonctions, des unions et des jointures

Activités :

- Exploration d'un jeu de données e-commerce avec SQL dans Google BigQuery
- Dépannage des erreurs SQL courantes avec BigQuery
- Dépannage et résolution des pièges de jointure de données

Module 03: Nettoyer et transformer vos données

Sujets :

- Vue d'ensemble
- Cinq principes d'intégrité d'un jeu de données
- Nettoyer et transformer des données avec SQL
- Nettoyer et transformer des données : autres options

Objectifs :

- Identifier les critères d'un bon jeu de données
- Nettoyer et transformer des données avec SQL
- Nettoyer et transformer des données avec d'autres options

Module 04: Ingestion et stockage de nouveaux jeux de données BigQuery

Sujets :

- Vue d'ensemble
- Tables de données permanentes vs temporaires
- Ingestion de nouveaux jeux de données
- Sources de données externes

Objectifs :

- Passer en revue les différences entre les tables permanentes et temporaires
- Ingest et stocker de nouveaux jeux de données BigQuery
- Discuter des options pour les sources de données externes

Activités :

- Création de nouvelles tables permanentes
- Ingestion et interrogation de nouveaux jeux de données

Module 05: Visualiser vos insights issus de BigQuery

Sujets :

- Vue d'ensemble
- Principes de visualisation de données
- Connected Sheets
- Pièges courants en visualisation de données
- Looker Studio
- Analyse dans un notebook

Objectifs :

- Passer en revue les différences entre les tables permanentes et temporaires
- Ingest et stocker de nouveaux jeux de données BigQuery
- Discuter des options pour les sources de données externes

Activités :

- Démarrage rapide de Connected Sheets
- Explorer et créer des rapports avec Looker Studio

Module 06: Développer des pipelines de transformation de données évolutifs dans BigQuery avec Dataform

Sujets :

- Vue d'ensemble
- Qu'est-ce que Dataform ?
- Prise en main de Dataform

Objectifs :

- Utiliser Dataform pour développer des pipelines de transformation de données évolutifs dans BigQuery
- Apprendre à démarrer avec Dataform en créant un dépôt et un espace de développement

- Créer et exécuter un workflow SQL dans Dataform

Activités :

- Démo
- Créer et exécuter un workflow SQL dans Dataform

Module 07: BigQuery Studio

Sujets :

- BigQuery Studio : quoi et pourquoi ?
- Analytique unifiée
- Gestion des assets
- Assistance intégrée

Objectifs :

- Présenter BigQuery Studio
- Utiliser Duet AI dans BigQuery pour expliquer et générer des requêtes SQL
- Découvrir les nouvelles fonctionnalités d'ergonomie et les intégrations avec Dataform et Dataplex dans la nouvelle interface BigQuery Studio

Activités :

- Démo
- Analyser des données avec l'assistance de Duet AI
- Générer du contenu d'email personnalisé avec les requêtes continues BigQuery et Gemini