



Création et utilisation de la base de données

Stéphane JAILLY



Laplace Immo

Contexte du projet

- Le groupe Laplace Immo souhaite développer un **modèle** pour se démarquer de la concurrence
- Projet dénommé "**DATAimmo**"
- Modification de la **base de données** relative aux transactions immobilières et foncières avec **analyses du marché**



La stratégie de sauvegarde et la conformité RGPD

- Dans le fichier valeur-foncières.xlsx, la colonne « **Nom de l'acquéreur** » indique des **données à caractère personnel** :
 - **Cette donnée n'a pas été intégrée dans la base de données**
 - L'adresse détaillée a été conservée, car disponible dans les différents référentiels de l'état relatifs aux données cadastrales
- Les données utilisées respectent le RGPD



Les données initiales

- Les données utilisées pour la création de cette base de données proviennent du data.gouv.fr et etalab.fr
- Données disponibles :
 - ✓ **Référentiel géographique** qui contient toutes les informations géographiques et administratives relatives aux communes
 - ✓ **Demande de valeurs foncières 2020** : mutations immobilières et foncières, qui depuis 2018 sont mises à la disposition du public



L'extrait du dictionnaire des données

DICTIONNAIRE DES DONNÉES - Valeurs foncières

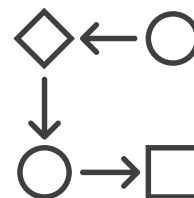
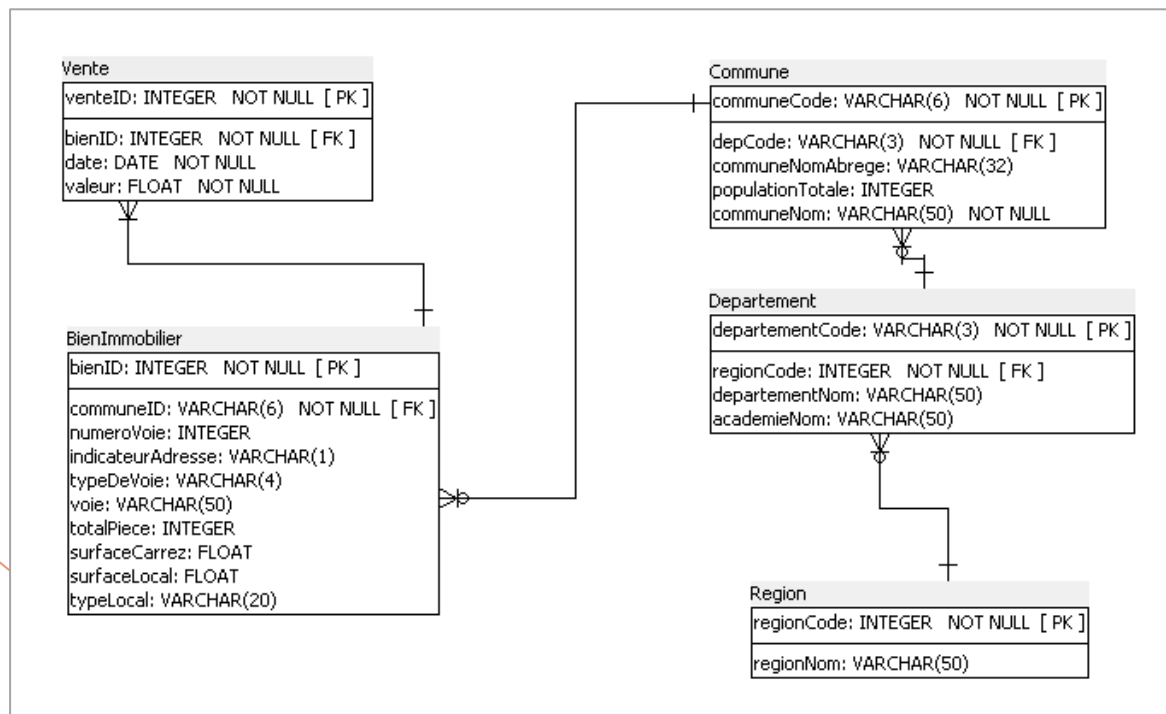
CODE	SIGNIFICATION	TYPE	LONGUEUR	NATURE	REGLE DE GESTION	REGLE DE CALCUL
Id_bien	ID dans la base de données	Integer	NC	Concaténer	Ne doit pas être nul	Auto incrémenter
No_voie	Numéro des rues	Integer	NC	Elémentaire		
Type_de_voie	Plusieurs valeurs (rue, avenue, chemin, etc.)	Varchar	4	Elémentaire	Rue, Bd,...	
Voie	Nom de la rue	Varchar		Elémentaire		
Code postal	Code postal	Integer	NC	Elémentaire		
Commune	Libellé de la commune	Varchar	50	Elémentaire		
Code departement	Code departement	Varchar	3	Elémentaire	Ne doit pas être nul	
Code commune	Code commune	Integer	NC	Elémentaire	Ne doit pas être nul	
Préfixe de section	Préfixe de section	Integer	NC	Elémentaire	Ne doit pas être nul	

DICTIONNAIRE DES DONNÉES - Données communes

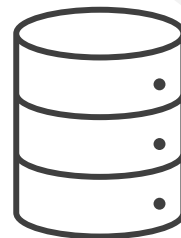
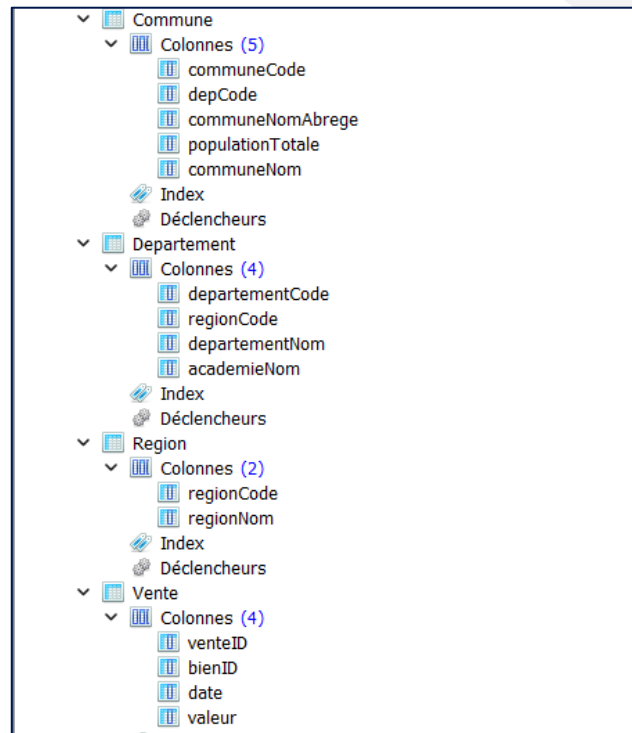
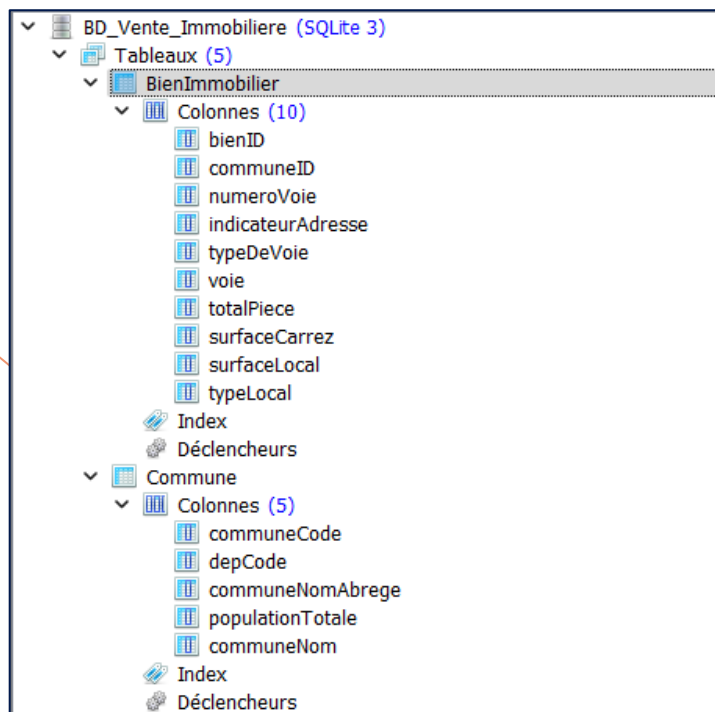
CODE	SIGNIFICATION	TYPE	LONGUEUR	NATURE	REGLE DE GESTION	REGLE DE CALCUL
CODREG	Code de la région administrative	integer		élémentaire	Non Null	
CODDEP	Code du département	CHAR	3	élémentaire	Non Null	
CODARR	Code de l'arrondissement	integer		élémentaire		
CODCAN	Code du canton	integer		élémentaire		
CODCOM	Code commune	integer		élémentaire	Non Null	
COM	Nom de la commune	VARCHAR	50	élémentaire		
PMUN	Population municipale en habitants	integer		élémentaire		
PCAP	Population comptée à part	integer		élémentaire		
PTOT	Population totale	integer		calculé		PCAP + PMUN = PTOT



Le schéma relationnel normalisé



La base de données avec les tables créées et les données chargées



Les screenshots qui permettent de démontrer le bon chargement des données

Commune (BD_Vente_Immobiliere)

Structure Données Contraintes Index Déclencheurs DDL

Table Formulaire

Filtre de donn... Nombre de lignes chargées : 38916

	communeCode	depCode	communeNomAbrege	populationTotale	communeNom
1	01001	1	L ABERGEMENT CLEMENCIAT	798	L'ABERGEMENT-CLEMENCIAT
2	01002	1	L ABERGEMENT DE VAREY	257	L'ABERGEMENT-DE-VAREY
3	01003	1	AMAREINS		AMAREINS
4	01004	1	AMBERIEU EN BUGEY	14514	AMBERIEU-EN-BUGEY
5	01005	1	AMBERIEUX EN DOMBES	1776	AMBERIEUX-EN-DOMBES
6	01006	1	AMBLEON	118	AMBLEON
7	01007	1	AMBRONAY	2915	AMBRONAY

Departement (BD_Vente_Immobiliere)

Structure Données Contraintes Index Déclencheurs DDL

Table Formulaire

Filtre de donn... Nombre de lignes chargées : 109

	departementCode	regionCode	departementNom	academieNom
1	1	84	Ain	Lyon
2	2	32	Aisne	Amiens
3	3	84	Allier	Clermont-Ferrand
4	4	93	Alpes-de-Haute-Provence	Aix-Marseille
5	5	93	Hautes-Alpes	Aix-Marseille
6	6	93	Alpes-Maritimes	Nice
7	7	84	Ardèche	Grenoble

Region (BD_Vente_Immobiliere)

Structure Données Contraintes Index Déclencheurs DDL

Table Formulaire

Filtre de donn... Nombre de lignes chargées : 19

	regionCode	regionNom
1	0	Collectivités d'outre-mer
2	1	Guadeloupe
3	2	Martinique
4	3	Guyane
5	4	La Réunion
6	6	Mayotte
7	11	Ile-de-France

BienImmobilier (BD_Vente_Immobiliere)

Structure Données Contraintes Index Déclencheurs DDL

Table Formulaire

Filtre de donn... Nombre de lignes chargées : 34169

	bienID	commune	numeroV	indicateur	typeDeVc	voie	totalPiec	surfaceC	surfaceL	typeLocal
1	B1	01103	347		RUE	DU CHATEAU	3	48,22		Appartement
2	B2	06004	4		BD	EDOUARD BAUDOUIN	1	39,11		Appartement
3	B3	06088	20	B	RUE	MARCEAU	3	80,25		Appartement
4	B4	06123	550		RTE	DES VESPINS RN7	1	27,51		Appartement
5	B5	13005	9300		RES	LES ARPEGES BD DES ABA	2	47,33		Appartement
6	B6	13028	27		RUE	DU GRAND MADIER	1	25,31		Appartement
7	B7	13208	360		AV	DU PRADO	3	70,84		Appartement
8	B8	13212	5076	F	PARC	DESSUARD	3	67,19		Appartement
9	B9	14338	1194		RUE	DE NORMANDIE	1	18,89		Appartement
10	B10	14366	30		ALL	DES NOISETIERS	4	105,37		Maison
11	B11	17300	11		RUE	ROUGET DE L ISLE	2	31,99		Appartement
12	B12	25056	13		RUE	BERTHE MORISOT	5	96,21		Maison
13	B13	29232	1		RUE	DU POHER	1	30,86		Appartement
14	B14	29260	2		RUE	DES JARDINS	3	66,21		Appartement

Vente (BD_Vente_Immobiliere)

Structure Données Contraintes Index Déclencheurs DDL

Table Formulaire

Filtre de donn... Nombre de lignes chargées : 34169

	venteID	bienID	date	valeur
1	1	B1	2020-01-02	165000
2	2	B2	2020-01-02	355680
3	3	B3	2020-01-02	229500
4	4	B4	2020-01-02	125000
5	5	B5	2020-01-02	90000
6	6	B6	2020-01-02	93000
7	7	B7	2020-01-02	298100
8	8	B8	2020-01-02	163500
9	9	B9	2020-01-02	53000
10	10	B10	2020-01-02	136000
11	11	B11	2020-01-02	125900
12	12	B12	2020-01-02	234000
13	13	B13	2020-01-02	46210
14	14	B14	2020-01-02	129000





Requêtes SQL et résultats

Analyse du marché

Requête 1

Nombre total d'appartements vendus au 1er semestre 2020

```
1 SELECT COUNT(v.venteid) AS nbVenteAppartementSemestre1
2
3 FROM Vente as v
4 JOIN BienImmobilier as b ON v.bienID = b.bienID
5
6 WHERE
7     v.date BETWEEN "2020-01-01" AND "2020-06-30"
8     AND b.typeLocal="Appartement"
```



	nbVenteAppartementSemestre1
1	31378

Requête 2

Nombres de ventes d'appartements / région - 1er semestre 2020

```
1 SELECT
2     r.regionNom AS region,
3     COUNT(v.venteid) AS nbVentes
4
5 FROM Vente as v
6 JOIN BienImmobilier as b ON v.bienID = b.bienID
7 JOIN Commune AS c ON c.communeCode = b.communeID
8 JOIN Departement AS d ON d.departementCode = c.depCode
9 JOIN Region AS r ON r.regionCode = d.regionCode
10
11 WHERE
12     v.date BETWEEN "2020-01-01" AND "2020-06-30"
13     AND b.typeLocal="Appartement"
14
15 GROUP BY r.regionCode, r.regionNom
16 ORDER BY nbVentes DESC
```

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



Requête 3

Proportion des ventes d'appartements par le nombre de pièces

```
1 WITH
2   Table1 AS (
3       SELECT * FROM Vente AS v
4       JOIN BienImmobilier AS b ON b.bienID = v.bienID
5       WHERE b.typeLocal="Appartement")
6
7 SELECT
8   t.totalPiece as nbPiece,
9   ROUND(CAST(COUNT(t.venteID) as FLOAT) *100 /
10  (SELECT COUNT(venteID) FROM Table1),2) as
11   proportionVente
12 FROM Table1 AS t
13 GROUP BY t.totalPiece
```

	nbPiece	proportionVente
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



Requête 4

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

```
1 SELECT
2     d.departementNom,
3     ROUND(AVG(v.valeur / b.surfaceCarrez),0) AS prixMetreCarre
4
5 FROM Vente AS v
6 JOIN BienImmobilier AS b ON b.bienID = v.bienID
7 JOIN Commune AS c ON c.communeCode = b.communeID
8 JOIN Departement AS d ON d.departementCode = c.depCode
9
10 GROUP BY d.departementCode, d.departementNom
11 ORDER BY prixMetreCarre DESC
12
13 LIMIT 10
```

	departementNom	prixMetreCarre
1	Paris	12084
2	Hauts-de-Seine	7301
3	Val-de-Marne	5428
4	Haute-Savoie	4781
5	Alpes-Maritimes	4756
6	Seine-Saint-Denis	4386
7	Yvelines	4276
8	Rhône	4100
9	Corse-du-Sud	4063
10	Gironde	3807

Requête 5

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

```
1 SELECT round(AVG(v.valeur / b.surfaceCarrez),0)
2       AS prixMoyenMetreCarreIDF
3
4 FROM Vente AS v
5 JOIN BienImmobilier AS b ON b.bienID = v.bienID
6 JOIN Commune AS c ON c.communeCode = b.communeID
7 JOIN Departement AS d ON d.departementCode = c.depCode
8 JOIN Region AS r ON r.regionCode = d.regionCode
9
10 WHERE    r.regionCode=11
11        AND b.typeLocal = "Maison"
```

	prixMoyenMetreCarreIDF
1	3765



Requête 6

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

```
1 SELECT
2     v.bienID,
3     r.regionNom,
4     round(v.valeur,0) AS prixVente,
5     cast(b.surfaceCarrez AS INTEGER) AS surface
6
7 FROM Vente AS v
8 JOIN BienImmobilier AS b ON b.bienID = v.bienID
9 JOIN Commune AS c ON c.communeCode = b.communeID
10 JOIN Departement AS d ON d.departementCode = c.depCode
11 JOIN Region AS r ON r.regionCode = d.regionCode
12 WHERE b.typeLocal="Appartement"
13 ORDER BY prixVente DESC
14 LIMIT 10
```

	bienID	regionNom	prixVente	surface
1	B30603	Ile-de-France	9000000.0	9
2	B5261	Ile-de-France	8600000	64
3	B3625	Ile-de-France	8577713	20
4	B7602	Ile-de-France	7620000	42
5	B9988	Ile-de-France	7600000	253
6	B17823	Ile-de-France	7535000	139
7	B410	Ile-de-France	7420000	360
8	B16357	Ile-de-France	7200000	595
9	B1924	Ile-de-France	7050000	122
10	B19161	Ile-de-France	6600000	79



Requête 7

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

```
1 WITH
2 table1 AS (
3     SELECT CAST(COUNT(venteID) AS FLOAT) as venteT1
4     FROM Vente
5     WHERE date BETWEEN "2020-01-01" AND "2020-03-31"),
6
7 table2 AS (
8     SELECT CAST(COUNT(venteID) AS FLOAT) as venteT2
9     FROM Vente
10    WHERE date BETWEEN "2020-04-01" AND "2020-06-30")
11
12 SELECT round((venteT2 - venteT1)*100/venteT1,2) AS tauxEvolutionT1T2
13 FROM
14     table1,
15     table2
```

tauxEvolutionT1T2
3.68



Requête 8

Classement des régions par rapport au prix au mètre carré des appartement de plus de 4 pièces

```
1 SELECT
2   r.regionNom,
3   round(AVG(v.valeur / b.surfaceCarrez),0) AS prixMetreCarre
4 FROM Vente AS v
5 JOIN BienImmobilier AS b ON b.bienID = v.bienID
6 JOIN Commune AS c ON c.communeCode = b.communeID
7 JOIN Departement AS d ON d.departementCode = c.depCode
8 JOIN Region AS r ON r.regionCode = d.regionCode
9 WHERE
10    b.typeLocal="Appartement"
11    AND b.totalPiece > 4
12 GROUP BY r.regionCode, r.regionNom
13 ORDER BY prixMetreCarre DESC
```

	regionNom	prixMetreCarre
1	Ile-de-France	8807
2	La Réunion	3660
3	Provence-Alpes-Côte d'Azur	3617
4	Corse	3118
5	Auvergne-Rhône-Alpes	2904
6	Nouvelle-Aquitaine	2477
7	Bretagne	2427
8	Pays de la Loire	2329
9	Hauts-de-France	2200
10	Occitanie	2107
11	Normandie	2026
12	Grand Est	1561
13	Centre-Val de Loire	1460
14	Bourgogne-Franche-Comté	1261
15	Martinique	575

Requête 9

Liste des communes ayant eu au moins 50 ventes au 1er trimestre

```
1 SELECT
2     c.communeNom,
3     COUNT(v.venteId) AS nbVentes
4 FROM Vente as v
5 JOIN BienImmobilier as b ON v.bienID = b.bienID
6 JOIN Commune AS c ON c.communeCode = b.communeID
7 WHERE
8     v.date BETWEEN "2020-01-01" AND "2020-03-31"
9 GROUP BY c.communeCode, c.communeNom
10 HAVING nbVentes >= 50
11 ORDER BY nbVentes DESC
```

	communeNom	nbVentes
1	PARIS 17E ARRONDISSEMENT	228
2	PARIS 15E ARRONDISSEMENT	215
3	PARIS 18E ARRONDISSEMENT	209
4	NICE	173
5	PARIS 11E ARRONDISSEMENT	169
6	PARIS 16E ARRONDISSEMENT	165
7	BORDEAUX	157
8	PARIS 14E ARRONDISSEMENT	146
9	PARIS 20E ARRONDISSEMENT	127
10	NANTES	119
11	PARIS 19E ARRONDISSEMENT	116
12	PARIS 12E ARRONDISSEMENT	110
13	PARIS 10E ARRONDISSEMENT	109
14	GRENOBLE	106
15	PARIS 9E ARRONDISSEMENT	106

Requête 10

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

```
1 WITH
2 Table1 AS (
3     SELECT AVG(v.valeur/b.surfaceCarrez) AS prixM2piece2
4     FROM Vente as v
5     JOIN BienImmobilier AS b ON v.bienID = b.bienID
6     WHERE
7         b.typeLocal = "Appartement"
8         AND b.totalPiece = 2),
9
10 Table2 AS (
11     SELECT AVG(v.valeur/b.surfaceCarrez) AS prixM2piece3
12     FROM Vente as v
13     JOIN BienImmobilier AS b ON v.bienID = b.bienID
14     WHERE
15         b.typeLocal = "Appartement"
16         AND b.totalPiece = 3)
17
18 SELECT round((prixM2piece3-prixM2piece2)/prixM2piece2*100,2) AS differencePrix
19 FROM
20     Table1,
21     Table2
```

differencePrix	
1	-12.68

Requête 11

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

```
1 WITH
2 table1 AS(
3     SELECT
4         RANK() OVER(PARTITION BY departementCode ORDER BY ROUND(AVG(v.valeur),0) DESC) AS rank,
5         d.departementCode,
6         c.communeNom,
7         ROUND(AVG(v.valeur),0) AS prixVenteMoyen
8
9     FROM Vente AS v
10    JOIN BienImmobilier AS b ON b.bienID = v.bienID
11    JOIN Commune AS c ON c.communeCode = b.communeID
12    JOIN Departement AS d ON d.departementCode = c.depCode
13
14    WHERE d.departementCode IN (6,13,33,59,69)
15
16    GROUP BY c.communeCode, c.communeNom
17    ORDER BY prixVenteMoyen DESC
18 )
19
20 SELECT *
21 FROM Table1
22 WHERE rank IN (1,2,3)
23 ORDER BY departementCode, rank
```

	rank	departementCode	communeNom	prixVenteMoyen
1	1	13	GIGNAC-LA-NERTHE	330000
2	2	13	SAINT-SAVOURNIN	314425
3	3	13	CASSIS	313417
4	1	33	LEGE-CAP-FERRET	549501
5	2	33	VAYRES	335000
6	3	33	ARCACHON	307436
7	1	59	BERSEE	433202
8	2	59	CYSOING	408550
9	3	59	HALLUIN	322250
10	1	6	SAINT-JEAN-CAP-FERRAT	968750
11	2	6	EZE	655000
12	3	6	MOUANS-SARTOUX	476898
13	1	69	VILLE-SUR-JARNIOUX	485300
14	2	69	LYON 2E ARRONDISSEMENT	455217
15	3	69	LYON 6E ARRONDISSEMENT	426968

Requête 12

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

```
1 WITH
2 Table1 AS (
3     SELECT *
4     FROM Commune
5     WHERE populationTotale > 10000)
6
7 SELECT
8     t1.communeNom,
9     COUNT(v.venteID) AS nbVente,
10    t1.populationTotale,
11    round(CAST(COUNT(v.venteID)*1000 AS FLOAT)/CAST(t1.populationTotale AS
12    FLOAT),2)
13    AS nbVentePour1000Habitant
14 FROM Vente AS v
15 JOIN BienImmobilier AS b ON v.bienID = b.bienID
16 JOIN Table1 AS t1 ON b.communeID = t1.communeCode
17 ORDER BY nbVentePour1000Habitant DESC
18 LIMIT 20
```

	communeNom	nbVente	populationTotale	nbVentePour1000Habitant
1	PARIS 2E ARRONDISSEMENT	127	21735	5.84
2	PARIS 1ER ARRONDISSEMENT	79	16055	4.92
3	PARIS 3E ARRONDISSEMENT	161	34306	4.69
4	ARCACHON	55	11898	4.62
5	LA BAULE-ESCOUBLAC	77	16797	4.58
6	PARIS 4E ARRONDISSEMENT	120	29390	4.08
7	ROQUEBRUNE-CAP-MARTIN	52	13041	3.99
8	PARIS 8E ARRONDISSEMENT	139	36250	3.83
9	SANARY-SUR-MER	60	17160	3.5
10	LA LONDE-LES-MAURES	37	10776	3.43
11	PARIS 9E ARRONDISSEMENT	208	60563	3.43
12	PARIS 6E ARRONDISSEMENT	139	41171	3.38
13	SAINT-CYR-SUR-MER	38	11725	3.24
14	CHANTILLY	35	11178	3.13
15	SAINT-MANDE	69	22576	3.06
16	PORNICHET	35	11440	3.06
17	PARIS 10E ARRONDISSEMENT	264	86863	3.04
18	MENTON	91	30981	2.94
19	SAINT-HILAIRE-DE-RIEZ	33	11501	2.87
20	VINCENNES	141	50230	2.81



Merci !