



# Requêtez une base de données avec SQL

Formation Data Analyst

Projet 3

# Etape 1 > Dictionnaire de données

	Nom des colonnes	Type de données	Taille	Clé	Description
CONTRAT.CSV	Contrat_ID	VARCHAR	15	Clé primaire	Id unique pour les contrats
	No_voie	VARCHAR	100		Numéro dans la voie pour l'adresse du logement assuré
	B_T_Q	VARCHAR	1		Indicateur éventuel de répétition pour l'adresse du logement assuré sur un caractère
	Type_de_voie	VARCHAR	100		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
	Commune	VARCHAR	100		Libellé de la commune de l'adresse du logement
	Code_departement	VARCHAR	3		Code département pour l'adresse du logement assuré
	Surface	INT			Surface du logement assuré
	Type_local	VARCHAR	100		Type du logement assuré (maison, appartement)
	Occupation	VARCHAR	100		Type d'occupation du logement assuré (propriétaire, locataire)
	Type_contrat	VARCHAR	100		Type de contrat du logement assuré (résidence principale, résidence secondaire, mise en location)
	Formule	VARCHAR	100		Formule du contrat du logement assuré (classique, intégral)
	Valeur_declaree_biens	VARCHAR	100		Valeur des biens déclarés pour le logement assuré, par tranche
	Prix_cotisation_mensuel	INT			Prix de la cotisation mensuelle pour le logement assuré
REGION.CSV	Code_dep_code_commune	VARCHAR	10	Clé primaire	Concaténation du code département et code commune pour avoir une clé unique
	reg_code	VARCHAR	2		Code de la région
	reg_nom	VARCHAR	100		Nom de la région
	aca_nom	VARCHAR	100		Nom des villes académiques qui composent la région
	dep_nom	VARCHAR	100		Nom du département dans lequel se situe la ville académique
	com_nom_maj_court	VARCHAR	100		Nom des communes situées dans le département
	dep_code	VARCHAR	3		Code départemental
	dep_nom_num	VARCHAR	100		Nom et numéro du code départemental

- Recherche des typologies de données pour les deux tables
- Ajout des « contraintes » dans le dictionnaire

## Etape 2 > Schéma relationnel de BDD normalisé en 3NF

Contrat_client	
Contrat_ID: VARCHAR(15)	NOT NULL [ PK ]
<hr/>	
No_voie: VARCHAR(100)	NOT NULL
B_T_Q: VARCHAR(1)	
Type_de_voie: VARCHAR(100)	NOT NULL
Voie: VARCHAR(100)	NOT NULL
Code_dep_code_commune: VARCHAR	NOT NULL [ FK ]
Code_postal: VARCHAR(5)	NOT NULL
Commune: VARCHAR(100)	NOT NULL
Code_departement: VARCHAR(3)	NOT NULL
Surface: INTEGER	NOT NULL
Type_local: VARCHAR(100)	NOT NULL
Occupation: VARCHAR(100)	NOT NULL
Type_contrat: VARCHAR(100)	NOT NULL
Formule: VARCHAR(100)	NOT NULL
Valeur_declaree_bien: VARCHAR(100)	NOT NULL
Prix_cotisation_mensuel: INTEGER	NOT NULL

Referentiel_geographique	
Code_dep_code_commune: VARCHAR(10)	NOT NULL [ PK ]
<hr/>	
Reg_code: VARCHAR(2)	NOT NULL
Reg_nom: VARCHAR(100)	NOT NULL
Aca_nom: VARCHAR(100)	NOT NULL
Dep_nom: VARCHAR(100)	NOT NULL
Com_nom_maj_court: VARCHAR(100)	NOT NULL
Dep_code: VARCHAR(3)	NOT NULL
Dep_nom_num: VARCHAR(100)	NOT NULL

- Réalisation sur SQL Power Architect
- Intégration de l'ensemble des données des deux fichiers
- Respect de la norme 3NF

## Etape 3.a > Création de la BDD

```
mysql>
mysql> SHOW DATABASES;
+-----+
| Database |
+-----+
| foodly   |
| information_schema |
| moviz    |
| mysql    |
| performance_schema |
| projet_3 |
+-----+
6 rows in set (0.00 sec)

mysql> USE projet_3;
Database changed
mysql>
mysql> SHOW TABLES;
+-----+
| Tables_in_projet_3 |
+-----+
| contrat_client     |
| referentiel_geographique |
+-----+
2 rows in set (0.00 sec)

mysql>
```

```
mysql>
mysql> SHOW COLUMNS FROM contrat_client;
+-----+-----+-----+-----+-----+-----+
| Field | Type | Null | Key | Default | Extra |
+-----+-----+-----+-----+-----+-----+
| Contrat_ID | varchar(15) | NO | PRI | NULL | |
| No_voie | varchar(100) | YES | | NULL | |
| B_T_Q | varchar(1) | YES | | NULL | |
| Type_de_voie | varchar(100) | YES | | NULL | |
| Voie | varchar(100) | NO | | NULL | |
| Code_dep_code_commune | varchar(10) | NO | | NULL | |
| Code_postal | varchar(5) | NO | | NULL | |
| Commune | varchar(100) | NO | | NULL | |
| Code_departement | varchar(3) | NO | | NULL | |
| Surface | int | NO | | NULL | |
| Type_local | varchar(100) | NO | | NULL | |
| Occupation | varchar(100) | NO | | NULL | |
| Type_contrat | varchar(100) | NO | | NULL | |
| Formule | varchar(100) | NO | | NULL | |
| Valeur_declaree_biens | varchar(100) | NO | | NULL | |
| Prix_cotisation_mensuel | int | NO | | NULL | |
+-----+-----+-----+-----+-----+-----+
16 rows in set (0.00 sec)

mysql> SHOW COLUMNS FROM referentiel_geographique;
+-----+-----+-----+-----+-----+-----+
| Field | Type | Null | Key | Default | Extra |
+-----+-----+-----+-----+-----+-----+
| Code_dep_code_commune | varchar(10) | NO | PRI | NULL | |
| reg_code | varchar(2) | NO | | NULL | |
| reg_nom | varchar(100) | NO | | NULL | |
| aca_nom | varchar(100) | NO | | NULL | |
| dep_nom | varchar(100) | NO | | NULL | |
| com_nom_maj_court | varchar(100) | NO | | NULL | |
| dep_code | varchar(3) | NO | | NULL | |
| dep_nom_num | varchar(100) | NO | | NULL | |
+-----+-----+-----+-----+-----+-----+
8 rows in set (0.00 sec)

mysql>
```

- Création de la BDD projet\_3 dans MySQL
- Création des tables contrat\_client et referentiel\_geographique

## Etape 3.b > Chargement de la BDD

```
mysql> LOAD DATA INFILE 'C:/ProgramData/MySQL/MySQL Server 8.0/Uploads/Contrat_client.csv'  
ED BY '\n' IGNORE 1 rows;  
Query OK, 30335 rows affected (2.00 sec)  
Records: 30335 Deleted: 0 Skipped: 0 Warnings: 0
```

```
mysql>  
mysql> LOAD DATA INFILE 'C:/ProgramData/MySQL/MySQL Server 8.0/Uploads/referentiel_geographique.csv'  
Y '' LINES TERMINATED BY '\n' IGNORE 1 rows;  
Query OK, 38916 rows affected (1.66 sec)  
Records: 38916 Deleted: 0 Skipped: 0 Warnings: 0
```

- Chargement de la BDD contrat\_client : 30 335 lignes
- Chargement de la BDD referentiel\_geographique : 38 916 lignes

# Etape 4 > Premières requêtes SQL

**Requête 1 : Lister les numéros de contrats (contrat\_ID) avec leur surface pour la commune de Caen.**

```
mysql> SELECT Contrat_ID, Surface FROM contrat_client WHERE Commune = 'Caen';
```

```
mysql>
mysql> SELECT Contrat_ID, Commune, Surface FROM contrat_client WHERE Commune = 'Caen';
+-----+-----+-----+
| Contrat_ID | Commune | Surface |
+-----+-----+-----+
| 103791     | CAEN    | 35      |
| 103792     | CAEN    | 99      |
| 103793     | CAEN    | 40      |
| 103794     | CAEN    | 20      |
+-----+-----+-----+
4 rows in set (0.02 sec)
```

**Requête 2 : Lister les numéros de contrats avec le type de contrat et leur formule pour les maisons du département 71.**

```
mysql> SELECT Contrat_ID, Type_contrat, Formule, Type_local, Code_departement FROM contrat_client WHERE Code_departement = '71' AND Type_local = 'Maison';
```

```
mysql>
mysql> SELECT Contrat_ID, Type_contrat, Formule, Type_local, Code_departement FROM contrat_client WHERE Code_departement = '71' AND Type_local = 'Maison';
+-----+-----+-----+-----+-----+
| Contrat_ID | Type_contrat | Formule | Type_local | Code_departement |
+-----+-----+-----+-----+-----+
| 114768     | Residence principale | Integral | Maison | 71                |
| 114779     | Residence principale | Classique | Maison | 71                |
| 114782     | Residence principale | Classique | Maison | 71                |
| 114812     | Residence principale | Integral | Maison | 71                |
+-----+-----+-----+-----+-----+
4 rows in set (0.02 sec)
```

**Requête 3 : Lister le nom des régions de France.**

```
mysql> SELECT DISTINCT reg_nom FROM referentiel_geographique;
```

```
mysql>
mysql> SELECT DISTINCT reg_nom FROM referentiel_geographique;
+-----+
| reg_nom |
+-----+
| Grand Est |
| Auvergne-Rhone-Alpes |
| Occitanie |
| Provence-Alpes-Cote d'Azur |
| Normandie |
| Nouvelle-Aquitaine |
| Centre-Val de Loire |
| Hauts-de-France |
| Bourgogne-Franche-Comte |
| Bretagne |
| Corse |
| Pays de la Loire |
| Ile-de-France |
| Guadeloupe |
| Martinique |
| Guyane |
| La Reunion |
| Collectivites d'outre-mer |
| Mayotte |
+-----+
19 rows in set (0.03 sec)
```

# Etape 5 > Ensemble des requêtes SQL

## 1. Combien existe-t-il de contrats sur les résidences principales ?

```
mysql> SELECT COUNT(*) FROM contrat_client WHERE Type_contrat LIKE "Residence principale";
```

```
mysql>
mysql> SELECT COUNT(*) FROM contrat_client WHERE Type_contrat LIKE "Residence principale";
+-----+
| COUNT(*) |
+-----+
|      25620 |
+-----+
1 row in set (0.04 sec)

mysql>
```

## 2. Quelle est la surface moyenne des logements avec un contrat à Paris ?

```
mysql> SELECT AVG (Surface) FROM contrat_client WHERE Commune LIKE "Paris%";
```

```
mysql>
mysql> SELECT AVG (Surface) FROM contrat_client WHERE Commune LIKE "Paris%";
+-----+
| AVG (Surface) |
+-----+
|          51.7695 |
+-----+
1 row in set (0.02 sec)

mysql>
```



### 3. Quels sont les 5 contrats qui ont les surfaces les plus élevées ?

```
mysql> SELECT * FROM contrat_client ORDER BY Surface DESC LIMIT 5;
```

```
mysql> SELECT * FROM contrat_client ORDER BY Surface DESC LIMIT 5;
```

Contrat_ID	No_voie	B_T_Q	Type_de_voie	Voie	Code_dep_code_commune	Code_postal	Commune	Code_departement	Surface	Type_local
Occupation	Type_contrat		Formule	Valeur_declaree_biens	Prix_cotisation_mensuel					
104211	34	B	CRS	VICTOR HUGO	17421	17600	SAUJON	17	815	Appartement
Locataire	Residence principale		Classique	100000+		240				
105463	5229		RUE	DE LA PRIVADIERE	30126	30190	GARRIGUES STE EULALIE	30	742	Maison
Proprietaire	Mise en location		Integral	50000-100000		189				
130878	23		BD	DE BEAUSEJOUR	75116	75016	PARIS 16	75	595	Appartement
Proprietaire	Residence principale		Classique	100000+		361				
100822	5	A	RUE	FERNAND BAZIN	2201	2600	COEUVRES ET VALSERY	2	570	Maison
Proprietaire	Residence principale		Integral	100000+		111				
109872	18		RUE	PAUL BELLAMY	44109	44000	NANTES	44	559	Appartement
Locataire	Residence principale		Integral	50000-100000		85				

5 rows in set (0.04 sec)

```
mysql>
```

### 4. Quel est le prix moyen de la cotisation mensuelle ?

```
mysql> SELECT AVG(Prix_cotisation_mensuel) FROM contrat_client;
```

```
mysql> SELECT AVG(Prix_cotisation_mensuel) FROM contrat_client;
```

AVG(Prix_cotisation_mensuel)
19.3287

1 row in set (0.01 sec)

```
mysql>
```



5. Quel est le nombre de contrats pour chaque catégorie de prix de la valeur déclarée des biens ?

```
mysql> SELECT Valeur_declaree_biens, COUNT(*) AS Contrat_ID FROM contrat_client GROUP BY  
Valeur_declaree_biens ORDER BY Valeur_declaree_biens ASC;
```

```
mysql> SELECT Valeur_declaree_biens, COUNT(*) AS Contrat_ID FROM contrat_client GROUP BY Valeur_declaree_biens ORDER BY Valeur_declaree_biens ASC;
```

Valeur_declaree_biens	Contrat_ID
0-25000	22720
100000+	104
25000-50000	6815
50000-100000	696

```
4 rows in set (0.04 sec)
```

```
mysql>
```

6. Quel est le classement des 10 départements où le prix moyen de la cotisation est le plus élevé ?

```
mysql> SELECT referentiel_geographique.dep_nom, AVG(contrat_client.Prix_cotisation_mensuel) AS  
Moyenne_cotisation FROM contrat_client INNER JOIN referentiel_geographique ON  
contrat_client.Code_dep_code_commune = referentiel_geographique.Code_dep_code_commune GROUP BY  
referentiel_geographique.dep_nom ORDER BY Moyenne_cotisation DESC LIMIT 10;
```

```
mysql> SELECT referentiel_geographique.dep_nom, AVG(contrat_client.Prix_cotisation_mensuel) AS Moyenne_cotisation FROM contrat_client INNER JOIN referentiel_geographique ON  
contrat_client.Code_dep_code_commune = referentiel_geographique.Code_dep_code_commune GROUP BY referentiel_geographique.dep_nom ORDER BY Moyenne_cotisation DESC LIMIT 10;
```

dep_nom	Moyenne_cotisation
Paris	35.3428
Hauts-de-Seine	26.4882
Allier	20.0000
Val-de-Marne	19.7327
Rhone	18.9572
Yvelines	18.1831
Alpes-Maritimes	18.1221
Ain	17.9526
Charente-Maritime	17.2957
La Reunion	17.2500

```
10 rows in set (0.16 sec)
```

```
mysql>
```

7. Quel est le nombre de contrats avec des formules "Intégral" pour la région Pays de la Loire ?

```
mysql> SELECT COUNT(*) AS Nombre_contrats  
-> FROM contrat_client  
-> INNER JOIN referentiel_geographique ON contrat_client.Code_dep_code_commune =  
referentiel_geographique.Code_dep_code_commune  
-> WHERE referentiel_geographique.reg_nom = "Pays de la Loire"  
-> AND contrat_client.Formule = "Integral";
```

```
mysql> SELECT COUNT(*) AS Nombre_contrats  
mysql> SELECT COUNT(*) AS Nombre_contrats  
-> FROM contrat_client  
-> INNER JOIN referentiel_geographique ON contrat_client.Code_dep_code_commune = referentiel_geographique.Code_dep_code_commune  
-> WHERE referentiel_geographique.reg_nom = "Pays de la Loire"  
-> AND contrat_client.Formule = "Integral";
```

Nombre_contrats
561

```
1 row in set (0.06 sec)
```

```
mysql>
```

#### 8. Quelle est la liste des communes ayant au moins 150 contrats ?

```
mysql> SELECT Commune FROM contrat_client GROUP BY Commune HAVING COUNT(*) >= 150;
```

```
mysql> SELECT Commune FROM contrat_client GROUP BY Commune HAVING COUNT(*) >= 150;
+-----+
| Commune |
+-----+
| NICE    |
| TOULOUSE |
| BORDEAUX |
| GRENOBLE |
| NANTES  |
| LILLE   |
| TOULON  |
| COURBEVOIE |
| PARIS 3 |
| PARIS 10 |
| PARIS 9 |
| PARIS 11 |
| PARIS 12 |
| PARIS 14 |
| PARIS 15 |
| PARIS 16 |
| PARIS 17 |
| PARIS 18 |
| PARIS 20 |
| PARIS 19 |
+-----+
20 rows in set (0.04 sec)

mysql>
```

#### 9. Quel est le nombre de contrats pour chaque région ?

```
mysql> SELECT rg.reg_nom, COUNT(cc.Contrat_ID) AS Nombre_contrats
```

```
-> FROM contrat_client AS cc
```

```
-> INNER JOIN referentiel_geographique AS rg ON cc.Code_dep_code_commune = rg.Code_dep_code_commune
```

```
-> GROUP BY rg.reg_nom;
```

```
mysql>
mysql> SELECT rg.reg_nom, COUNT(cc.Contrat_ID) AS Nombre_contrats
  -> FROM contrat_client AS cc
  -> INNER JOIN referentiel_geographique AS rg ON cc.Code_dep_code_commune = rg.Code_dep_code_commune
  -> GROUP BY rg.reg_nom;
+-----+-----+
| reg_nom | Nombre_contrats |
+-----+-----+
| Auvergne-Rhone-Alpes | 2972 |
| Hauts-de-France | 1327 |
| Provence-Alpes-Cote d'Azur | 3287 |
| Grand Est | 806 |
| Occitanie | 1837 |
| Normandie | 898 |
| Nouvelle-Aquitaine | 2097 |
| Centre-Val de Loire | 804 |
| Bourgogne-Franche-Comte | 402 |
| Bretagne | 945 |
| Corse | 247 |
| Pays de la Loire | 1138 |
| Ile-de-France | 13474 |
| Martinique | 60 |
| Guyane | 37 |
| La Reunion | 4 |
+-----+-----+
16 rows in set (0.12 sec)

mysql>
```