

# AWS – Data Warehousing on Amazon Web Services

## Description

Ce cours vous présente les concepts, les stratégies et les meilleures pratiques pour la conception d'une solution d'entreposage de données dans le cloud à l'aide d'Amazon Redshift, l'entrepôt de données à l'échelle du pétaoctet dans AWS. Ce cours montre comment collecter, stocker et préparer des données pour l'entrepôt de données à l'aide de services AWS tels qu'Amazon DynamoDB, Amazon EMR, Amazon Kinesis et Amazon S3. En outre, ce cours montre comment utiliser Amazon QuickSight pour effectuer une analyse sur vos données.

### Prix de l'inscription en Présentiel (CHF)

2500

### Prix de l'inscription en Virtuel (CHF)

2500

### Contenu du cours

#### Module 1 : Introduction à l'entreposage de données

- Bases de données relationnelles
- Concepts d'entreposage de données
- L'intersection de l'entreposage de données et du big data
- Vue d'ensemble de la gestion des données dans AWS

#### Module 2 : Introduction à Amazon Redshift

- Aperçu conceptuel
- Cas d'utilisation du monde réel

#### Module 3 : Lancement des clusters

- Construire le cluster
- Connexion au cluster
- Contrôle d'accès
- Sécurité de la base de données
- Charger des données

#### Module 4 : Conception du schéma de base de données

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

#### Module 5 : Identifier les sources de données

- Vue d'ensemble des sources de données
- Amazon S3
- Amazon DynamoDB
- Amazon EMR
- Amazon Kinesis Data Firehose

- AWS Lambda Database Loader pour Amazon Redshift

## **Module 6 : Chargement des données**

- Préparation des données
- Chargement des données à l'aide de COPY
- Maintien des tables
- Opérations d'écriture simultanées
- Dépannage des problèmes de charge

## **Module 7 : Rédaction de requêtes et optimisation des performances**

- Amazon Redshift SQL
- Fonctions définies par l'utilisateur (UDF)
- Facteurs qui affectent les performances des requêtes
- La commande EXPLAIN et les plans de requêtes
- Gestion de la charge de travail (WLM)

## **Module 8 : Amazon Redshift Spectrum**

- Spectre Amazon Redshift
- Configuration des données pour Amazon Redshift Spectrum
- Requêtes de spectre Amazon Redshift

## **Module 9 : Maintenance des clusters**

- Journalisation d'audit
- Suivi de la performance
- Événements et notifications
- Redimensionner les clusters
- Sauvegarde et restauration des clusters
- Balisage des ressources et limites et contraintes

## **Module 10 : Analyse et visualisation des données**

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

## **Lab / Exercices**

- Laboratoires officiels AWS

## **Documentation**

- Support de cours numérique inclus

## **Profils des participants**

- Architectes de bases de données
- Administrateurs de base de données
- Développeurs de bases de données
- Analystes de données
- Scientifiques exploitant des données numériques

## **Connaissances Préalables**

- Connaissance des bases de données relationnelles et des concepts de conception de bases de données
- Avoir suivi ou maîtriser les notions incluses dans le cours suivant : [AWS Technical Essentials](#)

## Objectifs

- Discutez des concepts de base de l'entreposage de données et de l'intersection entre l'entreposage de données et les solutions Big Data
- Lancez un cluster Amazon Redshift et utilisez les composants, fonctionnalités et fonctionnalités pour implémenter un entrepôt de données dans le cloud
- Utilisez 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
- Architecte de l'entrepôt de données
- Identifiez les problèmes de performances, optimisez les requêtes et ajustez la base de données pour de meilleures performances
- Utilisez Amazon Redshift Spectrum pour analyser les données directement à partir d'un compartiment Amazon S3
- Utilisez Amazon QuickSight pour effectuer des tâches d'analyse et de visualisation des données par rapport à l'entrepôt de données

## Niveau

Intermédiaire

## Durée (Nombre de Jours)

3

## Reference

AWS-12