



PROGRAMME DE FORMATION

Google Cloud Platform — Data

Construisez votre plateforme data moderne sur Google Cloud

Organisme de formation : **Lior Agency**

Durée : **1 jour (7 heures)**

Modalité : **Présentiel ou distanciel synchrone**

VISION ET OBJECTIF PRINCIPAL

Vision et objectif principal

L'objectif de cette journée est de fournir aux équipes data une maîtrise opérationnelle de la plateforme data Google Cloud. De l'ingestion des données à leur valorisation en dashboards et modèles analytiques, les participants découvriront comment BigQuery, Dataflow et Looker s'articulent pour construire une plateforme data unifiée, scalable et gouvernée. À l'issue de cette formation, les participants seront capables de concevoir une architecture data moderne sur GCP et de démarrer leurs premiers pipelines et analyses en autonomie.

PUBLIC CIBLE	Data engineers, data analysts, architectes data, responsables BI et analytics
PRÉREQUIS	Connaissances de base en SQL. Une familiarité avec les concepts de data warehouse et de pipelines de données est recommandée.

MODULE 1

Module 1 : L'écosystème data Google Cloud

Durée : 1h | Format : Présentation + démonstration

Cartographier l'ensemble des services data GCP et comprendre comment ils s'articulent dans une architecture data moderne. Positionner BigQuery comme pièce maîtresse de l'écosystème.

Objectifs clés

- Comprendre l'architecture de référence data sur GCP (ingestion → stockage → traitement → analyse → visualisation)
- Positionner les services clés : Cloud Storage, BigQuery, Dataflow, Pub/Sub, Dataproc, Looker
- Identifier les patterns d'architecture courants : data lake, data warehouse, lakehouse
- Comprendre le modèle de facturation data (stockage vs. compute, à la demande vs. forfait)

MODULE 2

Module 2 : BigQuery — Le data warehouse serverless

Durée : 1h30 | Format : Présentation + atelier pratique

Maîtriser BigQuery, le data warehouse serverless de Google Cloud : de la structuration des datasets aux requêtes analytiques avancées, en passant par l'optimisation des coûts et des performances.

Objectifs clés

- Créer des datasets, tables et vues dans BigQuery
 - Écrire des requêtes analytiques avancées (fenêtrage, CTE, requêtes imbriquées)
 - Charger des données depuis Cloud Storage (batch) et Pub/Sub (streaming)
 - Optimiser les performances et les coûts : partitionnement, clustering, requêtes à la demande vs. slots
 - Comprendre BigQuery ML pour créer des modèles directement en SQL
-

MODULE 3

Module 3 : Ingestion et pipelines de données

Durée : 1h30 | Format : Présentation + atelier pratique

Construire des pipelines d'ingestion robustes pour alimenter la plateforme data. Couvrir les patterns batch et streaming avec les outils natifs GCP.

Objectifs clés

- Mettre en place une ingestion batch avec Cloud Storage et les jobs de chargement BigQuery
 - Comprendre Pub/Sub pour l'ingestion de données en temps réel (événements, logs, IoT)
 - Découvrir Dataflow (Apache Beam) pour le traitement de données en batch et streaming
 - Savoir quand utiliser Dataflow vs. Dataproc (Spark/Hadoop) vs. les fonctions SQL de BigQuery
 - Orchestrer des pipelines avec Cloud Composer (Apache Airflow managé)
-

MODULE 4

Module 4 : Gouvernance, qualité et sécurité des données

Durée : 1h | Format : Présentation + démonstration

Mettre en place une gouvernance data solide : catalogage, lignage, contrôle d'accès fin et conformité.

Objectifs clés

- Utiliser Dataplex pour le catalogage, le lignage et la gestion de la qualité des données
 - Configurer les contrôles d'accès à BigQuery : IAM, row-level security, column-level security
 - Comprendre les mécanismes de protection des données sensibles (DLP API, masquage dynamique)
 - Mettre en place des politiques de rétention et de cycle de vie des données
-

MODULE 5

Module 5 : Visualisation, analytics et prochaines étapes

Durée : 1h30 | Format : Présentation + atelier pratique

Valoriser les données en dashboards interactifs avec Looker et Looker Studio. Explorer les capacités analytiques avancées et définir la feuille de route data.

Objectifs clés

- Créer un dashboard dans Looker Studio connecté à BigQuery
- Comprendre le positionnement de Looker (modélisation sémantique LookML) vs. Looker Studio (BI en libre-service)
- Explorer les capacités d'analyse avancée : BigQuery ML, fonctions analytiques, Vertex AI
- Identifier les bonnes pratiques d'architecture data (Well-Architected Data Framework)
- Définir une roadmap data : prochaines étapes et quick wins pour l'organisation

MÉTHODE PÉDAGOGIQUE

Méthode pédagogique

-
- Alternance de présentations théoriques et de démonstrations live sur la console GCP
 - Ateliers pratiques guidés sur un projet GCP dédié avec des jeux de données réalistes
 - Support de formation remis aux participants (slides + guide des ateliers + requêtes SQL)
 - Évaluation des acquis en fin de formation (QCM + exercice de synthèse)