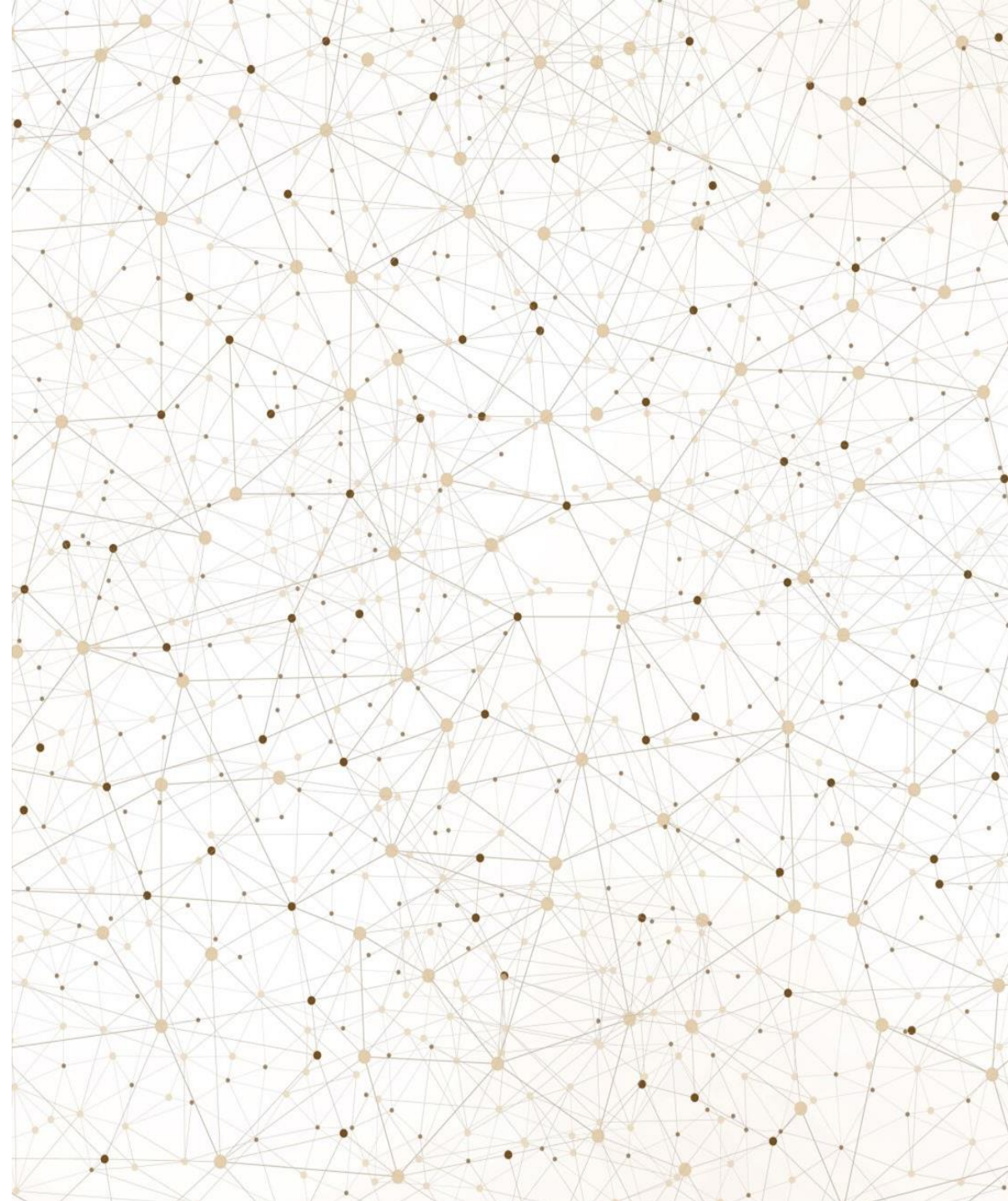


METHODOLOGIE

Analyse marché des assurances
habitation



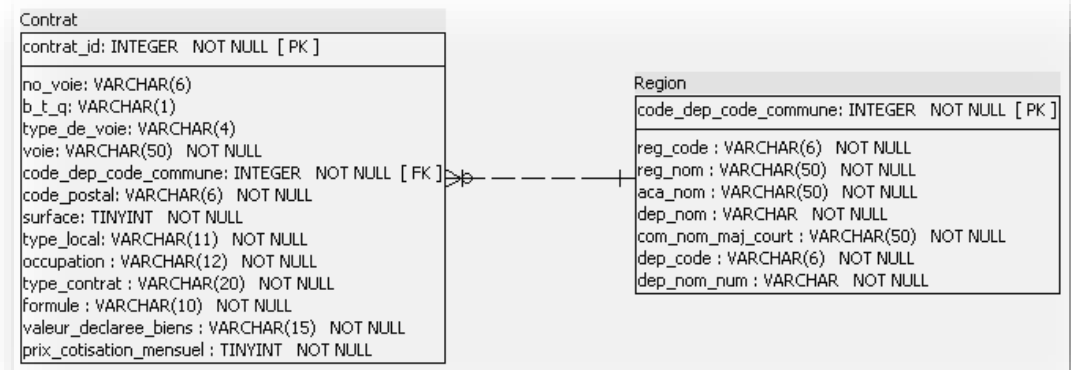
Etape 1 : Dictionnaire de données

- Observer les données des documents
 - Retranscrire sous forme de Tableau
 - Identifier les caractéristiques qui formeront les colonnes du Tableau
 - Associer à chacune ses contraintes
 - sa taille → quantité de données
 - valeur (null ou non),
 - son unicité
- Décrire chaque caractéristique, pour faciliter sa compréhension

	Nom des colonnes	Type de données	Taille	Clé	Description
CONTRAT.CSV	contrat_id	INT		Clé primaire	Id unique pour les contrats
	no_voie	VARCHAR	6		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	50		Libellé de la voie pour l'adresse du logement assuré
	code_dep_code_commune	INT		Clé secondaire	Concaténation du code département et code commune pour avoir une clé unique
	code_postal	VARCHAR	6		Code postal pour l'adresse du logement assuré
	surface	TINYINT			Surface du logement
	type_local	VARCHAR	11		Type de local (Appartement / Maison)
	occupation	VARCHAR	12		Occupation du logement (Propriétaire / Locataire)
	type_contrat	VARCHAR	20		Type de contrat du logement (Résidence principale, Résidence secondaire, Mis en location)
	formule	VARCHAR	10		Formule du contrat (Integral ou Classique)
	valeur_declaree_biens	VARCHAR	15		Catégorie de valeur déclaré du bien (0-25000, 25000-50000,50000-100000,100000+)
REGION.CSV	prix_cotisation_mensuel	TINYINT			Prix de la cotisation mensuel du contrat
	code_dep_code_commune	INT		Clé primaire	Concaténation du code département et code commune pour avoir une clé unique
	reg_code	VARCHAR	6		Code de la région
	reg_nom	VARCHAR	50		Nom de la région
	aca_nom	VARCHAR	50		Nom de la capitale de la région
	dep_nom	VARCHAR			Nom du département
	com_nom_maj_court	VARCHAR	50		Nom de la commune en majuscule
	dep_code	VARCHAR	6		Code du département
	dep_nom_num	VARCHAR			Concaténation du nom du département et du code departement

Etape 2 : Schéma Relationnel

- Retranscrire le dictionnaire de données sous forme de Schéma Relationnel
- Identifier la/les relation(s) qui lie(nt) les différentes tables.



Etape 3 : Création de la BDD

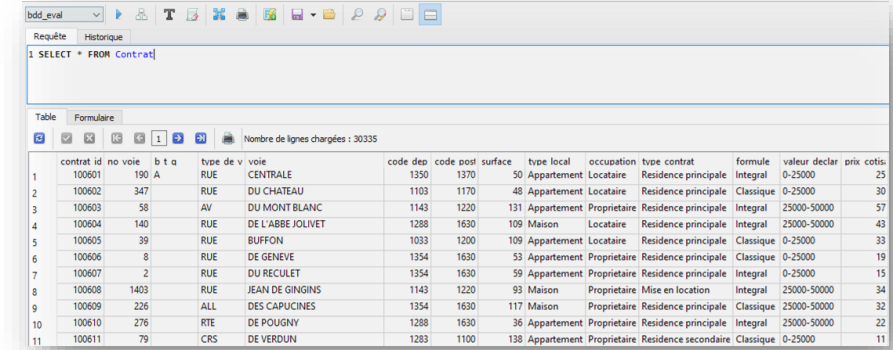
- Grâce au Schéma Relationnel précédemment établie définir les requêtes des créations de table
- Au sein du SGBDR utiliser ses requêtes pour créer nos tables dans notre BDD

```
CREATE TABLE Region (  
    code_dep_code_commune INT NOT NULL,  
    reg_code VARCHAR(6) NOT NULL,  
    reg_nom VARCHAR(50) NOT NULL,  
    aca_nom VARCHAR(50) NOT NULL,  
    dep_nom VARCHAR NOT NULL,  
    com_nom_maj_court VARCHAR(50) NOT NULL,  
    dep_code VARCHAR(6) NOT NULL,  
    dep_nom_num VARCHAR NOT NULL,  
    PRIMARY KEY (code_dep_code_commune)  
);
```

```
CREATE TABLE Contrat (  
    contrat_id INT NOT NULL,  
    no_voie VARCHAR(6),  
    b_t_q VARCHAR(1),  
    type_de_voie VARCHAR(4),  
    voie VARCHAR(50) NOT NULL,  
    code_dep_code_commune INT NOT NULL REFERENCES Region (code_dep_code_commune) ON  
DELETE NO ACTION ON UPDATE NO ACTION,  
    code_postal VARCHAR(6) NOT NULL,  
    surface TINYINT NOT NULL,  
    type_local VARCHAR(11) NOT NULL,  
    occupation VARCHAR(12) NOT NULL,  
    type_contrat VARCHAR(20) NOT NULL,  
    formule VARCHAR(10) NOT NULL,  
    valeur_declaree_biens VARCHAR(15) NOT NULL,  
    prix_cotisation_mensuel TINYINT NOT NULL,  
    PRIMARY KEY (contrat_id)  
);
```

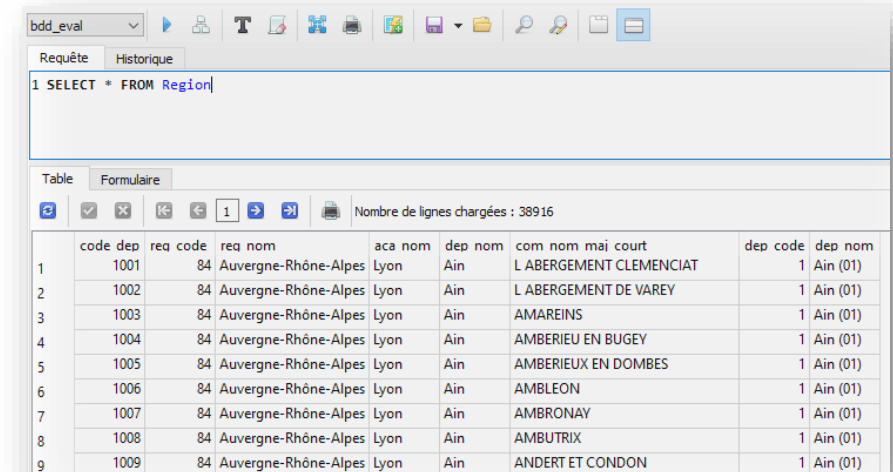
Etape 4 : Importation des données

- Au sein d'un SGBDR, utiliser la fonction d'import des données au format .csv pour chaque document afin de remplir les Tables précédemment créés
- Il sera ensuite possible de requêter ces tables, exemple :



The screenshot shows a database interface with a query window containing the SQL statement: `1 SELECT * FROM Contrat`. Below the query window, a table view displays 11 rows of data. The table has 14 columns: `contrat id`, `no voie`, `b t q`, `type de v`, `voie`, `code dep`, `code post`, `surface`, `type local`, `occupation`, `type contrat`, `formule`, `valeur declar`, and `prix cotis`.

	contrat id	no voie	b t q	type de v	voie	code dep	code post	surface	type local	occupation	type contrat	formule	valeur declar	prix cotis
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	Propriétaire	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
6	100606	8		RUE	DE GENEVE	1354	1630	53	Appartement	Propriétaire	Residence principale	Classique	0-25000	19
7	100607	2		RUE	DU RECULET	1354	1630	59	Appartement	Propriétaire	Residence principale	Integral	0-25000	15
8	100608	1403		RUE	JEAN DE GINGINS	1143	1220	93	Maison	Propriétaire	Mise en location	Integral	25000-50000	34
9	100609	226		ALL	DES CAPUCINES	1354	1630	117	Maison	Propriétaire	Residence principale	Classique	25000-50000	32
10	100610	276		RTE	DE POUIGNY	1288	1630	36	Appartement	Propriétaire	Residence principale	Integral	25000-50000	22
11	100611	79		CRS	DE VERDUN	1283	1100	138	Appartement	Propriétaire	Residence secondaire	Classique	0-25000	11



The screenshot shows a database interface with a query window containing the SQL statement: `1 SELECT * FROM Region`. Below the query window, a table view displays 9 rows of data. The table has 8 columns: `code dep`, `req code`, `req nom`, `aca nom`, `dep nom`, `com nom mai court`, `dep code`, and `dep nom`.

	code dep	req code	req nom	aca nom	dep nom	com nom mai court	dep code	dep nom
1	1001	84	Auvergne-Rhône-Alpes	Lyon	Ain	L ABERGEMENT CLEMENCIAT	1	Ain (01)
2	1002	84	Auvergne-Rhône-Alpes	Lyon	Ain	L ABERGEMENT DE VAREY	1	Ain (01)
3	1003	84	Auvergne-Rhône-Alpes	Lyon	Ain	AMAREINS	1	Ain (01)
4	1004	84	Auvergne-Rhône-Alpes	Lyon	Ain	AMBERIEU EN BUGEV	1	Ain (01)
5	1005	84	Auvergne-Rhône-Alpes	Lyon	Ain	AMBERIEUX EN DOMBES	1	Ain (01)
6	1006	84	Auvergne-Rhône-Alpes	Lyon	Ain	AMBLEON	1	Ain (01)
7	1007	84	Auvergne-Rhône-Alpes	Lyon	Ain	AMBRONAY	1	Ain (01)
8	1008	84	Auvergne-Rhône-Alpes	Lyon	Ain	AMBUTRIX	1	Ain (01)
9	1009	84	Auvergne-Rhône-Alpes	Lyon	Ain	ANDERT ET CONDON	1	Ain (01)

Etape 5 : Requêter la BDD

- Effectuer l'analyse des données au sein de la BDD grâce aux requêtes suivantes

Requête 1 : Lister les numéros de contrats (contrat_ID) avec leur surface pour la commune de Caen.

```
SELECT
    c.contrat_id,
    c.surface,
    r.com_nom_maj_court
FROM Contrat c
JOIN Region r
    ON c.code_dep_code_commune = r.code_dep_code_commune
    AND r.com_nom_maj_court = 'CAEN';
```

Table		Formulaire		
		Nombre de lignes chargées : 4		
	contrat id	surface	com nom maj court	
1	103791	35	CAEN	
2	103792	99	CAEN	
3	103793	40	CAEN	
4	103794	20	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,
    c.type_local,
    r.dep_code
FROM Contrat c
JOIN Region r
    ON c.code_dep_code_commune = r.code_dep_code_commune
    AND r.dep_code = 71
WHERE type_local = 'Maison';
```

Table		Formulaire			
		Nombre de lignes chargées : 4			
	contrat id	type contrat	formule	type local	dep code
1	114768	Residence principale	Integral	Maison	71
2	114779	Residence principale	Classique	Maison	71
3	114782	Residence principale	Classique	Maison	71
4	114812	Residence principale	Integral	Maison	71

Requête 3 : Lister le nom des régions de France.

```
SELECT
    DISTINCT reg_nom
FROM Region
ORDER BY reg_nom ASC;
```

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

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

```
SELECT
    *
FROM Contrat
ORDER BY surface DESC
LIMIT 0,5;
```

Table

Formulaire

1

Nombre de lignes chargées : 5

	contrat id	no voie	b t a	type de v	voie	code dep	code posta	surface	type local	occupation	type contrat	formule	valeur declar	prix cotisation
1	104211	34	B	CRS	VICTOR HUGO	17421	17600	815	Appartement	Locataire	Residence principale	Classique	100000+	240
2	105463	5229		RUE	DE LA PRIVADIERE	30126	30190	742	Maison	Proprietaire	Mise en location	Integral	50000-100000	189
3	130878	23		BD	DE BEAUSEJOUR	75116	75016	595	Appartement	Proprietaire	Residence principale	Classique	100000+	361
4	100822	5	A	RUE	FERNAND BAZIN	2201	2600	570	Maison	Proprietaire	Residence principale	Integral	100000+	111
5	109872	18		RUE	PAUL BELLAMY	44109	44000	559	Appartement	Locataire	Residence principale	Integral	50000-100000	85

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

SELECT

ROUND(AVG(prix_cotisation_mensuel),0) **as** moyenne_cotisation_mensuelle

FROM Contrat;

Table		Formulaire	
		Nombre de lignes chargées : 1	
	moyenne cotisation mensuelle		
1		19	

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

SELECT

valeur_declaree_biens as categorie_prix,

COUNT(contrat_id) as nombre_contrat

FROM Contrat

GROUP BY valeur_declaree_biens

Table		Formulaire	
		Nombre de lignes chargées : 4	
	categorie prix	nombre contrat	
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
    c.formule,
    COUNT(c.formule) as nombre_formule,
    r.reg_nom
FROM Contrat c
JOIN Region r
ON c.code_dep_code_commune = r.code_dep_code_commune
AND c.formule = 'Integral'
AND r.reg_nom like 'Pays%';
```

Table Formulaire			
Nombre de lignes chargées : 1			
	formule	nombre formule	req nom
1	Integral	589	Pays de la Loire

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,
    c.type_local
FROM Contrat c
JOIN Region r
ON c.code_dep_code_commune = r.code_dep_code_commune
AND r.dep_code=71
AND c.type_local like'Maison';
```

Table Formulaire				
Nombre de lignes chargées : 4				
	contrat id	type contrat	formule	type local
1	114768	Residence principale	Integral	Maison
2	114779	Residence principale	Classique	Maison
3	114782	Residence principale	Classique	Maison
4	114812	Residence principale	Integral	Maison

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

SELECT

ROUND(AVG(c.surface)) **as** surface_moyenne

FROM Contrat c

JOIN Region r

ON c.code_dep_code_commune = r.code_dep_code_commune

AND r.aca_nom='Paris';

Table		Formulaire	
Nombre de lignes chargées : 1			
surface moyenne			
1		52	

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

SELECT

r.dep_nom,

ROUND(AVG(c.prix_cotisation_mensuel)) **as** prix_moyen_cotisation

FROM Contrat c

JOIN Region r

ON c.code_dep_code_commune = r.code_dep_code_commune

GROUP BY r.dep_code

ORDER BY prix_moyen_cotisation **DESC**

LIMIT 10

Table		Formulaire	
Nombre de lignes chargées : 10			
	dep nom	prix moyen cotisation	
1	Paris	36	
2	Hauts-de-Seine	26	
3	Val-de-Marne	20	
4	Yvelines	19	
5	Rhône	18	
6	Alpes-Maritimes	18	
7	Ain	18	
8	Corse-du-Sud	17	
9	Seine-Saint-Denis	17	
10	Haute-Savoie	17	

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

SELECT

r.com_nom_maj_court,
COUNT(c.contrat_id) **as** nb_contrat

FROM Contrat c

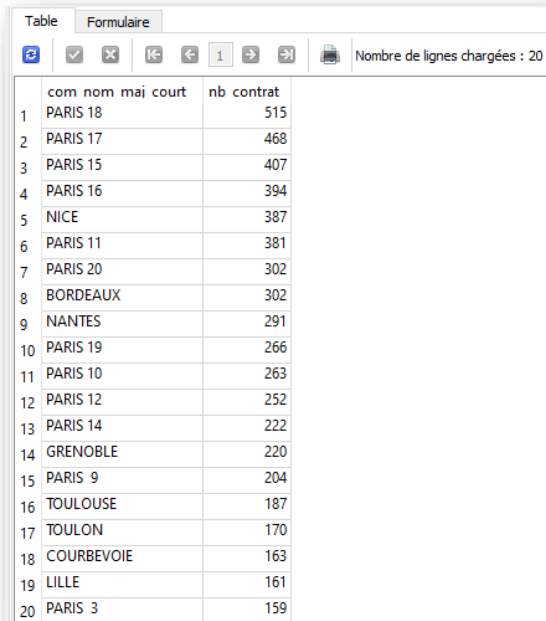
JOIN Region r

ON c.code_dep_code_commune = r.code_dep_code_commune

GROUP BY r.com_nom_maj_court

HAVING(COUNT(c.contrat_id)>150)

ORDER BY nb_contrat DESC



	com nom maj court	nb contrat
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

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

SELECT

r.reg_nom,
COUNT(c.contrat_id) as nb_contrat

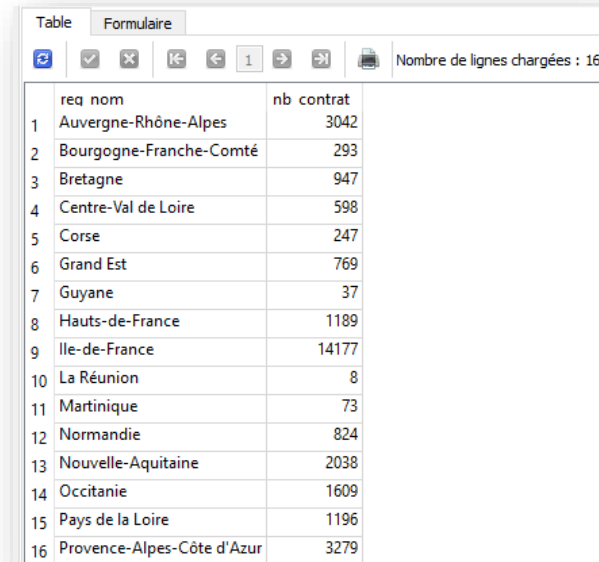
FROM Contrat c

JOIN Region r

ON c.code_dep_code_commune = r.code_dep_code_commune

GROUP BY r.reg_code

ORDER BY r.reg_nom ASC



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