EduGroupe
Accompagner pour réussir

Data Warehousing on Amazon Web Services (AWS)

Référence : **AWS-WAREH** Durée : **3 jours (21 heures)** Certification : **Aucune**

Connaissances préalables

- 1. Avoir suivi la formation "Amazon Web Services (AWS) - Fondamentaux techniques" (CC311) ou connaissances équivalentes
- 2. Être familiarisé avec les bases de données relationnelles et les concepts de conception de bases de données
- 3. Avoir des connaissances de base en langue anglaise car le support de cours est en langue anglaise

Profil des stagiaires

- Administrateurs de bases de données
- Analystes de données et data scientists
- Architectes de bases de données
- Développeurs de bases de données

Objectifs

- Discuter des concepts de base de l'entreposage de données et de l'intersection entre l'entreposage de données et les solutions de Big Data
- Être capable de lancer un cluster Amazon Redshift et d'utiliser les composants, fonctionnalités et fonctionnalités pour implémenter un entrepôt de données dans le cloud
- Comprendre comment utiliser d'autres services de données et d'analyse AWS, tels qu'Amazon DynamoDB, Amazon EMR, Amazon Kinesis et Amazon S3, pour contribuer à la solution d'entreposage de données
- Savoir architecturer l'entrepôt de données
- Apprendre à identifier les problèmes de performances, à optimiser les requêtes et à ajuster la base de données pour de meilleures performances
- Pouvoir utiliser Amazon Redshift Spectrum pour analyser les données directement à partir d'un compartiment Amazon S3
- Comprendre comment utiliser Amazon QuickSight pour effectuer des tâches d'analyse et de visualisation des données par rapport à l'entrepôt de données

Certification préparée

- Aucune

Méthodes pédagogiques

- 6 à 12 personnes maximum par cours, 1 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 et de réflexions

Formateur

- Consultant-Formateur expert Amazon Web Services

Méthodes d'évaluation des acquis

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

Contenu du cours

1. Introduction à l'entreposage de données

- Bases de données relationnelles
- Concepts d'entreposage de données
- L'intersection entre l'entreposage de données et le big data
- Présentation de la gestion des données dans AWS
- Atelier : Introduction à Amazon Redshift

2. Introduction à Amazon Redshift

- Aperçu du concept
- Cas d'utilisation réels
- Atelier : Lancement d'un cluster Amazon Redshift

3. Lancer des clusters

- Construire le cluster
- Connexion au cluster
- Contrôle d'accès
- Sécurité de la base de données
- Charger les données
- Atelier : Optimisation des schémas de base de données

4. Conception du schéma de la base de données

- Schémas et types de données
- Compression colonnaire
- Styles de distribution des données
- Méthodes de tri des données

5. Identification des sources de données

- Présentation des sources de données
- Amazon S3
- Amazon DynamoDB
- Amazon DME
- Amazon Kinesis Data Firehose
- Chargeur de base de données AWS Lambda pour Amazon Redshift
- Atelier : Chargement de données en temps réel dans une base de données Amazon Redshift

6. Chargement des données

- Préparation des données
- Chargement des données à l'aide de COPY
- Tenue des tableaux
- Opérations d'écriture simultanées
- Dépannage des problèmes de charge
- Atelier : Charger des données avec la commande COPY

7. Écriture de requêtes et optimisation des performances

- Amazon Redshift SQL
- Fonctions définies par l'utilisateur (UDF)
- Facteurs affectant les performances des requêtes
- La commande EXPLAIN et les plans de requête
- Gestion de la charge de travail (WLM)
- Atelier : Configuration de la gestion de la charge de travail

8. Amazon Redshift Spectrum

- Spectre Amazon Redshift
- Configuration des données pour Amazon Redshift Spectrum
- Requêtes Amazon Redshift Spectrum
- Utilisation d'Amazon Redshift Spectrum

9. Maintenance des clusters

- Journalisation des audits
- Suivi des performances
- Événements et notifications
- Redimensionnement des clusters
- Sauvegarde et restauration des clusters
- Étiquetage des ressources et limites et contraintes
- Atelier 1 : Audit et monitoring des clusters
- Atelier 2 : Sauvegarde, restauration et redimensionnement des clusters

10. Analyser et visualiser les données

- Puissance des visualisations
- Création de tableaux de bord
- Éditions et fonctionnalités d'Amazon QuickSight

Notre référent handicap se tient à votre disposition au 01.71.19.70.30 ou par mail à referent.handicap@edugroupe.com pour recueillir vos éventuels besoins d'aménagements, afin de vous offrir la meilleure expérience possible.