

## Google Cloud Platform : les fondamentaux du Big Data et du Machine Learning (CPB100)

Durée: 1 Jour    Réf de cours: GO8325    Méthodes d'apprentissage: Virtual Learning

### Résumé:

**Cette formation est délivrée en synchrone à distance tout en garantissant l'accès à un environnement d'apprentissage complet!**

Ce cours d'une journée initie les participants aux capacités de Big Data de Google Cloud Platform. Grâce à une combinaison de présentations, de démonstrations et de travaux pratiques, ils obtiendront un aperçu de la plate-forme Google Cloud et une vue détaillée des fonctionnalités de traitement des données et d'apprentissage automatique. Ce cours présente la facilité, la flexibilité et la puissance des solutions Big Data sur Google Cloud Platform.

### Public visé:

Cette formation s'adresse à de nombreux publics : Les analystes de données qui souhaitent se familiariser avec Google Cloud Platform; les personnes responsables de la conception de pipelines et d'architectures pour le traitement de données, la création et la maintenance de modèles d'apprentissage automatique et de statistiques, l'interrogation de jeux de données, la visualisation des résultats de requêtes et la création de rapports; les cadres et décideurs informatiques évaluant Google Cloud Platform à l'usage des data scientists.

### Objectifs pédagogiques:

- Découvrir les objectifs et la valeur des principaux produits Big Data et Machine Learning disponibles dans la plate-forme GoogleCloud
- Utiliser Cloud SQL et Cloud Dataproc pour migrer les workloads MySQL et Hadoop / Pig / Spark / Hive existants vers Google Cloud Platform
- Utiliser BigQuery et Cloud Datalab pour effectuer une analyse interactive des données
- Former et utiliser un réseau de neurones en utilisant TensorFlow
- Employer des API ML
- Choisir parmi différents produits de traitement de données sur Google Cloud Platform

### Pré-requis:

- Posséder des compétence de base en langage de requête commun tel que SQL
- Avoir une expérience de la modélisation, de l'extraction, de la transformation et du chargement de données
- Savoir développer des applications en utilisant un langage de programmation commun tel que Python
- Etre familier avec l'apprentissage automatique et / ou les statistiques

## Contenu:

### Présentation de Google Cloud Platform

- Présentation des principes fondamentaux de la plate-forme Google
- Produits et technologies de données Google Cloud Platform
- Scénarios d'utilisation

### Calcul et bases de stockage

- CPU à la demande (Compute Engine)
- Un système de fichiers global (Cloud Storage)
- CloudShell

### L'analyse de données sur le cloud

- Stepping-stones vers le cloud
- CloudSQL: votre base de données SQL sur le cloud
- Atelier: Importation de données dans CloudSQL et exécution de requêtes
- Spark sur Dataproc

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

- Accès aléatoire rapide
- Datalab
- BigQuery
- Apprentissage automatique avec TensorFlow
- Des modèles entièrement construits pour des besoins communs

### Architectures de traitement de données

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

### Résumé

- Pourquoi Google Cloud Platform?
- Où aller en partant d'ici

### Ressources additionnelles

- Lab 1: S'inscrire à Google Cloud Platform
- Atelier 2: Configurer un pipeline de traitement de données Ingest-Transform-Publish
- Lab 3: Recommandations d'apprentissage automatique avec SparkML
- Lab 4: Construire un ensemble de données d'apprentissage automatique
- Lab 5: Former et utiliser le réseau de neurones
- Lab 6: Employer des API ML

## Méthodes pédagogiques :

Pour le suivi de cette formation à distance depuis un site client équipé, il suffit d'avoir une bonne connexion internet, un casque avec micro et d'être dans un endroit au calme pour en profiter pleinement. Une fiche explicative est adressée en amont aux participants pour leur permettre de vérifier leur installation technique et de se familiariser avec la solution technologique utilisée.

L'accès à l'environnement d'apprentissage (support de cours officiel, labs), ainsi qu'aux feuilles d'émargement et d'évaluation est assuré. En savoir plus : <https://www.globalknowledge.com/fr-fr/solutions/methodes-d-apprentissage/classe-a-distance>

## Autres moyens pédagogiques et de suivi:

- Compétence du formateur : Les experts qui animent la formation sont des spécialistes des matières abordées et ont au minimum cinq ans d'expérience d'animation. Nos équipes ont validé à la fois leurs connaissances techniques (certifications le cas échéant) ainsi que leur compétence pédagogique.
- Suivi d'exécution : Une feuille d'émargement par demi-journée de présence est signée par tous les participants et le formateur.
- Modalités d'évaluation : le participant est invité à s'auto-évaluer par rapport aux objectifs énoncés.
- Chaque participant, à l'issue de la formation, répond à un questionnaire de satisfaction qui est ensuite étudié par nos équipes pédagogiques en vue de maintenir et d'améliorer la qualité de nos prestations.

### Délais d'inscription :

- Vous pouvez vous inscrire sur l'une de nos sessions planifiées en inter-entreprises jusqu'à 5 jours ouvrés avant le début de la formation sous réserve de disponibilité de places et de labs le cas échéant.
- Votre place sera confirmée à la réception d'un devis ou ""booking form"" signé. Vous recevrez ensuite la convocation et les modalités d'accès en présentiel ou distanciel.
- Attention, si vous utilisez votre Compte Personnel de Formation pour financer votre inscription, vous devrez respecter un délai minimum et non négociable fixé à 11 jours ouvrés.