

Analyse et explication des fichiers CSV.

1. Fichier ventes.csv :

- **Structure** : Ce fichier contient des informations sur les transactions de vente individuelles. Chaque ligne représente une vente unique.
- **Colonnes** :
 - **Date**: Indique la date de la vente. Le format semble être YYYY-MM-DD. La période couverte s'étend du 2023-05-27 au 2023-06-25, soit environ 30 jours comme mentionné dans le scénario.
 - **ID Référence produit**: Identifiant unique du produit vendu. Il s'agit d'une clé étrangère potentielle reliant ce fichier au fichier produits.csv.
 - **Quantité**: Le nombre d'unités du produit vendu lors de cette transaction. Il s'agit d'une valeur numérique.
 - **ID Magasin**: Identifiant unique du magasin où la vente a eu lieu. Il s'agit d'une clé étrangère potentielle reliant ce fichier au fichier magasin.csv.
- **Relations potentielles** :
 - **ID Référence produit** => produits.csv (pour obtenir le nom, le prix et le stock du produit).
 - **ID Magasin** => magasin.csv (pour obtenir la ville et le nombre de salariés du magasin).

2. Fichier produits.csv :

- **Structure** : Ce fichier contient des informations sur les différents produits vendus. Chaque ligne représente un produit unique.
- **Colonnes** :
 - **Nom**: Le nom du produit (e.g., Produit A, Produit B).
 - **ID Référence produit**: Identifiant unique du produit. Il s'agit de la clé primaire de cette table et correspond à la colonne ID Référence produit dans ventes.csv.
 - **Prix**: Le prix unitaire du produit. Il s'agit d'une valeur numérique (probablement en euros).
 - **Stock**: Le nombre d'unités disponibles en stock pour ce produit. Il s'agit d'une valeur numérique.
- **Relations potentielles** :
 - **ID Référence produit** <= ventes.csv (pour lier les ventes aux informations des produits).

3. Fichier magasin.csv :

- **Structure** : Ce fichier contient des informations sur les différents magasins où les ventes ont lieu. Chaque ligne représente un magasin unique.
- **Colonnes** :
 - **ID Magasin**: Identifiant unique du magasin. Il s'agit de la clé primaire de cette table et correspond à la colonne ID Magasin dans ventes.csv.
 - **Ville**: La ville où se situe le magasin (Paris, Marseille, Lyon, etc...).

- **Nombre de salariés:** Le nombre de personnes travaillant dans ce magasin. Il s'agit d'une valeur numérique.
- **Relations potentielles :**
 - **ID Magasin** <= ventes.csv (pour lier les ventes aux informations des magasins et de leur localisation).

Modélisation des données en tables et relations :

Sur la base de cette analyse, vous pouvez modéliser ces données en trois tables dans votre base de données SQLite :

1. Table ventes :

- date (TEXT)
- ref_produit (TEXT, clé étrangère référençant produits)
- quantite (INTEGER)
- id_magasin (INTEGER, clé étrangère référençant magasins)
- **Clés étrangères potentielles :** (date, ref_produit, id_magasin) pour assurer l'unicité de chaque vente (bien que l'énoncé suggère de vérifier les doublons lors de l'import).

2. Table produits :

- nom_produit (TEXT)
- reference (TEXT, clé primaire)
- prix (REAL)
- stock (INTEGER)

3. Table magasins :

- id_magasin (INTEGER, clé primaire)
- ville (TEXT)
- nombre_salaries (INTEGER)

Relations :

- La table **ventes** a une relation "**ManyToOne**" avec la table **produits** via la colonne **ref_produit**. Un produit peut apparaître dans plusieurs ventes, mais chaque vente concerne un seul produit.
- La table **ventes** a une relation "**ManyToOne**" avec la table **magasins** via la colonne **id_magasin**. Un magasin peut être le lieu de plusieurs ventes, mais chaque vente a lieu dans un seul magasin.