

Building Data Lakes on AWS

Durée: 1 Jour Réf de cours: GK7377

Résumé:

Cette formation permet aux participants d'acquérir les connaissances et compétences nécessaires pour créer un data lake opérationnel qui prend en charge l'analyse des données structurées et non structurées. Ils apprendront à utiliser les composants et les fonctionnalités des services impliqués dans la création d'un lac de données et utiliseront AWS Lake Formation pour créer un lac de données et AWS Glue pour créer un catalogue de données et Amazon Athena pour analyser les données.

Niveau de la formation : Intermédiaire

Formation éligible au CPF (cf rubrique Certification)

Mise à jour : 12.05.2023

Public visé:

Cette formation s'adresse aux ingénieurs de plate-forme de données, architectes de solutions et professionnels de l'informatique

Objectifs pédagogiques:

- A l'issue de la formation, les participants seront capables de :
- Appliquer les méthodologies de lac de données dans la planification et la conception d'un lac de données
- Articuler les composants et les services requis pour créer un lac de données AWS
- Sécuriser un lac de données avec les autorisations appropriées
- Ingérer, stocker et transformer des données dans un lac de données
- Interroger, analyser et visualiser des données dans un lac de données

Pré-requis:

Nous recommandons aux participants d'avoir un an d'expérience dans la création de pipelines d'analyse de données ou d'avoir suivi le cours en ligne Data Analytics Fundamentals.

Cours à suivre en pré-requis :

- GK4501 - AWS : Notions techniques de base Amazon Web Services

Test et certification

Afin de se préparer efficacement à l'obtention de la certification **AWS Certified Data Analytics Speciality**, nous vous recommandons de suivre également les formations suivantes :

[GK7378](#) - Building Batch Data Analytics Solutions on AWS

[GK7379](#) - Building Data Analytics Solutions using Amazon Redshift

[GK7400](#) - Building Streaming Data Analytics Solutions on AWS

**

Cette formation mène également à la Certification professionnelle inscrite au Répertoire Spécifique de France Compétences RS6222 "Créer et maintenir des solutions d'analyse de données efficaces et sécurisées sur le cloud AWS". Plus de détails sur :

<https://www.francecompetences.fr/recherche/rs/6222/>

Après cette formation, nous vous conseillons le(s) module(s) suivant(s):

- GK7378 - Building Batch Data Analytics Solutions on AWS
- GK7379 - Building Data Analytics Solutions Using Amazon Redshift

Contenu:

Module 1 : Introduction aux Data Lakes

- Décrire la valeur des data lakes
- Comparer les lacs de données et les entrepôts de données
- Décrire les composants d'un lac de données
- Reconnaître les architectures courantes construites sur des lacs de données

Module 2 : Ingestion, catalogage et préparation des données

- Décrire la relation entre le stockage du lac de données et l'ingestion de données
- Décrire les robots d'exploration AWS Glue et comment ils sont utilisés pour créer un catalogue de données

Identifier le formatage, le partitionnement et la compression des données pour un stockage et une requête efficaces

- Atelier 1 : Configurer un lac de données simple

Module 3 : Traitement et analyse des données

- Reconnaître comment le traitement des données s'applique à un lac de données
- Utiliser AWS Glue pour traiter les données dans un lac de données
- Décrire comment utiliser Amazon Athena pour analyser les données dans un lac de données

Module 4 : Construire un lac de données avec AWS Lake Formation

- Décrire les fonctionnalités et les avantages d'AWS Lake Formation
- Utiliser AWS Lake Formation pour créer un lac de données

Comprendre le modèle de sécurité AWS Lake Formation

- Atelier 2 : Créer un lac de données à l'aide d'AWS Lake Formation

Module 5 : Configurations supplémentaires de Lake Formation

- Automatiser AWS Lake Formation à l'aide de plans et de workflows
- Appliquer des contrôles de sécurité et d'accès à AWS Lake Formation
- Faire correspondre les enregistrements avec AWS Lake Formation FindMatches

Visualiser les données avec Amazon QuickSight

Atelier 3 : Automatiser la création d'un lac de données à l'aide des plans AWS Lake Formation

- Atelier 4 : Visualisation des données à l'aide d'Amazon QuickSight

Module 6 : Architecture et synthèse

- Vérification des connaissances après le cours
- Revue de l'architecture
- Synthèse du cours

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.