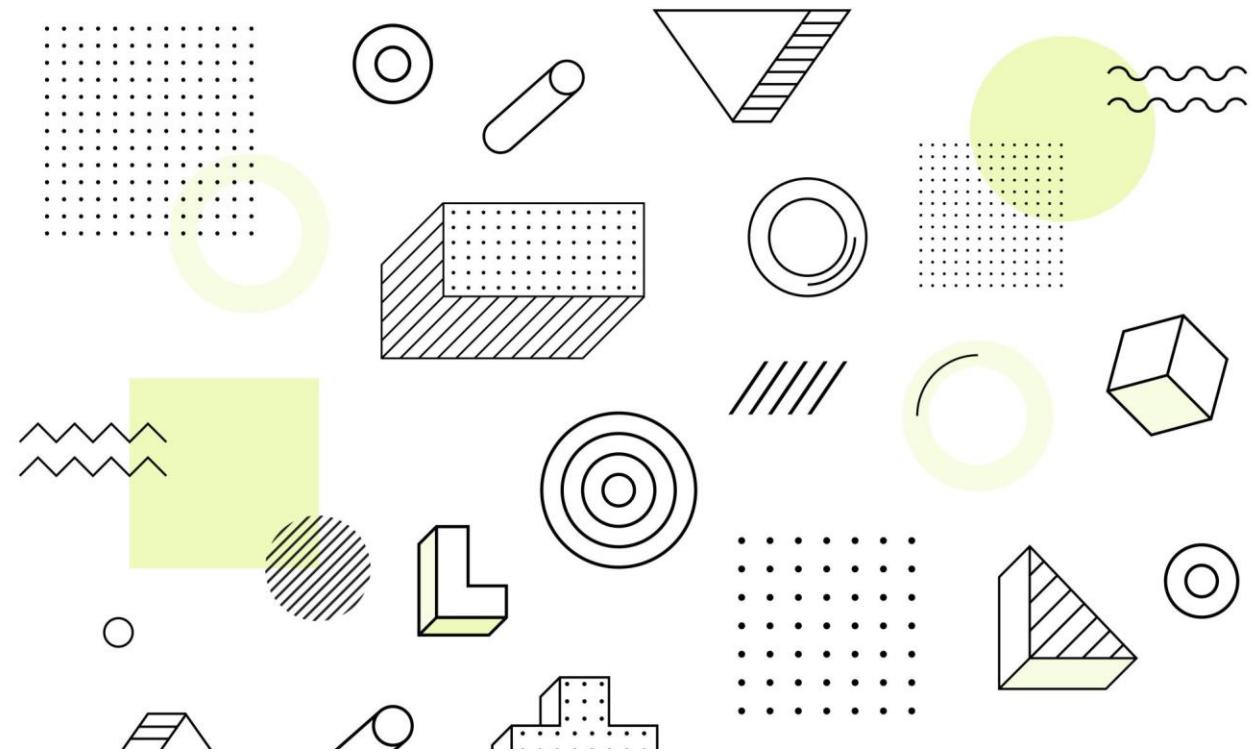


Requêter une base de données avec SQL

DÉMARCHE COMPLÈTE DE LA CONCEPTION À L'ANALYSE DES DONNÉES D'UNE BASE DE DONNÉES RELATIONNELLE.



Contexte et objectifs

- **Objectif** : implémenter une base de données relationnelle à partir de données brutes (region.csv et contrat.csv)
- **Finalité** : Interroger la base de données à travers différentes requêtes SQL (sélection, filtrage, agrégations, comptage et statistiques)
- **Outils** : PostgreSQL et SQL Power Architect
- **Compétences mobilisées** : Modélisation, création de tables, intégrité référentielle, import de données, requêtes SQL, analyse.

Analyse des fichiers CSV

Caractéristique fichier Contrat															
	Contrat_ID	No_voie	B_T_Q	Type_de_voie	Voie	Code_dep_code_commune	Code_postal	Surface	Type_local	Occupation	Type_contrat	Formule	Valeur_declaree_biens	Prix_cotisation_mensuel	
Type			Chaine	Entier	Chaine	Chaine	Entier	Entier	Chaine	Chaine	Chaine	Chaine	Chaine	Entier	
NBR caractères			1	0 à 4		4 à 6	4 à 5								
lignes vides		oui	oui	oui											
Autres	Clé primaire Auto Incrémantion					Clé étrangère dont la référence est la clé primaire de la table region			Deux choix	Deux choix	Trois choix	Deux choix	Fourchette de prix		

Caractéristique fichier region								
	Code_dep_code_commune	reg_code	reg_nom	aca_nom	dep_nom	com_nom_maj_court	dep_code	dep_nom_num
Type	Chaine	Entier	Chaine	Chaine	Chaine	Chaine	Chaine	Chaine
NBR caractères	6	1 à 2					1 à 3	
lignes vides								
Autres	clé primaire							

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	150		Libellé de la voie pour l'adresse du logement assuré
	Code_dep_code_commune	VARCHAR	6	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
	Surface	INT			Surface du logement assuré
	Type_local	ENUM('Appartement', 'Maison')			Type de local. Liste limitée à deux choix pour éviter les erreurs
	Occupation	ENUM('Locataire', 'Propriétaire')			Nature d'occupation. Liste limitée à deux choix pour éviter les erreurs
	Type_contrat	ENUM('Mise en location', 'Residence principale', 'Residence secondaire')			Type de contrat. Liste limitée à trois choix pour éviter les erreurs
	Formule	ENUM('Classique', 'Integrale')			formule souhaitée. Liste limitée à deux choix pour éviter les erreurs
	Valeur_declaree_biens	VARCHAR	30		Fourchettes de prix du biens
	Prix_cotisation_mensuel	INT			Cotisation versée par mois
REGION.CSV	Code_dep_code_commune	VARCHAR	6	Clé primaire	Concaténation du code département et code commune pour avoir une clé unique
	reg_code	INT			Identifiant numérique pour la région.
	reg_nom	VARCHAR	50		Nom de la région
	aca_nom	VARCHAR	50		Nom de l'académie
	dep_nom	VARCHAR	50		Nom du département
	com_nom_maj_court	VARCHAR	50		Nom de la commune en majuscule
	dep_code	VARCHAR	3		Code département
	dep_nom_num	VARCHAR	50		Le nom numérique du département

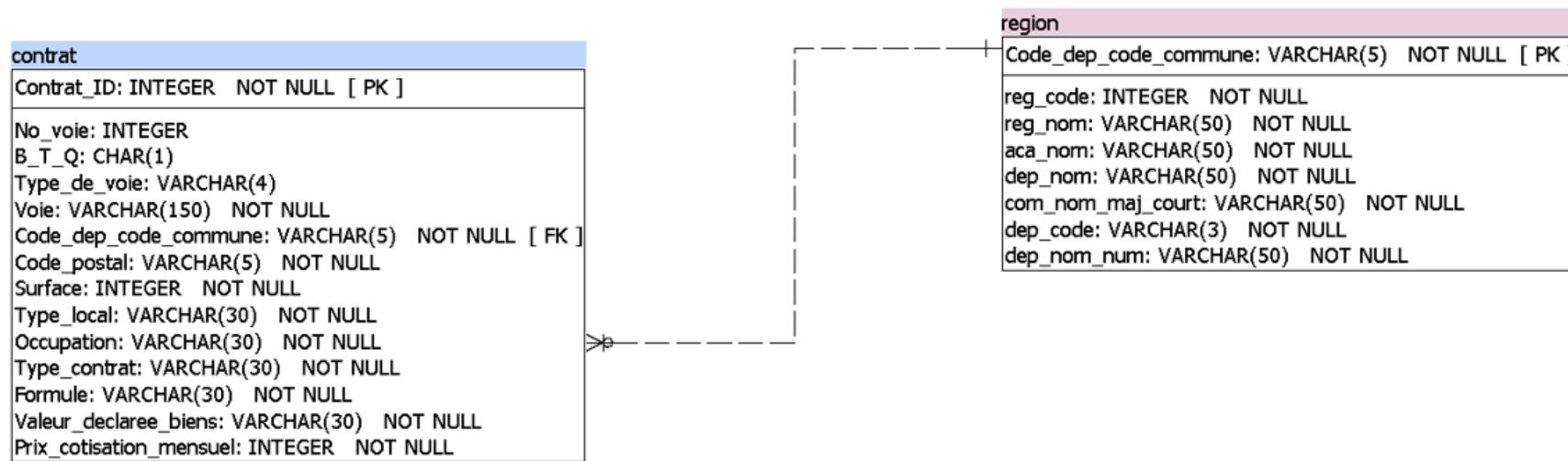
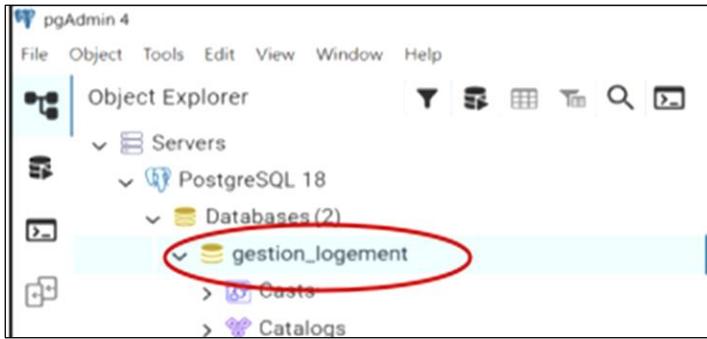


Schéma relationnel UML

Implémentation de la base de données sous PostgreSQL

Création de la base de données gestion_logement



Les contraintes

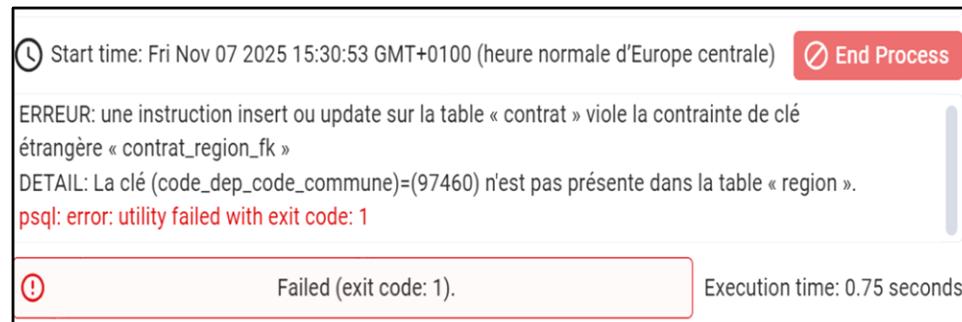


Création des tables : region et contrat



Import des données brutes

Erreur d'import



Analyse

Contrat_ID	No_voie	B_T_Q	Type_de_voie	Voie	Code_dep_com	Code_postal	Surface	Type_local	Occupation	Type_contrat	Formule	Valeur_declar	Prix_cotisation_mensuel
128054	2	RUE	JACQUOT	97460	97460	50	Appartement	Locataire	Residence pri	Classique	0-25000	12	
128056	9	RUE	DESMOUETTE	97434	97434	36	Appartement	Proprietaire	Residence pri	Integral	0-25000	17	
128059	8	LOT	PIGNOLET DE	97470	97470	85	Maison	Proprietaire	Residence pri	Integral	0-25000	9	
128061	26 B	IMP	YLANG YLANC	97460	97460	84	Maison	Proprietaire	Residence pri	Classique	25000-50000	21	
128064	212	RUE	DESARGONA	97434	97434	76	Appartement	Proprietaire	Residence pri	Classique	0-25000	17	
128068	40	RUE	DES ENGAGE	97434	97434	48	Appartement	Locataire	Residence pri	Classique	0-25000	16	
128070	1 C	RUE	DES MUREX	97434	97434	46	Appartement	Locataire	Residence pri	Integral	0-25000	15	
128077	34	RUE	DE LA CAVERN	97460	97460	55	Appartement	Proprietaire	Mise en locati	Integral	0-25000	15	
128082	7	RUE	DESSAPHIRS	97460	97460	59	Appartement	Locataire	Residence pri	Classique	0-25000	11	

Décision

› Suppression des 9 contrats existants dans le fichier contrat.csv et absents dans le fichier region.csv

Impact

- › Import fichier contrat.csv s'est effectué avec succès.
- › Le nombre total des enregistrements de la table contrat passe de 30335 lignes à 30326 lignes soit 9 contrats de moins.

Interrogation bdd: Sélection, Jointure et filtrage des contrats

Script SQL : Lister les contrats avec leur surface de la commune de CAEN

```
SELECT c.contrat_id, c.surface, r.com_nom_maj_court  
FROM contrat AS c  
JOIN region AS r  
ON c.code_dep_code_commune = r.code_dep_code_commune  
WHERE r.com_nom_maj_court = 'CAEN';
```



Data Output			Messages Notifications
	contrat_id	surface	com_nom_maj_court
1	103791	35	CAEN
2	103792	99	CAEN
3	103793	40	CAEN
4	103794	20	CAEN

Total rows: 4 Query complete 00:00:00.071

Script SQL : Lister le nom des régions de France

```
SELECT DISTINCT reg_nom  
FROM region;
```



Data Output	
	reg_nom
1	Hauts-de-France
2	La Réunion
3	Bretagne
4	Mayotte
5	Auvergne-Rhône-Alpes
6	Ile-de-France
7	Grand Est
8	Corse
9	Collectivités d'outre-mer
10	Guyane
11	Bourgogne-Franche-Comté
12	Provence-Alpes-Côte d'A...

Total rows: 19 Query complete 00:00:00.133

Interrogation bdd : Agrégations et statistiques

Script SQL : Calculer la moyenne de la cotisation mensuelle

```
SELECT  
    ROUND(AVG(prix_cotisation_mensuel), 2) AS "Moyenne de  
    cotisation mensuelle"  
  
    COUNT(*) AS "Nombre de contrats"  
  
FROM contrat;
```



Data Output		
	Messages	Notifications
Moyenne de cotisation mensuelle numeric 1	19.33	Nombre de contrats bigint 30326

Script SQL : Trouver le nombre de formules “Integral” sur la région Pays de la Loire

```
SELECT COUNT(c.contrat_id) AS "Nombre de formule Intégrale"  
  
FROM contrat AS c  
  
JOIN region AS r  
  
ON c.code_dep_code_commune = r.code_dep_code_commune  
  
WHERE c.formule = 'Integral'  
  
AND reg_nom = 'Pays de la Loire';
```



Data Output		
	Messages	Notifications
Nombre de formule Intégrale bigint 1	589	

Interrogation bdd: Comptage et filtre avancé

Script SQL : Top 10 départements

```
SELECT ROUND(AVG(prix_cotisation_mensuel), 2 ) AS  
"Cotisation mensuelle moyenne",  
r.dep_nom AS "Département"  
FROM contrat AS c  
JOIN region AS r  
ON c.code_dep_code_commune = r.code_dep_code_commune  
GROUP BY r.dep_nom  
ORDER BY "Cotisation mensuelle moyenne" DESC  
LIMIT 10;
```



	Cotisation mensuelle moyenne numeric	Département character varying (50)
1	36.40	Paris
2	26.27	Hauts-de-Seine
3	19.82	Val-de-Marne
4	18.89	Yvelines
5	18.49	Rhône
6	18.24	Ain
7	18.14	Alpes-Maritimes
8	17.32	Charente-Maritime
9	17.15	Haute-Savoie
10	17.07	Corse-du-Sud

Script SQL : Les 5 contrats dont les surfaces les plus élevées

```
SELECT *  
FROM contrat  
ORDER BY surface DESC  
LIMIT 5;
```



	contrat_id [PK] integer	no_voie integer	b_t_q character (1)	type_de_voie character varying	voie character varying (15)	code_dep_code_commu character varying (6)	code_postal character varying (5)	surface integer	type_local character varying (30)
1	104211	34	B	CRS	VICTOR HUGO	17421	17600	815	Appartement
2	105463	5229	[null]	RUE	DE LA PRIVADIÈRE	30126	30190	742	Maison
3	130878	23	[null]	BD	DE BEAUSEJOUR	75116	75016	595	Appartement
4	100822	5	A	RUE	FERNAND BAZIN	2201	2600	570	Maison
5	109872	18	[null]	RUE	PAUL BELLAMY	44109	44000	559	Appartement

Bilan et apprentissages

- **Compétences acquises :**
 - Concevoir une base de données (UML, génération script SQL)
 - Implémenter une base relationnelle (base, tables, contraintes...)
 - Nettoyer les fichiers CSV et importer les données brutes via pgAdmin de PostgreSQL
 - Rédiger des requêtes SQL pour extraire les données de la bdd
- **Découverte des outils :**
 - SQL Power Architect, PostgreSQL
- **Conclusion :**
 - Gestion de projet base de données relationnelle de la conception à l'exploitation de données