Requêter une base de données avec SQL

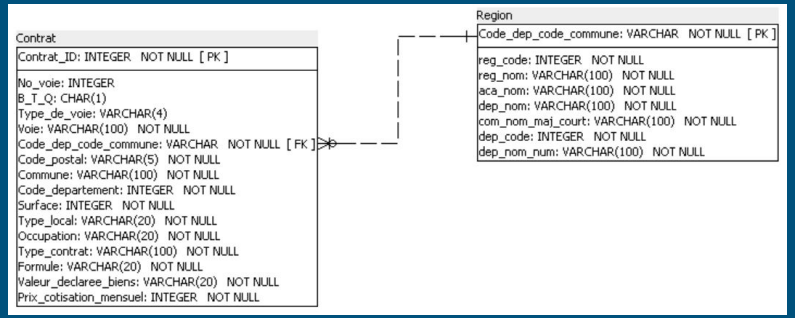
Projet 3

Création d’une base de donnée

1. Création du dictionnaire de donnée

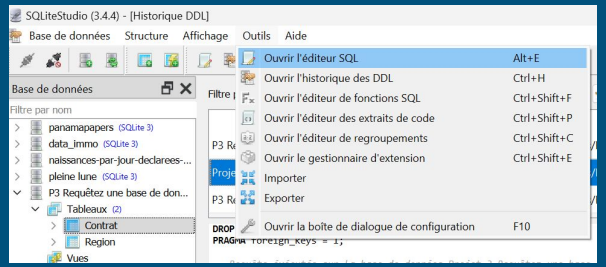
	Nom des colonnes	Type de données	Taille	Clé	Description
CONTRAT.CSV	Contrat_ID	INT		Clé primaire	Id unique pour les contrats
	No_voie	INT			Numéro dans la voie pour l'adresse du logement assuré
	B_T_Q	CHAR	1		Indicateur éventuel de répétition pour l'adresse du logement assuré sur un caractère
	Type_de_voie	VARCHAR	4		Type de voie pour l'adresse du logement assuré: rue, av (Avenue), rte (Route), ...
	Voie	VARCHAR	100		Libellé de la voie pour l'adresse du logement assuré
	Code_dep_code_commune	VARCHAR		Clé secondaire	Concaténation du code département et code commune pour avoir une clé unique
	Code_postal	INT	5		Code postal pour l'adresse du logement assuré
	Commune	VARCHAR	100		Libellé de la commune de l'adresse du logement
	Code_departement	INT			Numéro du département concerné par l'adresse du logement assuré
	Surface	INT			Taille en m² du logement assuré
	Type_local	VARCHAR	20		Type du logement assuré : Maison ou Appartement
	Occupation	VARCHAR	20		Relation entre la personne et le logement assurée : Propriétaire ou locataire
	Type_de_contrat	VARCHAR	100		Type d'utilisation du logement assurée : Résidence principale / secondaire ou mise en location
	Formule	VARCHAR	20		Offre souscrite par l'assurée : Classique ou Intégral
	Valeur_declaree_biens	VARCHAR	20		Estimation de la valeur du logement assuré (fourchette de prix)
	Prix_cotisation_mensuel	INT			Prix payé mensuellement pour le logement assuré
REGION.CSV	Code_dep_code_commune	VARCHAR		Clé primaire	Concaténation du code département et code commune pour avoir une clé unique
	reg_code	INT			Code de la région
	reg_nom	VARCHAR	100		Libellé de la région
	aca_nom	VARCHAR	100		Libellé de l'académie
	dep_nom	VARCHAR	100		Libellé du département
	com_nom_maj_court	VARCHAR	100		Libellé court de la commune en majuscule
	dep_code	INT			Code du département
	dep_nom_num	VARCHAR	100		Libellé du département et code

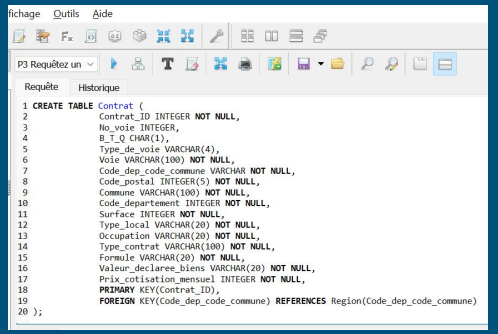
2. Etablir un schéma relationnel normalisé



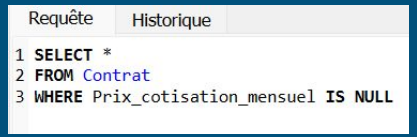
Création d’une base de donnée

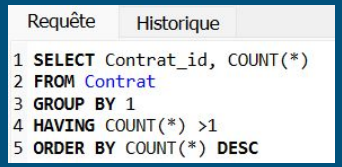
3. Création des tables via sqlite

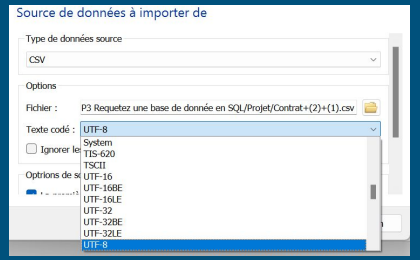




4. Chargement des données dans les tables

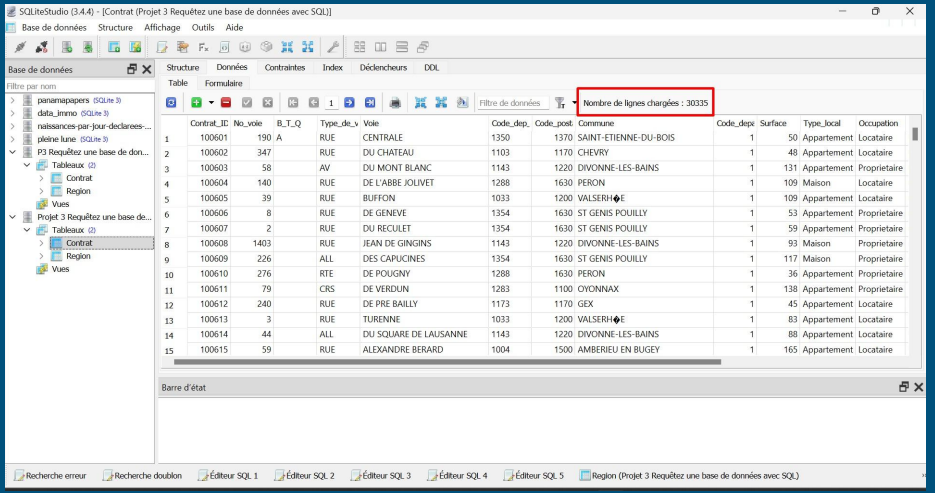






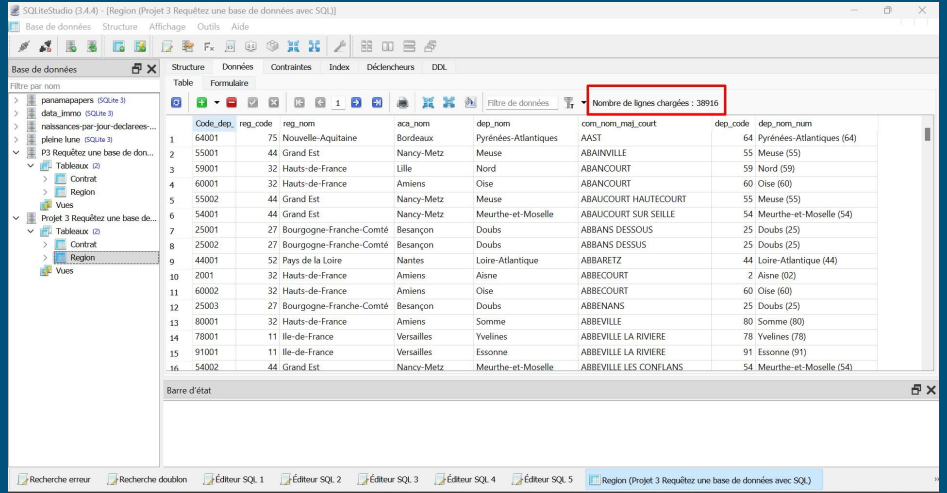
Création d’une base de donnée

Base de données chargée



	Contrat_ID	No_voie	B_T_Q	Type_de_v	Voie	Code_dep_	Code_post	Commune	Code_depa	Surface	Type_local	Occupation
1	100601	190	A	RUE	CENTRALE	1350	1370	SAINT-ETIENNE-DU-BOIS	1	50	Appartement	Locataire
2	100602	347		RUE	DU CHATEAU	1103	1170	CHEVRY	1	48	Appartement	Locataire
3	100603	58		AV	DU MONT BLANC	1143	1220	DIVONNE-LES-BAINS	1	131	Appartement	Proprietaire
4	100604	140		RUE	DE L'ABBE JOLIVET	1288	1630	PERON	1	109	Maison	Locataire
5	100605	39		RUE	BUFFON	1033	1200	VALSERH◆E	1	109	Appartement	Locataire
6	100606	8		RUE	DE GENEVE	1354	1630	ST GENIS POUILLY	1	53	Appartement	Proprietaire
7	100607	2		RUE	DU RECULET	1354	1630	ST GENIS POUILLY	1	59	Appartement	Proprietaire
8	100608	1403		RUE	JEAN DE GINGINS	1143	1220	DIVONNE-LES-BAINS	1	93	Maison	Proprietaire
9	100609	226		ALL	DES CAPUCINES	1354	1630	ST GENIS POUILLY	1	117	Maison	Proprietaire
10	100610	276		RTE	DE POUGNY	1288	1630	PERON	1	36	Appartement	Proprietaire
11	100611	79		CRS	DE VERDUN	1283	1100	OYONNAX	1	138	Appartement	Proprietaire
12	100612	240		RUE	DE PRE BAILLY	1173	1170	GEX	1	45	Appartement	Locataire
13	100613	3		RUE	TURENNE	1033	1200	VALSERH◆E	1	83	Appartement	Locataire
14	100614	44		ALL	DU SQUARE DE LAUSANNE	1143	1220	DIVONNE-LES-BAINS	1	88	Appartement	Locataire
15	100615	59		RUE	ALEXANDRE BERARD	1004	1500	AMBERIEU EN BUGEY	1	165	Appartement	Locataire

Table des contrats



	Code_dep_	reg_code	reg_nom	aca_nom	dep_nom	com_nom_maj_court	dep_code	dep_nom_num
1	64001	75	Nouvelle-Aquitaine	Bordeaux	Pyrénées-Atlantiques	AAST	64	Pyrénées-Atlantiques (64)
2	55001	44	Grand Est	Nancy-Metz	Meuse	ABAINVILLE	55	Meuse (55)
3	59001	32	Hauts-de-France	Lille	Nord	ABANCOURT	59	Nord (59)
4	60001	32	Hauts-de-France	Amiens	Oise	ABANCOURT	60	Oise (60)
5	55002	44	Grand Est	Nancy-Metz	Meuse	ABAUCOURT HAUTECOURT	55	Meuse (55)
6	54001	44	Grand Est	Nancy-Metz	Meurthe-et-Moselle	ABAUCOURT SUR SEILLE	54	Meurthe-et-Moselle (54)
7	25001	27	Bourgogne-Franche-Comté	Besançon	Doubs	ABBANS DESSOUS	25	Doubs (25)
8	25002	27	Bourgogne-Franche-Comté	Besançon	Doubs	ABBANS DESSUS	25	Doubs (25)
9	44001	52	Pays de la Loire	Nantes	Loire-Atlantique	ABBARETZ	44	Loire-Atlantique (44)
10	2001	32	Hauts-de-France	Amiens	Aisne	ABBECOURT	2	Aisne (02)
11	60002	32	Hauts-de-France	Amiens	Oise	ABBECOURT	60	Oise (60)
12	25003	27	Bourgogne-Franche-Comté	Besançon	Doubs	ABBENANS	25	Doubs (25)
13	80001	32	Hauts-de-France	Amiens	Somme	ABBEVILLE	80	Somme (80)
14	78001	11	Ile-de-France	Versailles	Yvelines	ABBEVILLE LA RIVIERE	78	Yvelines (78)
15	91001	11	Ile-de-France	Versailles	Essonne	ABBEVILLE LA RIVIERE	91	Essonne (91)
16	54002	44	Grand Est	Nancy-Metz	Meurthe-et-Moselle	ABBEVILLE LES CONFLANS	54	Meurthe-et-Moselle (54)

Table des régions

Les requêtes SQL avec les résultats

Contrat avec la surface pour la commune de Caen

Requête | Historique

```sql
--Lister les numéros de contrats (contrat_ID) avec leur surface pour la commune de Caen.
SELECT
    Contrat_ID,
    Surface,
    Commune
FROM Contrat
WHERE Commune = 'CAEN'
```

Table | Formulaire

Nombre de lignes chargées : 4

	Contrat ID	Surface	Commune
1	103791	35	CAEN
2	103792	99	CAEN
3	103793	40	CAEN
4	103794	20	CAEN

Liste des contrats avec le type et la formule pour les maisons dans le 71

Requête | Historique

```sql
--Lister les numéros de contrats avec le type de contrat et leur formule pour les maisons du département 71.
SELECT
    Contrat_ID,
    Formule,
    Type_contrat,
    Commune,
    Code_postal
FROM Contrat
WHERE Type_local ='Maison' AND Code_departement = 71
```

Table | Formulaire

Nombre de lignes chargées : 4

	Contrat ID	Formule	Type contrat	Commune	Code postal
1	114768	Integral	Residence principale	MACON	71000
2	114779	Classique	Residence principale	GIVRY	71640
3	114782	Classique	Residence principale	SAINT-REMY	71100
4	114812	Integral	Residence principale	BLANZY	71450

Dans chaque requête on doit traduire le besoin :

- Qu'est ce qu'on veut?
- Quelle table, quelle colonne avons nous besoin?
- Est ce que cela nécessite d'utiliser une fonction…

Il est primordiale de bien connaître et comprendre ses tables et leurs relations

Liste des régions de France

Requête | Historique

```sql
--Lister le nom des régions de France.
SELECT Distinct(reg_nom) as Region
FROM Region
```

Table | Formulaire

Nombre

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

Les requêtes SQL avec les résultats

Nombre de contrats sur les résidences principales

Requête | Historique

```sql
--Combien existe-t-il de contrats sur les résidences principales

SELECT COUNT(*) AS Nbre_residence_principale
FROM Contrat
WHERE Type_contrat = 'Residence principale'
```

Table | Formulaire

Nombre de lignes chargées : 1

	Nbre_residence_principale
1	25620

Combien = COUNT

Critère = WHERE

Calcul de la surface moyenne des contrats sur Paris

```sql
--Quelle est la surface moyenne des logements avec un contrat à Paris ?
SELECT ROUND(AVG(Surface), 2) as surface_moyenne_paris
FROM Contrat
WHERE Code_departement = "75"
```

Table | Formulaire

Nombre de lignes chargées : 1

	surface_moyenne_paris
1	51.77

Moyenne = AVG

Top 5 des surfaces les plus élevées

```sql
--Quels sont les 5 contrats qui ont les surfaces les plus élevées ?
SELECT contrat_id, surface
FROM Contrat
ORDER BY Surface DESC
LIMIT 5
```

Table | Formulaire

Nombre de lignes chargées : 5

	contrat_id	surface
1	104211	815
2	105463	742
3	130878	595
4	100822	570
5	109872	559

Classement = Order By + Limit

Les requêtes SQL avec les résultats

Calcul de la cotisation moyenne mensuelle

```sql
--Quel est le prix moyen de la cotisation mensuelle ?
SELECT ROUND(AVG(Prix_cotisation_mensuel), 2) AS Cotisation_moyenne
FROM Contrat
```

Table | Formulaire

Nombre de lignes chargées : 1

	Cotisation_moyenne
1	19.33

ROUND pour arrondir

Nombre de contrats par catégorie de valeur déclarée

```sql
--Quel est le nombre de contrats pour chaque catégorie de prix de la valeur déclarée des biens ?
SELECT
    Valeur_declaree_biens AS categorie_valeur,
    count(contrat_id) as nbre_contrat
FROM Contrat
GROUP BY 1
ORDER BY nbre_contrat DESC
```

Table | Formulaire

Nombre de lignes chargées : 4

	categorie_valeur	nbre_contrat
1	0-25000	22720
2	25000-50000	6815
3	50000-100000	696
4	100000+	104

Pour chaque catégorie = GROUP BY catégorie

Quel est le nombre = count

Classement des 10 départements où la cotisation est la plus élevée

```sql
--Quel est le classement des 10 départements où le prix moyen de la cotisation est le plus élevé ?
SELECT
    Code_departement,
    ROUND(AVG(Prix_cotisation_mensuel),2) AS cotisation_moyenne
FROM Contrat
GROUP BY 1
ORDER BY cotisation_moyenne DESC
LIMIT 10
```

Table | Formulaire

Nombre de lignes chargées : 10

	Code_departement	cotisation_moyenne
1	75	36.4
2	92	26.27
3	94	19.82
4	78	18.88
5	69	18.46
6	1	18.24
7	6	18.14
8	17	17.32
9	74	17.16
10	20	17.03

Classement = ORDER BY + Limit

Par département = GROUP BY

moyenne des cotisations = AVG

Les requêtes SQL avec les résultats

Nombre de formules intégrales dans le Pays de la Loire

```sql
--Quel est le nombre de contrats avec des formules "Intégral" pour la région Pays de la Loire ?
SELECT
    r.reg_nom as Region,
    count(*) as nbre_formule_integral
FROM Contrat as c
LEFT JOIN Region as r
ON c.Code_dep_code_commune = r.Code_dep_code_commune
WHERE c.Formule = 'Integral' AND r.reg_code = 52
```

Table | Formulaire

Nombre de lignes chargées : 1

	Region	nbre_formule_integral
1	Pays de la Loire	561

Count + Jointure + Condition

Récupérer les régions = LEFT JOIN Region

Liste des communes avec minimum 150 contrats

```sql
--Quelle est la liste des communes ayant au moins 150 contrats ?
SELECT
    Code_dep_code_commune,
    Commune,
    COUNT(contrat_id) as nbre_contrats
FROM Contrat
GROUP BY 1
HAVING nbre_contrats >= 150
ORDER BY nbre_contrats DESC
```

Table | Formulaire

Nombre de lignes chargées : 20

	Code_dep_code_commune	Commune	nbre_contrats
1	75118	PARIS 18	515
2	75117	PARIS 17	468
3	75115	PARIS 15	407
4	75116	PARIS 16	394
5	6088	NICE	387
6	75111	PARIS 11	381
7	75120	PARIS 20	302
8	33063	BORDEAUX	302
9	44109	NANTES	291
10	75119	PARIS 19	266
11	75110	PARIS 10	263
12	75112	PARIS 12	252
13	75114	PARIS 14	222
14	38185	GRENOBLE	220
15	75109	0	204
16	31555	TOULOUSE	187
17	83137	TOULON	170
18	92026	COURBEVOIE	163
19	59350	LILLE	161
20	75103	0	159

HAVING Pour filtrer les groupes après l'agrégation (#WHERE)

Nombre de contrats par région

```sql
--Quel est le nombre de contrats pour chaque région ?
SELECT
    r.reg_nom AS Region,
    count(c.Contrat_ID) AS nbre_contrats
FROM contrat as c
LEFT JOIN Region as r
ON c.Code_dep_code_commune = r.Code_dep_code_commune
GROUP BY 1
ORDER BY nbre_contrats DESC
```

Table | Formulaire

Nombre de lignes chargées : 16

	Region	nbre_contrats
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

Chaque région = GROUP BY
Infos à récupérer dans la table région = LEFT JOIN