

TP 3 : ETL, Alimentation du DW

L'objectif de ce TP3 est de faire tout le travail intermédiaire entre la base de données transactionnelle et la base de données d'analyse (DW), en utilisant l'outil SSIS de Visual Studio Data Tool.

Phase 1 : Création de la base de données intermédiaire Staging DB.

Dans le **SQL Server Management Studio**, créez une nouvelle base de données nommée « **GestionDBStaging** ».

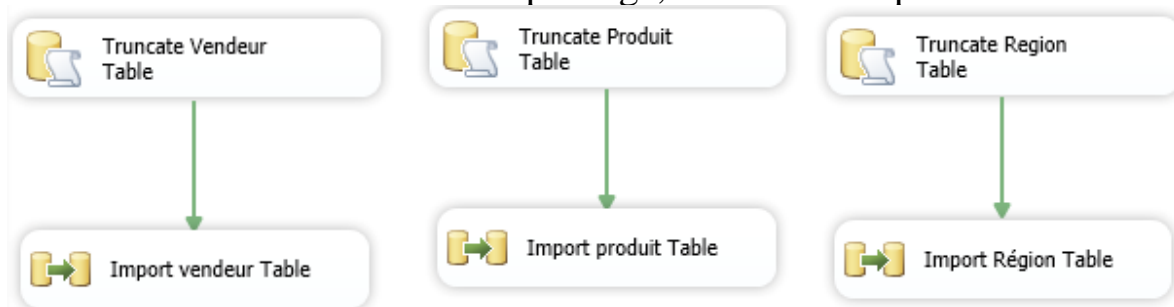
Phase 2 : Mapping GestionDB et GestionDBStaging.

Lancer **Visual Studio Data Tool**.

Créez un nouveau projet, de template Business Intelligence / Integration Services / Integration Services Project for Business Intelligence. Nommez **ssis_db_to_dw_mapping_prj**.

Renommez le package par défaut en « **import_data_to_staging_db.dtsx** ».

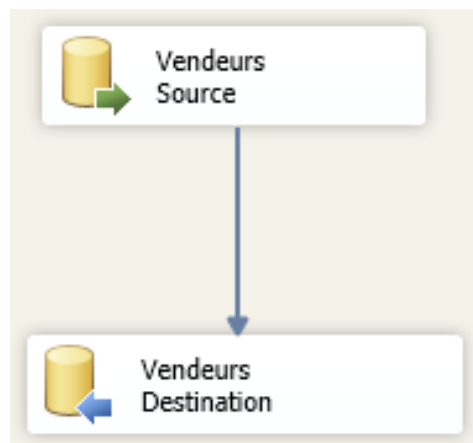
Dans « **le contrôle flow** » de ce package, créez les composants suivants :



Les premiers sont de type « **Execute SQL Task** » et renommés en « **Truncate Vendeur Table** », « **Truncate Produit Table** », « **Truncate**

Region Table ». Ces composants **vident** les tables de Vendeur, Produit et Region de la base de données **GestionDBStaging**, ces composants doivent être connectés à cette base de données et la requête à exécuter est de genre : **TRUNCATE TABLE Vendeur ;**

Les deuxièmes sont de type « **Data Flow Task** » et renommés en « Import vendeur Table », « Import produit Table », « Import region Table ». Ces composants sont faits chacun de **Control Flow** qui extrait les données de la table **GestionDB.Vendeur** et les insère dans la table **GestionDBStaging.Vendeur** :



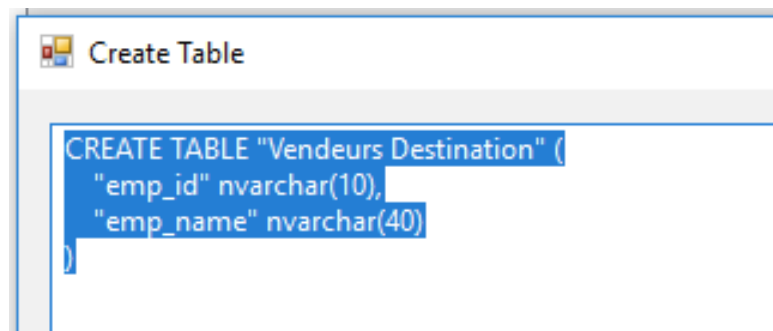
Etant donné que la table **GestionDBStaging.Vendeur** n' existe pas à ce stade, alors vous avez le choix entre la créer via le **SQL Server Management Studio** ou ici via le **composant de Destination** du SSIS. Pour ce faire via ce composant, au moment du choix de la table de destination, cliquez sur **New** au lieu de choisir une existante :

Specify a connection manager, data source, or data source view, and select the table or the view into which the data is copied. Click New to create a new table or view.

Connection manager:
LocalHost.GestionDBStaging ▼ New...

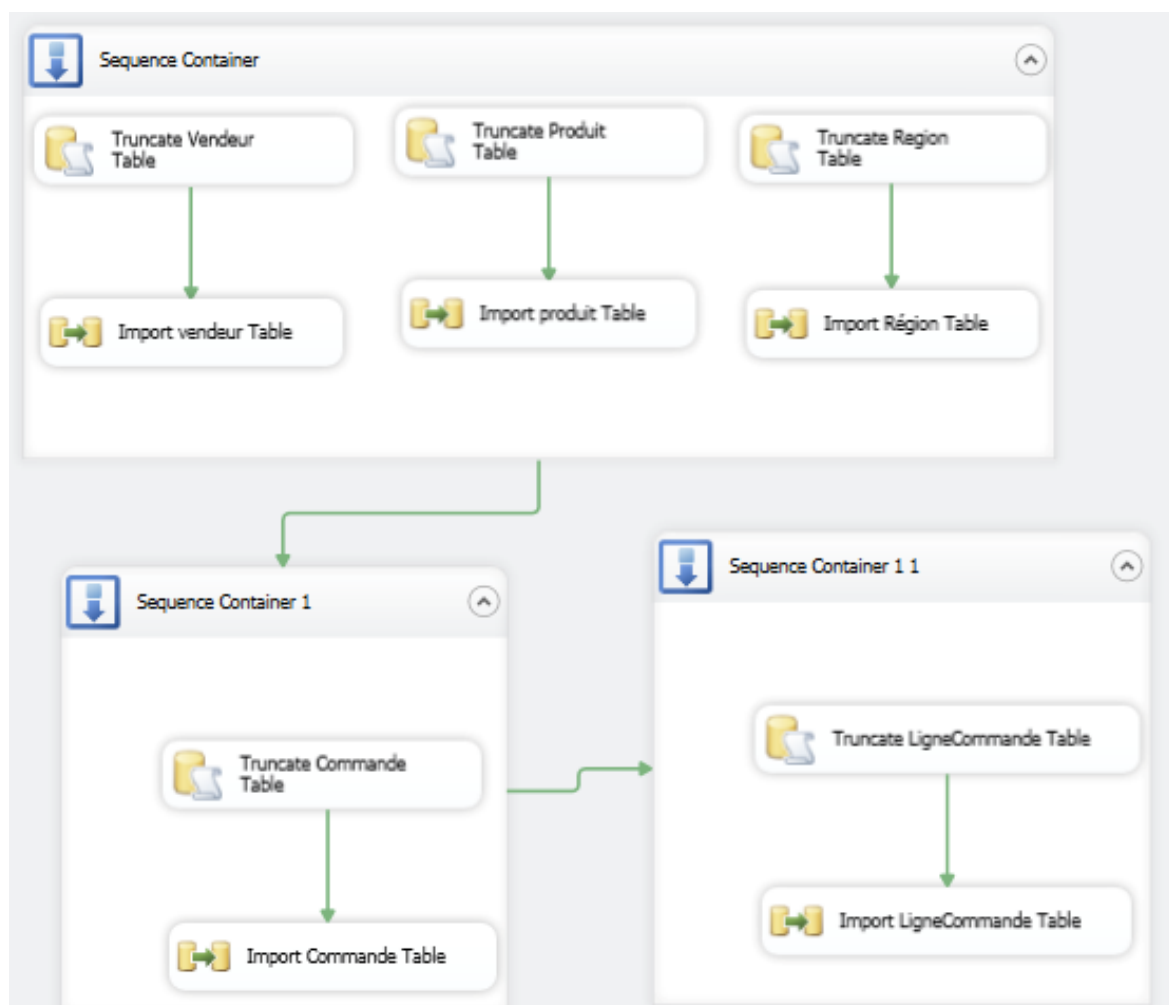
Use a table or view:
Loading... ▼ New...

Puis modifiez la requête SQL à votre convenance :



Surtout n'oubliez pas de changer le nom de la table en **Vendeur** et assurez-vous que les **attributs sont bons**.

Ce que vous avez fait pour ces 3 tables, vous pouvez le faire pour les deux autres. Puisque La table **Commande** contient **des clés étrangères**, il serait plus logique de l'importer après ces trois tables, et la table **LigneCommande** après la table Commande. Un composant SSIS appelé le **Sequence Container** peut assurer ces séquences d'exécution. Ajoutez à votre espace de travail trois instances de ce composant et mettez dedans les blocs nécessaires pour l'import des données vers les cinq tables.



Phase 3 : Mapping GestionDBStaging et DataWarehouseDB.

Créez un nouveau package « **import_data_into_dw.dtsx** ».

Dans ce package, créez un 4 composants Data Flow Task pour importer les les 3 tables de dimensions ainsi que la table des faits.

Le composant 1 : import DimProduit doit simplement copier le contenu de la table produit dans la table de dimension.

Le composant 2 : import DimRegion doit faire la jointure des trois tables Vendeur, Region et Commande pour extraire les données nécessaires et alimenter la table DimProduit. Attention, la clé de cette table de dimension doit être composée d' une concaténation des clés des deux tables Region et Vendeur.

Le composant 3 : import DimTime doit extraire les données nécessaires depuis la table Commande et alimenter la table DimTime. Attention, il ne doit pas y avoir des doublons et la clé de cette table de dimension doit correspondre à la date lue à l' envers. Pour la date 13/02/2019, la clé doit être 20190213.

Le composant 4 : import FactVentes doit extraire les données nécessaires depuis la table Commande et LigneCommande et alimenter la table FactVentes. Attention, faites les agrégations nécessaires.