

Projet B1 Exploration de Données

Une chaîne de magasins souhaite automatiser la gestion de ses offres commerciales basées sur des relevés de prix quotidiens provenant de fichiers CSV. Chaque fichier correspond à une journée précise et contient les prix observés pour certains produits dans divers magasins. L'objectif principal est de maintenir à jour une table **Offer**, où chaque ligne représente une offre active (produit, magasin, prix, période de validité, etc.). Si une offre n'est pas observée un jour, elle doit être considérée comme expirée.

Objectif du projet :

1 – Créer un processus automatisé pour l'importation des données :

- Lire les fichiers CSV contenant les relevés de prix.
- Valider les données et supprimer les anomalies (prix aberrants / prix égale à 0 ...).

2 – Mettre à jour la table **Offer** :

- Créer de nouvelles offres si elles apparaissent dans les relevés et qu'elles n'existent pas encore dans la table **Offer**.
- Mettre à jour les offres existantes (màj de la `VerifiedDate`).
- Clôturer les offres qui ne sont plus actives en fonction des relevés (attention plusieurs cas possibles).

3 – Ajouter un mécanisme de log :

- Créer une table log avec les informations que vous pensez nécessaire à suivre le processus.
- Capturer et enregistrer les étapes du processus (ex : nombre d'offres créées, fermées ...).

4 – Vérification des données :

- Développer un script qui vérifie l'exactitude des mises à jour apportés à la table **Offer**.

Informations complémentaires :

Dans le script **Data.sql** vous avez :

- La création et le remplissage de la table **Store** avec des informations relatives aux magasins.
- La création de la table **Offer** avec les colonnes nécessaire à remplir (**/!\ ne pas ajouter ou supprimer de colonnes**).

- La création et le remplissage de la table EANItem, qui lis un EAN avec son code article et sa description d'article.

Vous avez à disposition 5 csv (DATA20241104.csv, DATA20241111.csv, DATA20241118.csv, DATA20241125.csv, DATA20241202.csv) qui contiennent les relevés de plusieurs magasins du Maine et Loire avec comme colonnes :

- L'EAN du produit,
- Le StoreNumberID (= NumberID de la table Store),
- Le prix recommandé,
- Le prix appliqué en magasin.

Point notable :

- Le script doit permettre d'intégrer les fichiers CSV **un à un**.
- La Date du fichier correspond à laquelle les offres doivent être créées, prolongées ou fermées.
- Une offre active à une ClosingDate au '9999-12-31'
- Faites attention au nombre de ligne pour vérifier la cohérence des données.
- Un malus sera appliqué si le Script n'est pas exécutable.
- Innover au maximum, n'hésitez pas à proposer des initiatives pour améliorer les performances, la façon d'utiliser le script ... ce sont toujours des points gagnés !
- Il faut me remettre les scripts à exécuter et utiles ainsi qu'une petit dossier expliquant le mode opératoire, et l'explication de vos différents choix de conceptions
- Le projet est à rendre maximum **le 13 janvier**.
- Le projet peut être fait en binôme ou solo.

Au besoin n'hésitez pas à me contacter :

cdeletre@uco.fr ou par Teams