

Méthodologie: Requêtez une base de données avec SQL





Introduction

SQL (Structured Query Language), est un langage de programmation standardisé utilisé pour gérer et manipuler des bases de données relationnelles. La modélisation relationnelle permet de stocker des données facilement représentables sous forme de tableau.

L'objectif de la présentation est de comprendre la méthodologie pour formuler des requêtes efficaces.

Exploration des types de données

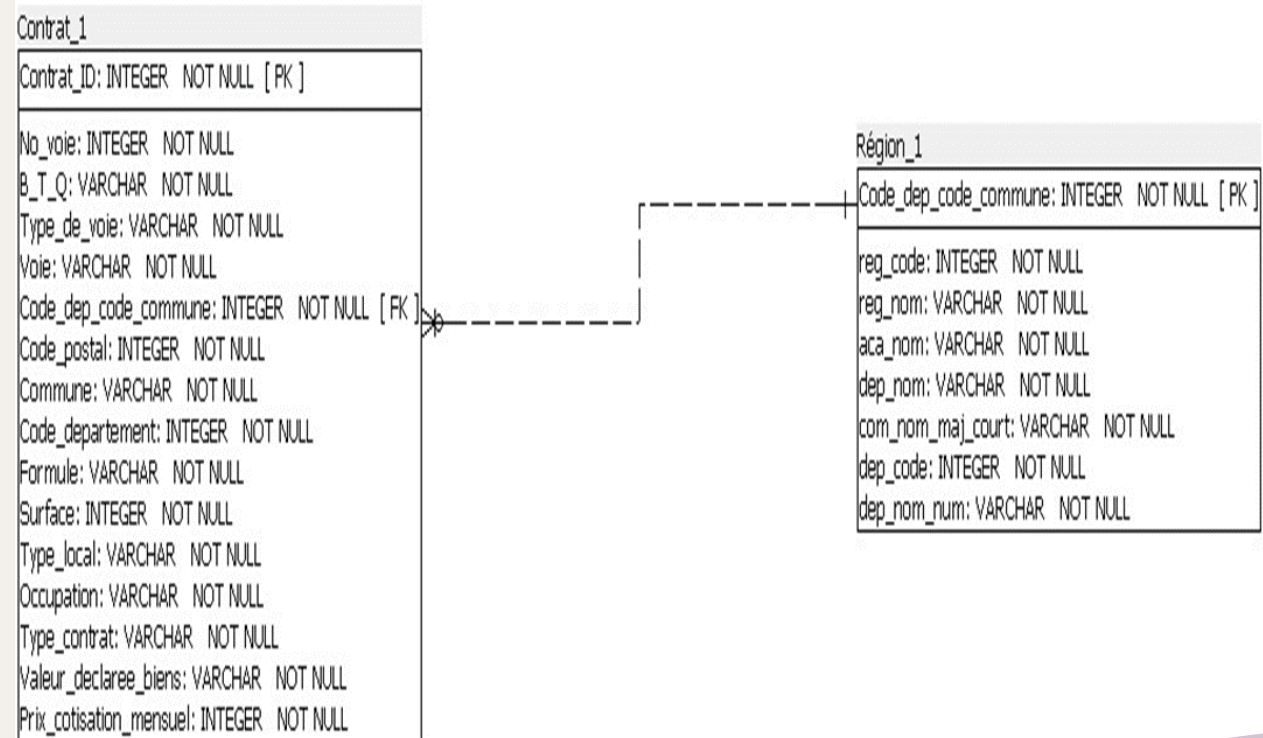
Cette étape a pour but d'observer le contenu de chaque colonne pour comprendre les données afin de rechercher les typologies de données pour les tables contrat et région.

Pour cela, nous avons utilisé les fichiers  Contrat  Region afin de constituer le 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	CVARCHAR			Type de voie pour l'adresse du logement assuré: rue, av (Avenue), rte (Route), ...
	Voie	VARCHAR			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	INT			Code postal pour l'adresse du logement assuré
	Commune	VARCHAR			Libellé de la commune de l'adresse du logement
	Code_département	INT			Code département
	Surface	INT			Surface local
	Type_local	VARCHAR			Type local
	Occupation	VARCHAR			Statut occupation
	Type_contrat	VARCHAR			Type de contrat
	Formule	VARCHAR			Formule de contrat
REGION.CSV	Valeur_declaree_biens	VARCHAR			Valeur déclarée des biens
	Prix_cotisation_mensuel	INT			Prix cotisation mensuel
	Code_dep_code_commune	INT		Clé primaire	Concaténation du code département et code commune pour avoir une clé unique
	reg_code	INT			Code région
	reg_nom	VARCHAR			Nom région
	aca_nom	VARCHAR			Nom académie
	dep_nom	VARCHAR			Nom département
	com_nom_maj_court	VARCHAR			Nom commune
	dep_code	INT			Code département
	dep_nom_num	VARCHAR			Concaténation du nom département et me code département

Schéma Relationnel 3NF

A travers le logiciel SQL Power Architects, nous avons créé le schéma relationnel 3NF qui satisfait aux exigences de la troisième forme normale, ce qui signifie que tous les attributs non-clés sont directement dépendants de la clé primaire, et qu'il n'y a pas de dépendances transitives.



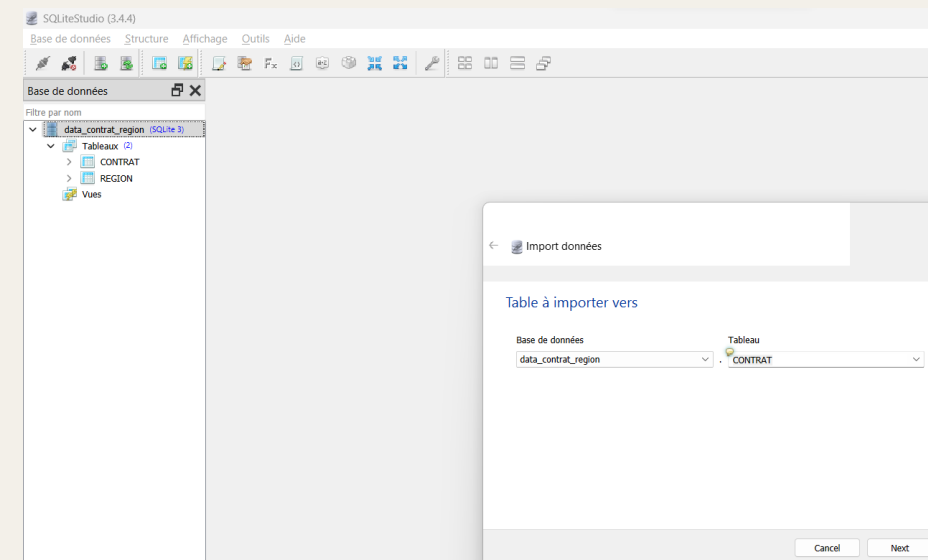
Création et chargement de la base de données

Nous avons utilisé le logiciel **SQLite studio** pour la création des tables CONTRAT et REGION à travers l'export des 2 fichiers CSV au niveau de la base.

Aussi, nous pouvons créer les tables en utilisant CREATE TABLE en suivant les étapes suivantes:

CREATE TABLE nom de la table (définit le nom de chaque colonne avec le type de variable et la fin on définit la clé primaire avec PRIMARY KEY (nom de la colonne).

```
CREATE TABLE Région (
    Code_dep_code_commune INTEGER NOT NULL,
    reg_code INTEGER NOT NULL,
    reg_nom VARCHAR NOT NULL,
    aca_nom VARCHAR NOT NULL,
    dep_nom VARCHAR NOT NULL,
    com_nom_maj_court VARCHAR NOT NULL,
    dep_code INTEGER NOT NULL,
    dep_nom_num VARCHAR NOT NULL,
    CONSTRAINT rgion_pk PRIMARY KEY (Code_dep_code_commune));
```



```
CREATE TABLE Contrat (
    Contrat_ID INTEGER NOT NULL,
    No_voie INTEGER NOT NULL,
    B_T_Q VARCHAR NOT NULL,
    Type_de_voie VARCHAR NOT NULL,
    Voie VARCHAR NOT NULL,
    Code_dep_code_commune INTEGER NOT NULL,
    Code_postal INTEGER NOT NULL,
    Commune VARCHAR NOT NULL,
    Code_departement INTEGER NOT NULL,
    Formule VARCHAR NOT NULL,
    Surface INTEGER NOT NULL,
    Type_local VARCHAR NOT NULL,
    Occupation VARCHAR NOT NULL,
    Type_contrat VARCHAR NOT NULL,
    Valeur_declaree_biens VARCHAR NOT NULL,
    Prix_cotisation_mensuel INTEGER NOT NULL,
    CONSTRAINT contrat_1_pk PRIMARY KEY (Contrat_ID));
```


Après l'export des fichiers CSV, nous constatons que la base contient deux tables avec l'ensemble des données:

❑ Table CONTRAT : 30 335 lignes

❑ Table REGION : 38 916 lignes

SQLiteStudio (3.4.4) - [CONTRAT (Base de données SGBD)]

Base de données Structure Données Contraintes Index Déclencheurs DDL

Table Formulaire

Nombre de lignes chargées : 30335

	Contrat_ID	No_vois	R_T_Q	Type_de_voie	Voie	Code_dep_s	Code_posta	Commune	Code_depar	Surface	Type_local	Occupation	Type_contrat	Formule	Valeur_declaree	P
1	100773	151		RTE	DE BELLEVILLE	01258	1090	MONTCEAUX	1	87	Appartement	Propriétaire	Residence principale	Classique	0-25000	1
2	100611	79		CRS	DE VERDUN	01283	1100	OYONNAX	1	138	Appartement	Propriétaire	Residence secondaire	Classique	0-25000	1
3	100645	10		RUE	AMPERE	01283	1100	OYONNAX	1	68	Appartement	Propriétaire	Residence principale	Integral	0-25000	1
4	100646	1		RUE	GERARD DE NERVAL	01031	1100	BELLIGNAT	1	71	Appartement	Propriétaire	Residence principale	Classique	0-25000	9
5	100691	14		RUE	CASTELLON	01283	1100	OYONNAX	1	88	Appartement	Propriétaire	Residence principale	Classique	0-25000	1
6	100705	1		ESPA	CHARLES DE GAULLE	01283	1100	OYONNAX	1	64	Appartement	Locataire	Residence principale	Integral	0-25000	1
7	100710	47		RUE	DU CHANE	01014	1100	ARBENT	1	48	Appartement	Propriétaire	Residence principale	Integral	0-25000	9
8	100744	86		RUE	GUSTAVE EIFFEL	01031	1100	BELLIGNAT	1	47	Appartement	Propriétaire	Residence secondaire	Integral	0-25000	9
9	100638	25		RUE	DE LA ROCHE SAMUEL	01060	1110	BRENOD	1	64	Appartement	Propriétaire	Residence principale	Integral	0-25000	9
10	100774	9		RUE	NEUVE	01262	1120	MONTLUEL	1	42	Appartement	Propriétaire	Residence principale	Classique	0-25000	1
11	100778	1172		RTE	DE GENEVE	01142	1120	DAGNEUX	1	38	Appartement	Propriétaire	Residence principale	Integral	0-25000	1
12	100642	3		RUE	DU DOCTEUR GREZEL	01269	1130	NANTUA	1	108	Appartement	Locataire	Residence principale	Classique	0-25000	1
13	116086	40		DOM	DU BOIS DE LA GARENNE	77001	77760	ACHERES LA FORET	77	135	Maison	Propriétaire	Residence principale	Classique	25000-50000	2
14	100621	369		AV	CHARLES DE GAULLE	01202	1150	LAGNIEU	1	53	Appartement	Locataire	Residence principale	Integral	0-25000	1
15	100630	79		GR	GRANDE RUE DE BULLIEZ	01202	1150	LAGNIEU	1	70	Appartement	Locataire	Residence principale	Classique	0-25000	1
16	100634	956		RTE	DE BOURG	01202	1150	LAGNIEU	1	94	Maison	Propriétaire	Residence principale	Integral	0-25000	1
17	100663	5		PL	DE L EGLISE	01202	1150	LAGNIEU	1	38	Appartement	Propriétaire	Residence principale	Classique	0-25000	9
18	100689	1143		RUE	CLAIRES FONTAINES	01390	1150	SAINT-VULBAS	1	43	Appartement	Propriétaire	Residence principale	Integral	0-25000	1
19	100712	15		CHE	DU LAVOIR	01390	1150	SAINT-VULBAS	1	32	Appartement	Propriétaire	Residence principale	Classique	0-25000	1
20	100732	44		RUE	PASTEUR	01202	1150	LAGNIEU	1	71	Appartement	Propriétaire	Residence principale	Classique	0-25000	1
21	100767	399		RUE	DU COTEILLON	01202	1150	LAGNIEU	1	104	Maison	Propriétaire	Residence principale	Classique	0-25000	1
22	100602	347		RUE	DU CHATEAU	01103	1170	CHEVRY	1	48	Appartement	Locataire	Residence principale	Classique	0-25000	3
23	100612	240		RUE	DE PRE BAILLY	01173	1170	GEX	1	45	Appartement	Locataire	Residence principale	Classique	0-25000	1
24	100616	282		CHE	DES LONGES RAYES	01071	1170	CESSY	1	42	Appartement	Propriétaire	Residence principale	Classique	0-25000	1
25	100619	17		LOT	LES JARDINS DE CHEVRY	01103	1170	CHEVRY	1	30	Appartement	Propriétaire	Residence secondaire	Classique	0-25000	1
26	100620	329		RUE	DES CARRIERES	01399	1170	SENGY	1	25	Appartement	Locataire	Residence principale	Integral	0-25000	1
27	100622	6		RUE	DES USINIERS	01173	1170	GEX	1	46	Appartement	Propriétaire	Residence secondaire	Classique	0-25000	1

SQLiteStudio (3.4.4) - [REGION (Base de données SGBD)]

Base de données Structure Données Contraintes Index Déclencheurs DDL

Table Formulaire

Nombre de lignes chargées : 38916

	Code_dep_s	reg_code	reg_nom	aca_nom	dep_nom	com_nom_maj_court	dep_code	dep_nom_n
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 BUGEY	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)
10	1010	84	Auvergne-Rhône-Alpes	Lyon	Ain	ANGLEFORT	1	Ain (01)
11	1011	84	Auvergne-Rhône-Alpes	Lyon	Ain	APREMONT	1	Ain (01)
12	1012	84	Auvergne-Rhône-Alpes	Lyon	Ain	ARANC	1	Ain (01)
13	1013	84	Auvergne-Rhône-Alpes	Lyon	Ain	ARANDAS	1	Ain (01)
14	1014	84	Auvergne-Rhône-Alpes	Lyon	Ain	ARBENT	1	Ain (01)
15	1015	84	Auvergne-Rhône-Alpes	Lyon	Ain	ARBIGNIEU	1	Ain (01)
16	1016	84	Auvergne-Rhône-Alpes	Lyon	Ain	ARBIGNY	1	Ain (01)
17	1017	84	Auvergne-Rhône-Alpes	Lyon	Ain	ARGIS	1	Ain (01)
18	1018	84	Auvergne-Rhône-Alpes	Lyon	Ain	ARLOD	1	Ain (01)
19	1019	84	Auvergne-Rhône-Alpes	Lyon	Ain	ARMIX	1	Ain (01)
20	1020	84	Auvergne-Rhône-Alpes	Lyon	Ain	ARNANS	1	Ain (01)
21	1021	84	Auvergne-Rhône-Alpes	Lyon	Ain	ARS SUR FORMANS	1	Ain (01)
22	1022	84	Auvergne-Rhône-Alpes	Lyon	Ain	ARTEMARE	1	Ain (01)
23	1023	84	Auvergne-Rhône-Alpes	Lyon	Ain	ASNIERES SUR SAONE	1	Ain (01)
24	1024	84	Auvergne-Rhône-Alpes	Lyon	Ain	ATTIGNAT	1	Ain (01)
25	1025	84	Auvergne-Rhône-Alpes	Lyon	Ain	BAGE LA VILLE	1	Ain (01)
26	1026	84	Auvergne-Rhône-Alpes	Lyon	Ain	BAGE LE CHATEL	1	Ain (01)
27	1027	84	Auvergne-Rhône-Alpes	Lyon	Ain	BALAN	1	Ain (01)

Les requêtes SQL avec les résultats obtenus






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

```
SELECT Contrat_ID, Surface
```

```
FROM CONTRAT
```

```
WHERE Commune='CAEN';
```

Cette requête sélectionne le 'Contrat_ID' et la 'Surface' venant de la table CONTRAT. La fonction WHERE permet de filtrer la colonne 'Commune' pour les enregistrements sont égales à 'CAEN'.

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

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

```
SELECT Contrat_ID, Type_contrat, Formule
```

```
FROM CONTRAT
```

```
WHERE Type_local='Maison' AND Code_departement='71';
```

Table		Formulaire	
1	Contrat_ID	Type_contrat	Formule
2	114768	Residence principale	Integral
3	114782	Residence principale	Classique
4	114812	Residence principale	Integral
5	114779	Residence principale	Classique

Cette requête sélectionne le 'Contrat_ID', 'Type_contrat' et la 'Formule' venant de la table CONTRAT. La fonction WHERE permet de filtrer la colonne 'Type_local' pour les enregistrements sont égales à 'Maison' et le 'Code_departement' est égal à 71.

3. Lister le nom des régions de France

```
SELECT DISTINCT reg_nom
```

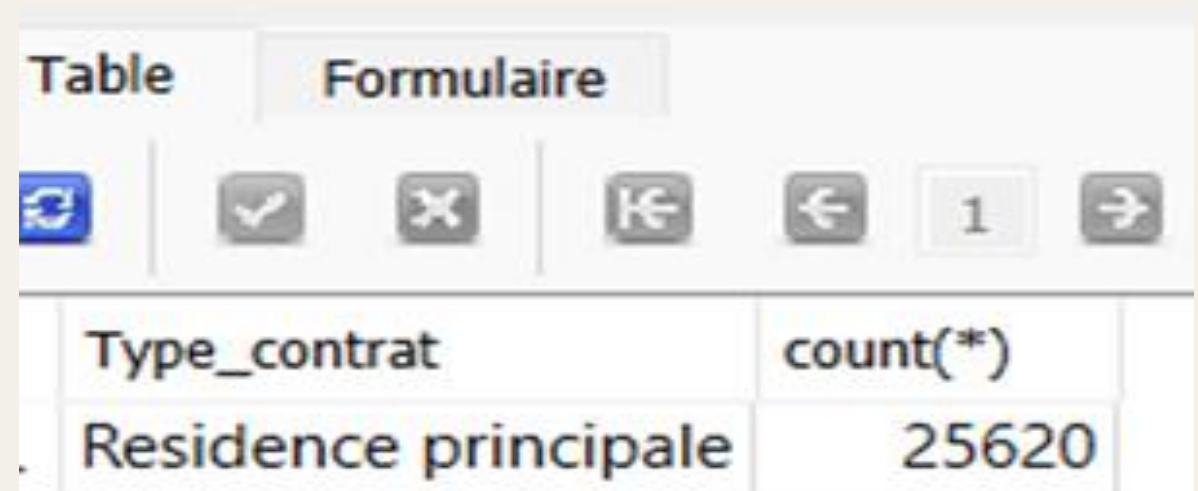
```
FROM REGION;
```

Cette requête renvoie la liste des noms de région sans aucune répétition, puisque `DISTINCT` élimine les doublons dans les résultats.

Table		Formulaire	
			
			<input type="text" value="1"/>
	reg_nom		
1	Auvergne-Rhône-Alpes		
2	Hauts-de-France		
3	Provence-Alpes-Côte d'Azur		
4	Grand Est		
5	Occitanie		
6	Normandie		

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

```
SELECT Type_contrat,  
count(*)  
FROM CONTRAT  
WHERE Type_contrat='Residence principale';
```



Type_contrat	count(*)
Residence principale	25620

Cette requête sélectionne le type de contrat et compte le nombre de contrat ayant le statut Residence principale dans la colonne "Type de contrat" venant de la table CONTRAT.






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

```
SELECT AVG(Surface)AS Surface_moyenne
```

```
FROM CONTRAT
```

```
WHERE Commune LIKE '%PARIS%';
```

Cette requête sélectionne la moyenne de la colonne "Surface" dans la table "CONTRAT" pour les enregistrements où la colonne "Commune" contient le mot "PARIS". La fonction LIKE est utilisée pour effectuer une recherche partielle sur la colonne "Commune" afin de trouver toutes les communes contenant le mot "PARIS".

Table		Formulaire		
				
1	Surface moyenne 50.67157263559553			

6. *Quels sont les 5 contrats qui ont les surfaces les plus élevées ?*






`SELECT Contrat_ID, Surface`

`FROM CONTRAT`

`ORDER BY Surface DESC`

`LIMIT 5;`






Cette requête sélectionne les colonnes "Contrat_ID" et "Surface" dans la table "CONTRAT", les ordonne en fonction de la colonne "Surface" de manière décroissante à l'aide de l'expression "ORDER BY Surface DESC", puis limite les résultats aux cinq premiers enregistrements avec la clause "LIMIT 5"

Table		Formulaire	
			
			
	Contrat_ID	Surface	
1	104211	815	
2	105463	742	
3	130878	595	
4	100822	570	
5	109872	559	

7. Quel est le prix moyen de la cotisation mensuelle ?

```
SELECT AVG(Prix_cotisation_mensuel) AS prix_moyen  
  
FROM CONTRAT;
```

Cette requête sélectionne la moyenne des valeurs de la colonne "Prix_cotisation_mensuel" dans la table "CONTRAT". La fonction AVG() calcule la moyenne des valeurs de cette colonne et la fonction AS renomme le nom de la colonne résultat en " prix moyen".

Table		Formulaire	
			
			
	prix_moyen		
1	19.32869622548212		








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

```
SELECT Valeur_declaree_biens, count(*)
```

```
FROM CONTRAT
```

```
GROUP BY Valeur_declaree_biens;
```

Cette requête sélectionne la colonne "Valeur_declaree_biens" ainsi que le nombre de contrats correspondant à chaque valeur déclarée des biens dans la table "CONTRAT". La clause GROUP BY est utilisée pour regrouper les résultats par valeur déclarée des biens.

Table		Formulaire
		     1  
	Valeur_declaree_biens	count(*)
1	0-25000	22720
2	100000+	104
3	25000-50000	6815
4	50000-100000	696

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

```
SELECT Code_departement, AVG(Prix_cotisation_mensuel)AS prix_moyen  
FROM CONTRAT  
GROUP BY Code_departement  
ORDER BY prix_moyen DESC  
LIMIT 10;
```

Table		Formulaire	
	Code_departement	prix_moyen	
1	75	36.39908732628086	
2	92	26.26935840707965	
3	94	19.82416396979504	
4	78	18.880187416332	
5	69	18.45681818181818	
6	1	18.23888888888889	
7	6	18.13762287756926	
8	17	17.31730769230769	
9	74	17.1583850931677	
10	20	17.02834008097166	

Cette requête sélectionne le code de département et la moyenne des valeurs de la colonne "Prix_cotisation_mensuel" dans la table "CONTRAT", regroupées par code de département. Ensuite, elle ordonne les résultats par prix moyen de manière décroissante à l'aide de l'expression "ORDER BY prix_moyen DESC", puis limite les résultats aux dix premiers départements avec la clause "LIMIT 10".

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

```
SELECT REGION.reg_nom, formule, count(*)AS Nombre_contrats
FROM CONTRAT
JOIN REGION ON
CONTRAT.Code_dep_code_commune=REGION.Code_dep_code_commune
GROUP BY reg_nom
HAVING reg_nom='Pays de la Loire' AND formule='Integral';
```

Table Formulaire











	rea nom	formule	Nombre contrat:
1	Pays de la Loire	Integral	1138

Cette requête effectue une jointure entre les tables "CONTRAT" et "REGION" sur la clé "Code_dep_code_commune", sélectionne les contrats pour la région "Pays de la Loire" avec la formule "Integral" à travers la fonction HAVING, puis compte le nombre de ces contrats. La clause WHERE est utilisée pour filtrer les enregistrements avant le regroupement avec GROUP BY.

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

`SELECT Commune, count(*)AS Nombre_contrats`

`FROM CONTRAT`

`GROUP BY Commune`

`HAVING Nombre_contrats>=150;`

Cette requête regroupe les contrats par commune, compte le nombre de contrats pour chaque commune, puis filtre les résultats pour inclure uniquement les communes ayant 150 contrats ou plus à l'aide de la clause HAVING

Table		Formulaire	
	Commune	Nombre_contrats	
1	0	4021	
2	BORDEAUX	302	
3	COURBEVOIE	163	
4	GRENOBLE	220	
5	LILLE	161	
6	NANTES	291	
7	NICE	387	
8	PARIS 10	263	
9	PARIS 11	381	
10	PARIS 12	252	
11	PARIS 14	222	
12	PARIS 15	407	
13	PARIS 16	394	
14	PARIS 17	468	
15	PARIS 18	515	
16	PARIS 19	266	
17	PARIS 20	302	
18	TOULON	170	
19	TOULOUSE	187	

12. Quel est le nombre de contrats pour chaque région ?

```
SELECT REGION.reg_nom, count(*)AS Nombre_contrats
FROM CONTRAT
JOIN REGION ON
CONTRAT.Code_dep_code_commune=REGION.Code_dep_code_commune
GROUP BY reg_nom
ORDER BY Nombre_contrats DESC;
```

Cette requête effectue une jointure entre les tables "CONTRAT" et "REGION" sur la clé "Code_dep_code_commune", sélectionne les noms des régions dans la table « REGION" puis compte le nombre de ces contrats. La clause GROUP BY est utilisée pour regrouper le nom des régions et ORDER BY permet d'ordonner les régions par ordre décroissant.

Table		Formulaire
	reg nom	Nombre contrat:
1	Ile-de-France	13474
2	Provence-Alpes-Côte d'Azur	3287
3	Auvergne-Rhône-Alpes	2972
4	Nouvelle-Aquitaine	2097
5	Occitanie	1837
6	Hauts-de-France	1327
7	Pays de la Loire	1138
8	Bretagne	945
9	Normandie	898
10	Grand Est	806
11	Centre-Val de Loire	804
12	Bourgogne-Franche-Comté	402
13	Corse	247
14	Martinique	60
15	Guyane	37
16	La Réunion	4

Conclusion

Une base de données sert à stocker des données pour qu'elles soient accessibles pour tous et la modélisation relationnelle permet de stocker des données facilement représentables sous forme de tableau.

L'objet principal du modèle relationnel est la TABLE qui contient des lignes et des colonnes.

Avant d'appliquer le LMD (Language de manipulation des données), il est très important de comprendre la structure des données afin de pouvoir définir les types de données, contraintes et identifier les différentes clés primaires qui doivent répondre à la contrainte d'unicité.

Merci



The image features a hand in the lower-left foreground, pointing its index finger towards a central digital element. This element is a large, glowing blue oval with a complex, concentric, and pixelated internal structure. The letters 'SQL' are prominently displayed in the center of this oval in a large, white, bold, sans-serif font. Surrounding this central oval are several smaller, circular nodes, each containing a different icon (such as a gear, a document, or a network symbol) and connected to the central node by thin, glowing blue lines. The entire scene is set against a dark, deep blue background with a subtle, out-of-focus pattern of light and lines, creating a high-tech, digital atmosphere.

SQL