Concevoir et Optimiser l’Entrepôt de Données dans BigQuery

Lien :
https://innov-systems.com/formation/concevoir-et-optimiser-lentrepot-de-donnees-dans-bigquery

DURÉE	RÉFÉRENCE	CATÉGORIE
4 jours (28h)	**VSC311**	**Google Cloud**

Innov Systems

OBJECTIFS DE LA FORMATION

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

- Comprendre en profondeur l’architecture, les principes de fonctionnement et les capacités de BigQuery
- Concevoir des structures de stockage et des schémas performants et évolutifs
- Mettre en œuvre des stratégies d’ingestion, de transformation et d’optimisation adaptées aux volumes de données massifs
- Gérer et automatiser les charges de travail analytiques tout en maîtrisant les coûts
- Sécuriser, superviser et industrialiser les environnements BigQuery dans une approche DataOps
- Exploiter BigQuery ML pour créer et déployer des modèles de machine learning intégrés

POUR QUI ?

- Data Engineers
- Data Analysts
- Architectes Cloud
- Responsables BI / Data Managers
- Développeurs spécialisés en données

Programme détaillé

1 / INTRODUCTION ET POSITIONNEMENT DE BIGQUERY

- Rôle de BigQuery dans un écosystème data moderne
- Architecture serverless et séparation stockage/calcul
- Modèle de facturation et optimisation des coûts
- Évolution et limites par rapport aux autres solutions de data warehouse cloud

2 / STRUCTURE DU STOCKAGE ET ARCHITECTURE INTERNE

- Fonctionnement du stockage en colonnes et de Dremel
- Compréhension du service de shuffling et de l'exécution distribuée
- Rôle du cache, des partitions et du clustering
- Bonnes pratiques pour l'organisation logique des datasets

3 / CONCEPTION DU SCHÉMA OPTIMISÉ

- Schéma en étoile vs schéma en flocon : cas d'usage
- Utilisation des champs imbriqués et répétés
- Gestion de l'expiration, des vues et des tables partitionnées
- Stratégies de modélisation orientées performance et coûts

4 / INGESTION DE DONNÉES DANS BIGQUERY

- Chargement batch vs streaming : différences et cas pratiques
- Connexion à des sources externes (Cloud Storage, Pub/Sub, Cloud SQL)
- Utilisation de Data Transfer Service et de l'API Storage Write
- Contrôle de la latence, fiabilité et validation des données

5 / TRANSFORMATION ET TRAITEMENT DES DONNÉES

Innov Systems

- Écriture et exécution d’instructions DML et DDL
- Gestion des Slowly Changing Dimensions (SCD)
- Création de scripts et procédures stockées pour les workflows complexes
- Intégration avec Dataform, Cloud Composer ou Dataflow pour l’orchestration

6 / AUTOMATISATION ET GESTION DES CHARGES DE TRAVAIL

- Planification des requêtes et automatisation des traitements
- Gestion de la capacité avec Reservations et Slot Commitments
- Optimisation du coût et suivi de la consommation
- Gouvernance et gestion des environnements multi-projets

7 / OPTIMISATION DES REQUÊTES ET PERFORMANCE

- Lecture du plan d’exécution et interprétation des métriques
- Optimisation des jointures, agrégations et fonctions analytiques
- Création de vues matérialisées et exploitation du cache
- Utilisation de BI Engine et Storage Read API pour accélérer les analyses

8 / SURVEILLANCE ET DÉPANNAGE

- Utilisation des tables INFORMATION_SCHEMA pour l’audit et le suivi
- Supervision via Cloud Monitoring et BigQuery Admin Panel
- Détection des requêtes coûteuses et identification des goulots d’étranglement
- Automatisation de la gouvernance via Cloud Logging et alertes

9 / SÉCURITÉ, GOUVERNANCE ET CONFORMITÉ

- Gestion des accès avec IAM, vues autorisées et politiques de ligne
- Découverte et classification des données via Data Catalog
- Application des politiques DLP (Data Loss Prevention)
- Chiffrement, audit et conformité réglementaire (RGPD, ISO, etc.)

Inov Systems

10 / MACHINE LEARNING AVEC BIGQUERY ML

- Introduction à BigQuery ML et intégration avec SQL
- Création et déploiement de modèles (régression, classification, clustering)
- Utilisation de AutoML Tables pour des modèles optimisés
- Analyse des performances et suivi du cycle de vie des modèles ML

11 / INTÉGRATION ET INDUSTRIALISATION DES PIPELINES ANALYTIQUES

- Interconnexion de BigQuery avec Looker, Data Studio, Vertex AI
- Stratégies DataOps pour le déploiement et la supervision continue
- Cas pratiques : pipeline complet de la donnée brute au tableau de bord
- Synthèse et recommandations pour l'exploitation à grande échelle

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

01 au 04 Déc. 2026	Distanciel
26 au 29 Jan. 2027	Présentiel - Casablanca
23 au 26 Mars 2027	Présentiel - Casablanca
18 au 21 Mai 2027	Présentiel - Casablanca
13 au 16 Juil. 2027	Présentiel - Casablanca

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

Inno Systems

Réservation & Renseignements

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

Document généré le 13/11/2026 — Réf : VSC311
Innov Systems — Tous droits réservés

Innov Systems