

# METHODOLOGIE POUR REQUETTER UNE BASE DE DONNEES AVEC SQL

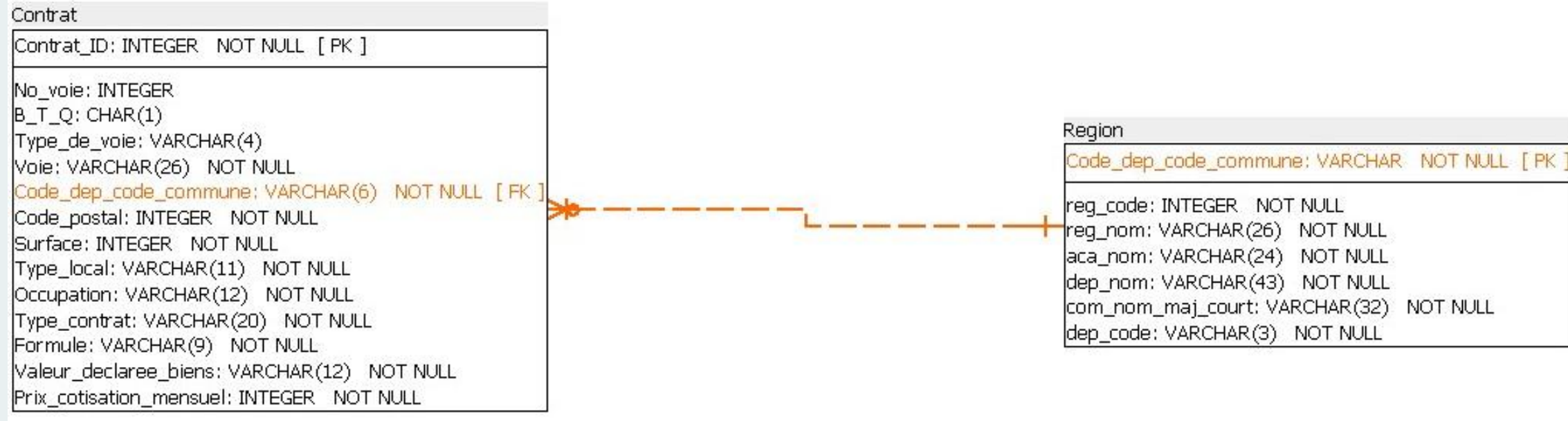


# Rédaction du dictionnaire des données

	Nom des colonnes	Type de données	Taille	Clé	Description
CONTRAT.CSV	Contrat_ID	INT		Clé primaire	Id unique pour les contrats
	No_voie	INT			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	26		Libellé de la voie pour l'adresse du logement assuré
	Code_dep_code_commune	VARCHAR	6	Clé étrangère	Concaténation du code département et code commune pour avoir une clé unique
	Code_postal	INT			Code postal pour l'adresse du logement assuré
	Surface	INT			Surperficie du logement assuré (en m2)
	Type_local	VARCHAR	11		Type de logement assuré
	Occupation	VARCHAR	12		statut juridique de l'assuré
	Type_contrat	VARCHAR	20		Usage du logement assuré
	Formule	VARCHAR	9		Niveau de couverture du logement assuré
	Valeur_declarée_biens	VARCHAR	12		Tranche de la valeur déclarée du logement assuré
REGION.CSV	Prix_cotisation_mensuel	INT			prime d'assurance mensuelle
	Code_dep_code_commune	VARCHAR	6	Clé primaire	Concaténation du code département et code commune pour avoir une clé unique
	reg_code	INT			Code de la région
	reg_nom	VARCHAR	26		Libellé de la région
	aca_nom	VARCHAR	24		Libellé de l'académie
	dep_nom	VARCHAR	43		Libellé du département
	com_nom_maj_court	VARCHAR	32		Libellé court de la commune en majuscule
	dep_code	VARCHAR	3		Code du département
	dep_nom_num	VARCHAR	49		Nom du département associé au code département entre parenthèses

- ✓ Deux fichiers au format csv : **liste des contrats des clients** et **caractéristiques des régions**
- ✓ Traitement effectué:
  - Identification de la taille et du type de chaque variable
  - Identification des variables avec des observations manquantes
  - Description des variables

# Conception du Schéma relationnel entre les tables Contrat et Region



- ✓ Utilisation de **SQL Power Architects**
- ✓ Création des variables et définition des contraintes
- ✓ Définition du lien entre les deux tables : la variable **Code\_dep\_code\_commune**, clé primaire (PK) de la table **Region** et clé étrangère (FK) de la table **Contrat**.

# Création de la base de données composée de deux tables









(1/2)

## Table des régions

```
CREATE TABLE Region (  
    Code_dep_code_commune VARCHAR (6) NOT NULL,  
    reg_code INT NOT NULL,  
    reg_nom VARCHAR (26) NOT NULL,  
    aca_nom VARCHAR (24) NOT NULL,  
    dep_nom VARCHAR (43) NOT NULL,  
    com_nom_maj_court VARCHAR (32) NOT NULL,  
    dep_code VARCHAR (3) NOT NULL,  
    PRIMARY KEY (Code_dep_code_commune)  
);
```

```
CREATE TABLE nom de la table (  
    variable 1, type (taille) contrainte,  
    variable 1, type (taille) contrainte,  
    ...  
    définition de la clé primaire );
```

Grid view Form view

     1    Total rows loaded: 38916

	Code dep	req code	req nom	aca nom	dep nom	com nom maj court	dep code
1	1001	84	Auvergne-Rhône-Alpes	Lyon	Ain	L ABERGEMENT CLEMENCIAT	1
2	1002	84	Auvergne-Rhône-Alpes	Lyon	Ain	L ABERGEMENT DE VAREY	1
3	1003	84	Auvergne-Rhône-Alpes	Lyon	Ain	AMAREINS	1
4	1004	84	Auvergne-Rhône-Alpes	Lyon	Ain	AMBERIEU EN BUGEY	1
5	1005	84	Auvergne-Rhône-Alpes	Lyon	Ain	AMBERIEUX EN DOMBES	1










# Création de la base de données composée de deux tables

(2/2)

## Table des contrats

```
CREATE TABLE Contrat (  
  Contrat_ID INT NOT NULL,  
  No_voie INT,  
  B_T_Q CHAR (1),  
  Type_de_voie VARCHAR (4),  
  Voie VARCHAR (26) NOT NULL,  
  Code_dep_code_commune VARCHAR (6) NOT NULL,  
  Code_postal INT NOT NULL,  
  Surface INT NOT NULL,  
  Type_local VARCHAR (11) NOT NULL,  
  Occupation VARCHAR (12) NOT NULL,  
  Type_contrat VARCHAR (20) NOT NULL,  
  Formule VARCHAR (9) NOT NULL,  
  Valeur_declaree_biens VARCHAR (12) NOT NULL,  
  Prix_cotisation_mensuel INT NOT NULL,  
  PRIMARY KEY (Contrat_ID)  
  FOREIGN KEY (Code_dep_code_commune)  
    REFERENCES Region (Code_dep_code_commune));
```

**CREATE TABLE** *nom de la table* (  
 variable 1, **type** (taille) **contrainte**,  
 variable 1, **type** (taille) **contrainte**,  
 ...  
 définition de la clé primaire  
 définition de la clé étrangère);

Grid view		Form view																	
																		Total rows loaded: 30335	
	Contrat ID	No voie	B T Q	Type de voie	Voie	Code dep	Code pos	Surface	Type local	Occupation	Type contrat	Formule	Valeur declare	Prix cotisa					
1	100601	190	A	RUE	CENTRALE	1350	1370	50	Appartement	Locataire	Residence principale	Integral	0-25000	25					
2	100602	347		RUE	DU CHATEAU	1103	1170	48	Appartement	Locataire	Residence principale	Classique	0-25000	30					
3	100603	58		AV	DU MONT BLANC	1143	1220	131	Appartement	Proprietaire	Residence principale	Integral	25000-50000	57					
4	100604	140		RUE	DE L'ABBE JOLIVET	1288	1630	109	Maison	Locataire	Residence principale	Integral	25000-50000	43					
5	100605	39		RUE	BUFFON	1033	1200	109	Appartement	Locataire	Residence principale	Classique	0-25000	33					



# Requêtes des données

(1/7)

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

```
SELECT
  C.Contrat_ID,
  C.Surface
FROM Contrat C
LEFT JOIN Region R
  ON C.Code_dep_code_commune = R.Code_dep_code_commune
WHERE R.com_nom_maj_court = "CAEN";
```

- ✓ Jointure (à gauche) des tables Contrat et Region
- ✓ Clé de jointure : **Code\_dep\_code\_commune**
- ✓ Filtre sur : le nom de la commune (requête 1) et le code du département (requête 2)
- ✓ Utilisation des alias pour renommer les tables

	Contrat_ID	Surface
1	103791	35
2	103792	99
3	103793	40
4	103794	20

4 contrats à Caen

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

```
SELECT
  C.Contrat_ID,
  C.Type_contrat,
  C.Formule
FROM Contrat C
LEFT JOIN Region R
  ON C.Code_dep_code_commune = R.Code_dep_code_commune
WHERE R.dep_code = 71;
```

Grid view		Form view	
		Total rows loaded: 48	
	Contrat_ID	Type_contrat	Formule
1	114767	Residence secondaire	Classique
2	114768	Residence principale	Integral
3	114769	Residence principale	Classique
4	114770	Residence principale	Classique
5	114771	Residence principale	Integral
6	114772	Residence principale	Classique
7	114773	Residence principale	Integral

48 contrats dans le département 71

# Requêtes des données

(2/7)

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

```
SELECT  
  DISTINCT (reg_nom)  
FROM Region ;
```

	reg_nom
1	Auvergne-Rhône-Alpes
2	Hauts-de-France
3	Provence-Alpes-Côte d'Azur
4	Grand Est
5	Occitanie
6	Normandie
7	Nouvelle-Aquitaine
8	Centre-Val de Loire
9	Bourgogne-Franche-Comté
10	Bretagne
11	Pays de la Loire
12	Ile-de-France
13	Guadeloupe
14	Martinique
15	Guyane
16	La Réunion
17	Collectivités d'outre-mer
18	Mayotte
19	Corse

Sélection unique du nom du département avec la clause **DISTINCT**

**Requête 4 :** Quels sont les 5 contrats qui ont les surfaces les plus élevées ?

```
SELECT  
  Contrat_ID  
FROM Contrat  
ORDER BY Surface DESC  
LIMIT 5 ;
```

- ✓ Sélection de la variable Contrat\_id de la table Contrat
- ✓ Classement des contrats selon la surface par ordre décroissant
- ✓ Limitation de l'affichage des résultats aux 5 premières lignes

	Contrat_ID
1	104211
2	105463
3	130878
4	100822
5	109872

# Requêtes des données

(3/7)

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

```
SELECT
  AVG(Prix_cotisation_mensuel) AS prix_moyen_cotisation_mensuelle
FROM Contrat ;
```

	prix moyen cotisation mensuelle
1	19.33

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

```
SELECT
  Valeur_declaree_biens,
  COUNT(Contrat_ID) AS Nombre_de_contrats
FROM Contrat
GROUP BY Valeur_declaree_biens ;
```

- ✓ Regroupement selon les modalités de la variable valeur\_declaree\_biens avec **GROUP BY**
- ✓ Comptage du nombre de contrat dans chaque groupe avec **COUNT**

	Valeur declaree biens	Nombre de contrats
1	0-25000	22720
2	100000+	104
3	25000-50000	6815
4	50000-100000	696



**Requête 7 :** Quel est le nombre de formules "integral" sur la région Pays de la Loire ?

```
SELECT
  COUNT(Contrat_ID)
FROM Contrat C
LEFT JOIN Region R
  ON C.Code_dep_code_commune = R.Code_dep_code_commune
WHERE Formule = "Integral" AND R.reg_nom = "Pays de la Loire";
```

- ✓ Jointure (à gauche) des tables Contrat et Region, avec Code\_dep\_code\_commune la clé de jointure
- ✓ Utilisation de la clause **WHERE** pour sélectionner les contrats avec la formule «*Intégral*» et dont le logement assuré se trouve au «*Pays de la Loire*»
- ✓ Comptage du nombre de contrat qui vérifient la contrainte ci-dessus

	COUNT(Contrat_ID)
1	589

**Requête 8 :** Lister les numéros de contrats avec le type de contrat et leur formule pour les maisons du département 71.

```
SELECT
  C.Contrat_ID,
  C.Type_contrat,
  C.Formule
FROM Contrat C
LEFT JOIN Region R
  ON C.Code_dep_code_commune = R.Code_dep_code_commune
WHERE C.Type_local = "Maison" AND R.dep_code = 71;
```

- ✓ Jointure (à gauche) des tables Contrat et Region, avec Code\_dep\_code\_commune la clé de jointure
- ✓ Utilisation de la clause **WHERE** pour sélectionner les contrats qui assurent les maisons dans le département 71

	Contrat_ID	Type_contrat	Formule
1	114768	Residence principale	Integral
2	114779	Residence principale	Classique
3	114782	Residence principale	Classique
4	114812	Residence principale	Integral

4 contrats assurent les maisons dans le département 71

# Requêtes des données

(5/7)

**Requête 9 :** Quelle est la surface moyenne des contrats à Paris ?

```
SELECT
  AVG(Surface) AS surf_moy_contrat_paris
FROM Contrat C
LEFT JOIN Region R
  ON C.Code_dep_code_commune = R.Code_dep_code_commune
WHERE R.com_nom_maj_court LIKE "PARIS%";
```

- ✓ Utilisation de la clause **LIKE** pour sélectionner toutes les communes dont le nom commence par la chaîne de caractère «*PARIS*»
- ✓ Calcul de la surface moyenne des logements assurés qui vérifient cette condition, avec la clause **AVG**

	surf moy contrat paris
1	51.77

- ✓ Jointure (à gauche) des tables Contrat et Region
- ✓ Regroupement selon le département avec **GROUP BY**
- ✓ Classement du prix de cotisation moyen dans chaque département par ordre décroissant
- ✓ Limitation de l'affichage des résultats aux 10 premières lignes

**Requête 10 :** Classement des 10 départements où le prix moyen de la cotisation est le plus élevé.

```
SELECT
  R.dep_code,
  R.dep_nom,
  ROUND(AVG(Prix_cotisation_mensuel), 2) AS Prix_moyen_cotisation
FROM Contrat C
LEFT JOIN Region R
  ON C.Code_dep_code_commune = R.Code_dep_code_commune
GROUP BY R.dep_code, R.dep_nom
ORDER BY Prix_moyen_cotisation DESC
LIMIT 10 ;
```

	dep code	dep nom	Prix moyen cotisation
1	75	Paris	36.4
2	92	Hauts-de-Seine	26.27
3	94	Val-de-Marne	19.82
4	78	Yvelines	18.89
5	69	Rhône	18.49
6	1	Ain	18.24
7	6	Alpes-Maritimes	18.14
8	17	Charente-Maritime	17.32
9	74	Haute-Savoie	17.15
10	2A	Corse-du-Sud	17.07

# Requêtes des données

(6/7)

**Requête 11 :** Liste des communes ayant eu au moins 150 contrats.

```
SELECT
  R.com_nom_maj_court,
  COUNT(Contrat_ID) AS Nombre_de_contrats
FROM Contrat C
LEFT JOIN Region R
  ON C.Code_dep_code_commune = R.Code_dep_code_commune
GROUP BY R.com_nom_maj_court
HAVING Nombre_de_contrats >= 150
ORDER BY Nombre_de_contrats ;
```

- ✓ Jointure (à gauche) des tables Contrat et Region
- ✓ Regroupement selon le nom de la commune avec **GROUP BY**
- ✓ Comptage du nombre de contrat dans chaque commune avec **COUNT**
- ✓ Sélection des communes ayant au minimum 150 contrats, avec la clause **HAVING**

			Total rows loaded: 20
	com nom maj court	Nombre de contrats	
1	PARIS 18	515	
2	PARIS 17	468	
3	PARIS 15	407	
4	PARIS 16	394	
5	NICE	387	
6	PARIS 11	381	
7	PARIS 20	302	
8	BORDEAUX	302	
9	NANTES	291	
10	PARIS 19	266	
11	PARIS 10	263	
12	PARIS 12	252	
13	PARIS 14	222	
14	GRENOBLE	220	
15	PARIS 9	204	
16	TOULOUSE	187	
17	TOULON	170	
18	COURBEVOIE	163	
19	LILLE	161	
20	PARIS 3	159	

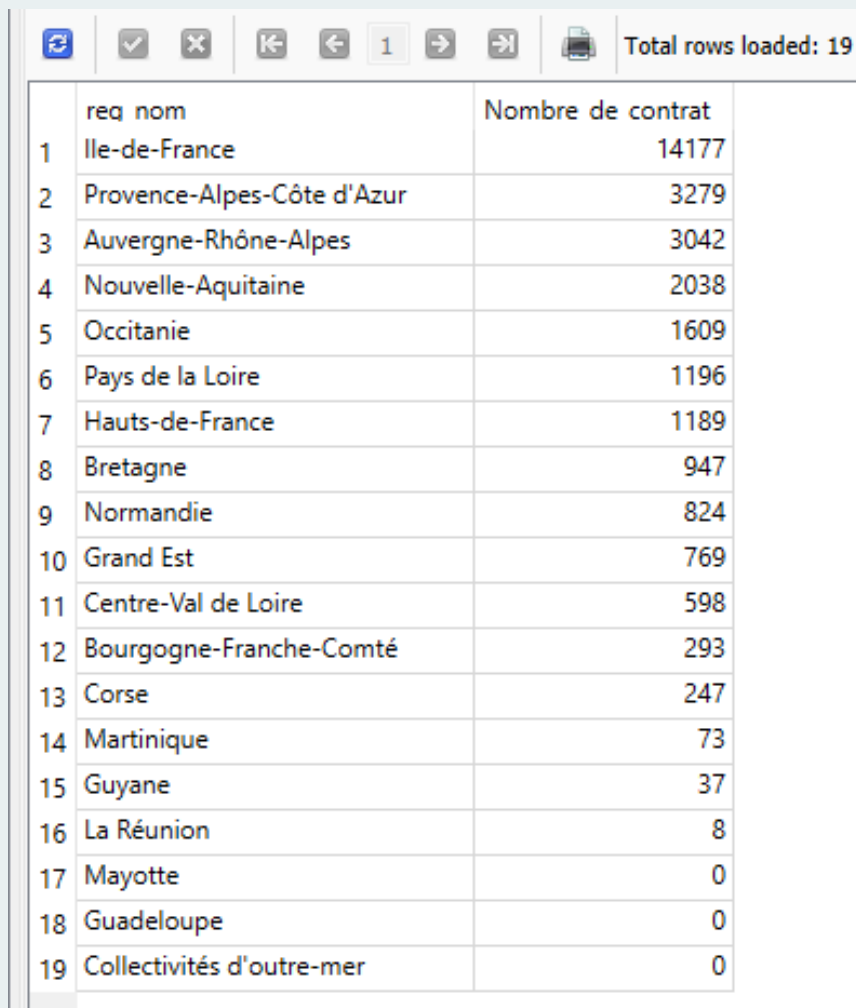
**Les communes de Paris détiennent le plus grand nombre de contrats**

**Requête 12 :** Quel est le nombre de contrats pour chaque région ?

```
SELECT
  R.reg_nom,
  COUNT(Contrat_ID) AS Nombre_de_contrat
FROM Contrat C
RIGHT JOIN Region R
  ON C.Code_dep_code_commune = R.Code_dep_code_commune
GROUP BY R.reg_nom;
```

- ✓ **Jointure (à droite)** des tables Contrat et Région
- ✓ Regroupement selon la région avec **GROUP BY**
- ✓ Comptage du nombre de contrat dans chaque région avec **COUNT**

L'Ile-de-France est la région ayant le plus grand nombre de contrats.



	req nom	Nombre de contrat
1	Ile-de-France	14177
2	Provence-Alpes-Côte d'Azur	3279
3	Auvergne-Rhône-Alpes	3042
4	Nouvelle-Aquitaine	2038
5	Occitanie	1609
6	Pays de la Loire	1196
7	Hauts-de-France	1189
8	Bretagne	947
9	Normandie	824
10	Grand Est	769
11	Centre-Val de Loire	598
12	Bourgogne-Franche-Comté	293
13	Corse	247
14	Martinique	73
15	Guyane	37
16	La Réunion	8
17	Mayotte	0
18	Guadeloupe	0
19	Collectivités d'outre-mer	0