



PROJET

CRÉEZ ET UTILISEZ UNE BASE DE DONNÉES
IMMOBILIÈRE - DATA IMMO AVEC SQL



CONTENU

01

INTRODUCTION

02

LA STRATÉGIE DE SAUVEGARDE ET LA CONFORMITÉ RGPD

03

LES DONNÉES INITIALES

04

UN EXTRAIT DU DICTIONNAIRE DES DONNÉES

05

LE SCHÉMA RELATIONNEL NORMALISÉ

06

BASE DE DONNÉES

07

LES REQUÊTES PAR SQL

1.INTRODUCTION

CONTEXT

- Le projet "DATAImmo" vise à développer un modèle prédictif pour estimer les prix de vente des biens immobiliers

OBJETIF DE PROJET

- Modifier de le base de données permettant de collecter les transaction immobilières et foncières en France
- Analyser le marché et aider les différentes agences régionales à mieux accompagner leurs clients.

2.LA STRATÉGIE DE SAUVEGARDE ET LA VÉRIFICATION DE LA CONFORMITÉ RGPD



LA STRATÉGIE DE SAUVEGARDE

- Établit une fréquence de sauvegarde (quotidienne, hebdomadaire, mensuelle) en fonction de la fréquence des mises à jour des données.
- Protéger les données contre les pertes et assurer la récupération en cas de besoin.



LA VÉRIFICATION DE LA CONFORMITÉ RGPD

- Vérifier s'il y a des informations personnellement identifiables
- Assurer que des mesures appropriées sont mises en place pour la protection des données et la vie privée.



3.LES DONNÉES INITIALES

Données sur les valeurs foncières

Les informations sur les transactions immobilières:

- Date de vente
- Valeur de transaction
- Type de bien
- Surface

...

Données communes

Données de recensement de L'INSEE:

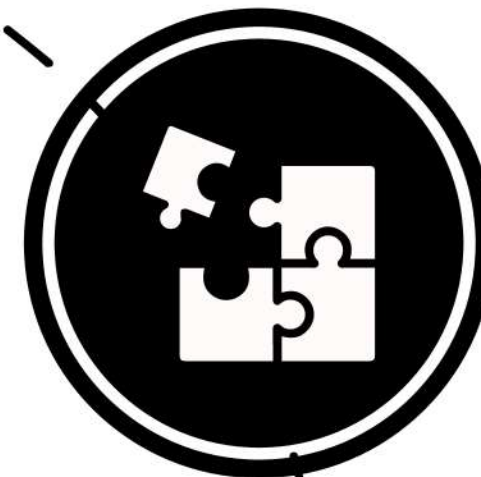
- Nom de commune
- Code de commune
- Population

Données géographique

Données de Data.gouv concernant:

- Les régions
- Communes
- L'unités urbaines
- Aires urbaines
- Départements

...



4. UNE EXTRAIT DU DICTIONNAIRE DES DONNÉES



LAPLACE IMMO.

DICTIONNAIRE DES DONNÉES - Valeurs foncières

CODE	SIGNIFICATION	TYPE	LONGUEUR	NATURE	REGLE DE GESTION	REGLE DE CALCUL
Id_bien	Identifiant de bien	Integer	NC	Elémentaire	Ne doit pas être nul	
Date mutation	Date de la transaction	Date	5	Elémentaire	Format de la date (aaaa/mm/jj)	
Valeur fonciere	Valeur de la transaction	Integer	10	Elémentaire	Ne doit pas être nul	
No voie	Numéro d'adresse d'un bâtiment ou terrain	Integer	5	Elémentaire		
B/T/Q	Bâtiment, tour, ou quartier (spécifie une partie d'un complexe ou d'une	CHAR	5	Elémentaire		
Type de voie	Type de route ou de passage (rue, avenue, boulevard, etc...)	Varchar	5	Elémentaire		
Voie	Nom de la rue ou du passage	Varchar	30	Elémentaire	Ne doit pas être nul	
Code postal	Code numérique qui représente une zone postale au sein d'une commune	Integer	5	Elémentaire	Ne doit pas être nul	
Commune	Nom de la commune	Varchar	30	Elémentaire	Ne doit pas être nul	
Code departement	Code du département	Integer	5	Elémentaire	Ne doit pas être nul	
Code commune	Code de la commune	Integer	5	Elémentaire	Ne doit pas être nul	
Type local	Type de local (maison, appartement...)	Varchar	15	Elémentaire	Ne doit pas être nul	
Surface carrez		Float		Elémentaire	Ne doit pas être nul	
Surface reelle bati	Surface réelle bâtiment	Integer	5	Elémentaire	Ne doit pas être nul	
Nombre pieces principale	Nombre de pièces principales	Integer	5	Elémentaire	Ne doit pas être nul	
Id_vente	Identifiant de vente	Integer		Elémentaire	Ne doit pas être nul	

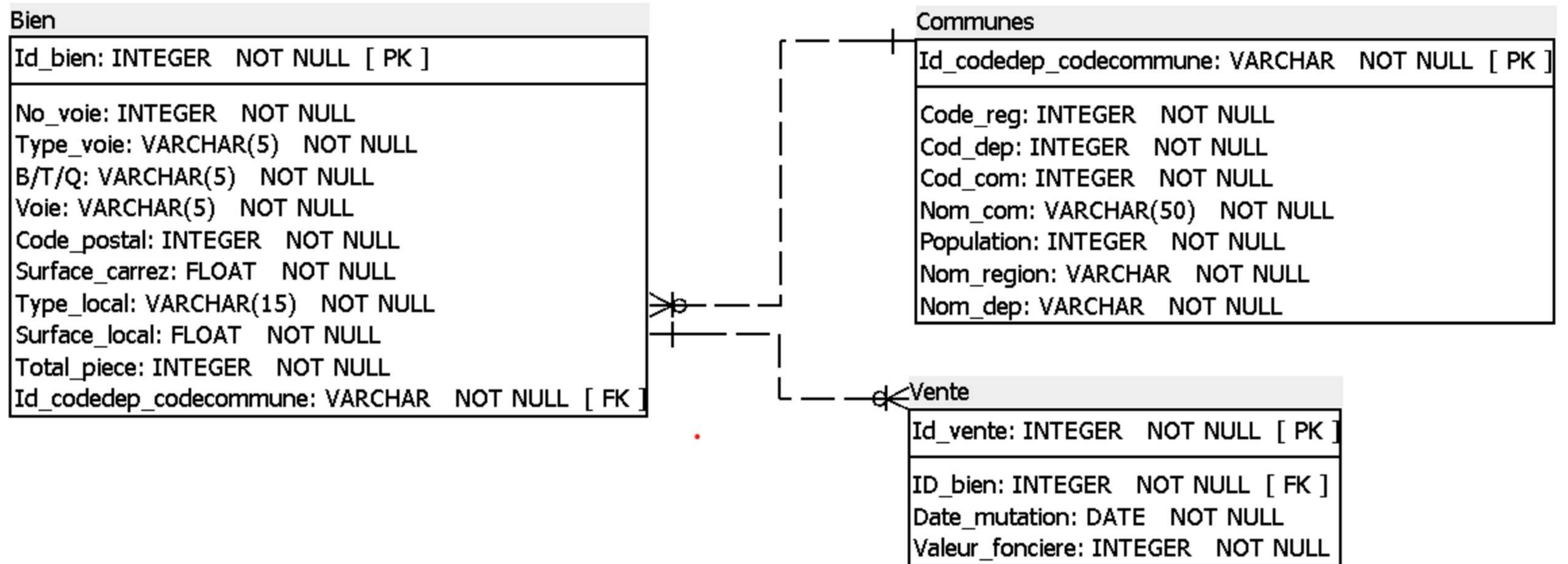
DICTIONNAIRE DES DONNÉES - Référentiel géographique

CODE	SIGNIFICATION	TYPE	LONGUEUR	NATURE	REGLE DE GESTION	REGLE DE CALCUL
reg_nom	Nom de la région	Varchar	30	E	Ne doit pas être nul	
dep_nom	Nom du département	Varchar	45	E	Ne doit pas être nul	
com_code	Code département et de la commune	Integer	10	E	Ne doit pas être nul	
dep_code	Code du département	Integer	5	E	Ne doit pas être nul	
dep_num_nom	Code du département avec nom du département	Varchar	50	Co	Ne doit pas être nul	dep_nom_num (invers)
reg_code	Code de la région actuel	Integer	5	E	Ne doit pas être nul	

DICTIONNAIRE DES DONNÉES - Données communes

CODE	SIGNIFICATION	TYPE	LONGUEUR	NATURE	REGLE DE GESTION	REGLE DE CALCUL
Id_codedep_codecommune	Identifiant de code département et code commune	Integer	10	Ca	Ne doit pas être nul	CODDEP+ CODECOM
CODREG	Code de la région	Integer	5	E	Ne doit pas être nul	
CODDEP	Code du département	Integer	5	E	Ne doit pas être nul	
CODARR	Code de l'arrondissement	Integer	5	E	Ne doit pas être nul	
CODCOM	Code de la commune	Integer	5	E	Ne doit pas être nul	
COM (Nom_commune)	Nom de la commune	Varchar	45	E	Ne doit pas être nul	
PMUN	Population municipale	Integer	10	E	Ne doit pas être nul	
PCAP	Population comptée à part	Integer	5	E	Ne doit pas être nul	
PTOT	Population totale	Integer	10	Ca	Ne doit pas être nul	PMUN + PCAP

5. LE SCHÉMA RELATIONNEL NORMALISÉ



6. BASE DE DONNÉES

BIEN

Projet 5 Table name: Bien ☐ WITHOUT ROWID ☐ STRICT

	Name	Data type	Primary Key	Foreign Key	Unique	Check	Not NULL	Collate	Generated
1	Id_bien	INTEGER							
2	No voie	INTEGER							
3	B/T/Q	INTEGER							
4	Type_voie	TEXT							
5	Voie	TEXT							
6	Code postal	INTEGER							
7	Surface_carrez	NUMERIC							
8	Type_local	INTEGER							
9	Surface_local	INTEGER							
10	Total_piece	INTEGER							
11	Id_codedep_codecommune	TEXT							



Grid view Form view

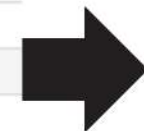
Total rows loaded: 34169

	Id bien	No voie	B/T/Q	Type voie	Voie	Code postal	Surface carrez	Type local	Surface local	Total piece	Id codedep	codecommune
1	1	347		RUE	DU CHATEAU	1170	48,22	Appartement	48	3	1103	
2	2	4		BD	EDOUARD BAUDOUIN	6160	39,11	Appartement	40	1	6004	
3	3	20	B	RUE	MARCEAU	6000	80,25	Appartement	82	3	6088	
4	4	550		RTE	DES VESPINS RN7	6700	27,51	Appartement	27	1	6123	
5	5	9300		RES	LES ARPEGES BD DES ABA	13400	47,33	Appartement	47	2	13005	
6	6	27		RUE	DU GRAND MADIER	13600	25,31	Appartement	24	1	13028	
7	7	360		AV	DU PRADO	13008	70,84	Appartement	70	3	13208	
8	8	5076	F	PARC	DESSUARD	13012	67,19	Appartement	66	3	13212	
9	9	1194		RUE	DE NORMANDIE	14510	18,89	Appartement	19	1	14338	
10	10	30		ALL	DES NOISETIERS	14100	105,37	Maison	99	4	14366	
11	11	11		RUE	ROUGET DE L ISLE	17000	31,99	Appartement	34	2	17300	

COMMUNES

Projet 5 Table name: Communes ☐ WITHOUT ROWID ☐ STRICT

	Name	Data type	Primary Key	Foreign Key	Unique	Check	Not NULL	Collate	Generated
1	Code_reg	INTEGER							
2	Code_dep	INTEGER							
3	Code_com	INTEGER							
4	Nom_com	TEXT							
5	Population	INTEGER							
6	Id_codedep_codecommune	TEXT							
7	Nom_reg	TEXT							
8	Nom_dep	INTEGER							



Grid view Form view

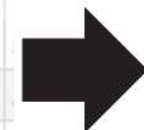
Total rows loaded: 34991

	Code reg	Code dep	Code com	Nom com	Population	Id codedep	codecommune	Nom reg	Nom dep
1	84	1	1	L'Abergement-Clémenciat	798	1001		Auvergne-Rhône-Alpes	Ain
2	84	1	2	L'Abergement-de-Varey	257	1002		Auvergne-Rhône-Alpes	Ain
3	84	1	4	Ambérieu-en-Bugey	14514	1004		Auvergne-Rhône-Alpes	Ain
4	84	1	5	Ambérieux-en-Dombes	1776	1005		Auvergne-Rhône-Alpes	Ain
5	84	1	6	Ambléon	118	1006		Auvergne-Rhône-Alpes	Ain
6	84	1	7	Ambronay	2915	1007		Auvergne-Rhône-Alpes	Ain

VENTE

Projet 5 Table name: Vente ☐ WITHOUT ROWID ☐ STRICT

	Name	Data type	Primary Key	Foreign Key	Unique	Check	Not NULL	Collate	Generated
1	Id_vente	INTEGER							
2	Id_bien	INTEGER							
3	Date_mutation	DATE							
4	Valeur_fonciere	INTEGER							



Grid view Form view

Total rows loaded: 34169

	Id_vente	Id_bien	Date_mutation	Valeur_fonciere
1	1	1	02/01/2020	165000
2	2	2	02/01/2020	355680
3	3	3	02/01/2020	229500

7.LES REQUÊTES PAR SQL



LAPLACE IMMO.

1. NOMBRE TOTAL D'APPARTEMENTS VENDUS AU 1ER SEMESTRE 2020.

```
1 --1 Nombre total d'appartements vendus au 1er semestre 2020
2 SELECT COUNT(*) AS Total_appartements_vendus
3 FROM Vente
4 JOIN Bien ON Vente.Id_bien = Bien.Id_bien
5 WHERE Bien.Type_local = 'Appartement';
```

	Total appartements vendus
1	31378

2. LE NOMBRE DE VENTES D'APPARTEMENT PAR RÉGION POUR LE 1ER SEMESTRE 2020.

```
7 --2 Le nombre de ventes d'appartement par région pour le 1er semestre 2020
8 SELECT Communes.Nom_reg, COUNT(Vente.Id_vente) AS Nombre_de_vente
9 FROM Vente
10 JOIN Bien ON Bien.Id_bien= Vente.Id_bien
11 JOIN Communes ON Bien.Id_codedep_codecommune= Communes.Id_codedep_codecommune
12 WHERE Bien.Type_local='Appartement'
13 GROUP BY Communes.Nom_reg
14 ORDER BY nombre_de_vente;
```

	Nom reg	nombre de vente
1	Guadeloupe	2
2	Guyane	34
3	La Réunion	44
4	Martinique	94
5	Corse	223
6	Bourgogne-Franche-Comté	376
7	Centre-Val de Loire	696
8	Normandie	862
9	Bretagne	983
10	Grand Est	984
11	Hauts-de-France	1254
12	Pays de la Loire	1357
13	Occitanie	1640
14	Nouvelle-Aquitaine	1932
15	Auvergne-Rhône-Alpes	3253
16	Provence-Alpes-Côte d'Azur	3649
17	Ile-de-France	13995

7.LES REQUÊTES PAR SQL



LAPLACE IMMO.

3. PROPORTION DES VENTES D'APPARTEMENTS PAR LE NOMBRE DE PIÈCES.

```
16 --3 Proportion des ventes d'appartement par le nombre de pièce
17 SELECT Bien.Total_piece, ROUND (COUNT(*) * 100.0 / |
18     (SELECT COUNT(*)
19     FROM Vente AS V
20     JOIN Bien AS B ON V.Id_bien = B.Id_bien
21     WHERE B.Type_local = 'Appartement'
22     ),2) AS Proportion_des_ventes
23 FROM Vente
24 JOIN Bien ON Vente.Id_bien = Bien.Id_bien
25 WHERE Bien.Type_local = 'Appartement'
26 GROUP BY Bien.Total_piece
27 ORDER BY Bien.Total_piece;
```



	Total piece	Proportion des ventes
1	0	0.1
2	1	21.48
3	2	31.18
4	3	28.57
5	4	14.21
6	5	3.55
7	6	0.65
8	7	0.17
9	8	0.05
10	9	0.03
11	10	0.01
12	11	0

4. LISTE DES 10 DÉPARTEMENTS OÙ LE PRIX DU MÈTRE CARRÉ EST LE PLUS ÉLEVÉ

```
--4 Liste 10 départements où le prix du mètre carré est le plus élevé
SELECT Communes.Nom_dep AS Nom_departement,
       ROUND((Vente.Valeur_fonciere/Bien.Surface_carrez),2) AS Prix_metre_carré
FROM Vente
JOIN Bien ON Vente.Id_bien= Bien.Id_bien
JOIN Communes ON Bien.Id_codedep_codecommune= Communes.Id_codedep_codecommune
GROUP BY Communes.Nom_dep
ORDER BY prix_metre_carré DESC
LIMIT 10;
```



	Nom departement	Prix metre carré
1	Paris	11900
2	Alpes-Maritimes	9120
3	Hérault	8743
4	Savoie	8000
5	Ille-et-Vilaine	6428
6	Hautes-Alpes	5535
7	Hauts-de-Seine	5464
8	Yvelines	5314
9	Corse-du-Sud	5000
10	La Réunion	4801

7. LES REQUÊTES PAR SQL



LAPLACE IMMO.

5. PRIX MOYEN DU MÈTRE CARRÉ D'UNE MAISON EN ÎLE-DE-FRANCE.

```
39 --5 Prix moyen du mètre carrée d'une maison en île de France
40 SELECT ROUND(AVG(Vente.Valeur_fonciere/ Bien.Surface_carrez),2)AS Moyen_mètre_carrée_en_île_de_France
41 FROM Bien
42 JOIN Vente ON Bien.Id_bien= Vente.Id_bien
43 JOIN Communes ON Bien.Id_codedep_codecommune= Communes.Id_codedep_codecommune
44 WHERE Bien.Type_local='Maison' AND Communes.Nom_reg= 'Ile-de-France';
```

Moyen mètre carrée en île de France
3764.39

6. LISTE DES 10 APPARTEMENTS LES PLUS CHERS AVEC LA RÉGION ET LE NOMBRE DE MÈTRES CARRÉS.

```
46 --6 Liste des 10 appartements les plus chers avec la région et le nombre de mètres carrés
47 SELECT Communes.Nom_reg,
48        Bien.Surface_carrez,
49        Vente.Valeur_fonciere
50 FROM Vente
51 JOIN Bien ON Bien.Id_bien= Vente.Id_bien
52 JOIN Communes ON Bien.Id_codedep_codecommune= Communes.Id_codedep_codecommune
53 WHERE Bien.Type_local='Appartement'
54 ORDER BY Vente.Valeur_fonciere DESC
55 LIMIT 10;
```

	Nom reg	Surface carrez	Valeur fonciere
1	Occitanie	105,77	99109,61
2	Pays de la Loire	45,43	98762,59
3	Ile-de-France	54	97872,14
4	Auvergne-Rhône-Alpes	132,8	965159,4
5	Ile-de-France	28,24	95886,25
6	Bretagne	40,93	92423,43
7	Ile-de-France	66,03	91791,25
8	Centre-Val de Loire	58,43	89937,86
9	Auvergne-Rhône-Alpes	28,46	87132,2
10	Ile-de-France	20,7	86735,85

7. TAUX D'ÉVOLUTION DU NOMBRE DE VENTES ENTRE LE PREMIER ET LE SECOND TRIMESTRE DE 2020.

```
57 --7 Taux d'évolution du nombre de ventes entre le premier et le second trimestre de 2020
58 SELECT ROUND(((SELECT COUNT(Vente.Id_vente)
59                FROM Vente
60                WHERE Vente.Date_mutation BETWEEN '01/04/2020' AND '31/06/2020')-
61                (SELECT COUNT(Vente.Id_vente)
62                FROM Vente
63                WHERE Vente.Date_mutation BETWEEN '01/01/2020' AND '31/03/2020')) /
64                (SELECT COUNT(Vente.Id_vente)
65                FROM Vente
66                WHERE Vente.Date_mutation BETWEEN '01/01/2020' AND '31/03/2020') * 100,4)AS Taux_evolution;
```

Taux evolution
1 0

7. LES REQUÊTES PAR SQL



LAPLACE IMMO.

8. LE CLASSEMENT DES RÉGIONS PAR RAPPORT AU PRIX AU MÈTRE CARRÉ DES APPARTEMENT DE PLUS DE 4 PIÈCES.


```
--8 Le classement des régions par rapport au prix au mètre carré des appartement de plus de 4 pièce
SELECT Communes.Nom_reg AS Nom_region, ROUND(Vente.Valeur_fonciere/Surface_carrez) AS Mètre_carrez
FROM Vente
LEFT JOIN Bien ON Bien.Id_bien = Vente.Id_bien
LEFT JOIN Communes ON Bien.Id_codedep_codecommune = Communes.Id_codedep_codecommune
WHERE Bien.Type_local='Appartement'
AND Bien.Total_piece >4
GROUP BY Communes.Nom_reg|
ORDER BY Mètre_carrez DESC;
```



	Nom region	Mètre carrez
1	La Réunion	4801
2	Ile-de-France	4510
3	Corse	4198
4	Occitanie	3410
5	Nouvelle-Aquitaine	3103
6	Grand Est	2629
7	Normandie	2600
8	Provence-Alpes-Côte d'Azur	2530
9	Centre-Val de Loire	1860
10	Hauts-de-France	1286
11	Bretagne	1285
12	Bourgogne-Franche-Comté	1250
13	Pays de la Loire	893
14	Auvergne-Rhône-Alpes	776
15	Martinique	574

9. LISTE DES COMMUNES AYANT EU AU MOINS 50 VENTES AU 1ER TRIMESTRE

```
--9 Liste des communes ayant eu au moins 50 ventes au 1er trimestre
SELECT Communes.Nom_com AS Communes, COUNT (Bien.Id_bien) AS Nombre_vente
FROM Bien
JOIN Vente ON Bien.Id_bien= Vente.Id_bien
JOIN Communes ON Bien.Id_codedep_codecommune= Communes.Id_codedep_codecommune
WHERE Vente.Date_mutation >='01/01/2020' AND Vente.Date_mutation <='31/03/2020'
GROUP BY Communes.Nom_com
HAVING COUNT (Bien.Id_bien) >= 50
ORDER BY Nombre_vente DESC;
```



Grid view		Form view
		Total rows loaded: 152
	Communes	Nombre vente
1	Paris 18e Arrondissement	516
2	Paris 17e Arrondissement	470
3	Paris 15e Arrondissement	407
4	Paris 16e Arrondissement	394
5	Nice	393
6	Paris 11e Arrondissement	383
7	Paris 20e Arrondissement	303
8	Bordeaux	302
9	Nantes	299
10	Paris 19e Arrondissement	267
11	Paris 10e Arrondissement	264
12	Paris 12e Arrondissement	254
13	Grenoble	226
14	Paris 14e Arrondissement	222
15	Paris 9e Arrondissement	208
16	Toulouse	188

7. LES REQUÊTES PAR SQL

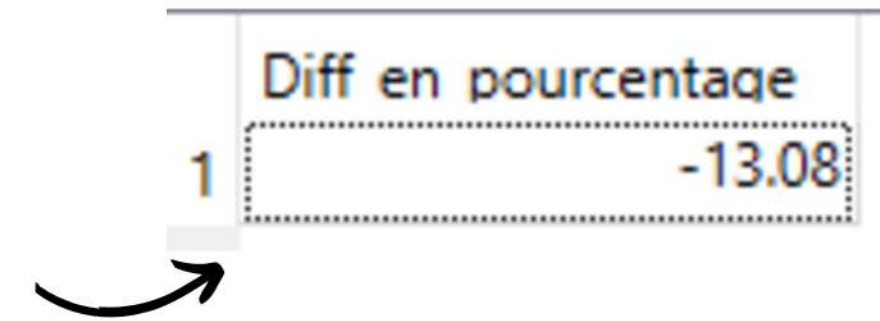


LAPLACE IMMO.

10. DIFFÉRENCE EN POURCENTAGE DU PRIX AU MÈTRE CARRÉ ENTRE UN APPARTEMENT DE 2 PIÈCES ET UN APPARTEMENT DE 3 PIÈCES.

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

```
SELECT ROUND(((SELECT AVG(Vente.Valeur_fonciere / Bien.Surface_carrez)
FROM Vente
JOIN Bien ON Vente.Id_bien = Bien.Id_bien
WHERE Bien.Total_piece = 3) -
(SELECT AVG(Vente.Valeur_fonciere / Bien.Surface_carrez)
FROM Vente
JOIN Bien ON Vente.Id_bien = Bien.Id_bien|
WHERE Bien.Total_piece = 2)) /
(SELECT AVG(Vente.Valeur_fonciere/ Bien.Surface_carrez)
FROM Vente
JOIN Bien ON Vente.Id_bien = Bien.Id_bien
WHERE Bien.Total_piece = 2) * 100, 2) AS Diff_en_pourcentage;
```



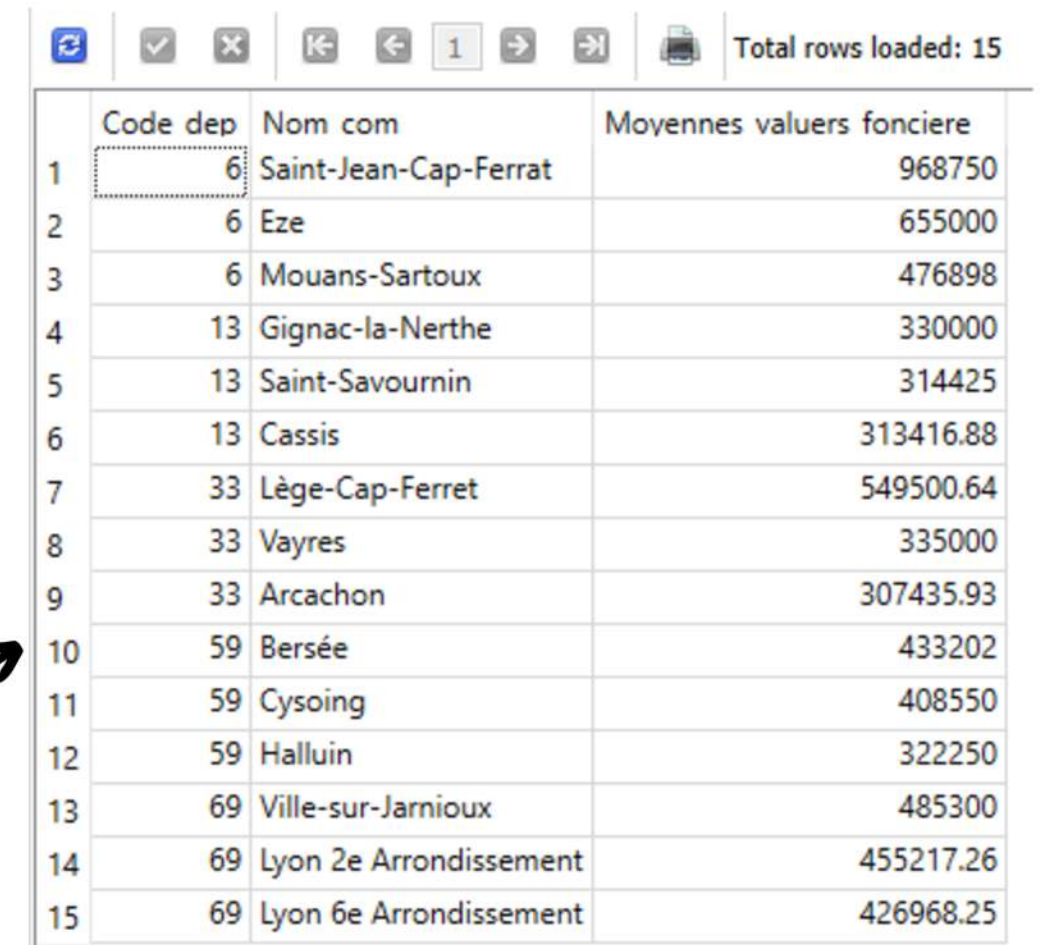
	Diff en pourcentage
1	-13.08

11. LES MOYENNES DE VALEURS FONCIÈRES POUR LE TOP 3 DES COMMUNES DES DÉPARTEMENTS 6, 13, 33, 59 ET 69.

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

```
WITH TOP3 AS (SELECT Communes.Code_dep,Communes.Nom_com, ROUND(AVG(Vente.Valeur_fonciere),2) AS Moyennes_valuers_fonciere,
RANK()OVER ( PARTITION BY Code_dep ORDER BY AVG(Vente.Valeur_fonciere) DESC) AS Valuer_rank
FROM Vente
JOIN Bien ON Vente.Id_bien= Bien.Id_bien
JOIN Communes ON Bien.Id_codedep_codecommune= Communes.Id_codedep_codecommune
WHERE Communes.Code_dep IN (6,13,33,59,69)
GROUP BY Communes.Nom_com
ORDER BY moyennes_valuers_fonciere DESC)

SELECT Code_dep,Nom_com,Moyennes_valuers_fonciere
FROM TOP3|
WHERE Valuer_rank <=3
ORDER BY Code_dep;
```



	Code dep	Nom com	Moyennes valuers fonciere
1	6	Saint-Jean-Cap-Ferrat	968750
2	6	Eze	655000
3	6	Mouans-Sartoux	476898
4	13	Gignac-la-Nerthe	330000
5	13	Saint-Savournin	314425
6	13	Cassis	313416.88
7	33	Lège-Cap-Ferret	549500.64
8	33	Vayres	335000
9	33	Arcachon	307435.93
10	59	Bersée	433202
11	59	Cysoing	408550
12	59	Halluin	322250
13	69	Ville-sur-Jarnioux	485300
14	69	Lyon 2e Arrondissement	455217.26
15	69	Lyon 6e Arrondissement	426968.25

7.LES REQUÊTES PAR SQL




LAPLACE IMMO.

12. LES 20 COMMUNES AVEC LE PLUS DE TRANSACTIONS POUR 1000 HABITANTS POUR LES COMMUNES QUI DÉPASSENT LES 10 000 HABITANTS

--12 Les 20 communes avec le plus de transactions pour 1000 habitants pour les communes qui dépassent les 10000 habitants

```
SELECT Communes.Nom_com,  
       Communes.Population,  
       COUNT (Bien.Id_bien) AS Nombre_transactions,  
       ROUND(COUNT (Bien.Id_bien)/(Communes.Population/1000),2) AS Transactions_pour_1000  
FROM Communes  
JOIN Bien ON Bien.Id_codedep_codecommune= Communes.Id_codedep_codecommune  
WHERE Communes.Population >10000  
GROUP BY Communes.Nom_com  
ORDER BY Transactions_pour_1000 DESC  
LIMIT 20;
```



	Nom com	Population	Nombre transactions	Transactions pour 1000
1	Paris 2e Arrondissement	21735	127	6
2	Arcachon	11898	55	5
3	Roquebrune-Cap-Martin	13041	52	4
4	Paris 4e Arrondissement	29390	120	4
5	Paris 3e Arrondissement	34306	161	4
6	Paris 1er Arrondissement	16055	79	4
7	La Baule-Escoublac	16797	77	4
8	Sanary-sur-Mer	17160	60	3
9	Saint-Mandé	22576	69	3
10	Saint-Hilaire-de-Riez	11501	33	3
11	Saint-Cyr-sur-Mer	11725	38	3
12	Pornichet	11440	35	3
13	Paris 9e Arrondissement	60563	208	3
14	Paris 8e Arrondissement	36250	139	3
15	Paris 6e Arrondissement	41171	139	3
16	Paris 10e Arrondissement	86863	264	3
17	Menton	30981	91	3
18	La Londe-les-Maures	10776	37	3
19	Chantilly	11178	35	3
20	Vincennes	50230	141	2



LAPLACE IMMO.

MERCI POUR VOTRE ATTENTION