Maîtrise Avancée de BigQuery pour la Performance et l’Optimisation des Coûts

Lien : https://innov-systems.com/formation/matrise-avancee-de-bigquery-pour-la-performance-et-loptimisation-des-cots

DURÉE
3 jours (21h)

RÉFÉRENCE
VSC307

CATÉGORIE
Google Cloud

Innov Systems

OBJECTIFS DE LA FORMATION

À l'issue de cette formation, vous serez capable de :

- Approfondir la maîtrise du langage SQL dans BigQuery à travers les fonctions avancées et les techniques d’optimisation
- Comprendre les mécanismes internes de BigQuery pour améliorer la performance des requêtes et la gestion des ressources
- Mettre en œuvre des stratégies concrètes de réduction des coûts de stockage et de requêtage
- Exploiter les capacités analytiques et d’intelligence artificielle intégrées à BigQuery (ML, IA générative)
- Concevoir et administrer des environnements BigQuery optimisés, sécurisés et gouvernés

POUR QUI ?

- Développeurs de données
- Analystes BI et data engineers
- Administrateurs de plateformes data
- Responsables data souhaitant optimiser les performances et coûts BigQuery

Programme détaillé

1 / APPROFONDISSEMENT DU SQL DANS BIGQUERY

- Manipulations complexes : arrays, STRUCT et JSON
- Gestion avancée des dates, heures et chaînes de caractères
- Utilisation du mot-clé QUALIFY et des fonctions analytiques
- Requêtes dynamiques et transformations de données avec PIVOT et UNPIVOT

2 / AUTOMATISATION ET TRANSFORMATION DES DONNÉES

- Création et usage des fonctions SQL et table functions
- Développement de User-Defined Functions (UDF) en JavaScript ou SQL
- Usage de MERGE pour les mises à jour incrémentales et synchronisations
- Bonnes pratiques de structuration et de réutilisation des requêtes complexes

3 / GESTION DES TRANSACTIONS ET DES CONTEXTES D'EXÉCUTION

- Transactions, sessions et gestion de la cohérence des données
- Gestion des droits et des labels pour la gouvernance des ressources
- Introduction à Dataplex Catalog pour la classification et la traçabilité des données

4 / MISE EN PRATIQUE : CONCEPTION DE REQUÊTES OPTIMISÉES

- Atelier : création d'un pipeline SQL complet et réutilisable
- Exercice : implémentation d'un système d'audit des requêtes BigQuery

5 / ARCHITECTURE ET FONCTIONNEMENT INTERNE DE BIGQUERY

- Architecture de stockage distribuée (Colossus)
- Traitement et exécution parallèle des requêtes (Dremel)
- Gestion des ressources et orchestration (Borg)

Inov Systems

- Lecture et interprétation du plan d’exécution

6 / OPTIMISATION DES REQUÊTES ET DES TABLES

- Lecture et analyse des statistiques d’exécution
- Optimisation des jointures et réduction des scans inutiles
- Techniques de partitionnement et clustering des tables
- Utilisation des vues matérialisées pour les pré-agrégations et le smart tuning

7 / OPTIMISATION DES PERFORMANCES EN PRATIQUE

- Approximations et calculs rapides avec HyperLogLog et APPROX_*
- Exploitation du cache et du moteur BI Engine pour les analyses en mémoire
- Choix et impact des formats de fichiers : Avro, Parquet, ORC, JSON
- Diagnostic et correction d’un problème de lenteur dans un cas réel

8 / MISE EN PRATIQUE : OPTIMISATION ET DÉBOGAGE

- Atelier : amélioration d’une requête lente (étude de cas guidée)
- Exercice : évaluation comparative avant/après optimisation

9 / OPTIMISATION DES COÛTS DE STOCKAGE ET DE REQUÊTAGE

- Stockage logique vs physique : principes et impacts
- Imbrication et politiques de rétention des données
- Stratégies d’exécution : on-demand vs réservation de slots
- Suivi des coûts et mise en place d’alertes et de budgets automatisés

10 / GOUVERNANCE, SÉCURITÉ ET AUDIT

- Gestion fine des accès et des rôles IAM dans BigQuery
- Audit et traçabilité des actions utilisateurs
- Qualité des données : validation, vérification et scans automatiques

Inov Systems

- Intégration de Dataplex pour la conformité et la gouvernance globale

11 / INNOVATION ET INTELLIGENCE ARTIFICIELLE

- Initiation à BigQuery ML : création et évaluation de modèles de ML intégrés
- Introduction aux embeddings et intégration de l'IA générative
- Exemples d'utilisation pratique : prédictions, recommandations et text embeddings
- Exploration des perspectives futures de BigQuery avec Vertex AI et Looker

12 / MISE EN PRATIQUE FINALE : PROJET D'OPTIMISATION GLOBALE

- Atelier complet : réduction des coûts et accélération d'un projet BigQuery réel
- Analyse de la performance et simulation d'un modèle BigQuery ML
- Restitution et discussion autour des meilleures pratiques

Approche pédagogique

- Support Ecrit et Projection
- Exposés Interactifs, Podcasts et Vidéos
- Brainstorming et Jeux de Rôle
- Cas Pratiques et Labs inclus pour leur impact opérationnel
- Test de Validation des Acquis des Connaissances

Prochaines dates programmées

09 au 11 Déc. 2026	Distanciel
13 au 15 Jan. 2027	Présentiel - Casablanca
10 au 12 Mars 2027	Présentiel - Casablanca
05 au 07 Mai 2027	Présentiel - Casablanca
30 Juin au 02 Juil. 2027	Présentiel - Casablanca

Autres dates possibles sur demande. Contactez-nous pour organiser une session intra-entreprise.

Innov Systems

Réservation & Renseignements

Téléphone : +212 522 247 210
Email : contact@innov-systems.com
Web : https//www.innov-systems.com

Document généré le 25/11/2026 — Réf : VSC307
Innov Systems — Tous droits réservés

Innov Systems