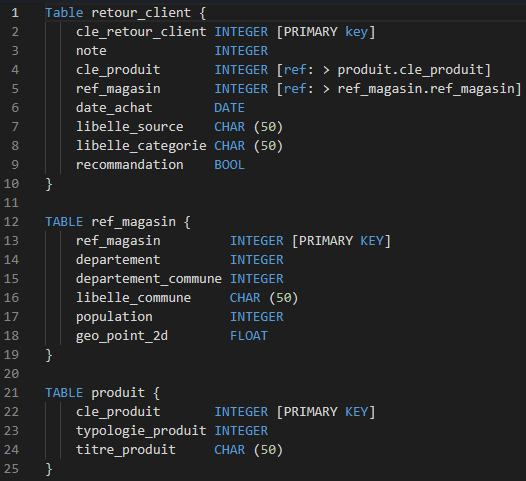
Documentation Technique

# Schéma de la base de données :

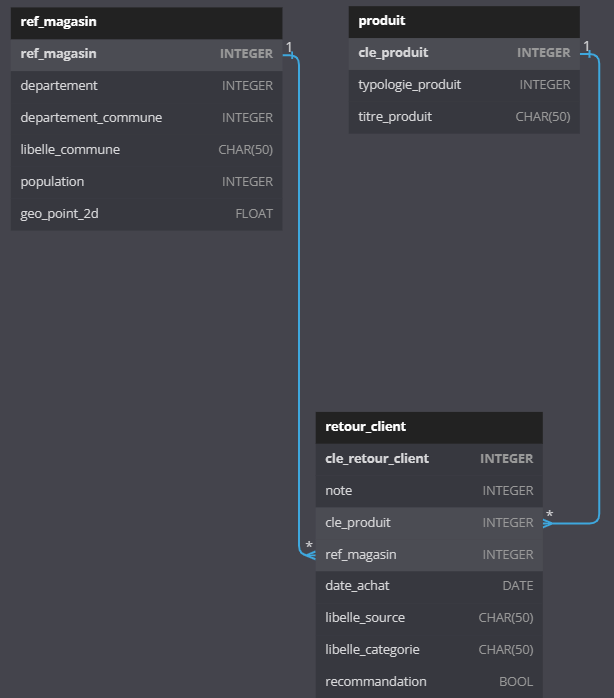
Préparation écrite de la base de données :



Ajout des tables « Produit » et « Ref\_magasin », les types de données, de clés et sa relation.

* Construire trois tables différentes :
  + retour\_client ;
  + ref\_magasin ;
  + produit.
* Le type de donnée :
  + Integers (Nombres entiers) ;
  + Float (Nombres décimaux) ;
  + Character[50] (Suite de caractères) ;
  + Date (Format AAAA-MM-DD) ;
  + Boolean (réponse binaire) ;
* Le type de clé et sa relation :
  + Clé primaire (toutes les tables) ;
  + Clé secondaire (relation one to many vers la table retour\_client) ;

Rendu au format schéma :

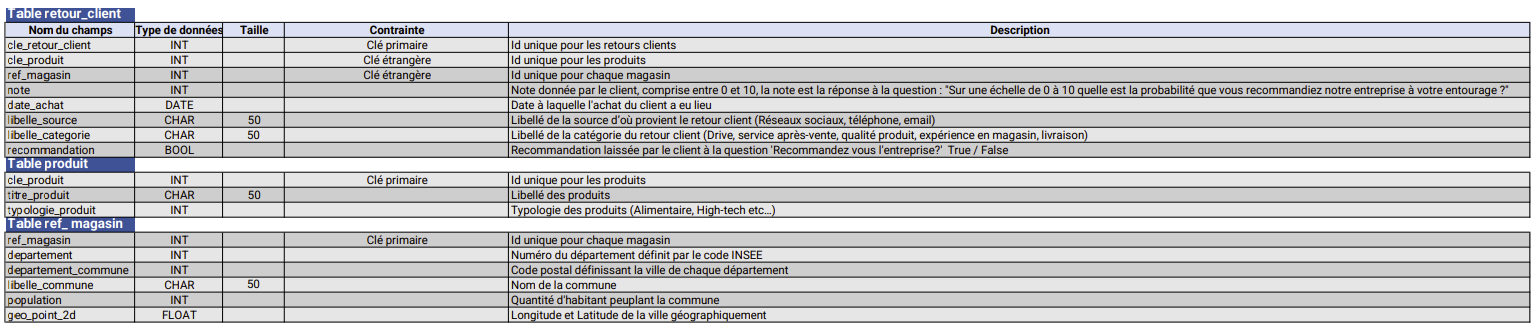


Effectué à partir de DBDiagram.io

Dans chacune des tables, la colonne de donnée en gras représente donc la clé primaire.

La relation est représentée en bleu, avec son type, 1, traduisant l’unicité, vers plusieurs, signifiant la possibilité de répétition.

# Le dictionnaire de données



Ajout des tables « Produit » et « Ref\_magasin », des types de données, et de clés.

# Le système de gestion de la base de données

Table retour\_client :

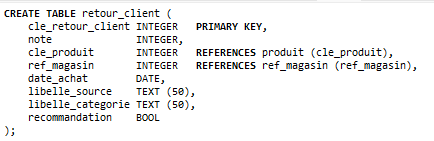


Table ref\_magasin :

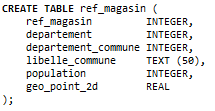
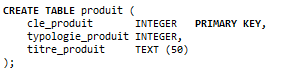
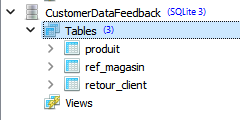


Table produit :



Assemblage des trois tables que sont « Retour\_client », « Produit » et « Ref\_magasin » dans la base de données « Customer\_data\_Feedback ». 

La feuille de données « ref\_magasin » au format xlsx, était mal organisée, le séparateur d’exportation étant la virgule, alors que le séparateur employé était le point-virgule.

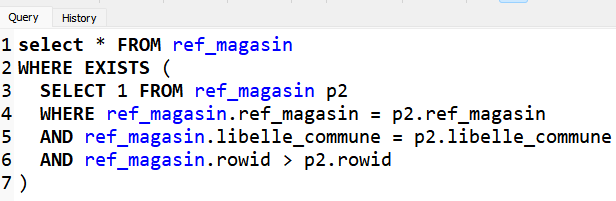
J’ai donc dû reformuler la feuille en concaténant les données latitude et longitude de la colonne « geo\_point\_2D », puis en exportant la feuille sous format csv avec comme séparateur, le point-virgule.

Après l’importation de la table « ref\_magasin », du format CSV, vers la base de données, je me suis retrouvé face à une duplication fois 3 de chaque ligne.

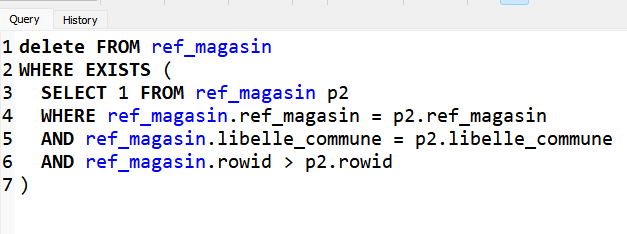
Peut-être une erreur de ma part.

## Nettoyage d’une table de données

J’ai donc effectué un nettoyage, en demandant l’éditeur, de pointer les lignes dupliquées, afin de vérifier les données que je compte évincer :



Puis j’ai demandé la suppression de ces données :



Je suis donc non seulement passé de 252 à 84 lignes, mais j’ai surtout pu assigner le type de clé de la colonne « ref\_magasin », en tant que clé primaire.

Dans la feuille de données « retour\_client », se trouve des valeurs vides, dans la colonne « recommandation ».

D’une part il faudrait les traiter comme étant équivalent à 0, d’autre part, cela représente une forme de passivité. Donc le mieux serait de ne pas corriger ces valeurs nulles, afin de laisser les clients maîtres de leurs choix.

De même pour le calcul du NetPromoterScore, par exemple pour analyser la satisfaction client des magasins.

Certains magasins, n’ont aucun client ayant noté au deca de 7. Dans ce cas les résultats de soustraction des promoteurs face aux détracteurs sont nuls, car il ne s’y trouve aucun détracteur.

Par contre, lors de généralisations, tels que le calcul du NetPromoterScore, ou du pourcentage de recommandation, nous pouvons traiter les valeurs nulles comme étant égal à 0, par l’usage de la fonction coalesce.

