

Building Data Analytics Solutions Using Amazon Redshift

Durée: 1 Jour Réf de cours: GK7379

Résumé:

Cette formation permet aux participants d'acquérir les compétences et connaissances nécessaires pour construire une solution d'analyse de données à l'aide d'Amazon Redshift, un service d'entrepôt de données Cloud. La formation se concentre sur les composants de collecte, d'ingestion, de catalogage, de stockage et de traitement des données du pipeline analytique. Les participants apprendront à intégrer Amazon Redshift à un lac de données pour prendre en charge les charges de travail d'analyse et d'apprentissage automatique, et également à appliquer les meilleures pratiques de sécurité, de performance et de gestion des coûts au fonctionnement d'Amazon Redshift.

Formation éligible au CPF (cf rubrique Certification)

Mise à jour : 12.05.2023

Public visé:

Cette formation s'adresse aux ingénieurs d'entrepôts de données, aux ingénieurs de plateformes de données, ainsi qu'aux architectes et opérateurs qui construisent et gèrent des pipelines d'analyse de données.

Objectifs pédagogiques:

- A l'issue de la formation, les participants seront capables de :
- Comparer les caractéristiques et les avantages des entrepôts de données, des lacs de données et des architectures de données modernes.
- Concevoir et mettre en œuvre une solution d'analyse d'entrepôt de données
- Identifier et appliquer les techniques appropriées, notamment la compression, pour optimiser le stockage des données
- Sélectionner et déployer les options appropriées pour ingérer, transformer et stocker les données.
- Définir les types d'instances et de nœuds, les clusters, la mise à l'échelle automatique et la topologie réseau appropriés pour un cas d'utilisation particulier.
- Examiner comment le stockage et le traitement des données affectent les mécanismes d'analyse et de visualisation nécessaires pour obtenir des informations commerciales exploitables.
- Identifier comment sécuriser les données au repos et en transit
- Identifier et résoudre les problèmes en surveillant les charges de travail analytiques
- Expliquer les meilleures pratiques de gestion des coûts

Pré-requis:

Avoir idéalement une expérience d'au moins un an dans la gestion d'entrepôts de données.

Nous recommandons aux participants d'avoir les connaissances équivalentes aux programmes de ces cours :

- GK4501 - AWS : Notions techniques de base Amazon Web Services
- GK7377 - Building Data Lakes on AWS
- GK7378 - Building Batch Data Analytics Solutions on AWS

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 :

[GK7377](#) - Building Data Lakes on AWS

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

[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/>

Contenu:

Aperçu de l'analyse des données et du pipeline de données

- Cas d'utilisation de l'analyse de données
- Utilisation du pipeline de données pour l'analyse

Module 1 : Utilisation d'Amazon Redshift dans le pipeline d'analyse de données

- Pourquoi Amazon Redshift pour l'entreposage de données ?
- Présentation d'Amazon Redshift

Module 2 : Introduction à Amazon Redshift

- Architecture d'Amazon Redshift
- Démo interactive 1 : Visite de la console Amazon Redshift
- Caractéristiques d'Amazon Redshift
- Laboratoire pratique 1 : charger et interroger des données dans un cluster Amazon Redshift

Module 3 : Ingestion et stockage

- Ingestion
- Démo interactive 2 : connecter votre cluster Amazon Redshift à l'aide d'un carnet Jupyter avec Data API
- Distribution et stockage des données
- Démo interactive 3 : analyse de données semi-structurées à l'aide du type de données SUPER
- Interrogation de données dans Amazon Redshift
- Laboratoire pratique 2 : analyse de données à l'aide d'Amazon Redshift Spectrum

Module 4 : Traitement et optimisation des données

- Transformation des données
- Interrogation avancée
- Labo pratique 3 : Transformation et interrogation de données dans Amazon Redshift
- Gestion des ressources
- Démo interactive 4 : application de la gestion des charges de travail mixtes sur Amazon Redshift
- Automatisation et optimisation
- Démo interactive 5 : Redimensionnement du cluster Amazon Redshift du cluster dc2.large au cluster ra3.xplus

Module 5 : Sécurité et surveillance des clusters Amazon Redshift

- Sécurisation du cluster Amazon Redshift
- Surveillance et dépannage des clusters Amazon Redshift

Module 6 : Conception de solutions analytiques d'entrepôt de données

- Examen des cas d'utilisation des entrepôts de données
- Activité : Conception d'un workflow d'analyse d'entrepôt de données

Développement d'architectures de données modernes sur AWS

Architectures de données modernes

Méthodes pédagogiques :

Support de cours officiel remis aux participants.

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