

Atelier 3 - Transformation

Pour terminer notre architecture décisionnelle, vous transformez les données pour les faire passer de la structure brute obtenue en atelier 2 à la modélisation proposée dans l'atelier 1.

Étape 1 - Installation dbt

Dans le dépôt de l'atelier précédent, installer les packages dbt avec `pip install dbt-core dbt-duckdb`.

Créer un projet dbt dans le répertoire de votre projet avec `dbt init`.

Créer le fichier `profiles.yml` un fichier `models/sources.yml` référençant la ou les tables du schéma `raw`.

Étape 2 - Normalisation, nettoyage et enrichissement

Dans un schéma `cleansed`, définir un modèle par table du schéma `raw` (atelier 2) qui, selon le modèle cible (atelier 1) :

- garde les colonnes nécessaires,
- normalise les colonnes qui le nécessitent,
- filtre les données inutiles,
- consolide les données ayant plusieurs sources.

Vérifier avec `dbt build` et `dbt run` la bonne exécution de vos transformations

Versionnez.

Étape 3 - Dimensions et faits

Dans un schéma `application`, créer au les modèles d'au moins un fait, en commençant par les dimensions dont il dépend.

Vérifier avec `dbt build` et `dbt run` la bonne exécution de vos transformations

Versionnez.

Étape 4 - Documentation et graphe

Renseignez les descriptions de vos sources et modèles et visualisez la documentation grâce aux commandes `dbt docs build` et `dbt docs serve`.

Faites une copie d'écran de la partie la plus représentative de votre documentation et du graphe DAG de vos modèles.

Remise

Si vous n'intégrez pas les copies d'écran au readme du dépôt remis à l'atelier 2, envoyez un pdf avec les copies d'écran de l'étape 4 par mail à sylvain.labasse@mail-formateur.net avant le 27/01.