

Mis à jour le 11/04/2024

S'inscrire

Formation Préparation à la certification GCP Professional Data Engineer

3 jours (21 heures)

Présentation

Notre formation GCP Professional Data Engineer est conçue pour vous doter des compétences essentielles nécessaires à la création, à la gestion et à l'optimisation de solutions de données à grande échelle sur Google Cloud Platform.

Au cours de cette formation, vous développerez une expertise approfondie dans le domaine de l'ingénierie des données, en explorant en détail les fonctionnalités et les concepts clés de la plateforme GCP.

Vous apprendrez à appliquer ces connaissances de manière pratique pour concevoir des solutions de données efficaces et hautement évolutives.

Cette formation vous permettra d'acquérir une compréhension approfondie de concepts fondamentaux tels que le stockage des données, le traitement en temps réel, et les architectures de données distribuées.

Vous serez formé à configurer un environnement de développement optimal et à maîtriser le déploiement de solutions de données sur Google Cloud. Tout au long de la formation, nous mettrons l'accent sur l'utilisation efficace des outils et services GCP spécifiques à l'ingénierie des données.

Comme dans toutes nos formations, celle-ci vous présentera les [dernières ressources de GCP](#)

Objectifs

- Acquérir une compréhension approfondie du cycle de vie des données, des aspects techniques et des types de structures
- Comprendre les services GCP et les technologies hébergées

- Apprendre à concevoir et à opérationnaliser des systèmes de stockage efficaces
- Développer une expertise dans la conception de pipelines de données GCP
- Approfondir la conception et le choix de la structure de traitement des données

Public visé

- Professionnels du cloud
- Ingénieurs de données

Pré-requis

Des connaissances de base dans le domaine de l'ingénierie des données et une familiarité avec les services cloud seraient un atout.

Programme de notre formation GCP professional Data Engineer

Selection des technologies de stockages GCP

- Data Lifecycle (Ingest, Store, Process and Analyze, Explore and Visualize)
- Technical Aspects of Data: Volume, Velocity, Variation, Access, and Security
- Types of Structure
 - Structured
 - Semi-Structured
 - Unstructured
- Schema Design Considerations (Relational Design & NoSQL Design)

Construire et rendre opérationnels les systèmes de stockage

- Cloud SQL
- Cloud Spanner
- Cloud Bigtable
- Cloud Firestore
- BigQuery
- Cloud Memorystore
- Cloud Storage
- Unmanaged Databases (non autogérées)

Conception de pipelines de données

- Un aperçu des pipelines de données (Étapes / Types de données)
- les composantes / services des pipelines GCP
- Migrer Hadoop et Spark sur GCP

Processing / Traitement des données

- Conception et choix de la structure
- Conception d'un traitement distribué
- Migrer un DataWarehouse (Entrepôt de données)

Configuration de l'infrastructure de traitement

- Provisionnement et ajustement des ressources de traitement (Compute Engine, Kubernetes, Cloud Bigtable, Dataproc)
- Supervision des ressources de traitement (Stackdriver Monitoring, Stackdriver Logging, Stackdriver Trace)

La sécurité / Administration sur GCP

- Gestion des identités et des accès avec Cloud IAM
- Utiliser IAM avec les services de stockage et de traitement
- Data Security
- La confidentialité et l'API Data Loss Prevention

Conception de bases de données pour la fiabilité, l'évolutivité, et la disponibilité

- Conception de bases de données Bigtable (Scalability and Reliability)
- Conception de bases de données Cloud Spanner
- Conception de bases de données BigQuery

Comprendre les opérations de données pour plus de flexibilité et la portabilité

- Catalogue et découverte avec Data Catalog
- Traitement des données avec Data Prep
- Visualisation des données avec Data Studio (Looker studio)
- Exploration des données avec Cloud Datalab
- Orchestrer les Workflows avec Cloud Composer

Déployer des pipelines d'apprentissage automatique

- Structure des pipelines ML
 - Data Ingestion
 - Data Preparation
 - Data Segregation
 - Model Training
 - Model Evaluation
 - Model Deployment
 - Model Monitoring
- Options GCP pour déploiement des pipelines de ML
 - Cloud AutoML
 - BigQuery ML
 - Kubeflow
 - Spark Machine Learning

Choisir l'infrastructure d'entraînement

- Accélérateurs matériels (GPUs, TPUs, CPUs)
- Infrastructure distribuée et machine unique
- Edge Computing avec GCP (Cloud IOT, Edge TPU)

La supervision et les alertes

- Les trois types d'algorithmes d'apprentissage automatique
- Deep Learning
- Ingénierie des modèles d'apprentissage automatique
- Sources d'erreurs courantes dans les modèles d'apprentissage automatique

Exploiter les modèles préconstruits en tant que service

- Visio (Vision AI, Video AI)
- Conversation (Dialogflow, Cloud Text-to-Speech API, Cloud Speech-to-Text API)
- Langage (Traduction, Langage naturel)
- Données structurées (Recommendations AI API, Cloud Inference API)

Sociétés concernées

Cette formation s'adresse à la fois aux particuliers ainsi qu'aux entreprises, petites ou grandes, souhaitant former ses équipes à une nouvelle technologie informatique avancée ou bien à acquérir des connaissances métiers spécifiques ou des méthodes modernes.

Positionnement à l'entrée en formation

Le positionnement à l'entrée en formation respecte les critères qualité Qualiopi. Dès son inscription définitive, l'apprenant reçoit un questionnaire d'auto-évaluation nous permettant d'apprécier son niveau estimé sur différents types de technologies, ses attentes et objectifs personnels quant à la formation à venir, dans les limites imposées par le format sélectionné. Ce questionnaire nous permet également d'anticiper certaines difficultés de connexion ou de

sécurité interne en entreprise (intraentreprise ou classe virtuelle) qui pourraient être problématiques pour le suivi et le bon déroulement de la session de formation.

Méthodes pédagogiques

Stage Pratique : 60% Pratique, 40% Théorie. Support de la formation distribué au format numérique à tous les participants.

Organisation

Le cours alterne les apports théoriques du formateur soutenus par des exemples et des séances de réflexions, et de travail en groupe.

Validation

À la fin de la session, un questionnaire à choix multiples permet de vérifier l'acquisition correcte des compétences.

Sanction

Une attestation sera remise à chaque stagiaire qui aura suivi la totalité de la formation.