

Google Cloud Platform - Analyse et visualisation de données

Obtenez des informations via l'analyse et la visualisation de données grâce à Google Cloud Platform

 Présentiel ou en classe à distance

Durée : 3 jours (21 h)

Réf. : CC403

Prix inter : 2.550,00 € HT

Depuis ses débuts comme plateforme cloud, la Google Cloud Platform (GCP) a connu une expansion fulgurante. Du stockage à l'analyse des données, la GCP s'est étoffée, offrant une panoplie complète de services. Les Data Analysts, Business Analysts et Cloud Data Engineers ont vu leurs rôles se métamorphoser avec cette croissance.

Notre formation répond à cette évolution

En trois jours, vous pourrez acquérir les compétences vous permettant de tirer le meilleur de vos données. Que vous soyez un analyste chevronné ou un ingénieur de données cloud, cette formation vous offre les clés pour naviguer dans le monde complexe de l'analyse et de la visualisation des données au sein de la GCP.

Les objectifs de la formation

- Pouvoir obtenir des informations à partir de données à l'aide des outils d'analyse et de visualisation de Google Cloud Platform
- Savoir analyser des données à l'échelle avec BigQuery
- Comprendre comment charger, nettoyer et transformer les données à grande échelle à l'aide de Google Cloud Dataprep
- Être capable d'explorer et visualiser les données à l'aide de Google Data Studio
- Apprendre à optimiser et écrire des requêtes hautes performances et résoudre les problèmes associés
- Savoir dériver des informations à partir de données à l'aide du Machine Learning

A qui s'adresse cette formation ?

Pour qui

- Data Analysts, Business Analysts, Business Intelligence, Cloud Data Engineers qui s'associeront à des Data Analysts pour créer des solutions de données évolutives sur Google Cloud Platform

Prérequis

- Posséder une certaine maîtrise d'ANSI SQL

Programme

1 - Présentation des données dans Google Cloud Platform

- Mettre en avant les difficultés auxquelles font face les analystes de données
- Comparer le Big Data sur site et dans le Cloud
- Étudier des cas d'utilisation concrets d'entreprises qui se sont transformées grâce à l'analyse dans le Cloud
- Découvrir les principes de base du projet Google Cloud Platform
- Atelier : 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 9 fonctionnalités essentielles de Google BigQuery
- Comparer les outils GCP pour les analystes, les data scientists et les ingénieurs de données
- Atelier : Explorer des ensembles de données avec Google BigQuery

3 - Exploration de vos données avec SQL

- Comparer les techniques courantes d'exploration des données
- Apprendre à coder le langage SQL standard de haute qualité
- Explorer les ensembles de données publics Google BigQuery
- Aperçu de visualisation : Google Data Studio
- Atelier : Résoudre les erreurs SQL courantes

4 - Tarifs de Google BigQuery

- Présentation complète d'une tâche BigQuery
- Calculer les tarifs de BigQuery : coûts du stockage, des requêtes et de streaming
- Optimiser les coûts des requêtes
- Atelier : Calculer les tarifs de Google BigQuery

5 - Nettoyage et transformation de vos données

- Étudier les 5 principes de l'intégrité d'un ensemble de données
- Caractériser la forme et l'inclinaison des ensembles de données
- Nettoyer et transformer les données à l'aide de SQL
- Nettoyer et transformer les données à l'aide d'une nouvelle interface utilisateur : présentation de Cloud Dataprep
- Atelier : Explorer et mettre en forme les données avec Cloud Dataprep

6 - Stockage et exportation des données

- Comparer les tables permanentes aux tables temporaires
- Enregistrer et exporter les résultats de requêtes
- Aperçu des performances : cache des requêtes
- Atelier : Créer des tables permanentes

7 - Ingestion de nouveaux ensembles de données dans Google BigQuery

- Requête à partir de sources de données externes
- Éviter les pièges liés à l'ingestion de données
- Ingérer de nouvelles données dans des tables permanentes
- Discuter des insertions en streaming
- Atelier : Ingérer et interroger de nouveaux ensembles de données

8 - Visualisation des données

- Présentation des principes de visualisation des données
- Comparer l'analyse exploratoire à l'analyse explicative
- Démonstration : Interface utilisateur de Google Data Studio
- Connecter Google Data Studio à Google BigQuery
- Atelier : Explorer un ensemble de données dans Google Data Studio

9 - Regroupement et fusion des ensembles de données

- Fusionner les tables de données historiques avec UNION
- Intégrer des caractères génériques de table pour faciliter la fusion
- Étudier les schémas de données : associer les données entre plusieurs tables
- Présentation complète d'exemples JOIN et des pièges associés
- Atelier : Associer et regrouper des données provenant de plusieurs tables

10 - Fonctions avancées et clauses

- Passer en revue les instructions de cas SQL
- Présentation des fonctions de fenêtre d'analyse
- Sauvegarder les données avec le cryptage de champ unidirectionnel
- Discuter de la conception efficace des sous-requête et des CTE
- Comparer les fonctions définies par l'utilisateur dans SQL et JavaScript
- Atelier : Obtenir des informations grâce aux fonctions SQL avancées

11 - Conception de schémas et structures de données imbriquées

- Comparer Google BigQuery à l'architecture de données SGBDR traditionnelle
- Normalisation et dénormalisation : compromis de performance
- Étude du schéma : Le Bon, la Brute et le Truand
- Tableaux et données imbriquées dans Google BigQuery
- Atelier : Interroger des données imbriquées et répétées

12 - Visualisation améliorée avec Google Data Studio

- Créer des instructions de cas et des champs calculés
- Éviter les problèmes de performance grâce au cache
- Partager les tableaux de bord et discuter de l'accès aux données

13 - Optimisation des performances

- Éviter les problèmes de performance de Google BigQuery
- Empêcher les points d'accès dans vos données
- Diagnostiquer les problèmes de performances grâce au mappage Query Explanation
- Atelier : Optimisation et dépannage des performances de requête

14 - Accès aux données

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

15 - Notebooks dans le Cloud

- Cloud Datalab
- Compute Engine et Cloud Storage
- Atelier : Louer une VM pour traiter des données sur des séismes
- Analyse de données avec BigQuery

16 - Google et le Machine Learning

- Introduction au Machine Learning pour les analystes
- S'exercer avec des API de ML prédéfinies pour analyser les images et le texte

- Atelier : API de ML pré-entraînées

17 - Application du Machine Learning à vos ensembles de données (BQML)

- Créer des ensembles de données de Machine Learning et analyser les fonctionnalités
- Créer des modèles de classification et de prévision à l'aide de BQML
- Atelier : Prédire les achats des visiteurs avec un modèle de classification dans BQML
- Atelier : Prédire le prix d'une course en taxi à l'aide d'un modèle de prévision BigQuery ML

Evaluation

- Cette formation fait l'objet d'une évaluation formative.

Les points forts de la formation

- Cette formation permet aux participants d'acquérir des connaissances grâce à l'analyse et à la visualisation de données avec de Google Cloud Platform.
- À travers des scénarios interactifs et des ateliers pratiques, les participants explorent, exploitent, chargent, visualisent et extraient des informations à partir de divers ensembles de données Google BigQuery.
- Les consultants spécialistes de la technologie apportent leurs conseils et leur expérience.
- Une formation animée par un formateur certifié Google Cloud Platform.
- La qualité d'une formation officielle Google (support de cours en anglais).

Dates et villes 2024 - Référence CC403

A distance

du 15 mai au 17 mai

du 10 juil. au 12 juil.

du 25 sept. au 27 sept.

du 20 nov. au 22 nov.