

Concevoir, implémenter et gérer un Data Warehouse AWS

Date et durée
Code formation : DATA003FR Durée : 3 jours Nombre d'heures : 21 heures
Description
<p>Le Data Warehousing sur AWS s'articule principalement autour d'Amazon Redshift, un service d'entrepôt de données cloud rapide et évolutif. Il permet de stocker et d'analyser de vastes quantités de données structurées et semi-structurées pour obtenir des informations décisionnelles. Redshift s'intègre à d'autres services AWS comme S3 pour le stockage et QuickSight pour la visualisation, offrant une solution d'analyse complète et performante.</p> <p>Cette formation intensive vous fournira les compétences pratiques pour maîtriser la création de solutions d'entreposage de données avec Amazon. Vous explorerez en profondeur les outils et les services AWS essentiels, tels qu'Amazon Redshift, et les méthodes pour ingérer, stocker et transformer des données dans un entrepôt de données. À travers des ateliers pratiques, vous apprendrez à optimiser le chargement et la transformation des données, ainsi qu'à sécuriser et surveiller Amazon Redshift.</p> <p>À l'issue de ce programme de 3 jours, vous développerez l'expertise nécessaire pour concevoir des solutions d'analyse de données performantes et sécurisées sur AWS. Vous maîtriserez l'optimisation d'Amazon Redshift et les meilleures pratiques pour exploiter pleinement le potentiel de vos données dans le cloud.</p>

<p><i>En tant que partenaire de formation premium (ATP) agréé par Amazon Web Services, Oo2 vous propose des formations qualifiantes et certifiantes qui respectent les normes de qualité rigoureuses de l'organisme.</i></p>
Objectifs
<p>À l'issue de cette formation data warehouse on AWS, vous atteindrez les objectifs de compétences suivants :</p> <ul style="list-style-type: none">• décrire l'architecture d'Amazon Redshift et ses rôles dans une architecture de données moderne ;• concevoir et implémenter un data warehouse dans le cloud en utilisant Amazon Redshift ;• identifier et charger des données dans un data warehouse depuis diverses sources ;• analyser des données en utilisant les notebooks SQL QEV2 ;• concevoir et implémenter une stratégie de reprise après sinistre pour un data warehouse ;• effectuer la maintenance et l'optimisation des performances d'un data warehouse ;• sécuriser et gérer l'accès à un data warehouse ;• partager des données entre plusieurs clusters Redshift au sein d'une organisation ;• orchestrer des flux de travail dans le data warehouse en utilisant les machines d'état AWS Step Functions ;• créer un modèle de ML et configurer des prédicteurs en utilisant Amazon Redshift ML.

Points forts

- **Formateurs experts AWS certifiés** : bénéficiez de l'expertise de formateurs reconnus et certifiés par AWS, qui possèdent une connaissance approfondie des concepts de data warehousing et des outils et services AWS.
- **Mise en pratique interactive** : maîtrisez les outils et techniques de data warehousing sur AWS grâce à des démonstrations concrètes et des ateliers pratiques. Vous serez ainsi préparé à relever les défis réels de la conception de solutions de data warehousing, du chargement et de la transformation des données, et de l'optimisation des performances d'Amazon Redshift.
- **Acquisition de compétences clés** : le contenu de la formation est soigneusement conçu pour vous permettre d'acquérir les compétences essentielles en matière de conception et d'implémentation de data warehouses, de gestion des données, de sécurisation d'Amazon Redshift, et d'utilisation des fonctionnalités avancées comme Amazon Redshift ML.

Modalités d'évaluation

Travaux Pratiques

Pré-requis

Suivre cette formation AWS nécessite les prérequis recommandés suivants :

- avoir suivi la formation "Les fondamentaux de l'analyse sur AWS" ;
- avoir suivi la [formation "Construire un data lake performant avec AWS Lake Formation"](#) ;
- avoir suivi la [Formation "Créez des solutions d'analyse de données avec Amazon Redshift"](#).

Public

Cette formation s'adresse aux publics suivants :

- **les ingénieurs de données** chargés de la conception, de la construction et de la maintenance des entrepôts de données sur AWS ;
- **les architectes de données** responsables de la conception et de l'implémentation des solutions de data warehousing sur la plateforme AWS ;
- **les administrateurs de bases de données** chargés de la gestion, de l'optimisation et de la sécurisation des bases de données Amazon Redshift ;
- **les développeurs** qui souhaitent intégrer les services de data warehousing d'AWS dans leurs applications pour analyser efficacement les données.

Programme

Module 1 : découvrir les concepts du Data Warehouse

- Les bases de l'architecture moderne des données.
- Introduction au scénario du cours.
- Le Data Warehousing avec Amazon Redshift.
- L'architecture Amazon Redshift Serverless.

Lab :

- Lancer et configurer Amazon Redshift Serverless Data Warehouse.

Module 2 : configurer Amazon Redshift

- La gestion des modèles de données.
- La gestion des données.
- La gestion des permissions.

Lab :

- Configurer un Data Warehouse en utilisant Amazon Redshift Serverless.

Module 3 : charger des données

- Présentation générale des sources de données.
- L'importation des données depuis le service Amazon S3.
- L'extraction, la transformation et le chargement (ETL).
- L'extraction, le chargement et la transformation (ELT).
- Le chargement des données en streaming.
- Le chargement des données depuis des bases de données relationnelles.

Lab :

- Peupler le Data Warehouse.

Module 4 : explorer SQL Query Editor v2 et les Notebooks

- Les fonctionnalités d'Amazon Redshift Query Editor v2.
- Les requêtes avancées.

Lab :

- Utiliser Data Wrangling sur AWS.

Module 5 : sauvegarder et restaurer

- La reprise après sinistre.
- La sauvegarde et la restauration d'Amazon Redshift provisionné et Serverless.

Module 6 : optimiser les performances d'Amazon Redshift

- Les facteurs qui impactent les performances des requêtes.
- La maintenance des tables et les vues matérialisées.
- L'analyse des requêtes.
- La gestion de la charge de travail.
- Les conseils d'optimisation.
- Le monitoring d'Amazon Redshift.

Lab :

- Optimiser les performances du Data Warehouse.

Module 7 : sécuriser Amazon Redshift

- Les bases de la sécurité et de la conformité.
- L'authentification.
- Le contrôle d'accès.
- Le chiffrement des données.
- L'audit et la conformité.

Lab :

- Sécuriser Amazon Redshift.

Module 8 : orchestrer des datas

- Les bases de l'orchestration des données.
- L'orchestration avec AWS Step Functions.
- L'orchestration avec Amazon Managed Workflows for Apache Airflow (MWAA).

Lab :

- Orchestrer le pipeline du Data Warehouse.

Module 9 : utiliser Amazon Redshift ML

- Les bases du Machine Learning avec Amazon Redshift ML.
- Les éléments de démarrage.
- Les scénarios de flux de travail.
- L'utilisation d'Amazon Redshift ML.

Lab :

- Prédire le désabonnement des clients avec Amazon Redshift ML.

Module 10 : partager des données avec Amazon Redshift

- Les bases du partage de données dans Amazon Redshift.
- L'utilisation d'Amazon DataZone pour le Data as a service.

Module 11 : conclure la formation

- Révision sur les points clés du cours.
- Lab challenge.

AWS, Amazon Redshift, Amazon S3, et les autres marques AWS sont des marques déposées d'[Amazon.com, inc.](https://www.amazon.com) ou de ses filiales.