EduGroupe
Accompagner pour réussir

Google Cloud Platform - Form Data to Insights

Référence : **GCP200DA** Durée : **3 jours (21 heures)** Certification : **Aucune**

Connaissances préalables

- Maîtrise de base d'ANSI SQL (référence)

Profil des stagiaires

- 1-Data Analysts, Business Analysts, Business Intelligence professionals
- 2-Cloud Data Engineers who will be partnering with Data Analysts to build scalable data solutions on Google Cloud

Objectifs

- Aucun objectif spécifique

Certification préparée

- Aucune

Méthodes pédagogiques

- 6 à 12 personnes maximum par cours, 1 poste de travail par stagiaire
- Remise d'une documentation pédagogique papier ou numérique pendant le stage
- La formation est constituée d'apports théoriques, d'exercices pratiques et de réflexions

Formateur

- Consultant-Formateur expert Cloud

Méthodes d'évaluation des acquis

- Auto-évaluation des acquis par le stagiaire via un questionnaire
- Attestation des compétences acquises envoyée au stagiaire
- Attestation de fin de stage adressée avec la facture

Contenu du cours

1. 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

2. 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
- Atelier: Principes de base de BigQuery

3. Explorer votre 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
- Atelier: Explorer votre ensemble de données de commerce électronique avec SQL dans Google BigQuery

4. Nettoyage et transformation de vos 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
- Atelier: Création d'un pipeline de transformation de données avec Cloud Dataprep

5. Visualisation des informations et création de requêtes planifiées

- Présentation des principes de visualisation des données
- Approches exploratoires vs approches explicatives
- Démonstration: interface utilisateur de Google Data Studio
- Connecter Google Data Studio à Google BigQuery
- Atelier: Comment créer un tableau de bord BI à l'aide de Google Data Studio et BigQuery

6. 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
- Atelier: Ingestion de nouveaux jeux de données dans BigQuery

7. Enrichir votre 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
- Atelier: Dépannage et résolution des problèmes liés aux jointures

8. Partitionnement de vos 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
- Atelier: Création de tables partitionnées par date dans BigQuery

9. 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
- Atelier: Interrogation de données imbriquées et répétées
- Atelier: Conception de schéma pour la performance: tableaux et structures dans BigQuery

10. Optimisation des requêtes pour la performance

- Exploration d'un job BigQuery
- Calculer les tarifs BigQuery: coûts de stockage, de requête et de streaming
- Optimiser les requêtes pour le coût

11. 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

12. Prédire les achats de retour des visiteurs avec BigQuery ML

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

13. Dériver 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
- Atelier: Entraînement avec des modèles ML prédéfinis à l'aide de l'API Cloud Vision et d'AutoML

14. Conclusion

- Résumé du cours

Notre référent handicap se tient à votre disposition au 01.71.19.70.30 ou par mail à referent.handicap@edugroupe.com pour recueillir vos éventuels besoins d'aménagements, afin de vous offrir la meilleure expérience possible.