

Documentation Technique

Mise en place et requêtes SQL

Titre du projet	<i>Projet "RetailInsight360"</i>
Contact Business	<i>Mélanie - lead data analyst</i> Olivier-responsable du service client

Version	Auteur	Description	Date
V1	<i>Liodenot julien</i>	<i>Business Intelligence Analyst</i>	<i>01.03.2025</i>

Objectif

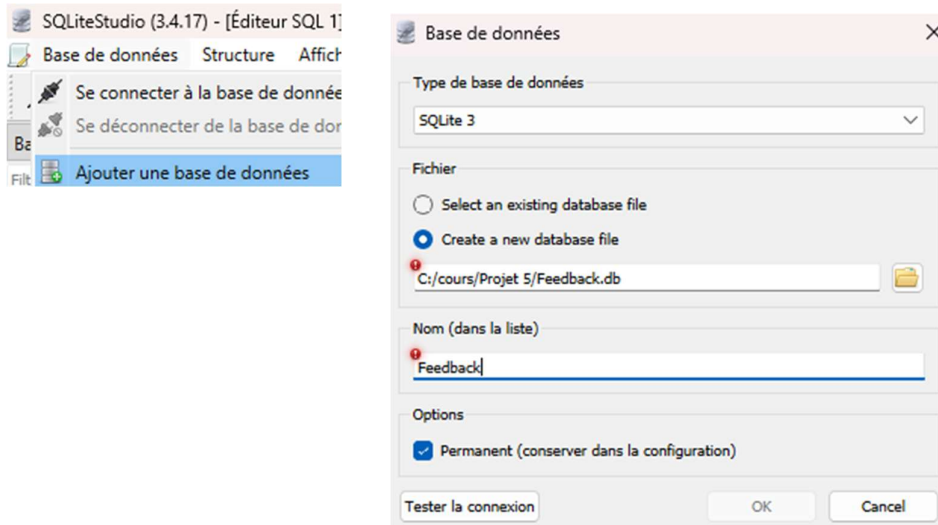
Dans le Cadre du projet stratégique interne "RetailInsight360", la société a besoin d'aide sur l'analyse des données de retours client. L'objectif est ici de détailler la mise en place technique des données récupérées, afin de pouvoir lancer les requêtes d'analyse, puis détailler l'ensemble des requetes

Table des matières

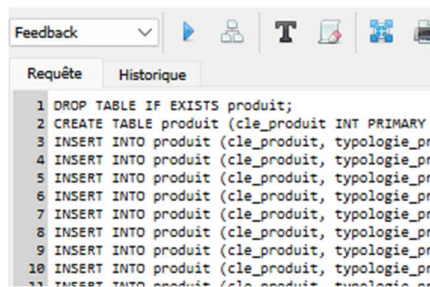
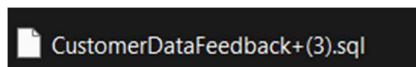
1.Création de la base SQLITE et chargement des données.....	2
2.Ajout d'une table.....	3
3.Ajout d'une clé étrangère	4
4.Requêtes SQL	5

1. Création de la base SQLITE et chargement des données

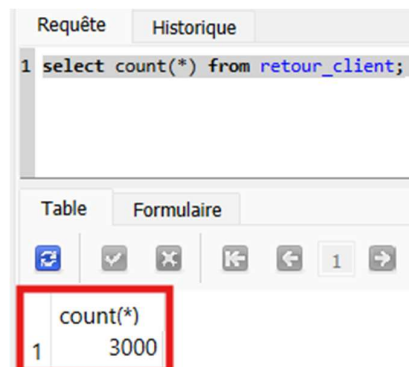
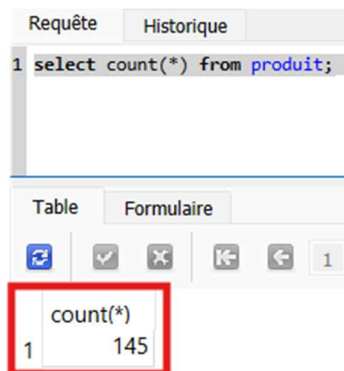
La première étape consiste à créer la base “Feedback” (vide) :



Une fois la base créée il ne reste qu'à lancer le script d'ajout des données qui contient le "create table".



On vérifie alors que l'ensemble des données sont bien présentes, ici produit doit contenir 145 lignes et retour_client 3000 :



2. Ajout d'une table

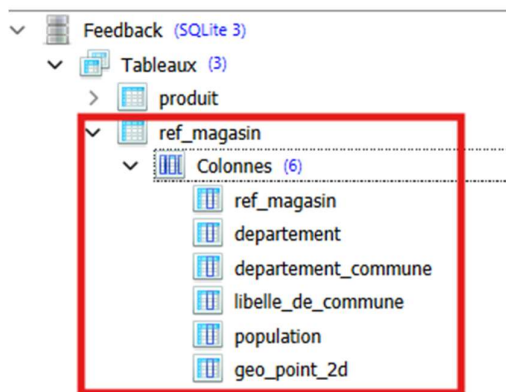
Création de la table :

Une nouvelle table doit être ajoutée : ref_magasin

Pour cela on va utiliser la requête suivante exportée depuis SQL power Architect :

```
CREATE TABLE ref_magasin (
    ref_magasin INT NOT NULL PRIMARY KEY,
    departement VARCHAR(2) NOT NULL,
    departement_commune VARCHAR(5) NOT NULL,
    libelle_de_commune VARCHAR(50) NOT NULL,
    population INT NOT NULL,
    geo_point_2d VARCHAR(50) NOT NULL
);
```

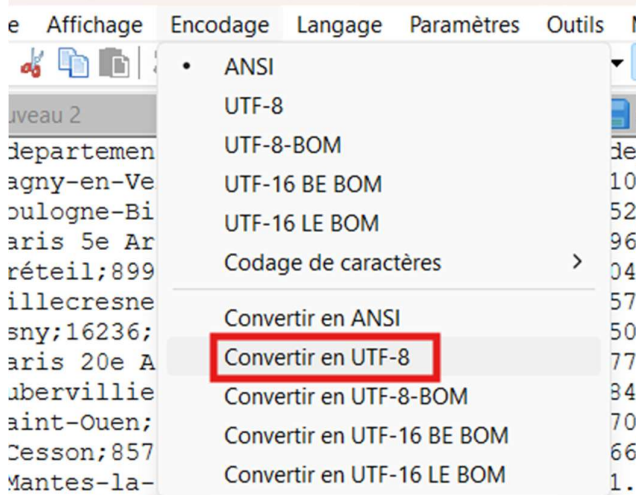
Une fois créée, la table apparait dans la structure de notre base :



Import du fichier de données de la nouvelle table :

Dans un premier temps on va convertir l'encodage de notre fichier à importer en UTF8 via notepad++

\\ref_magasin+(1).csv - Notepad++



On ajoute alors les données de la table fournies via la fonction d'import de SQLITE.

Table à importer vers

Base de données

Feedback

Tableau

ref_magasin

Source de données à importer de

Type de données source

CSV

Options

Fichier :

C:/cours/Projet 5/ref_magasin.csv

Texte codé :

UTF-8

☐ Ignorer les erreurs

Options de source de données

☒ La première ligne représente les noms de colonnes CSV

On vérifie ensuite que les 84 lignes présentes dans le fichier sont bien insérées :

```
1 select count(*) from ref_magasin;
```

Table		Formulaire	
count(*)			
1	84		

3. Ajout d'une clé étrangère

La nouvelle table ref_magasin contient une liaison avec la table retour_client. Il faut donc insérer une clé étrangère (FK) dans retour_client.

SQLite n'acceptant pas les modifications de structures sur une table préalablement remplie, Ci-après la technique utilisée pour éviter un nouvel import de données :

1/On renomme la table en retour_client_old:

```
ALTER TABLE retour_client RENAME TO retour_client_old;
```

2/on recrée la table avec la FK

```
CREATE TABLE retour_client (
  cle_retour_client INT PRIMARY KEY,
  note INT,
  cle_produit INT,
  ref_magasin INT,
  date_achat DATE,
  libelle_source CHAR(50),
  libelle_categorie CHAR(50),
  recommandation CHAR(50),
  FOREIGN KEY (cle_produit) REFERENCES produit(cle_produit)
  FOREIGN KEY (ref_magasin) REFERENCES ref_magasin(ref_magasin)
);
```

3/ on recopie les données dans retour_client

```
INSERT INTO retour_client
SELECT * from retour_client_old;
```

4/on supprime la table temporaire


```
DROP TABLE retour_client_old;
```

4. Requêtes SQL

Ci-après la liste des requêtes SQL issues des questions :

--Nombre de retour client sur la livraison

```
select count(*) as nb_retour_sur_livraison from retour_client where libelle_categorie like '%livraison%';
```

		Nombre de lignes chargées	1
	nb retour sur livraison		
1	639		


-- Quel est le nombre de retour clients par source ?

```
select libelle_source,count(*) as Nbr_retours from retour_client Group By libelle_source;
```

		Nombre de lignes chargées	3
	libelle source	Nbr retours	
1	email	1032	
2	réseaux sociaux	998	
3	téléphone	970	

--Quels sont les magasins qui ont plus de 12 feedbacks sur le drive ?

```
select mag.ref_magasin,mag.libelle_de_commune, count(*) as nb_feedback from ref_magasin mag
LEFT join retour_client rc on mag.ref_magasin=rc.ref_magasin
where libelle_categorie='drive'
group by mag.ref_magasin,mag.libelle_de_commune HAVING count(*)>12
Order by nb_feedback desc;
```






		Nombre de lignes chargées	3
	ref magasin	libelle de commune	nb feedback
1	67	Éragny	14
2	45	Paris 12e Arrondissement	13
3	63	Ivry-sur-Seine	13




-- Sur quel mois a-t-on le plus de retour sur le service après-vente ?

```
SELECT CASE strftime('%m', date_achat)
WHEN '01' THEN 'Janvier'
WHEN '02' THEN 'Février'
WHEN '03' THEN 'Mars'
WHEN '04' THEN 'Avril'
WHEN '05' THEN 'Mai'
WHEN '06' THEN 'Juin'
WHEN '07' THEN 'Juillet'
WHEN '08' THEN 'Août'
WHEN '09' THEN 'Septembre'
WHEN '10' THEN 'Octobre'
WHEN '11' THEN 'Novembre'
WHEN '12' THEN 'Décembre'
END AS mois, COUNT(*) AS Nbr_retour FROM retour_client
WHERE libelle_categorie = 'service après-vente'
GROUP BY mois
ORDER BY Nbr_retour DESC
LIMIT 1;
```

-- Quel est le pourcentage de recommandations client ?

```
select (CAST((select count(*) from retour_client where recommandation=1) as FLOAT)/(select count(*)
from retour_client )*100) as Pourcentage_reco_clients;
```




 Nombre de lignes chargées 1

-- Quel est le classement des départements par note ?

```
select departement,round(AVG(note),2) as Note_Moyenne from ref_magasin mag
LEFT JOIN retour_client RC on mag.ref_magasin=rc.ref_magasin
GROUP BY departement
Order by Note_Moyenne DESC;
```

Nombre de lignes chargées 8

	departement	Note Moyenne
1	95	8.14
2	75	8.11
3	94	8.06
4	91	8.05
5	77	8.04
6	92	8.03
7	78	8.02
8	93	7.94

-- Quel est le classement des jours de la semaine où l'expérience client est la meilleure expérience en magasin?

```
select CASE strftime('%w', date_achat)
WHEN '0' THEN 'Dimanche'
WHEN '1' THEN 'Lundi'
WHEN '2' THEN 'Mardi'
WHEN '3' THEN 'Mercredi'
WHEN '4' THEN 'Jeudi'
WHEN '5' THEN 'Vendredi'
WHEN '6' THEN 'Samedi'
END as jour,round(AVG(note),2) as Note_Moyenne from retour_client
GROUP BY jour Order by Note_Moyenne DESC;
```

Nombre de lignes chargées 7

	jour	Note Moyenne
1	Jeudi	8.19
2	Dimanche	8.14
3	Mardi	8.07
4	Vendredi	8.03
5	Lundi	8
6	Samedi	7.99
7	Mercredi	7.97

-- Quels sont les 5 magasins avec les meilleures notes moyennes ?

```
select mag.ref_magasin,mag.libelle_de_commune,round(AVG(note),2) as Note_Moyenne from
ref_magasin mag
LEFT JOIN retour_client rc on mag.ref_magasin=rc.ref_magasin
GROUP BY mag.ref_magasin
ORDER BY Note_Moyenne DESC
LIMIT 5;
```

Nombre de lignes chargées 5

	ref magasin	libelle de commune	Note Moyenne
1	75	Paris 14e Arrondissement	8.73
2	78	Saint-Pierre-du-Perray	8.55
3	62	Paris 19e Arrondissement	8.5
4	23	Paris 11e Arrondissement	8.48
5	19	Coulommiers	8.45

-- Quelle est la note moyenne pour chaque catégorie de produit ? (Classé de la meilleure à la moins bonne)

```
select typologie_produit, round(AVG(note), 2) as Note_Moyenne from produit prd
LEFT JOIN retour_client RC on prd.cle_produit=RC.cle_produit
GROUP BY typologie_produit
ORDER BY Note_Moyenne DESC;
```

Nombre de lignes chargées : 4

	typologie produit	Note Moyenne
1	High-Tech	8.16
2	Loisirs	8.09
3	Alimentaire	8.04
4	Maison	7.85

-- Quels sont les 5 magasins avec le plus de feedbacks ?

```
select mag.ref_magasin, mag.libelle_de_commune, count(*) as nb_feedback from ref_magasin mag
LEFT JOIN retour_client rc on mag.ref_magasin=rc.ref_magasin
GROUP BY mag.ref_magasin
ORDER BY nb_feedback DESC
LIMIT 5;
```

Nombre de lignes chargées : 5

	ref magasin	libelle de commune	nb feedback
1	29	Mareuil-lès-Meaux	55
2	6	Osny	49
3	80	Lognes	47
4	5	Villecresnes	45
5	63	Ivry-sur-Seine	44

-- Calcul du NPS

```
SELECT ROUND(
(COUNT(CASE WHEN note BETWEEN 9 AND 10 THEN 1 END) * 100.0 / COUNT(*))
-
(COUNT(CASE WHEN note BETWEEN 0 AND 6 THEN 1 END) * 100.0 / COUNT(*)) , 2)
AS NPS
FROM retour_client
```

Table Formulaire

Nombre de lignes chargées : 1

	NPS
1	30.97

-- Calcul du NPS par sources

```
SELECT libelle_source, ROUND(
(COUNT(CASE WHEN note BETWEEN 9 AND 10 THEN 1 END) * 100.0 / COUNT(*))
-
(COUNT(CASE WHEN note BETWEEN 0 AND 6 THEN 1 END) * 100.0 / COUNT(*)) , 2)
AS NPS
FROM retour_client GROUP BY libelle_source
```

Nombre de lignes chargées : 3

	libelle source	NPS
1	email	29.65
2	réseaux sociaux	29.56
3	téléphone	33.81

--Calcul du NPS par Département

```
SELECT departement,ROUND(
(COUNT(CASE WHEN note BETWEEN 9 AND 10 THEN 1 END) * 100.0 / COUNT(*))
-
(COUNT(CASE WHEN note BETWEEN 0 AND 6 THEN 1 END) * 100.0 / COUNT(*)) ,2)
AS NPS
FROM ref_magasin mag
LEFT JOIN retour_client rc on mag.ref_magasin=rc.ref_magasin
GROUP BY departement;
```

Nombre de lignes chargées 8

departement	NPS	Promoteurs	detracteurs	passif
75	32.89	184	37	225
77	29.65	177	43	232
78	29.18	180	42	251
91	31.93	173	36	220
92	32.72	141	35	148
93	26.46	96	28	133
94	32.99	84	20	90
95	31.76	165	30	230

-- Quels sont les 5 magasins avec le moins de feedbacks et le type de notes pour le NPS ?

```
select mag.ref_magasin,mag.libelle_de_commune,count(*) as nb_feedback , COUNT(CASE WHEN
note BETWEEN 9 AND 10 THEN 1 END) as Promoteurs,
COUNT(CASE WHEN note BETWEEN 0 AND 6 THEN 1 END) as detracteurs,
COUNT(CASE WHEN note BETWEEN 7 AND 8 THEN 1 END) as passif
from ref_magasin mag
INNER JOIN retour_client rc on mag.ref_magasin=rc.ref_magasin
GROUP BY mag.ref_magasin
ORDER BY nb_feedback ASC
LIMIT 5;
```

Nombre de lignes chargées 5

	ref magasin	libelle de commune	nb feedback	Promoteurs	detracteurs	passif
1	25	Conflans-Sainte-Honorine	23	5	1	17
2	21	Longjumeau	26	14	1	11
3	34	Asnières-sur-Seine	27	12	4	11
4	46	Paris 15e Arrondissement	27	7	2	18
5	65	Saint-Cyr-l'École	27	12	4	11

--Quelle est la Liste des notes des Clients sur les réseaux sociaux sur les TV ?

```
select distinct note from produit pdt
LEFT JOIN retour_client rc on rc.cle_produit=pdt.cle_produit
where libelle_source ='réseaux sociaux' and titre_produit='TV';
```

Nombre de lignes chargées 3

	note
1	8
2	9
3	10

--Quelle est la note moyenne sur l'ensemble des boissons ?

```
select ROUND(AVG(note)) as Note_Moyenne_Boissons from produit pdt
LEFT JOIN retour_client rc on rc.cle_produit=pdt.cle_produit
where titre_produit like 'Boissons';
```

Nombre de lignes chargées : 1

Note Moyenne Boisson
1 8

--Quels sont les magasins qui ont une note inférieure à la moyenne ?

```
select distinct mag.ref_magasin,mag.libelle_de_commune,round(AVG(note),1) as
note_moyMag,(select round(avg (note),1) from retour_client) as note_moyenneG from ref_magasin
mag
LEFT JOIN retour_client rc on mag.ref_magasin=rc.ref_magasin
GROUP BY mag.ref_magasin,mag.libelle_de_commune
HAVING note_moyMag<note_moyenneG
ORDER BY mag.ref_magasin;
```

ref magasin	libelle de commune	note moyMag	note moyenneG
1	3 Paris 5e Arrondissement	8	8.1
2	6 Osny	7.9	8.1
3	7 Paris 20e Arrondissement	8	8.1
4	8 Aubervilliers	7.7	8.1
5	13 Ballainvilliers	7.9	8.1
6	14 Paris 17e Arrondissement	7.9	8.1
7	18 Saint-Germain-lès-Arpajon	7.8	8.1
8	20 Provins	7.9	8.1
9	24 Levallois-Perret	7.6	8.1
10	25 Conflans-Sainte-Honorine	7.8	8.1
11	33 Paris 18e Arrondissement	7.9	8.1
12	34 Asnières-sur-Seine	7.9	8.1
13	38 Rueil-Malmaison	8	8.1
14	44 Fontenay-sous-Bois	7.7	8.1
15	45 Paris 12e Arrondissement	7.9	8.1
16	46 Paris 15e Arrondissement	7.6	8.1

17	47 Freneuse	8	8.1
18	51 Saint-Denis	7.8	8.1
19	53 Puteaux	8	8.1
20	54 Paris 2e Arrondissement	7.8	8.1
21	55 Rosny-sous-Bois	7.6	8.1
22	57 Herblay	7.7	8.1
23	58 Coignières	7.9	8.1
24	59 Paris 1er Arrondissement	8	8.1
25	60 Buchelay	7.4	8.1
26	63 Ivry-sur-Seine	7.9	8.1
27	64 Ozoir-la-Ferrière	7.9	8.1
28	65 Saint-Cyr-l'École	7.9	8.1
29	66 Les Ulis	7.8	8.1
30	68 Fleury-Mérogis	7.8	8.1
31	73 Viry-Châtillon	8	8.1
32	74 Chelles	7.7	8.1
33	76 Carrières-sur-Seine	7.7	8.1
34	79 Livry-Gargan	7.9	8.1
35	80 Lognes	7.6	8.1
36	81 Nanterre	7.4	8.1
37	82 Montgeron	7.5	8.1

-- Quelles sont les typologies de produits qui ont amélioré leur moyenne entre le 1 er et le 2ème trimestre 2021 ?

```
SELECT typologie_produit,
ROUND(AVG(CASE WHEN date_achat BETWEEN '2021-01-01' AND '2021-03-31' THEN note END),2) AS
"TRIM1",
ROUND(AVG(CASE WHEN date_achat BETWEEN '2021-04-01' AND '2021-06-30' THEN note END),2)
AS "TRIM2"
FROM produit pdt
LEFT JOIN retour_client rc on rc.cle_produit=pdt.cle_produit
GROUP BY typologie_produit
HAVING TRIM1-TRIM2< 0
```

Nombre de lignes chargées : 2

typologie produit	TRIM1	TRIM2
1 Alimentaire	7.99	8.06
2 Loisirs	8	8.34

-- Quelle est la typologie de produit qui apporte le meilleur service après-vente ?

```
select typologie_produit, ROUND(AVG(note),0) as note_moy
from produit pdt
LEFT JOIN retour_client rc on rc.cle_produit=pdt.cle_produit
WHERE libelle_categorie= 'service après-vente'
GROUP BY typologie_produit
ORDER BY note_moy DESC
LIMIT 1;
```

Nombre de lignes chargées : 1

typologie produit	note moy
1 Loisirs	9