



Rapport technique

Réalisé par : Soumare Djibril

Avril 2023

Contexte de la mission

Dat'Assur Habitation, une société spécialisée dans l'assurance habitation pour particuliers, dont elle conçoit et gère les contrats d'assurance pour les particuliers.

L'ambition du directeur général de l'entreprise est d'utiliser les données pour analyser le portefeuille des contrats actuels et prendre des décisions basées sur les données pour se démarquer de la concurrence. Cette démarche permettrait à l'entreprise de devenir **"Data-Driven"**. Le directeur général cherche donc à standardiser la politique tarifaire de l'entreprise.

Le but du projet est de revoir les prix des contrats d'assurance pour se distinguer de leurs concurrents. Cela implique de réviser les prix des contrats d'assurance pour les adapter au profil des clients et à leur situation géographique, afin de proposer aux clients des prix compétitifs et clairs.

C'est ainsi que le responsable technique de la stratégie de développement et d'innovation, m'a confié la mission en tant que consultant en Business Intelligence Analyst pour effectuer une analyse de l'état actuel des contrats.

Pour réaliser cette mission, certaines opérations préalables sont nécessaires, entre autres :

- ✓ Mise en place de la base de données dans un SGBDR;
- ✓ Complétion de dictionnaire de données ;
- ✓ Remplissage du modèle d'aide aux requêtes SQL.

1. Mise en place de la base de données dans un SGBDR

Dans un premier temps, on a procédé à la création de la base de données de l'entreprise sur le Système de Gestion de Base de Données relationnelle (**SGBDR**) SQLiteStudio.

Il existe plusieurs systèmes de gestion de base de données, notre choix est porté sur l'utilisation de SQLiteStudio, car c'est un SGBDR open source, simple et largement accessible qui stocke les données dans des fichiers plats, plutôt que dans un serveur dédié.

Ce qui rend SQLiteStudio particulièrement adapté aux applications mobiles et aux petites applications de bureau. La base de données créée pour l'entreprise est dénommée « **datassur_habitation.db** ».

Nous avons ensuite importé les deux tables au format CSV contenant les données de l'entreprise. La procédure ci-après est utilisée pour charger les données :

- Dans SQLiteStudio, cliquer sur "Import" dans la barre d'outils supérieure.
- Sélectionner "CSV file" dans la liste des types de fichiers d'importation.
- Sélectionner le fichier CSV à importer en cliquant sur le bouton "Choisir le fichier".
- Dans les options d'importation, sélectionner le séparateur ";" et définir les colonnes à importer en les associant aux noms des colonnes de la table créée précédemment.
- Cliquer sur "Import" pour lancer l'importation du fichier CSV vers la table dans la base de données SQLiteStudio.
- Ensuite, vérifier que les données ont été correctement importées en consultant la table dans la base de données.

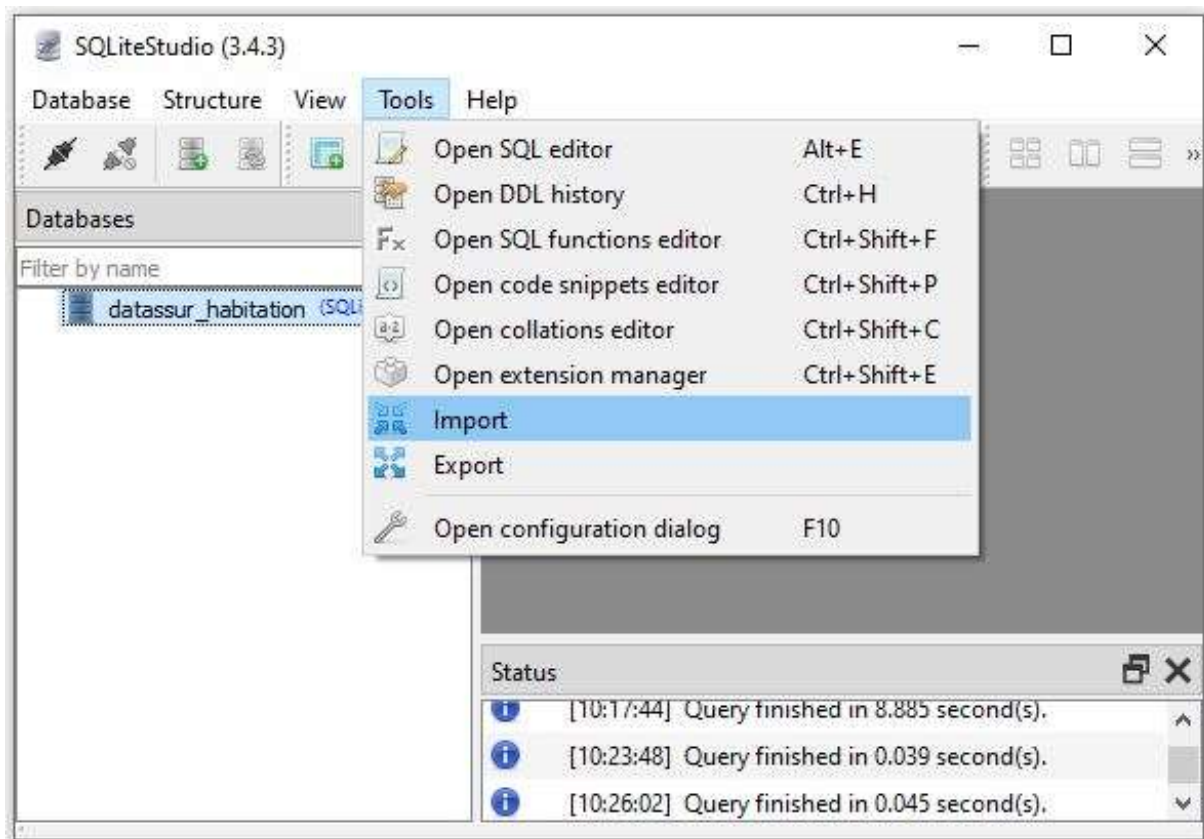


Figure 1: Menu Tools

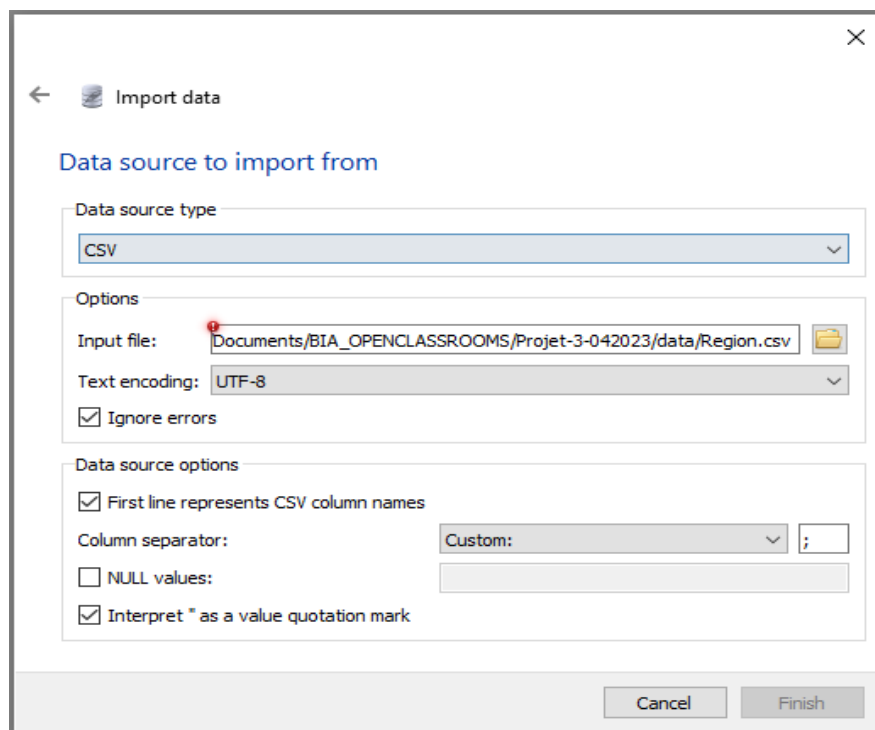


Figure 2: Paramètre import CSV comme table

Les figures ci-dessous représentent les images capturées de ces deux tables importées dans la base de données.

SQLiteStudio (3.4.3) - [ref_geo_regions (datassur_habitation)]

Database Structure View Tools Help

Databases

Filter by name

- datassur_habitation (SQLite 3)
 - Tables (2)
 - contrats_clients
 - ref_geo_regions
 - Views

Structure Data Constraints Indexes Triggers DDL

Grid view Form view

	code_postal	reg_code	reg_nom	aca_nom	dep_nom	com_nom_maj
1	1001	84	Auvergne-Rhône-Alpes	Lyon	Ain	L'ABERGEMENT-CLEMENCIA
2	1002	84	Auvergne-Rhône-Alpes	Lyon	Ain	L'ABERGEMENT-DE-VAREY
3	1003	84	Auvergne-Rhône-Alpes	Lyon	Ain	AMAREINS
4	1004	84	Auvergne-Rhône-Alpes	Lyon	Ain	AMBERIEU-EN-BUGEY
5	1005	84	Auvergne-Rhône-Alpes	Lyon	Ain	AMBERIEUX-EN-DOMBES
6	1006	84	Auvergne-Rhône-Alpes	Lyon	Ain	AMBLEON
7	1007	84	Auvergne-Rhône-Alpes	Lyon	Ain	AMBRONAY
8	1008	84	Auvergne-Rhône-Alpes	Lyon	Ain	AMBUTRIX
9	1009	84	Auvergne-Rhône-Alpes	Lyon	Ain	ANDERT-ET-CONDON
10	1010	84	Auvergne-Rhône-Alpes	Lyon	Ain	ANGLEFORT
11	1011	84	Auvergne-Rhône-Alpes	Lyon	Ain	APREMONT
12	1012	84	Auvergne-Rhône-Alpes	Lyon	Ain	ARANC
13	1013	84	Auvergne-Rhône-Alpes	Lyon	Ain	ARANDAS
14	1014	84	Auvergne-Rhône-Alpes	Lyon	Ain	ARBENT
15	1015	84	Auvergne-Rhône-Alpes	Lyon	Ain	ARBIGNIEU
16	1016	84	Auvergne-Rhône-Alpes	Lyon	Ain	ARBIGNIEU

Status

- [10:54:20] New updates are available. [Click here for details.](#)
- [11:25:50] Query finished in 0.001 second(s).

SQL editor 1 ref_geo_regions (datassur_habitation)

Figure 3: Table de contrats clients et Table référentiel géographique de régions

2. Complétion de dictionnaire de données

Dans cette phase, nous avons exploré l'ensemble de dictionnaire de données pour avoir une idée claire et précise sur le type des données.

Le type de données est une ligne directrice permettant à SQL de comprendre quel type de données est attendu à l'intérieur de chaque colonne, et il identifie également la manière dont SQL interagira avec les données stockées. Chaque valeur stockée dans une base de données SQLite possède l'une des classes de stockage suivantes :

- **NUL**, La valeur est une valeur NULL.
- **ENTIER**, La valeur est un entier signé, stocké sur 0, 1, 2, 3, 4, 6 ou 8 octets selon l'amplitude de la valeur.
- **RÉEL**, La valeur est une valeur à virgule flottante, stockée sous la forme d'un nombre à virgule flottante IEEE de 8 octets.
- **TEXTE**, La valeur est une chaîne de texte, stockée à l'aide du codage de la base de données (UTF-8, UTF-16BE ou UTF-16LE).
- **BLOB**, La valeur est un blob de données, stocké exactement tel qu'il a été saisi.

Le type de données d'une colonne définit la valeur que la colonne peut contenir. Les types de données et leur taille de chaque attribut de table contrat et région sont déterminés dans le dictionnaire de données selon le modèle de type SQLite.

La figure ci-dessous illustre le dictionnaire de données complété.

Nom du champs	Type de données	Taille	Contrainte	Description	Table
Contrat_ID	INT		Clé primaire	Id unique pour les contrats	Contrat
No_voie	INT	4		Numéro dans la voie pour l'adresse du logement assuré	
B_T_Q	CHAR	10		Indicateur éventuel de répétition pour l'adresse du logement assuré	
Type_de_voie	TEXT	4		Type de voie pour l'adresse du logement assuré: rue, av (Avenue), rte (Route), ...	
Voie	TEXT	50		Libellé de la voie pour l'adresse du logement assuré	
Code_postal	INT	5		Code postal pour l'adresse du logement assuré	
Commune	TEXT	50		Libellé de la commune de l'adresse du logement	
Code_departement	INT	3		Code du département pour l'adresse du logement	
Surface	NUMERIC	4		Surface du logement	
Type_local	TEXT	11		Type de logement (Maison, Appartement)	
Occupation	TEXT	12		Type d'occupation (Locataire ou Propriétaire)	
Type_contrat	TEXT	20		Type de contrat (Residence principale, secondaire ou mise en location)	
Formule	TEXT	9		Type de formule choisi par l'assuré (Integral ou Classique)	
Valeur_declare_mobilier	NUMERIC	12		Montant déclaré par l'assuré. Cette valeur sert de base à l'indemnisation en cas de sinistre.	Region
Prix_cotisation_mensuel	NUMERIC	3		La cotisation d'assurance correspond à la somme que doit payer l'assuré chaque mois,	
code_postal	INT		Clé primaire	code commune (referentiel-geographique-francais, source www.data.gouv.fr)	
reg_code	INT	2		Code région	
reg_nom	TEXT	50		Libellé région	
aca_nom	TEXT	25		Libellé de l'académie	
dep_nom	TEXT	50		Libellé du département	
com_nom_maj	TEXT	50		Libellé de la commune en majuscule	
dep_code	INT	3		Code département	Region
dep_nom_num	TEXT	50		Libellé du département et code	

Figure 4: Dictionnaire de données de tables

4. Remplissage du modèle d'aide aux requêtes SQL

Le modèle d'aide à la création de requêtes est conçu pour aider à formuler des requêtes plus efficaces dans le cadre de l'exploration de données ou de l'analyse de données.

C'est un outil très utile pour les utilisateurs qui ne sont pas experts en exploration de données ou en analyse de données, en leur permettant d'obtenir des résultats plus rapidement et plus facilement.

Dans le cadre de cette mission le modèle de requêtes ci-dessous est utilisé pour compléter nos trois premières requêtes. On s'en est inspiré également pour élaborer les neuf requêtes supplémentaires afin d'extraire des informations sur les données d'assurance de l'entreprise.

N°	Besoin	Informations recherchées	Clauses	Requete
1	Quels sont les contrats avec le prix de la cotisation et leur surface pour les appartements ?	Contrats, prix de la cotisation, surface	appartements	<pre>select contrat_id, prix_cotisation_mensuel, surface from contrats_clients where type_local = 'Appartement'</pre>
2	Lister les numéros de contrats (contrat_ID) avec leur surface pour la commune de Caen	contrat_ID, surface	Caen	<pre>select contrat_id, surface from contrats_clients where commune='CAEN'</pre>
3	Lister les numéros de contrats (contrat_ID), avec le type de contrat et leur formule pour les maisons du département de la Saône-et-Loire (Département 71)	contrat_ID, Type_contrat, Formule	Saône-et-Loire (71)	<pre>select c. contrat_ID, c.Type_contrat, c.Formule fromcontrats_clients c join ref_geo_regions r on c.Code_postal=r.code_postal where dep_nom_num='Saône-et-Loire (71)'</pre>
4	Lister le nom des régions de France	reg_nom		<pre>select distinct reg_nom from ref_geo_regions</pre>

Figure 5: Modèle de création des requêtes

Ce document technique représente la première phase de mission qui servira de base pour toutes les autres analyses ou des requêtes. La base de données relationnelle de Dat'Assur Habitation est maintenant créée et bien configurée. Elle est flexible, performant pour les opérations de sélection, de mise à jour et de suppression de données. Elle est donc prête pour répondre aux besoins de l'entreprise.