



**Marché des assurances habitation**

**Document technique**

**2024**

*Julien Séré*

## **SOMMAIRE :**

- I. Exploration du type de données**
  - 1) Types de données et contraintes**
  - 2) Nettoyage des données**
- II. Schéma relationnel**
- III. Code SQL**
- IV. Base de données MySQL**
  - 1) Chargement de la base de données**
  - 2) Quelques requêtes SQL...**

## I. Exploration du type de données

### 1) Types de données et contraintes

Les données ont été répertoriées et renseignées pour leurs types ainsi que leurs tailles.

Les types de données sont les suivantes :

**INTEGER** : entier long

**CHAR** : valeur alpha de longueur fixe

**VARCHAR** : valeur alpha de longueur maximale fixée (Les longueurs maximales sont ainsi renseignées dans la colonne taille)

Ensuite, il est question de définir les clés primaires de chaque table :

Client\_ID : Clé primaire de la table contrat

Code\_dep\_code\_commune : Clé primaire de la table région

Code\_dep\_code\_commune est donc une clé étrangère dans la table contrat

Voici le dictionnaire actualisé des données :

	Nom des colonnes	Type de données	Taille	Clé	Description
CONTRAT.CSV	Contrat_ID	INT		Clé primaire	Id unique pour les contrats
	No_voie	INT (NULL)	4		Numéro dans la voie pour l'adresse du logement assuré
	B_T_Q	CHAR	1		Indicateur éventuel de répétition pour l'adresse du logement assuré sur un caractère
	Type_de_voie	VARCHAR	4		Type de voie pour l'adresse du logement assuré: rue, av (Avenue), rte (Route), ...
	Voie	VARCHAR	100		Libellé de la voie pour l'adresse du logement assuré
	Code_dep_code_commune	VARCHAR	10	Clé secondaire	Concaténation du code département et code commune pour avoir une clé unique
	Code_postal	VARCHAR	5		Code postal pour l'adresse du logement assuré
	Surface	INT			Surface du logement assuré
	Type_Local	VARCHAR	15		Type de logement assuré : maison ou appartement
	Occupation	VARCHAR	15		Type de contrat : Propriétaire ou locataire
	Type_contrat	VARCHAR	20		Type de résidence : Résidence principale ou secondaire
	Formule	VARCHAR	10		Formule du contrat d'assurance
REGION.CSV	Valeur_declaree_biens	VARCHAR	20		Valeur déclarée du logement assuré
	Prix_cotisation_mensuel	INT			Valeur de la cotisation mensuelle client
	Code_dep_code_commune	VARCHAR	10	Clé primaire	Concaténation du code département et code commune pour avoir une clé unique
	reg_code	VARCHAR	3		Code régional
	reg_nom	VARCHAR	50		Nom de la région
	aca_nom	VARCHAR	50		Nom de la métropole
	dep_nom	VARCHAR	50		Nom du département
	com_nom_maj_court	VARCHAR	50		Nom de la commune
	dep_code	VARCHAR	3		Code départemental
	dep_nom_num	VARCHAR	50		Concaténation du nom du département et code département

	Nom des colonnes	Type de données	Taille	Clé	Description
CONTRAT	Contrat_ID	INT		PRIMAIRE	Id unique pour les contrats

	No_voie	INT (NULL)			Numéro dans la voie pour l'adresse du logement assuré
	B_T_Q	CHAR	1		Indicateur éventuel de répétition pour l'adresse du logement assuré sur un caractère
	Type_de_voie	VARCHAR	4		Type de voie pour l'adresse du logement assuré: rue, av (Avenue), rte (Route), ...
	Voie	VARCHAR	100		Libellé de la voie pour l'adresse du logement assuré
	Code_dep_code_commune	VARCHAR	10	ETRANGER	Concaténation du code département et code commune pour avoir une clé unique
	Code_postal	VARCHAR	5		Code postal pour l'adresse du logement assuré
	Surface	INT			Surface du logement assuré
	Type_Local	VARCHAR	15		Type de logement assuré : maison ou appartement
	Occupation	VARCHAR	15		Type de contrat : Propriétaire ou locataire
	Type_contrat	VARCHAR	20		Type de résidence : Résidence principale ou secondaire
	Formule	VARCHAR	10		Formule du contrat d'assurance
	Valeur_declaree_biens	VARCHAR	20		Valeur déclarée du logement assuré
	Prix_cotisation_mensuel	INT			Valeur de la cotisation mensuelle client
<b>REGION</b>	Code_dep_code_commune	VARCHAR	10	Clé primaire	Concaténation du code département et code commune pour avoir une clé unique
	reg_code	VARCHAR	3		Code régional
	reg_nom	VARCHAR	50		Nom de la région
	aca_nom	VARCHAR	50		Nom de la métropole
	dep_nom	VARCHAR	50		Nom du département
	com_nom_maj_court	VARCHAR	50		Nom de la commune
	dep_code	VARCHAR	4		Code départemental
	dep_nom_num	VARCHAR	50		Concaténation du nom du département et code département

Pour certaines colonnes, comme le N°Voie, il est possible également d'utiliser un type VARCHAR puisque le numéro de Voie sert à identifier une adresse et n'est pas destiné à une quelconque opération. Cependant, le dictionnaire étant déjà rempli par un type INTEGER, nous continuons avec un type INTEGER, en prenant soin d'attribuer la valeur NULL à toutes les cellules vides la colonne correspondante.

## 2) Nettoyage des données

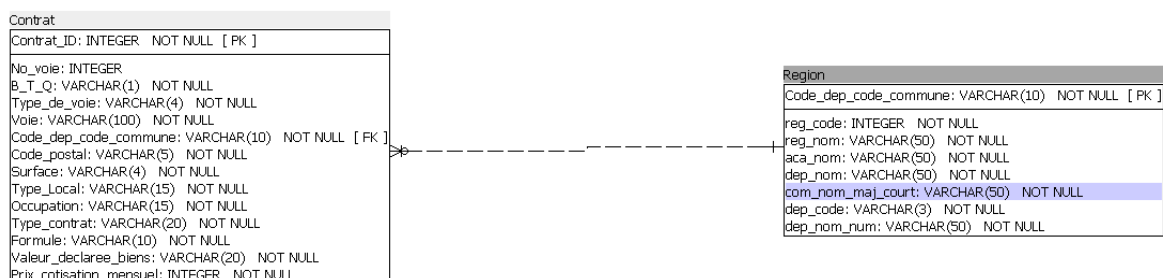
Après avoir défini les types de données qui seront traitées. Il est nécessaire de passer en revue les tableurs de données afin de déceler des possibles erreurs, des cellules vides qui pourraient poser problèmes, ou enfin des valeurs manquantes.

On remarque lors des imports que l'import de la table CONTRAT comprend des erreurs liées aux contrats passés dans le département de La Réunion. En effet, des codes postaux ont été renseignés dans les Code\_dep\_commune ce qui a conduit à une erreur lors de l'identification à la clé primaire de la table Région (données partiellement inexistantes). Ainsi, il convient de veiller à corriger les données erronées en se référant aux codes communes de l'INSEE disponibles en ligne, ou bien en utilisant la table Région pour laquelle les codes sont renseignés.

## II. Schéma relationnel

A l'aide de l'outil SQL Power Architect et du dictionnaire des données, nous pouvons déduire et construire le schéma relationnel des tables des données. Nous veillons bien à renseigner les contraintes pour chaque colonne et renseigner si celle-ci peut accepter les valeurs nulles.

Nous avons vu dans les données que la colonne N° voie contient des cellules vides, il convient alors de renseigner que cette colonne doit nécessairement accepter les valeurs nulles pour que l'import soit un succès.



**SCHEMA RELATIONNEL GENERE PAR POWER A**

L'association de tables est de type un à plusieurs puisque une région peut être associée à plusieurs contrats mais un contrat ne peut être associé qu'à une seule région.

Le code SQL peut donc être généré afin d'importer l'architecture des tables dans notre SGBD.

Pour ce projet, nous travaillons avec MySQL Workbench sur la database openclassrooms.

### III. Code SQL

Voici le code SQL permettant de créer les tables et associations dans notre SGBD MySQL Workbench :

```
CREATE TABLE Region (  
    Code_dep_code_commune VARCHAR(10) NOT NULL,  
    reg_code VARCHAR(3) NOT NULL,  
    reg_nom VARCHAR(50) NOT NULL,  
    aca_nom VARCHAR(50) NOT NULL,  
    dep_nom VARCHAR(50) NOT NULL,  
    com_nom_maj_court VARCHAR(50) NOT NULL,  
    dep_code VARCHAR(3) NOT NULL,  
    dep_nom_num VARCHAR(50) NOT NULL,  
    PRIMARY KEY (Code_dep_code_commune)  
);
```

```
CREATE TABLE Contrat (  
    Contrat_ID INT NOT NULL,  
    No_voie VARCHAR (3) NOT NULL  
    B_T_Q VARCHAR(1) NOT NULL,  
    Type_de_voie VARCHAR(4) NOT NULL,  
    Voie VARCHAR(100) DEFAULT 26 NOT NULL,  
    Code_dep_code_commune_contrat VARCHAR(10) NOT NULL,  
    Code_postal VARCHAR(5) NOT NULL,
```

```
Surface VARCHAR(4) NOT NULL,  
Type_Local VARCHAR(15) NOT NULL,  
Occupation VARCHAR(15) NOT NULL,  
Type_contrat VARCHAR(20) NOT NULL,  
Formule VARCHAR(10) NOT NULL,  
Valeur_declaree_biens VARCHAR(20) NOT NULL,  
Prix_cotisation_mensuel INT NOT NULL,  
PRIMARY KEY (Contrat_ID)  
);
```

```
ALTER TABLE Contrat ADD CONSTRAINT region_contrat_fk  
FOREIGN KEY (Code_dep_code_commune_contrat)  
REFERENCES Region (Code_dep_code_commune)  
ON DELETE NO ACTION  
ON UPDATE NO ACTION;
```

## V. Base de données MySQL

### 1) Chargement de la base de données

La base de données MySQL est désormais configuré et prête à l'emploi.

Les données sont chargées à l'aide de l'outil d'import fourni dans MySQL WorkBench « Table Import Data Wizard ».

Cet outil nous permet d'attribuer aisément les bonnes colonnes avec les bonnes données dans le tableur.

L'import peut également se faire via un code SQL dédié à l'importation :

```
LOAD DATA INFILE '/chemin/vers/monfichier.csv'
```

INTO TABLE nom\_de\_la\_table

FIELDS TERMINATED BY ',' -- Séparateur de colonnes (virgule ici)

ENCLOSED BY '"' -- Si les valeurs sont entourées de guillemets

LINES TERMINATED BY '\n' -- Séparateur de lignes (nouvelle ligne ici)

IGNORE 1 ROWS;

Nous voici avec les bases de données chargées :

1 • `SELECT COUNT(*) FROM contrat`

Result Grid	Filter Rows:	Export:	Wrap Cell Content:
COUNT(*)			
30335			

Result Grid

1 • `SELECT COUNT(*) FROM region`

Result Grid	Filter Rows:	Export:	Wrap Cell Content:
COUNT(*)			
38916			

Result Grid

Nous avons ici la certitude que toutes les lignes ont bien été importé avec la fonction COUNT. Pour pouvoir manipuler et analyser les données, il peut être plus pratique de générer un aperçu du contenu des tables. Il est possible de générer un aperçu des lignes présentes dans chaque table avec la requête suivante :



SELECT \* FROM contrat

SELECT \* FROM region

Contrat	No_vr	B_T_	Type_de_voie	Voie	Code_x	Code_l	Surf	Type_Local	Occupation	Type_contrat	Formule	Valeur_declan	Prix_cotis
100601	190	A	RUE	CENTRALE	1350	1370	50	Appartement	Locataire	Residence principale	Integral	0-25000	25
100602	347		RUE	DU CHATEAU	1103	1170	48	Appartement	Locataire	Residence principale	Classique	0-25000	30
100603	58		AV	DU MONT BL...	1143	1220	131	Appartement	Proprietaire	Residence principale	Integral	25000-50000	57
100604	140		RUE	DE L'ABBE JO...	1288	1630	109	Maison	Locataire	Residence principale	Integral	25000-50000	43
100605	39		RUE	BUFFON	1033	1200	109	Appartement	Locataire	Residence principale	Classique	0-25000	33
100606	8		RUE	DE GENEVE	1354	1630	53	Appartement	Proprietaire	Residence principale	Classique	0-25000	19
100607	2		RUE	DU RECULET	1354	1630	59	Appartement	Proprietaire	Residence principale	Integral	0-25000	15
100608	1403		RUE	JEAN DE GIN...	1143	1220	93	Maison	Proprietaire	Mise en location	Integral	25000-50000	34
100609	226		ALL	DES CAPUCI...	1354	1630	117	Maison	Proprietaire	Residence principale	Classique	25000-50000	32
100610	276		RTE	DE POUIGNY	1288	1630	36	Appartement	Proprietaire	Residence principale	Integral	25000-50000	22
100611	79		CRS	DE VERDUN	1283	1100	138	Appartement	Proprietaire	Residence secondaire	Classique	0-25000	11
100612	240		RUE	DE PRE BAILLY	1173	1170	45	Appartement	Locataire	Residence principale	Classique	0-25000	16
100613	3		RUE	TURENNE	1033	1200	83	Appartement	Locataire	Residence principale	Classique	0-25000	14
100614	44		ALL	DU SQUARE ...	1143	1220	88	Appartement	Locataire	Residence principale	Integral	25000-50000	34
100615	59		RUE	ALEXANDRE ...	1004	1500	165	Appartement	Locataire	Residence principale	Classique	25000-50000	24
100616	282		CHE	DES LONGES ...	1071	1170	42	Appartement	Proprietaire	Residence principale	Classique	0-25000	17
100617	54		GR	GRANDE RUE	1396	1150	68	Appartement	Proprietaire	Residence secondaire	Classique	0-25000	10
100618	1		RUE	DE GEX	1160	1210	83	Appartement	Proprietaire	Residence principale	Classique	0-25000	20
100619	17		LOT	LES JARDINS ...	1103	1170	30	Appartement	Proprietaire	Residence secondaire	Classique	0-25000	12

#### APERÇU DE LA TABLE CONTRAT

Code_dep_code_commune	reg_code	reg_nom	aca_nom	dep_nom	com_nom_maj_court	dep_code	dep_nom_num
10001	44	Grand Est	Reims	Aube	L' ABBAYE SOUS PLANCY	10	Aube (10)
10002	44	Grand Est	Reims	Aube	AILLEVILLE	10	Aube (10)
10003	44	Grand Est	Reims	Aube	AIX EN OTHE	10	Aube (10)
10004	44	Grand Est	Reims	Aube	ALLIBAUDIERES	10	Aube (10)
10005	44	Grand Est	Reims	Aube	AMANCE	10	Aube (10)
10006	44	Grand Est	Reims	Aube	ARCIS SUR AUBE	10	Aube (10)
10007	44	Grand Est	Reims	Aube	ARCONVILLE	10	Aube (10)
10008	44	Grand Est	Reims	Aube	ARGANCON	10	Aube (10)
10009	44	Grand Est	Reims	Aube	ARRELLES	10	Aube (10)
1001	84	Auvergne-Rhône-Alpes	Lyon	Ain	L ABERGEMENT CLEMENCIAT	1	Ain (01)
10010	44	Grand Est	Reims	Aube	ARREMBECOURT	10	Aube (10)
10011	44	Grand Est	Reims	Aube	ARRENTIERES	10	Aube (10)
10012	44	Grand Est	Reims	Aube	ARSONVAL	10	Aube (10)
10013	44	Grand Est	Reims	Aube	ASSENAY	10	Aube (10)
10014	44	Grand Est	Reims	Aube	ASSENCIERES	10	Aube (10)
10015	44	Grand Est	Reims	Aube	AUBETERRE	10	Aube (10)
10016	44	Grand Est	Reims	Aube	AUBIGNY	10	Aube (10)
10017	44	Grand Est	Reims	Aube	AULNAY	10	Aube (10)
10018	44	Grand Est	Reims	Aube	AUXON	10	Aube (10)

#### APERÇU DE LA TABLE REGION

## 2) Quelques requêtes SQL...

Le but d'une requête SQL est de permettre l'interaction avec une base de données pour effectuer diverses opérations, telles que la récupération, l'insertion, la mise à jour ou la suppression de données. SQL est le langage standard pour interroger et manipuler des bases de données relationnelles. Ici, nous pouvons mettre en avant des tendances, des éléments forts, ou des éléments faibles concernant le marché de l'assurance sur lequel la société opère.

### Requête 5 : Quel est le prix moyen de la cotisation mensuelle ?

```
1 SELECT AVG(prix_cotisation_mensuel) AS Moyenne_cotisation
2 FROM contrat
3
```

Result Grid		Filter Rows:	Export:	Wrap Cell Content:
	Moyenne_cotisation			
▶	19.3287			

### Requête 6 : Quel est le nombre de contrats pour chaque catégorie de prix de la valeur déclarée des biens ?

```
1 • SELECT
2     Valeur_declaree_biens AS categorie_prix,
3     COUNT(Contrat_ID) AS nombre_contrats
4 FROM
5     contrat
6 GROUP BY
7     Valeur_declaree_biens
8 ORDER BY
9     Valeur_declaree_biens;
```

	categorie_prix	nombre_contrats
▶	0-25000	22720
	100000+	104
	25000-50000	6815
	50000-100000	696