

Mis à jour le 12/07/2024

S'inscrire

Formation Starburst

3 jours (21 heures)

Présentation

Notre formation Starburst vous permettra de maîtriser les fonctionnalités de ce [data lakehouse](#) et d'améliorer grandement vos processus d'analyse de données. Starburst dispose d'une architecture de données haute performance qui permet un traitement et une analyse rapides.

Ce programme de formation vous permettra de comprendre l'ensemble de l'architecture et de la terminologie de StarBurst ainsi que son utilisation avec Trino, le moteur d'analyse SQL.

Notre formation vous permettra de mettre vos connaissances à l'épreuve avec des exercices pratiques. Vous disposerez d'un entrepôt de données évolutif et sécurisé ainsi que la gouvernance nécessaire pour garantir que vos équipes d'ingénieurs et d'analystes ne rencontrent jamais de problèmes.

Comme toutes nos formations, elle se déroulera sur la dernière version de l'outil : [Trino 448](#)

Objectifs

- Configurer Starburst pour vos flux de données
- Comprendre l'utilisation du SQL dans Starburst
- Gérer et exploiter votre base de données Starburst

Public visé

- **Data Analysts**
- Data scientists

Pré-requis

- Une connaissance de base en SQL et en manipulation de données
- Connaissance de base des commandes Linux/Unix

PROGRAMME DE NOTRE FORMATION STARBURST

INTRODUCTION À STARBURST

- Présentation générale de Starburst
- Composants clés et architecture
- Avantages et cas d'utilisation
- Introduction aux plateformes
- Vue d'ensemble des connecteurs Starburst

FONDAMENTAUX DU SQL ET DE LA PARALLÉLISATION

- Principes de base du SQL dans Starburst
- Techniques de parallélisation des requêtes
- Utilisation des agrégations et requêtes conditionnelles
- Syntaxe SQL et jointures
- Comprendre le langage de définition des données (DDL)

CARACTÉRISTIQUES AVANCÉES DE STARBURST

- Exploration des tutoriels sur les fonctionnalités de Starburst
- Exécution tolérante aux pannes et haute disponibilité
- Fédération des données et mise à l'échelle des clusters
- Introduction aux produits de données et contrôle d'accès
- Découverte de schémas et recherche universelle

GESTION DES BASES DE DONNÉES ET STOCKAGE

- Types de bases de données : relationnelles vs non relationnelles
- Gestion des Data Warehouses et des Data Lakes
- Formats de fichiers orientés ligne et colonne
- Stockage d'objets dans le cloud et métadonnées
- Comparaison des formats de table modernes : Hive vs Iceberg

CONCEPTION ET AUTOMATISATION DES PIPELINES DE DONNÉES

- Principes de l'ETL (Extract, Transform, Load)
- Automatisation et intégration des pipelines de données avec Starburst
- Construction de jeux de données organisés comme des produits
- Meilleures pratiques pour la gestion des pipelines de données

INTÉGRATION AVEC D'AUTRES SOLUTIONS

- Compréhension des outils ETL tels que dbt et Informatica
- Utilisation des outils BI comme Tableau et Power BI
- Notions de base sur les outils de gouvernance et de catalogage
- Intégration pratique avec des exemples de configuration

DÉPANNAGE ET SÉCURITÉ

- Identification et résolution des problèmes de performance des requêtes
- Stratégies de mise à l'échelle des clusters pour l'optimisation
- Sécurité des données et contrôle d'accès basé sur les rôles et les attributs
- Best practices pour la maintenance et la sécurité des systèmes

ATELIERS PRATIQUES ET ÉTUDES DE CAS

- Mise en place d'un environnement Starburst
- Exécution de requêtes complexes et optimisation
- Simulation de dépannage sur des scénarios réels
- Création de tableaux de bord d'analyse avec intégration BI

RÉVISION ET ÉVALUATION FINALE

- Révision des principaux concepts et fonctionnalités de Starburst
- Test de connaissances avec des questions d'examen basées sur des cas réels

Sociétés concernées

Cette formation s'adresse à la fois aux particuliers ainsi qu'aux entreprises, petites ou grandes, souhaitant former ses équipes à une nouvelle technologie informatique avancée ou bien à acquérir des connaissances métiers spécifiques ou des méthodes modernes.

Positionnement à l'entrée en formation

Le positionnement à l'entrée en formation respecte les critères qualité Qualiopi. Dès son inscription définitive, l'apprenant reçoit un questionnaire d'auto-évaluation nous permettant d'apprécier son niveau estimé sur différents types de technologies, ses attentes et objectifs personnels quant à la formation à venir, dans les limites imposées par le format sélectionné. Ce questionnaire nous permet également d'anticiper certaines difficultés de connexion ou de sécurité interne en entreprise (intraentreprise ou classe virtuelle) qui pourraient être problématiques pour le suivi et le bon déroulement de la session de formation.

Méthodes pédagogiques

Stage Pratique : 60% Pratique, 40% Théorie. Support de la formation distribué au format numérique à tous les participants.

Organisation

Le cours alterne les apports théoriques du formateur soutenus par des exemples et des séances de réflexions, et de travail en groupe.

Validation

À la fin de la session, un questionnaire à choix multiples permet de vérifier l'acquisition correcte des compétences.

Sanction

Une attestation sera remise à chaque stagiaire qui aura suivi la totalité de la formation.