

Data Warehousing on AWS

Durée: 3 Jours Réf de cours: GK4375 Méthodes d'apprentissage: Intra-entreprise & sur-mesure

Résumé:

L'entreposage de données sur AWS vous présente des concepts, des stratégies et des bonnes pratiques pour concevoir une solution d'entreposage de données basée sur le cloud à l'aide d'Amazon Redshift. Ce cours explique comment ingérer, stocker et transformer des données dans l'entrepôt de données. Les sujets abordés sont les suivants : l'objectif d'Amazon Redshift, la manière dont Amazon Redshift relève les défis commerciaux et techniques, les fonctionnalités et les capacités d'Amazon Redshift, la conception d'une solution d'entreposage de données sur AWS en appliquant les bonnes pratiques basées sur le cadre Well-Architected, l'intégration avec les produits et services AWS et non-AWS, l'optimisation des performances, l'orchestration, ainsi que la sécurisation et la surveillance d'Amazon Redshift.

Niveau du cours : Avancé

Durée : 3 jours

Activités

Ce cours comprend des présentations, des travaux pratiques et des démonstrations.

Formation intra-entreprise

Cette formation est délivrable en session intra-entreprise, dans vos locaux ou dans les nôtres. Son contenu peut être adapté sur-mesure pour répondre aux besoins de vos collaborateurs. Contactez votre conseiller formation Global Knowledge ou adressez votre demande à info@globalknowledge.fr.

Public visé:

Ce cours s'adresse à :

- Ingénieurs de données
- Architectes de données
- Architectes de bases de données
- Administrateurs de bases de données
- Développeurs de bases de données

Objectifs pédagogiques:

- Dans ce cours, vous apprenez à :
- Décrire l'architecture d'Amazon Redshift et ses rôles dans une architecture de données moderne
- Concevoir et mettre en œuvre un entrepôt de données dans le cloud à l'aide d'Amazon Redshift
- Identifier et charger des données dans un entrepôt de données Amazon Redshift à partir de diverses sources
- Analyser les données à l'aide de notebooks SQL QEV2
- Concevoir et mettre en œuvre une stratégie de reprise après sinistre pour un entrepôt de données Amazon Redshift
- Effectuer la maintenance et l'optimisation des performances sur un entrepôt de données Amazon Redshift
- Sécurisez et gérez l'accès à un entrepôt de données Amazon Redshift
- Partager des données entre plusieurs clusters Redshift au sein d'une organisation
- Orchestrez les flux de travail dans l'entrepôt de données à l'aide des machines d'état AWS Step Functions
- Création d'un modèle de ML et configuration de prédicteurs à l'aide d'Amazon Redshift ML

Pré-requis:

Nous recommandons aux participants de ce cours d'avoir suivi les cours suivants :

- Principes de base de l'analytique sur AWS – Partie 1 (cours numérique)

- Principes de base de l'analytique sur AWS – Partie 2 (cours numérique)
- Création de lacs de données sur AWS (formation dispensée par un instructeur)
- Création de solutions d'analyse de données à l'aide d'Amazon Redshift (formation dispensée par un instructeur)

Contenu:

Jour 1

Module 1 : Concepts d'entrepôt de données

- Architecture de données moderne
- Introduction à l'histoire du cours
- Entreposage de données avec Amazon Redshift
- Architecture sans serveur d'Amazon Redshift
- Atelier pratique : Lancement et configuration d'un entrepôt de données sans serveur Amazon Redshift

Module 2 : Configuration d'Amazon Redshift

- Modèles de données pour Amazon Redshift
- Gestion des données dans Amazon Redshift
- Gestion des autorisations dans Amazon Redshift
- Atelier pratique : Configuration d'un entrepôt de données à l'aide d'Amazon Redshift Serverless

Module 3 : Chargement des données

- Vue d'ensemble des sources de données
- Chargement de données à partir d'Amazon Simple Storage Service (Amazon S3)
- Extraire, transformer et charger (ETL) et extraire, charger et transformer (ELT)
- Chargement des données de streaming
- Chargement de données à partir de bases de données relationnelles
- Atelier pratique : Remplissage de l'entrepôt de données

Jour 2

Module 4 : Plongée dans SQL Query Editor v2 et les notebooks

- Fonctionnalités d'Amazon Redshift Query Editor v2
- Démonstration : Utilisation d'Amazon Redshift Query Editor v2
- Requêtes avancées
- Atelier pratique : Data Wrangling sur AWS

Module 5 : Sauvegarde et récupération

- Reprise après sinistre
- Sauvegarde et restauration d'Amazon Redshift provisionné
- Sauvegarde et restauration d'Amazon Redshift Serverless

Module 6 : Réglage des performances d'Amazon Redshift

- Facteurs qui ont un impact sur les performances des requêtes
- Maintenance des tables et vues matérialisées
- Analyse des requêtes
- Gestion de la charge de travail
- Conseils de réglage
- Surveillance Amazon Redshift
- Atelier pratique : Optimisation des performances de l'entrepôt de données

Module 7 : Sécurisation d'Amazon Redshift

- Présentation de la sécurité et de la conformité d'Amazon Redshift
- Authentification avec Amazon Redshift
- Contrôle d'accès avec Amazon Redshift
- Chiffrement des données avec Amazon Redshift
- Audit et conformité avec Amazon Redshift
- Atelier pratique : Sécurisation d'Amazon Redshift

Jour 3

Module 8 : Orchestration

- Vue d'ensemble de l'orchestration des données
- Orchestration avec AWS Step Functions
- Orchestration avec Amazon Managed Workflows for Apache Airflow (MWAA)
- Atelier pratique : Orchestration du pipeline d'entrepôt de données

Module 9 : Amazon Redshift ML

- Présentation de l'apprentissage automatique
- Mise en route d'Amazon Redshift ML
- Scénarios de flux de travail Amazon Redshift ML
- Utilisation d'Amazon Redshift ML
- Atelier pratique : Prévision de l'attrition des clients avec Amazon Redshift ML

Module 10 : Partage de données Amazon Redshift

- Présentation du partage de données dans Amazon Redshift
- Amazon DataZone pour les données en tant que service

Module 11 : Conclusion

- Laboratoire pratique : Laboratoire de défi de fin de cours

Méthodes pédagogiques :

Ce cours vous permet de tester de nouvelles compétences et d'appliquer vos connaissances à votre environnement de travail grâce à différents exercices pratiques.

Un support de cours officiel AWS est remis aux participants, au format électronique.

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.
- En fin de formation, le participant est invité à s'auto-évaluer sur l'atteinte des objectifs énoncés, et à répondre à un questionnaire de satisfaction qui sera 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 cette formation est éligible au Compte Personnel de Formation, vous devrez respecter un délai minimum et non négociable fixé à 11 jours ouvrés avant le début de la session pour vous inscrire via moncompteformation.gouv.fr.

Accueil des bénéficiaires :

- En cas de handicap : plus d'info sur globalknowledge.fr/handicap
- Le Règlement intérieur est disponible sur globalknowledge.fr/reglement