

# Google Cloud Platform - Fondamentaux Big Data and Machine Learning

Référence : **GCP100B**

Durée : **1 jour**

Certification : **Aucune**

## CONNAISSANCES PREALABLES

- Expérience avec la programmation en Python.
- Expérience avec un ETL.
- Expérience avec un langage de requête commun tel que SQL.
- Expérience de modélisation des données.
- Expérience en machine learning et / ou statistiques.

## PROFIL DES STAGIAIRES

- Participants désirant un aperçu des produits et services Google Cloud Platform orientés traitement des données et machine learning.

## OBJECTIFS

- Connaissance des produits et services de Google Cloud Platform, particulièrement ceux liés au traitement des données et au machine learning.
- Connaissance des produits et services fondamentaux concernant le calcul et le stockage.
- Connaissance de Cloud SQL et de Dataproc.
- Connaissance de Datalab et BigQuery.
- Connaissance de TensorFlow et des APIs Machine Learning.
- Connaissance de Pub / Sub et de Dataflow.

## CERTIFICATION PREPAREE

Aucune

## METHODES PEDAGOGIQUES

- Mise à disposition d'un 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, de réflexions et de retours d'expérience
- Le suivi de cette formation donne lieu à la signature d'une feuille d'émargement

## FORMATEUR

Consultant-Formateur expert Cloud

## METHODE D'EVALUATION DES ACQUIS

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

## CONTENU DU COURS

### Présentation de Google Cloud Platform

- Présentation des principes de base de la plateforme Google
- Produits Big Data de Google Cloud Platform

### Fondamentaux du calcul et du stockage

- CPU à la demande (Compute Engine)
- Un système de fichiers global (Cloud Storage)
- CloudShell
- Atelier : Configurer un pipeline de traitement des données Ingest-Transform-Publish

### Data Analytics sur le Cloud

- Tremplins vers le nuage

- Cloud SQL: votre base de données SQL sur le cloud
- Atelier: Importation de données dans CloudSQL et exécution de requêtes
- Spark sur Dataproc
- Travaux pratiques : recommandations d'apprentissage automatique avec Spark sur Dataproc

### Mise à l'échelle de l'analyse des données

- Accès aléatoire rapide
- Datalab
- BigQuery
- Atelier : Créer un ensemble de données d'apprentissage automatique

### **Apprentissage automatique**

- Apprentissage automatique avec TensorFlow
- Laboratoire : Effectuer ML avec TensorFlow
- Modèles pré-construits pour les besoins communs
- Atelier : Utiliser des API ML

### **Architectures de traitement des données**

- Architectures orientées message avec Pub / Sub
- Création de pipelines avec Dataflow
- Architecture de référence pour le traitement de données en temps réel et par lots

### **Résumé**

- Pourquoi GCP ?
- Où aller en partant d'ici
- Ressources supplémentaires