

DBT (Data Build Tool)

Description

Adopté par des milliers d'organisations dans le monde, DBT est devenu l'outil incontournable pour l'ingénierie des données en SQL. Que vous soyez data analyst, data engineer ou data scientist, vous repartirez avec des compétences directement applicables à vos projets.

Formation DBT 100 % pratique, animée par des experts de la data, avec des jeux de données concrets, des cas métiers et un mini-projet fil rouge.

Durée

2.00 jours (14.00 heures)

Profils des stagiaires et prérequis

Profils :

- Data analysts, data engineers, data scientists
- Professionnels de la donnée souhaitant industrialiser leurs transformations

Prérequis :

- Maîtrise de SQL (jointures, agrégations, CTE)
- Familiarité avec un entrepôt de données (BigQuery, Redshift, PostgreSQL, etc.)

Objectifs pédagogiques

- Comprendre l'architecture et les concepts fondamentaux de DBT
- Structurer et exécuter un projet DBT de bout en bout
- Créer des modèles SQL dynamiques et maintenables
- Mettre en place des tests, de la documentation et des snapshots
- Intégrer DBT dans une pipeline de données avec de bonnes pratiques

Programme détaillé

- Fondamentaux de DBT et mise en route
 - Présentation générale : modern data stack, rôle de DBT
 - Philosophie ELT vs ETL
 - Installation de DBT (CLI ou DBT Cloud)
 - Connexion à un data warehouse (BigQuery, Snowflake, PostgreSQL...)
 - Structure d'un projet DBT (arborescence, fichiers clés, YAML)
 - Exécution de base : dbt run, dbt debug, dbt compile
 - Premier modèle simple (staging d'une table brute)
 - Exercice pratique 1 : Créer un projet DBT connecté à un entrepôt, loader un fichier seeds et créer un modèle stg_clients.
- Modélisation SQL avec DBT et Jinja
 - Modèles : structure, bonnes pratiques de nommage
 - Utilisation de ref() et gestion des dépendances entre modèles

Stat4decision

37-39 avenue Ledru Rollin Paris 75012

Tel. +33172254082 | E-mail : info@stat4decision.com | www.stat4decision.com

Numéro SIRET: 81048985600015 | Numéro de déclaration d'activité: 11755352275 (auprès du préfet de région de: Ile de France)

Organisme de formation certifié Qualiopi pour ses actions de formation

- Création d'une vue intermédiaire et d'un modèle final
- Introduction à Jinja : variables, blocs, filtres
- Création de modèles dynamiques avec Jinja (ex : filtrage conditionnel, templates)
- Utilisation des seeds comme données de référence
- Exercice pratique 2 : Créer un flux staging → intermediate → marts sur un jeu de données CRM (clients, commandes, produits), avec logique Jinja.
- Tests, documentation, snapshots
 - Tests intégrés (not_null, unique, relationships, accepted_values)
 - Écriture de tests personnalisés
 - Création et utilisation de sources
 - Génération de documentation (dbt docs generate, serve)
 - Écriture de documentation dans les YAML : modèles, colonnes, sources
 - Création et usage des snapshots : suivi des évolutions de données
 - Exercice pratique 3 : Ajouter des tests à l'ensemble des modèles + générer la documentation complète + ajouter un snapshot sur la table clients.
- Industrialisation et bonnes pratiques
 - Structuration du projet : staging / intermediate / marts
 - Organisation des dossiers et conventions de nommage
 - Introduction à DBT Cloud : environnement, déclencheurs, scheduler
 - Intégration dans un pipeline CI/CD (Git, GitHub Actions, dbt Cloud jobs...)
 - Déploiement multi-environnement (dev/prod)
 - Introduction aux macros avancées et packages DBT Hub
- Mini-projet final (fil rouge) : Chaque participant (ou binôme) construit un mini-pipeline analytique
 - Chargement de données (seeds ou source)
 - Construction de modèles multi-niveaux
 - Tests + documentation + snapshot
 - Présentation du résultat (structure, dépendances, tests, doc)

Organisation de la formation

Moyens pédagogiques et techniques

- Support projeté
- Support PDF ou html distribué
- Accès à un environnement de travail durant la formation
- Outils utilisés : DBT / VSCode / Git / PostgreSQL

Dispositif de suivi de l'exécution et d'évaluation des objectifs de la formation

- Quiz et QCM à la fin de chaque module pour valider l'acquisition des connaissances théoriques.
- Mini-projets guidés
- Auto-évaluation sur la compréhension et la mise en application des concepts.
- Feedback individuel : chaque participant se voit proposer un bilan de ses points forts et axes d'amélioration.

Stat4decision

37-39 avenue Ledru Rollin Paris 75012

Tel. +33172254082 | E-mail : info@stat4decision.com | www.stat4decision.com

Numéro SIRET: 81048985600015 | Numéro de déclaration d'activité: 11755352275 (auprès du préfet de région de: Ile de France)

Organisme de formation certifié Qualiopi pour ses actions de formation