

Eléments du projet



Contexte projet DATA-Immo

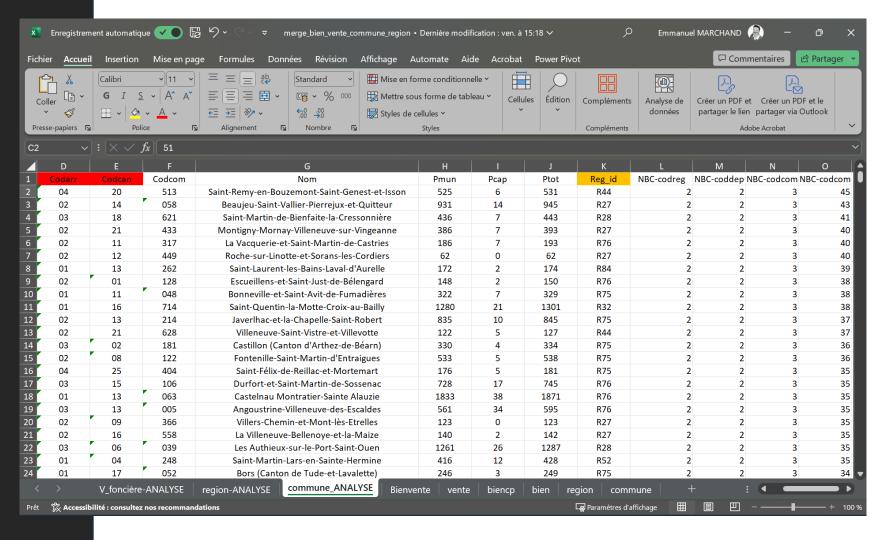
Mise en place d'un modèle de donnée complet pour le calcul prévisionnel des prix de vente de biens.

Un projet en **plusieurs étapes**

EXISTANT	COURT TERME	MOYEN TERME
		Modèle prévisionnel
		Analyse historique complet
	Premières analyses -2020	Premières analyses -2020
	Nouvelle base enrichie	Nouvelle base enrichie
	Nouveau model enrichi	Nouveau model enrichi
Base de données simplifiée	Base de données simplifiée	Base de données simplifiée
Etape 0	Etape 1	Etape 2



La stratégie de sauvegarde et la conformité RGPD



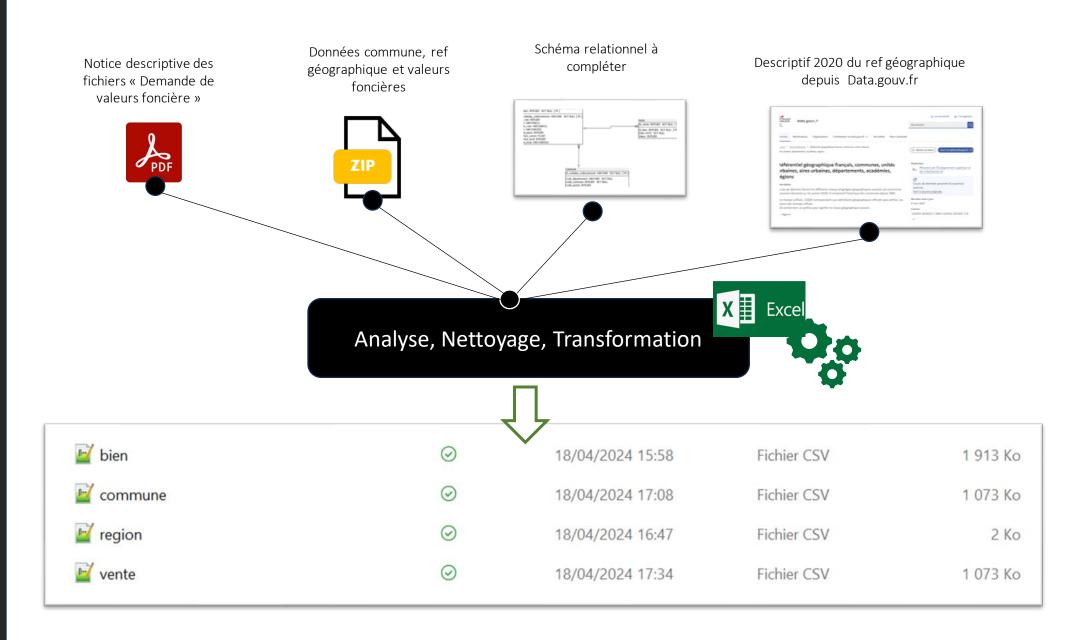
- Copie de toutes les données sources dans un seul fichier Excel . Cela autorise une jointure des données après analyse.
- Analyse de chacun des fichiers pour identifier les données à conserver.



Conformité RGPD : Dans le fichier Valeur foncière suppression de la colonne Nom_de_acquereur.



Les données initiales





L'extrait du dictionnaire des données (1/2)

DICTIONNAIRE DES DONNÉES - Vente

CODE	SIGNIFICATION	TYPE	LONGUEUR	NATURE	REGLE DE GESTION	REGLE DE CALCUL
						Clé primaire
Id_vente	Iddentifiant de la vente - Clé	Integer		Calculer	Ne doit pas être nul	Incrémentation auto par pas de 1 en partant de 1
No_disposition	Numéro de disposition dans le cas d'acte comprenant plusieurs mutations.	Integer		Elémentaire	Ne doit pas être nul	
					Ne doit pas être nul	
Date	Date de la vente.	Date		Elémentaire	Format de la date :	
					jj/mm/aaaa	
Valeur_fonciere	Montant de la mutation en euros	FLoat		Elémentaire		
						Incrémentation auto par pas de 1 pou
Id_bien	Identifiant unique du bien vendu	Integer		Calculer	Ne doit pas être nul	les biens identiques Clé étrangère avec la table Bien

DICTIONNAIRE DES DONNÉES - Bien

CODE	SIGNIFICATION	TYPE	LONGUEUR	NATURE	REGLE DE GESTION	REGLE DE CALCUL
Id_bien	ID dans la base de données	Integer		Elémentaire	Ne doit pas être nul	Clé primaire Incrémentation auto par pas de 1 en partant de 1
No_voie	Numéro de la voie	Integer		Elémentaire		
BTQ	Indice de répétition	Varchar	1			
Type_de_voie	Plusieurs valeurs (rue, avenue, chemin, etc.)	Varchar	4	Elémentaire		
Voie	Nom de la rue	Varchar	30	Elémentaire	Ne doit pas être nul	
Surface_Carrez	Surface en loi Carez du lot de copropriété	FLoat		Elémentaire	Ne doit pas être nul	
Type local	Libellé type local	Varchar	14	Elémentaire	Ne doit pas être nul	Il peut prendre deux valeurs : Appartement ou Maison
Surface_reelle_bati	Surface mesurée au sol entre les murs ou séparations et arrondie au mètre carré inférieur. Les surfaces des dépendances ne sont pas prises en compte.	Integer		Elémentaire	Ne doit pas être nul	
Nombre_pieces_principale	Les cuisines, salles d'eau et dépendances ne sont pas s prises en compte.	Integer		Elémentaire	Ne doit pas être nul	
Surface_terrain	Surface cadastrale du terrain	Integer		Elémentaire		
Codedep_codecom_bien	Code unique commune et département	Varchar	6	Concatener	Ne doit pas être nul	Code_departement et Code_commune Clé étrangère avec la table Commune



L'extrait du dictionnaire des données (2/2)

DICTIONNAIRE DES DONNÉES - Commune

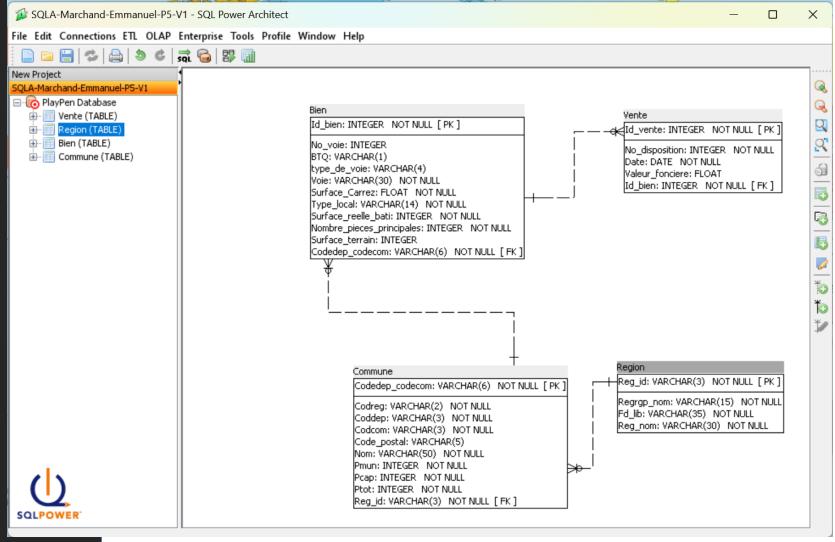
CODE	SIGNIFICATION	TYPE	LONGUEUR	NATURE	REGLE DE GESTION	REGLE DE CALCUL
Codedep_codecom	Code unique commune et département - Clé	Varchar	6	Concatener	Ne doit pas être nul	Clé primaire Concaténation de Coddep et Codcom
Codreg	Identifiant de région	Varchar	2	Elémentaire	Ne doit pas être nul	
Coddep	Code du département	Varchar	3	Elémentaire	Ne doit pas être nul	
Codcom	Code de la commune	Varchar	3	Elémentaire	Ne doit pas être nul	
Code_postal	Code postal	Varchar	5	Calculer		Faire une jointure avec la table Bien
Nom	Libellé de la commune	Varchar	50		Ne doit pas être nul	
Pmun	Population municipale. Elle comprend les personnes ayant leur résidence habituelle sur le territoire de la commune.	Integer		Elémentaire	Ne doit pas être nul	
Pcap	Population comptée à part comprend certaines personnes dont la résidence habituelle est dans une autre commune mais qui ont conservé une résidence sur le territoire de la commune	Interger		Elémentaire	Ne doit pas être nul	
Ptot	Population totale est la somme de la population municipale et de la population comptée à part.	Integer		Elémentaire	Ne doit pas être nul	Somme de PMUN et PCAP
Reg_id	Identifiant de région	Varchar	3	Concatener	Ne doit pas être nul	Clé étrangère avec la table Region 'R' + Codereg

DICTIONNAIRE DES DONNÉES - Région

CODE	SIGNIFICATION	ТҮРЕ	LONGUEUR	NATURE	REGLE DE GESTION	REGLE DE CALCUL
Reg_id	Identifiant de la région	Varchar	3	Elementaire	Ne doit pas être nul	Clé primaire
Regrgp_nom	Nom de regroupement de région	Varchar	15	Elementaire	Ne doit pas être nul	
Fd_lib	Nom du découpage administratif	Varchar	35	Elementaire	Ne doit pas être nul	
Reg_nom	Nom de la région	Varchar	30	Elementaire	Ne doit pas être nul	
neg_nom	Nom de la region	Varenar	30	Licincitane	The dore pus etre ridi	



Le schéma relationnel normalisé

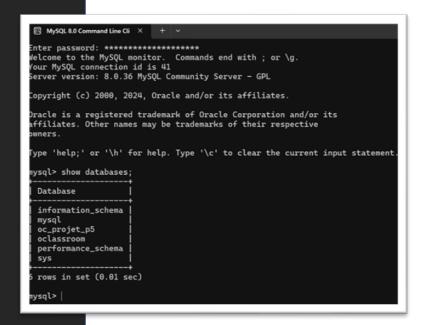


- Uilisation du logiciel SQL Power Architect.
- Le respect de la normalisation **3NF** implique :
 - Que chaque table soit constituée d'une clé primaire
 - La suppression des **redondances** de données :
 - Toutes les informations concernant une commune sont déplacées de la table « Bien » vers la table « Commune ».
 - Dans table « Bien » suppression des codes Voie inutiles.



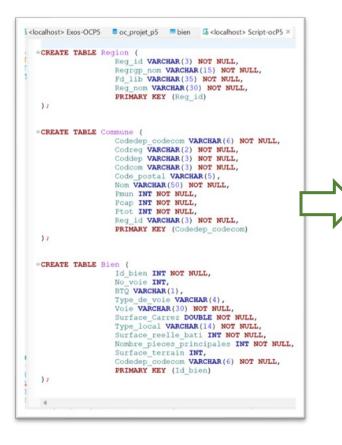
Création de la base de données





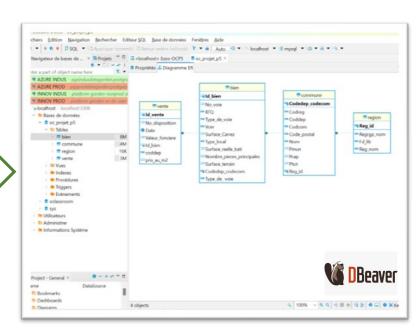
- Utilisation du SGBD MySQL
- Installation MySQL pour création une base en local.
- Dans MysSQL avec la commande CREATE DATABASE oc_projet_P5; , création d'une base dédiée au projet P5.





 Exécution du script SQL de création de base issue du logiciel de modélisation de base SQL Power Architecte. Dans le gestionnaire de base Dbeaver

CRÉATION DE LA BASE





Chargement de la base de données

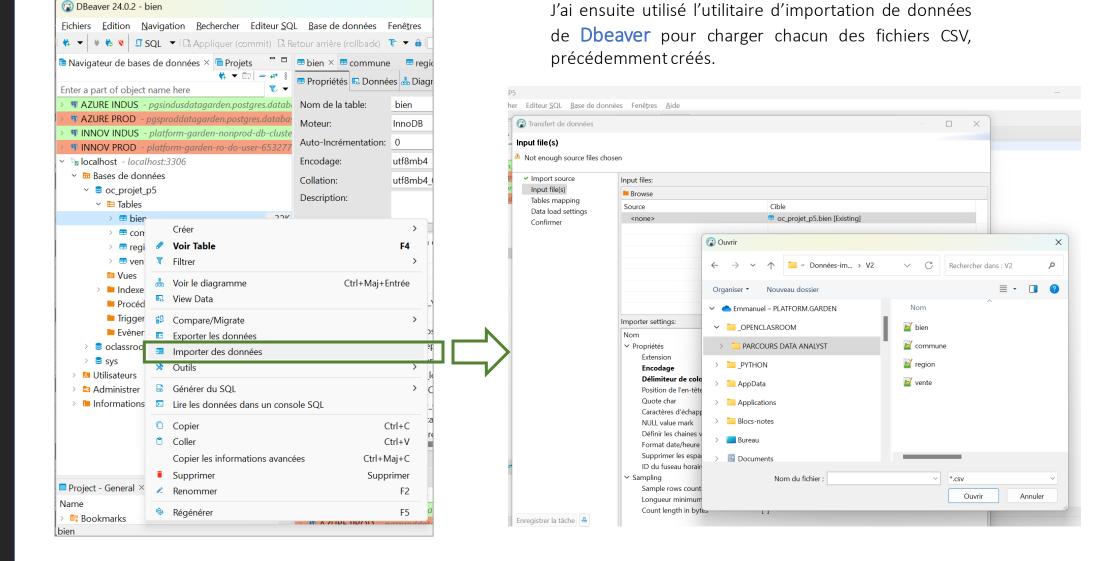




Table Vente chargée

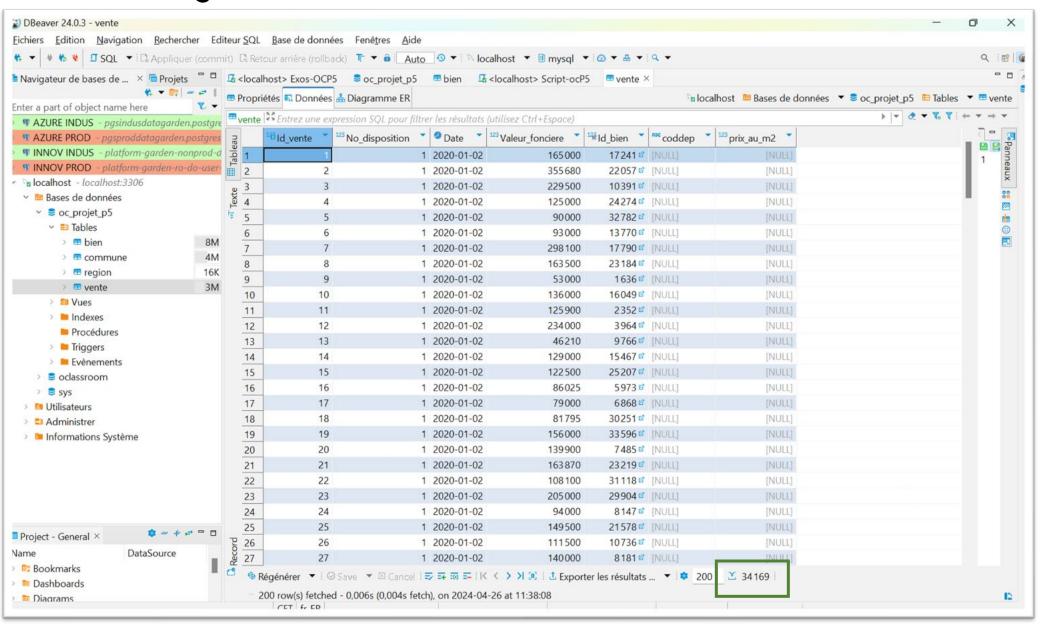




Table Bien chargée

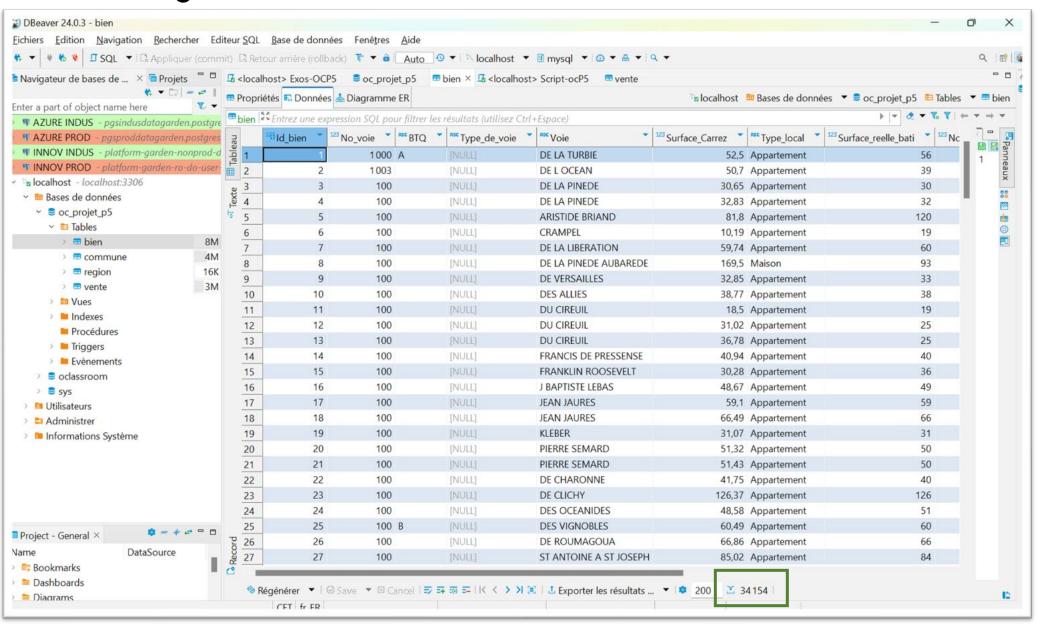




Table Commune chargée

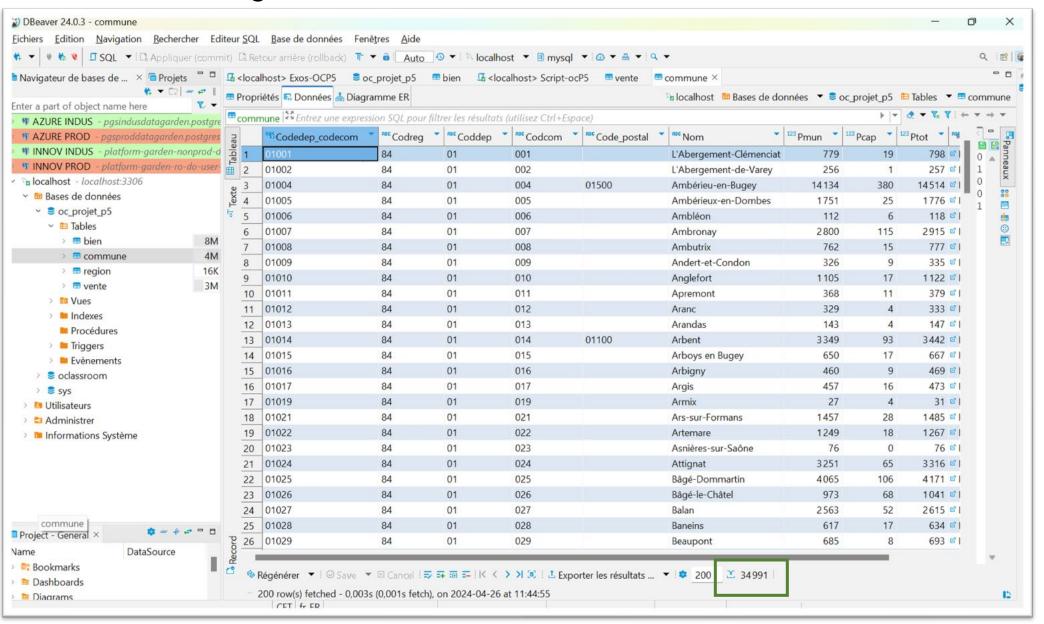
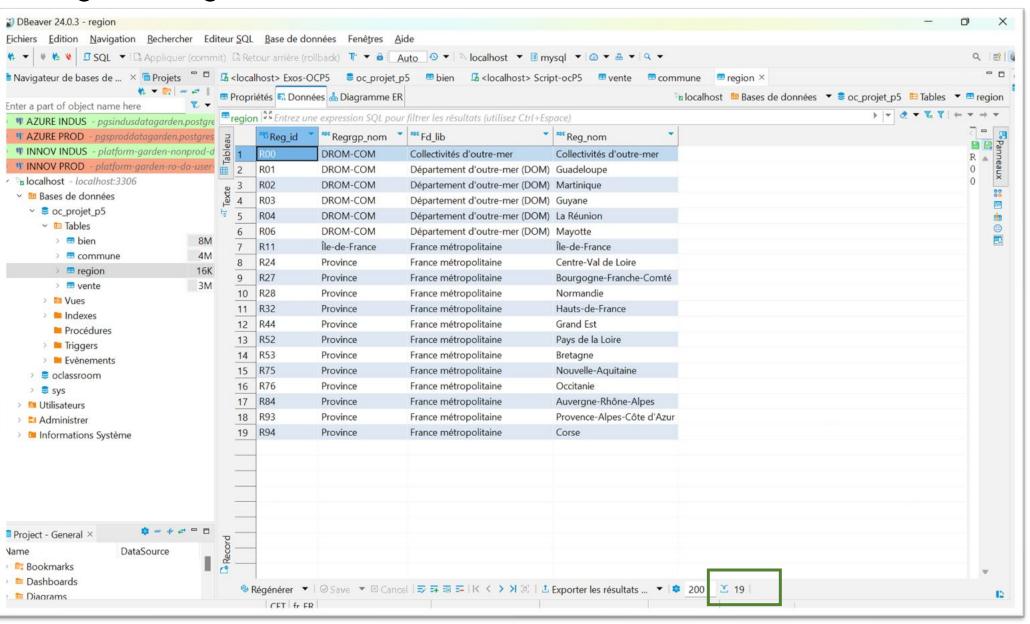
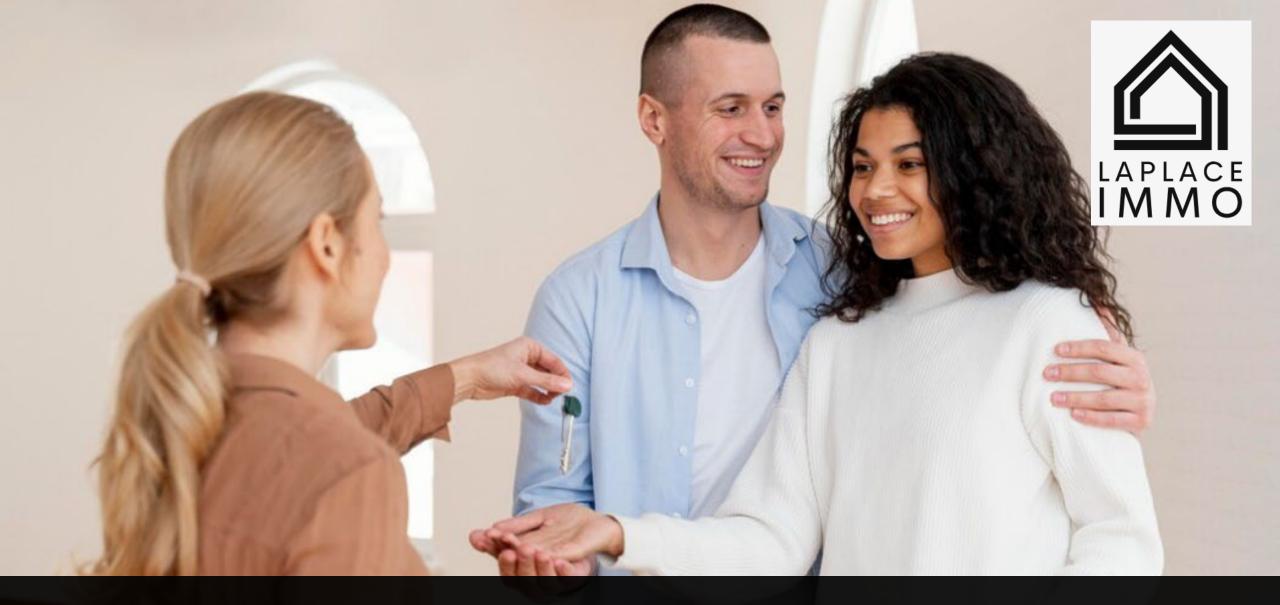




Table Région chargée

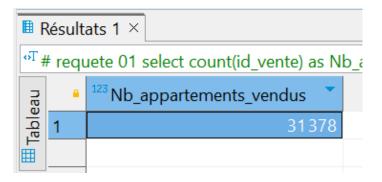




Requêtes SQL et résultats

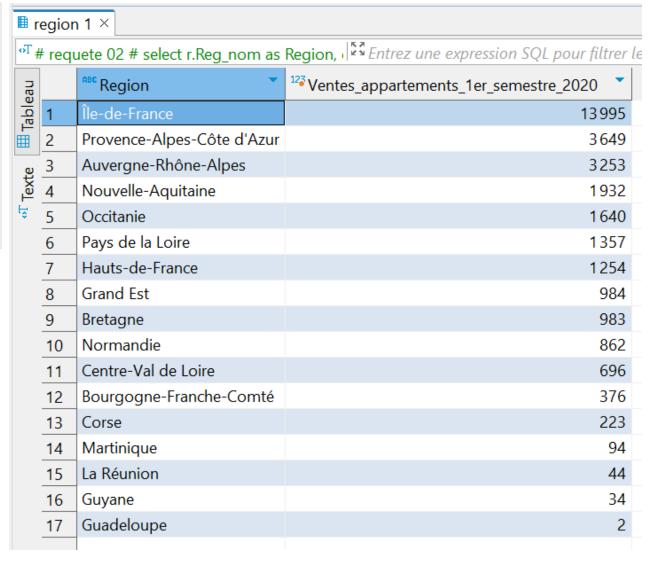
1 - Nombre total d'appartements vendus au 1er semestre 2020.

```
select count(id_vente) as Nb_appartements_vendus
from vente v
inner join bien b on v.Id_bien = b.Id_bien
where
  b.Type_local = 'Appartement'
  and v.`Date` >= '2020-01-01'
  and v.`Date` <= '2020-06-30'</pre>
```



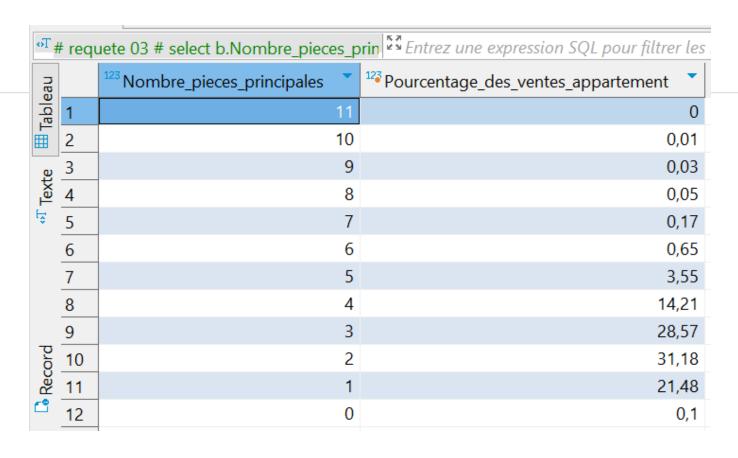
2 - Le nombre de ventes d'appartement par région pour le 1 er semestre 2020.

```
select r.Reg_nom as Region, count(id_vente) as
Ventes_appartements_ler_semestre_2020
from vente v
inner join bien b on v.Id_bien = b.Id_bien
inner join commune c on b.Codedep_codecom = c.Codedep_codecom
inner join region r on c.Reg_id = r.Reg_id
where
   b.Type_local = 'Appartement'
   and v.`Date` >= '2020-01-01'
   and v.`Date` <= '2020-06-30'
group by r.Reg_nom
order by 2 desc</pre>
```





3 - Proportion des ventes d'appartements par le nombre de pièces.



4 - Liste des 10 départements où le prix du mètre carré est le plus élevé.

```
select
c.coddep,
ROUND(avg(Valeur_fonciere/Surface_Carrez),2) as prix_au_m2
from vente v
inner join bien b on v.Id_bien = b.Id_bien
inner join commune c on b.Codedep_codecom = c.Codedep_codecom
group by 1
order by 2 desc
limit 10
```

⇔ Τ #	√T # requete 04 # select c.coddep, ROUND(a								
		coddep	prix_au_m2						
Tableau	1	75	12052,82						
	2	92	7219,39						
	3	94	5343,28						
[⊕] Texte	4	06	4700,26						
Ê	5	74	4667,13						
	6	93	4344,78						
	7	78	4225,25						
	8	69	4059,28						
	9	2A	4026,97						
cord	10	33	3764,14						



5 - Prix moyen du mètre carré d'une maison en Île-de-France.

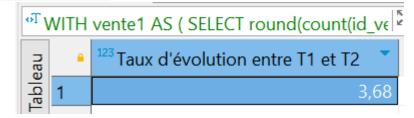
6 - Liste des 10 appartements les plus chers avec la région et le nombre de mètres carrés

```
select
b.Surface_reelle_bati as surface_m2, r.Reg_nom ,
Valeur_fonciere as prix_appartement
from vente v
inner join bien b on v.Id_bien = b.Id_bien
inner join commune c on b.Codedep_codecom = c.Codedep_codecom
inner join region r on c.Reg_id = r.Reg_id
where
   b.Type_local = 'Appartement'
order by 3 desc
limit 10
```

*T # requete 06 # select b.Surface_reelle_bati a Entrez une expressi						
1 Lableau		¹²³ surface_m2	Reg_nom	prix_appartement		
- -	1	10	Île-de-France	9000000		
ı.		62	Île-de-France	8600000		
	3 ر	289	Île-de-France	8577713		
4	4	42	Île-de-France	7620000		
Ê	5	200	Île-de-France	7600000		
	6	143	Île-de-France	7535000		
	7	357	Île-de-France	7420000		
	8	241	Île-de-France	7200000		
	9	310	Île-de-France	7050000		
2	10	76	Île-de-France	6600000		

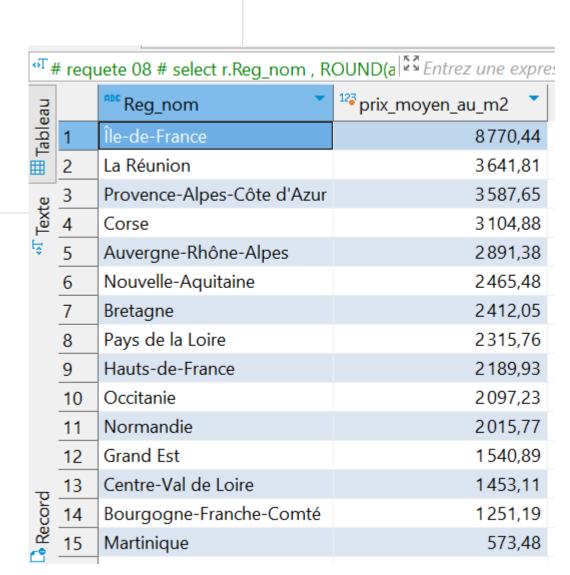


7 - Taux d'évolution du nombre de ventes entre le premier et le second trimestre de 2020



8 - Le classement des régions par rapport au prix au mètre carré des appartement de plus de 4 pièces.

```
select
   r.Reg_nom ,
   ROUND(avg(Valeur_fonciere/Surface_Carrez),2) as prix_moyen_au_m2
from vente v
inner join bien b on v.Id_bien = b.Id_bien
inner join commune c on b.Codedep_codecom = c.Codedep_codecom
inner join region r on c.Reg_id = r.Reg_id
where
   b.Type_local = 'Appartement'
   and b.Nombre_pieces_principales > 4
group by 1
order by 2 desc
```





9 - Liste des communes ayant eu au moins 50 ventes

au 1^{er} trimestre

```
WITH
vente trim1 as (
  select
 c. Nom as Nom commune,
 count(id vente) as nbv
  FROM
 vente v
 inner join bien b on v.id bien = b.id bien
 inner join commune c on b.Codedep codecom =
c.Codedep codecom
 WHERE V.date BETWEEN '2020-01-01' AND '2020-03-31'
 group by 1)
select *
from vente trim1
where nbv >= 50
order by 2 desc
```





[•] T # requete 09 # WITH vente_trim1 as (selec						
au	<u> </u>	Nom_commune	123 nbv	•		
Tableau	22	Courbevoie		80		
	23	Paris 5e Arrondissement		79		
e.	24	Paris 3e Arrondissement		79		
[⊹] T Texte	25	Toulouse		78		
Ê	26	Antibes		77		
	27	Marseille 4e Arrondisseme		72		
	28	Marseille 1er Arrondissem		71		
	29	Rueil-Malmaison		68		
	30	Vincennes		68		
	31	Lille		67		
	32	Marseille 9e Arrondisseme		66		
	33	Montreuil		65		
	34	Angers		64		
	35	Nîmes		63		
	36	La Ciotat		62		
	37	Sète		62		
	38	Paris 8e Arrondissement		62		
	39	Paris 2e Arrondissement		61		
	40	Rennes		61		
	41	Paris 4e Arrondissement		60		
	42	Toulon		59		
	43	Levallois-Perret		59		
	44	Saint-Maur-des-Fossés		56		
	45	Versailles		54		
	46	Ajaccio		54		
ord	47	Puteaux		53		
Secord	48	Issy-les-Moulineaux		50		

10 - Différence en pourcentage du prix au mètre carré entre un appartement de 2 pièces et un appartement de 3 pièces.

```
WITH
prix3pieces AS (
  select
 ROUND (avg (Valeur fonciere/Surface Carrez), 2) as prix m2 T3
  from vente v
  inner join bien b on v.Id bien = b.Id bien
  where
 b.Type local = 'Appartement'
  and b.Nombre pieces principales = 3
                                                                                     T# requete 10 # WITH prix3pieces AS ( sele
prix2pieces AS (
                                                                                      Tableau
                                                                                               Différence en pourcentage
  select
 ROUND (avg (Valeur fonciere/Surface Carrez), 2) as prix m2 T2
  from vente v
  inner join bien b on v.Id bien = b.Id bien
  where
 b.Type local = 'Appartement'
  and b.Nombre pieces principales = 2)
SELECT
round(((prix m2 T3 - prix m2 T2) / prix m2 T2 * 100), 2)
AS "Différence en pourcentage"
FROM prix3pieces, prix2pieces
```

11 - Les moyennes de valeurs foncières pour le top 3 des communes des départements 6, 13, 33, 59 et 69.

```
with
liste dep as
  select
  c. Coddep as departement,
  c. Nom as Nom commune,
  ROUND (AVG (v. Valeur fonciere), 0) as valeur fonciere moyenne
  from vente v
  inner join bien b on v.Id bien = b.Id bien
  inner join commune c on b. Codedep codecom = c. Codedep codecom
  where
  c.Coddep in ('06','13','33', '59', '69')
 group by 1,2
  order by 1,3 desc
ranking dep as (
select
departement, Nom commune, valeur fonciere moyenne,
RANK() over (partition by departement
order by valeur fonciere moyenne desc) vf rank
from liste dep
select
departement, Nom commune, valeur fonciere moyenne
from ranking dep
where vf rank in (1,2,3)
```

	*T #Autre version with liste_dep as (select c.Cc Entrez une expression SQL pour filtrer							
	an	<u> </u>	departement *	Nom_commune	valeur_fonciere_moyenne			
	Tableau	1	06	Saint-Jean-Cap-Ferrat	968750			
		2	06	Eze	655 000			
ľ	e	3	06	Mouans-Sartoux	476898			
	⁴T Texte	4	13	Gignac-la-Nerthe	330000			
	Ê	5	13	Saint-Savournin	314425			
		6 13		Cassis	313417			
		7	33	Lège-Cap-Ferret	549 501			
		8	33	Vayres	335 000			
		9	33	Arcachon	307436			
		10	59	Bersée	433 202			
		11	59	Cysoing	408550			
		12	59	Halluin	322250			
		13	69	Ville-sur-Jarnioux	485 300			
		14	69	Lyon 2e Arrondissemen	455 217			
		15	69	Lyon 6e Arrondissemen	426968			

12 - Les 20 communes avec le plus de transactions pour 1000 habitants pour les communes qui dépassent les 10 000 habitants

```
select
c.Coddep as departement, c.Nom,
ROUND (AVG (v. Valeur fonciere), 0) as valeur fonciere moyenne,
ROUND (avg (b. Surface Carrez), 2) as surface carrez moyenne,
Round((count(v.Id vente) * 1000) / max(c.Ptot) ,2) as
Transactio pour mille habitants
from vente v
inner join bien b on v.Id bien = b.Id bien
inner join commune c on b. Codedep codecom = c. Codedep codecom
where
                                                                   oT with cte as( select c.Codedep_codecom as ic Estatrez une expression SQL pour filtrer les résultats (utilisez Ctrl+Espace)
b.Surface Carrez != 0
                                                                          Code postal Nom_commune
                                                                                                         valeur_fonciere_moyenne valeur_fonciere_moyenne valeur_fonciere_moyenne valeur_fonciere_moyenne valeur_fonciere_moyenne
and c. Ptot > 10000
                                                                         75
                                                                                        Paris 2e Arrondissement
                                                                                                                              41.75
                                                                                                                                                      521065
                                                                                                                                                                                       5,84
group by 1, 2
                                                                                        Paris 1er Arrondissement
                                                                     2
                                                                         75
                                                                                                                              54.39
                                                                                                                                                    1013504
                                                                                                                                                                                       4,92
order by 5 desc
                                                                                        Paris 3e Arrondissement
                                                                                                                              48,51
                                                                                                                                                     655954
                                                                                                                                                                                       4,69
limit 20
                                                                         33
                                                                                        Arcachon
                                                                                                                              49,86
                                                                                                                                                     307436
                                                                                                                                                                                       4,62
                                                                                        La Baule-Escoublac
                                                                                                                              49,78
                                                                                                                                                                                       4,58
                                                                                                                                                     264745
                                                                         75
                                                                                        Paris 4e Arrondissement
                                                                                                                              47,72
                                                                                                                                                      650909
                                                                                                                                                                                       4,08
                                                                                        Roquebrune-Cap-Martin
                                                                                                                               47.1
                                                                                                                                                     274173
                                                                                                                                                                                       3,99
                                                                         75
                                                                                        Paris 8e Arrondissement
                                                                                                                              75,31
                                                                                                                                                    1010536
                                                                                                                                                                                       3.83
                                                                                                                                                                                        3,5
                                                                                        Sanary-sur-Mer
                                                                                                                              48,49
                                                                                                                                                     268489
                                                                         83
                                                                                        La Londe-les-Maures
                                                                                                                              37.13
                                                                                                                                                     151213
                                                                                                                                                                                       3,43
                                                                      10
                                                                         75
                                                                                        Paris 9e Arrondissement
                                                                                                                              49.89
                                                                                                                                                     583 177
                                                                                                                                                                                       3,43
                                                                      11
                                                                         75
                                                                                        Paris 6e Arrondissement
                                                                                                                              58.74
                                                                                                                                                    1005956
                                                                                                                                                                                       3.38
                                                                      12
                                                                         83
                                                                                        Saint-Cyr-sur-Mer
                                                                                                                              53.13
                                                                                                                                                     263837
                                                                                                                                                                                       3,24
                                                                      13
                                                                         60
                                                                                        Chantilly
                                                                                                                              60,55
                                                                                                                                                     238481
                                                                                                                                                                                       3,13
                                                                      14
                                                                                        Pornichet
                                                                                                                              43,67
                                                                      15
                                                                                                                                                      204 127
                                                                                                                                                                                       3,06
                                                                                        Saint-Mandé
                                                                                                                              54,93
                                                                      16
                                                                                                                                                      494061
                                                                                                                                                                                       3,06
                                                                      17
                                                                         75
                                                                                        Paris 10e Arrondissemen
                                                                                                                              49.01
                                                                                                                                                      532282
                                                                                                                                                                                       3,04
                                                                      18
                                                                                        Menton
                                                                                                                              45,44
                                                                                                                                                      208585
                                                                                                                                                                                       2,94
                                                                         85
                                                                                        Saint-Hilaire-de-Riez
                                                                                                                              35.99
                                                                                                                                                      85265
                                                                                                                                                                                       2.87
                                                                      20
                                                                                                                              49,65
                                                                                                                                                      433 108
                                                                                                                                                                                       2,81
                                                                                        Vincennes
```



Ressources



ATTRIBUTION images:

Image de freepik

lmage de rawpixel.com sur Freepik