

Snowflake, prise en main

Cours Pratique de 2 jours - 14h

Réf : SWF - Prix 2024 : 1 620€ HT

Snowflake premier data warehouse dans le cloud, allie la puissance des entrepôts de données et la flexibilité du cloud. Cette formation a pour objectif de vous faire découvrir l'outil Snowflake via ses différentes fonctionnalités et ses différences dans la construction d'un entrepôt de données.

OBJECTIFS PÉDAGOGIQUES

À l'issue de la formation l'apprenant sera en mesure de :

Découvrir les concepts clés et l'architecture de Snowflake

Comprendre et utiliser les différents rôles utilisateurs

Créer et transformer des données dans les bases de données Snowflake

Importer des données structurées et semi structurées complexes

Utiliser le cache pour optimiser ses requêtes et ses coûts

Créer des pipes et stream de données

MÉTHODES PÉDAGOGIQUES

L'expérience de nos formateurs permet d'enrichir la formation avec des cas concrets de mise en œuvre.

TRAVAUX PRATIQUES

Apports théoriques illustrés par des exercices pratiques pour chaque fonctionnalité, formant un fil rouge.

LE PROGRAMME

dernière mise à jour : 10/2022

1) Introduction

- Bref rappel historique sur les data warehouses.
- Rappel des fondamentaux d'un data warehouse.
- L'évolution des architectures traditionnelles.
- Le cloud computing.

2) Présentation de Snowflake : concepts clés et architecture

- Snowflake, un Software as a Service (SaaS) : avantages, instances de calcul virtuelles, service de stockage.
- Architecture à trois couches : data base storage, Query Processing, cloud services.
- Connexion à Snowflake : interface web, SnowSQL, pilotes, connecteurs.
- Hiérarchie de stockage : data bases et schémas, tables, views, stages.

Démonstration : L'écosystème de Snowflake. Connexion à Snowflake.

3) Prise en main de Snowflake WebUI

- Introduction et "user roles".
- Navigation data base : panneau de propriété bases de données et objets, utiliser les hyperliens, voir les privilèges.
- Exploration des données : navigation, volets d'entrée SQL et de résultats de la worksheet. Exploration avec Worksheet.
- Afficher des statistiques détaillées.

PARTICIPANTS

Chefs de projets, experts BI, ingénieurs système DBA.

PRÉREQUIS

Bonnes notions de SQL et des Systèmes de gestion de base de données (SGBD).

COMPÉTENCES DU FORMATEUR

Les experts qui animent la formation sont des spécialistes des matières abordées. Ils ont été validés par nos équipes pédagogiques tant sur le plan des connaissances métiers que sur celui de la pédagogie, et ce pour chaque cours qu'ils enseignent. Ils ont au minimum cinq à dix années d'expérience dans leur domaine et occupent ou ont occupé des postes à responsabilité en entreprise.

MODALITÉS D'ÉVALUATION

Le formateur évalue la progression pédagogique du participant tout au long de la formation au moyen de QCM, mises en situation, travaux pratiques...
Le participant complète également un test de positionnement en amont et en aval pour valider les compétences acquises.

MOYENS PÉDAGOGIQUES ET TECHNIQUES

- Les moyens pédagogiques et les méthodes d'enseignement utilisés sont principalement : aides audiovisuelles, documentation et support de cours, exercices pratiques d'application et corrigés des exercices pour les stages pratiques, études de cas ou présentation de cas réels pour les séminaires de formation.
- À l'issue de chaque stage ou séminaire, ORSYS fournit aux participants un questionnaire d'évaluation du cours qui est ensuite analysé par nos équipes pédagogiques.
- Une feuille d'émargement par demi-journée de présence est fournie en fin de formation ainsi qu'une attestation de fin de formation si le stagiaire a bien assisté à la totalité de la session.

MODALITÉS ET DÉLAIS D'ACCÈS

L'inscription doit être finalisée 24 heures avant le début de la formation.

ACCESSIBILITÉ AUX PERSONNES HANDICAPÉES

Vous avez un besoin spécifique d'accessibilité ? Contactez Mme FOSSE, référente handicap, à l'adresse suivante psh-accueil@orsys.fr pour étudier au mieux votre demande et sa faisabilité.

- Export des data via le clip board ou en utilisant un fichier.

Travaux pratiques : Création d'un compte. Affichage du panneau d'aide. Changement de rôle. Identifier la base de données en cours/objets/schéma. Créer une instruction SQL select. Explorer et afficher les données.

4) Warehouse Snowflake - Objets DB - Transformations

- Notion de warehouse au sens Snowflake.

- Utilisations du menu de contexte, d'un script SQL pour spécifier le warehouse.

- Fonctionnalité "Code Highlight".

- Les objets data base. Importer les données dans la base.

- Transformations : les fonctions SQL REPLACE, CONCAT. ETL & ELT. Import d'un fichier texte dans une worksheet.

Travaux pratiques : Création d'un warehouse, d'une BDD, d'une table, d'un format de fichier pour l'import. Effectuer des transformations.

5) Stockage et requêtage avancé

- Staging data : upload et organisation des données.

- Data storage structures : entité, attribut et relation, séquence.

- Données semi structurées : définition, formats utilisables avec Snowflake, stockage dans une colonne de type VARIANT.

- Données semi structurées complexes : requête, utilisation des fonctions FLATTEN, GET, GET_PATH.

Travaux pratiques : Charger des données dans un stage. Créer une séquence. Utiliser une séquence avec la fonction NEXT VALUE. Effectuer des requêtes.

6) Performances

- Les différents niveaux de cache : Result cache, Local Disk cache, Remote Disk.

- Optimiser l'utilisation du cache.

- Bonnes pratiques d'optimisation en management : des warehouses, des coûts.

- Bonnes pratiques en sécurité et gestion des rôles.

Travaux pratiques : Analyses du cache, optimiser son utilisation.

7) Design & Deploy

- Planifier des exécutions à l'aide des Tasks.

- Réplication des objets à l'aide du clonage et du time travel.

- Création de pipelines et streaming de données (change data capture).

Travaux pratiques : Planifier des exécutions. Faire du clonage. Créer des pipelines et des streaming.

8) Présentation d'un cas métier

- Présentation d'un cas métier : vision 360° d'une chaîne logistique.

- Conclusion.

LES DATES

CLASSE À DISTANCE

2024 : 08 juil., 14 oct.

PARIS

2024 : 01 juil., 07 oct.