

OPENCLASSROOMS

REQUÊTER UNE BASE DE DONNÉES AVEC SQL

- Février 2024 -

Caroline G
Data Analyst



REQUÊTER UNE BASE DE DONNÉES AVEC SQL

INTRODUCTION

❑ ETAPES DU PROJET

❑ METHODOLOGIE

❑ RAPPEL DES RÉSULTATS



Dictionnaire de données

- 1^{ère} étape -

	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			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	VARCHAR	5	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			Libellé de la commune de l'adresse du logement
	Code_département	VARCHAR			Code département pour l'adresse du logement assuré (Correspond à dep_code de Region.csv)
	Surface	INT			Surface du logement assuré
	Type_local	VARCHAR			Type du logement assuré (2 valeurs possibles : Maison/Appartement)
	Occupation	VARCHAR			Statut de l'occupation (2 valeurs possibles : Locataire/Propriétaire)
	Type_contrat	VARCHAR			Utilisation du logement pour ce contrat (3 valeurs possibles)
	Formule	VARCHAR			Formule de protection choisie (2 valeurs possibles)
REGION.CSV	Valeur_declaree_biens	VARCHAR			Tranche de la valeur des biens déclarée (4 valeurs possibles)
	Prix_cotisation_mensuel	INT			Prix de la cotisation pour un mois
	Code_dep_code_commune	VARCHAR	5	Clé primaire	Concaténation du code département et code commune pour avoir une clé unique
	reg_code	INT			Code la région dans laquelle se trouve la commune (19 valeurs)
	reg_nom	VARCHAR			Nom de la région dans laquelle se trouve la commune (19 valeurs)
	aca_nom	VARCHAR			Nom de l'académie (liée à la commune / département)
	dep_nom	VARCHAR			Nom du département dans lequel se trouve la commune
	com_nom_maj_court	VARCHAR			Nom de la commune - (Aurait pu correspondre à commune de Contrats.csv mais formatage différent)
	dep_code	VARCHAR			Code du département dans lequel se trouve la commune (Correspond à code_département de Contrats.csv)
	dep_nom_num	VARCHAR			Nom et n°du département dans lequel se trouve la commune

REQUÊTER UNE BASE DE DONNÉES AVEC SQL



Schéma relationnel normalisé 3NF

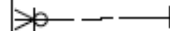
- 2^{ème} étape -

Contrat

Contrat_ID: INTEGER NOT NULL [PK]
No_voie: INTEGER NOT NULL
B_T_Q: VARCHAR(1) NOT NULL
Type_de_voie: VARCHAR NOT NULL
Voie: VARCHAR NOT NULL
Code_dep_code_commune: VARCHAR(5) NOT NULL [FK]
Code_postal: VARCHAR(5) NOT NULL
Commune: VARCHAR NOT NULL
Code_departement: VARCHAR NOT NULL
Surface: INTEGER NOT NULL
Type_local: VARCHAR NOT NULL
Occupation: VARCHAR NOT NULL
Type_contrat: VARCHAR NOT NULL
Formule: VARCHAR NOT NULL
Valeur_declaree_biens: VARCHAR NOT NULL
Prix_cotisation_mensuel: INTEGER NOT NULL

Region

Code_dep_code_commune: VARCHAR(5) NOT NULL [PK]
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: VARCHAR NOT NULL
dep_nom_num: VARCHAR NOT NULL



→ Code générant les tables et relation

REQUÊTER UNE BASE DE DONNÉES AVEC SQL



Charger les données

- 3^{ème} étape -

SQLiteStudio (3.4.4)

Table Contrat

Table Region

Region (caro_gou_P2_Etp3_BDD)

Structure | Data | Constraints | Indexes | Triggers | DDL

Grid view | Form view

Filter data | Total rows loaded: 38916

	code_de	reg_cod	reg_nom	aca_nom	dep_nom	com_nom_maj_court	dep_cod	dep_nom_num
1001	2542	32	Hauts-de-France	Amiens	Aisne	NEUFLIEUX	2	Aisne (02)
1002	2543	32	Hauts-de-France	Amiens	Aisne	NEUILLY ST FRONT	2	Aisne (02)
1003	2544	32	Hauts-de-France	Amiens	Aisne	NEUVE MAISON	2	Aisne (02)
1004	2545	32	Hauts-de-France	Amiens	Aisne	LA NEUVILLE BOSMONT	2	Aisne (02)
1005	2546	32	Hauts-de-France	Amiens	Aisne	LA NEUVILLE EN BEINE	2	Aisne (02)
1006	2547	32	Hauts-de-France	Amiens	Aisne	LA NEUVILLE HOUSSET	2	Aisne (02)
1007	2548	32	Hauts-de-France	Amiens	Aisne	LA NEUVILLE LES DORENGT	2	Aisne (02)
1008	2549	32	Hauts-de-France	Amiens	Aisne	NEUVILLE ST AMAND	2	Aisne (02)
1009	2550	32	Hauts-de-France	Amiens	Aisne	NEUVILLE SUR AILETTE	2	Aisne (02)
1010	2551	32	Hauts-de-France	Amiens	Aisne	NEUVILLE SUR MARGIVAL	2	Aisne (02)
1011	2552	32	Hauts-de-France	Amiens	Aisne	NEUVILLETTE	2	Aisne (02)
1012	2553	32	Hauts-de-France	Amiens	Aisne	NIZY LE COMTE	2	Aisne (02)
1013	2554	32	Hauts-de-France	Amiens	Aisne	NOGENTEL	2	Aisne (02)
1014	2555	32	Hauts-de-France	Amiens	Aisne	NOGENT L ARTAUD	2	Aisne (02)
1015	2556	32	Hauts-de-France	Amiens	Aisne	NOIRCOURT	2	Aisne (02)
1016	2557	32	Hauts-de-France	Amiens	Aisne	NOROY SUR OURCQ	2	Aisne (02)

Status

[12:30:15] Query finished in 0.001 second(s).

[12:30:23] Query finished in 0.000 second(s).

[12:30:28] Query finished in 0.000 second(s).

[12:38:31] Error while importing data: FOREIGN KEY constraint failed

[12:39:33] Error while importing data: FOREIGN KEY constraint failed

[12:39:44] Committed changes for table 'Contrat' successfully.

[12:39:58] Imported data to the table 'Contrat' successfully. Number of imported rows: 30335

[12:41:35] Error while importing data: FOREIGN KEY constraint failed

[12:41:55] Committed changes for table 'Region' successfully.

[12:42:22] Imported data to the table 'Region' successfully. Number of imported rows: 38916

30 335 lignes importées

38 916 lignes importées

REQUETER UNE BASE DE DONNÉES AVEC SQL



Résultats des requêtes SQL

- dernière étape -

Lister les n° de contrats avec la surface pour la ville de Caen

```
1
select contrat_id, type_contrat, formule
from contrat
where type_local = 'Maison' AND code_departement = '71'
```

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

Total rows loaded: 4

	contrat_id	type_contrat	formule
1	114768	Residence principale	Integral
2	114782	Residence principale	Classique
3	114812	Residence principale	Integral
4	114779	Residence principale	Classique

```
1
select DISTINCT reg_nom
from region
```

Lister le nom des régions de France

	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

REQUÊTER UNE BASE DE DONNÉES AVEC SQL



Résultats des requêtes SQL

```
1 SELECT count(*)
2 FROM contrat
3 WHERE type_contrat = 'Residence principale'
```

Nombre de contrats sur les résidences principales

Total rows loaded: 1

	count(*)
1	25620

```
1 SELECT AVG (surface)
2 FROM contrat
3 WHERE commune LIKE 'PARIS%'
```

Surface moyenne des logements avec un contrat à Paris

Total rows loaded: 1

	AVG(surface)
1	51.76954988591579

```
1 SELECT *
2 FROM contrat
3 ORDER BY surface DESC
4 LIMIT 5
```

les 5 contrats ayant les surfaces les plus élevées

Grid view Form view

Total rows loaded: 5

	Contrat_	No_voie	B_T_Q	Type_de	Voie	Code_de	Code_pc	Commune	Code_de	Surface
1	104211	34	B	CRS	VICTOR HUGO	17421	17600	SAUJON	17	819
2	105463	5229		RUE	DE LA PRIVADIERE	30126	30190	0	30	742
3	130878	23		BD	DE BEAUSEJOUR	75116	75016	PARIS 16	75	595
4	100822	5	A	RUE	FERNAND BAZIN	02201	2600	COEUVRES ET VALSERY	2	570
5	109872	18		RUE	PAUL BELLAMY	44109	44000	NANTES	44	559

```
1 SELECT AVG (prix_cotisation_mensuel)
2 FROM contrat
```

prix moyen de la cotisation mensuelle

Total rows loaded: 1

	AVG(prix_cotisation_mensuel)
1	19.32869622548212



Résultats des requêtes SQL

```
1 SELECT code_departement,
2 AVG(prix_cotisation_mensuel) AS PRIX_MOYEN
3 FROM contrat
4 GROUP BY 1
5 ORDER BY 2 DESC
6 LIMIT 10
```

10 départements où le prix moyen de la cotisation est le plus élevé

	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

```
4 SELECT valeur_declaree_biens, count(*)
5 FROM contrat
6 GROUP BY valeur_declaree_biens
```

Nombre de contrats par catégorie de valeur déclarée

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

Communes ayant au moins 150 contrats

```
1 SELECT commune, count(*)
2 FROM contrat
3 GROUP BY 1
4 HAVING count(*) >= 150
```

	commune	count(*)
1	BORDEAUX	302
2	COURBEVOIE	163
3	GRENOBLE	220
4	LILLE	161
5	NANTES	291
6	NICE	387
7	PARIS 3	159
8	PARIS 9	204
9	PARIS 10	263
10	PARIS 11	381
11	PARIS 12	252
12	PARIS 14	222
13	PARIS 15	407
14	PARIS 16	394
15	PARIS 17	468
16	PARIS 18	515
17	PARIS 19	266
18	PARIS 20	302
19	TOULON	170
20	TOULOUSE	187



Résultats des requêtes

- avec jointures -

Query History

```
1 SELECT count (*) FROM (SELECT a.*
2 FROM contrat a
3 LEFT JOIN region b
4 ON a.Code_dep_code_commune = b.code_dep_code_commune
5 WHERE (
6     formule = "Integral"
7     AND a.Code_dep_code_commune = b.code_dep_code_commune
8     AND reg_nom = "Pays de la Loire"
9 )
10 )
```

nombre de contrats avec des formules "Integral" pour la région
Pays de la Loire

Total rows loaded: 1

count(*)
561

```
1 SELECT reg_nom,count (*) FROM (SELECT a.*,b.reg_nom
2 FROM contrat a
3 LEFT JOIN region b
4 ON a.Code_dep_code_commune = b.code_dep_code_commune
5 )
6 GROUP BY reg_nom
```

Total rows loaded: 16

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

nombre de contrats
par région

REQUÊTER UNE BASE DE DONNÉES



CONCLUSION

- ✓ Créer un schéma d'une base de données
- ✓ Charger des données dans une base de données
- ✓ Rédiger des scripts simples en SQL pour aider à la prise de décision business