

Projet DATAimmo



Table des matières

- Dictionnaire des Données et choix effectués
- Modèle Conceptuel de Données (UML)
- Modèle Physique de Données
- Implémentation des Tables
- Interrogation de la Base de Données

Dictionnaire des données conservées



Dictionnaire des données

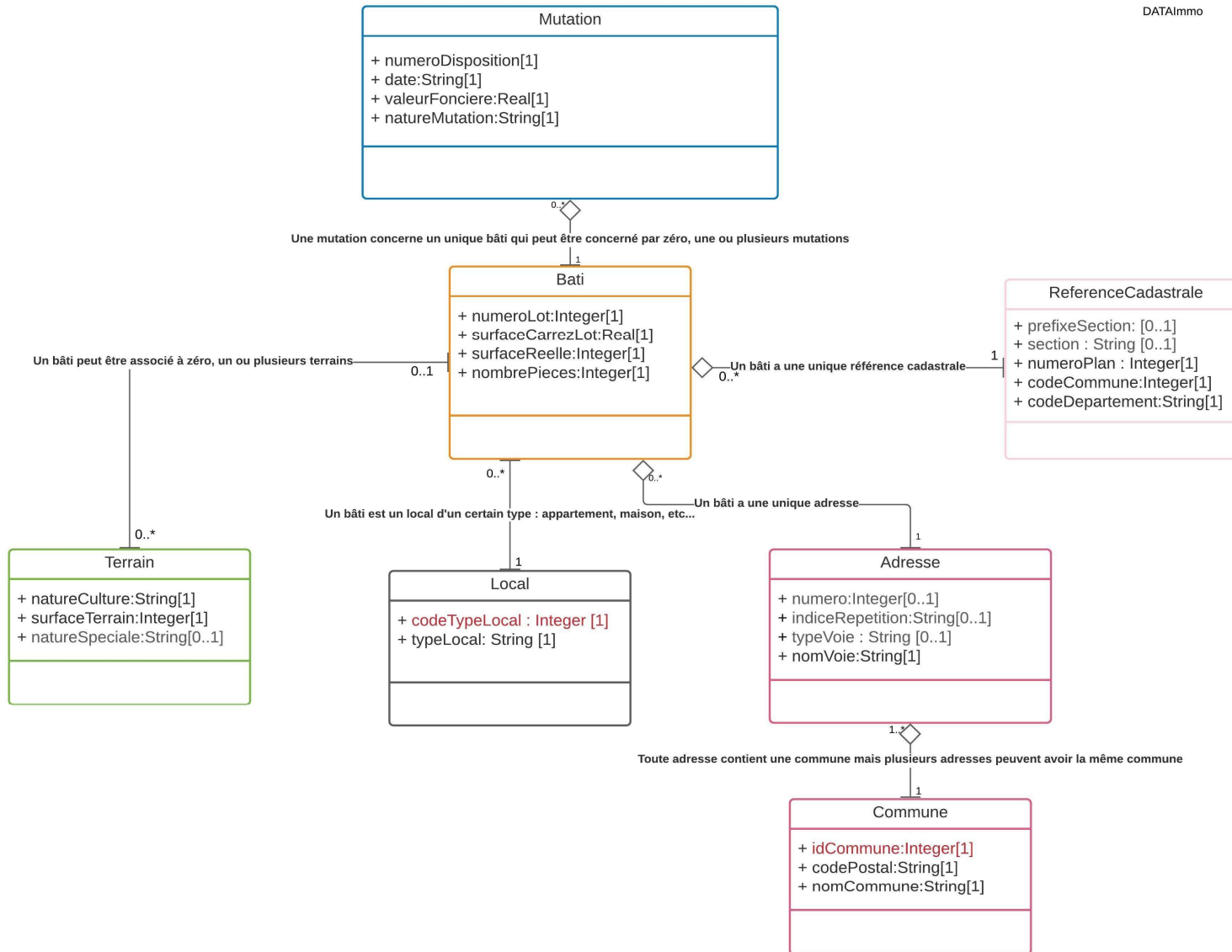
Numéro	Code propriété	Nom dans le fichier de données CSV	Signification/Description	Type	Observation
1	DateMut	Date mutation	Date de mutation - Date de signature de l'acte (au format JJ/MM/AAAA) –	Date	Longueur : 10
2	NatMut	Nature mutation	Nature de la mutation - Il peut s'agir de vente, vente en l'état futur d'achèvement, vente de terrain à bâtir, adjudication, expropriation ou échange	Texte	Longueur : 34
3	ValFonc	Valeur foncière	Valeur foncière - La valeur foncière est le prix net vendeur (TVA non incluse, ni frais de notaire)	Numérique	
4	NumVoie	No voie	Numéro dans la voie - Numéro de voirie (dans la rue/avenue... ou autre voie)	Numérique	Longueur : 4
5	IndRep	B/T/Q	Indice de répétition (B/T/Q/...) - En cas de numéro de voirie redondant, ajout de bis/ter/ etc...	Alphanumérique	Longueur : 1
6	TyVoie	Type de voie	Type de voie - Abbréviation normalisée de la dénomination de la voie (RUE, AV, PL, BD...)	Alphanumérique	Longueur : 4
7	NomVoie	Voie	Nom de la Voie - Nom de la rue, de l'avenue, ou autre voie	Texte	Longueur : 30
8	ColdCom	Code ID commune	Code ID de la commune - Nombre identifiant la commune de manière unique.	Numérique	Identifiant
9	CoPost	Code postal	Code postal - Code Postal	Alphanumérique	Longueur : 5
10	Com	Commune	Commune - Nom de la commune	Texte	Longueur : 50
11	CoDep	Code département	Code du département - [Référence cadastrale de la parcelle] Numéro du département	Alphanumérique	Identifiant
12	CoCom	Code commune	Code de la commune - [Référence cadastrale de la parcelle] En règle générale, il s'agit du code INSEE. Plusieurs communes peuvent avoir le même code.	Numérique	
13	PréfSec	Prefixe de section	Préfixe de section - [Référence cadastrale de la parcelle] : Il contient le code commune des communes absorbées pour les fusions de communes. Pour les autres communes, y compris Paris et Lyon, la zone est à blanc.	Numérique	
14	Sec	Section	Section - [Référence cadastrale de la parcelle] : Elle identifie (par préfixe ou commune) le plan cadastral. Cette zone peut recevoir 1 lettre cadrée à droite ou 2 lettres, ou, en Alsace Moselle, 2 chiffres.	Alphanumérique	Longueur : 2
15	NumPl	No plan	Numéro de plan - [Référence cadastrale de la parcelle] : Il identifie la parcelle dans la section cadastrale	Numérique	
16	Lot	1er lot	Numéro du lot : Un lot de copropriété est constitué d'une partie privative (appartement, cave, etc.) et d'une quote-part de partie commune (tantièmes).	Alphanumérique	Longueur : 10
17	SurfCarLot	Surface Carrez du 1er lot	Surface Loi Carrez du lot : diagnostic immobilier portant sur la surface habitable d'un bien immobilier, c'est-à-dire l'espace qui, au final, est réellement disponible pour les occupants.	Numérique	Décimal
18	CoTyLoc	Code type local	Code du type de local - Identifiant numérique du type de local (cf Type Local)	Numérique	Identifiant
19	TyLoc	Type local	Type de local - 1 : maison ; 2 : appartement ; 3 : dépendance (isolée) ; 4 : local industriel et commercial ou assimilés	Texte	Longueur : 43
20	SurfReBa	Surface réelle bati	Surface réelle du bâti - La surface réelle est la surface mesurée au sol entre les murs ou séparations et arrondie au mètre carré inférieur (hors dépendances).	Numérique	
21	NbPiePri	Nombre pièces principales	Nombre de pièces principales - Hors cuisines, salles d'eau et dépendances.	Numérique	
22	CoNatCult	Nature culture	Code de nature de culture - Nature de culture identifiée par un code (ex : T : Terrain, S : Sol, B bois BF futaies feuillues)	Alphanumérique	
23	CoNatCultSpec	Nature culture spéciale	Code de nature de culture spéciale - Nature de culture spéciale identifiée par un code : ex FLOR Jardin floral FONT Fontaine	Alphanumérique	
24	SurfTer	Surface terrain	Surface du terrain - Contenance du terrain par nature de culture	Numérique	

Choix des données à conserver

- Pour chaque mutation, les n° de disposition, date, valeur foncière et nature de la mutation doivent être renseignés.
- Seuls les biens bâtis sont pris en compte (avec ou sans terrain)
- Une transaction ne concernera qu'un seul bâti lié à un seul lot, afin d'avoir un renseignement exact en terme de surface habitable (surface loi Carrez)
- Un bien bâti peut avoir un seul terrain ou plusieurs types de terrain (ou aucun).
- Les références cadastrales sont conservées.
- L'adresse complète du bien est conservée.
- Seuls les bâtis dont la valeur foncière a été renseignée sont pris en compte.

Modèle Conceptuel de Données

Modélisation avec UML réalisé avec l'outil online Lucidchart : [Lien vers le diagramme online](#)



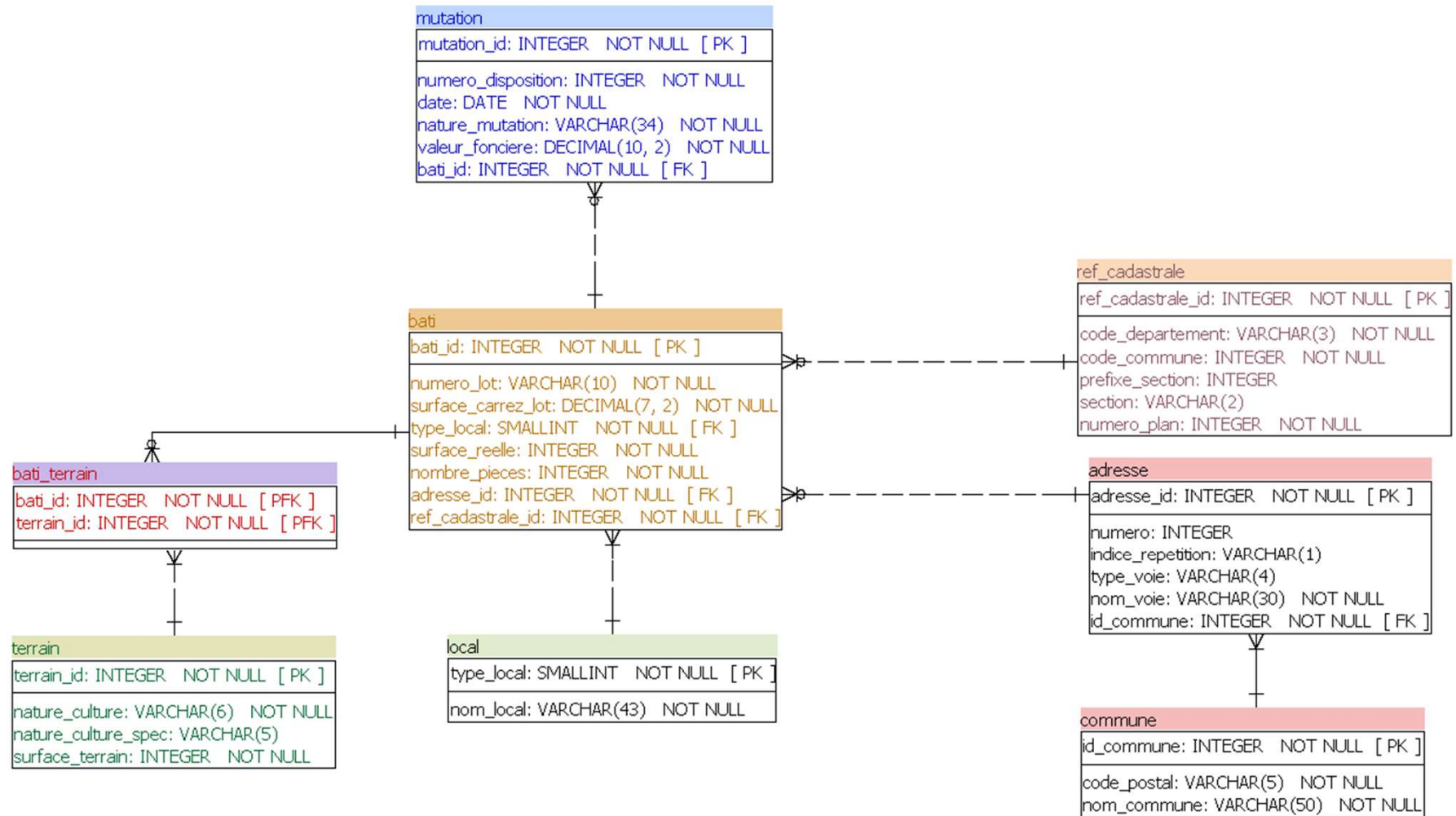
Modèle Physique de Données

Réalisé avec SQL Power Architect



P3_DATAIMMO_MPD.architect



Modèle Physique de Données




Implémentation des tables

Sous IBM Db2 on Cloud


Implémentation des tables


 IBM Db2 on Cloud  Storage: 16% Préférences de cookies Discover

TABLES

Filter by schema name or table name 

Schemas

 Select All

 New implicit schema

☒ NLG66392 8 tables

☐ AUDIT 0 table

☐ DB2INST1 0 table


☐ ERRORSHEMA 0 table

☐ SQL54324 0 table

☐ SQL63196 0 table

☐ ST_INFORMTN_SCHEMA 0 table

Tables

 New table

NAME	SCHEMA	PROPERTIES
<input type="checkbox"/> ADRESSE	NLG66392	...
<input type="checkbox"/> BATI	NLG66392	...
<input type="checkbox"/> BATI_TERRAIN	NLG66392	...
<input type="checkbox"/> COMMUNE	NLG66392	...
<input type="checkbox"/> LOCAL	NLG66392	...
<input type="checkbox"/> MUTATION	NLG66392	...
<input type="checkbox"/> REF_CADASTRALE	NLG66392	...
<input type="checkbox"/> TERRAIN	NLG66392	...

Signed in as:

sophie.piekarec@gmail.com

About

My Profile

Sign out

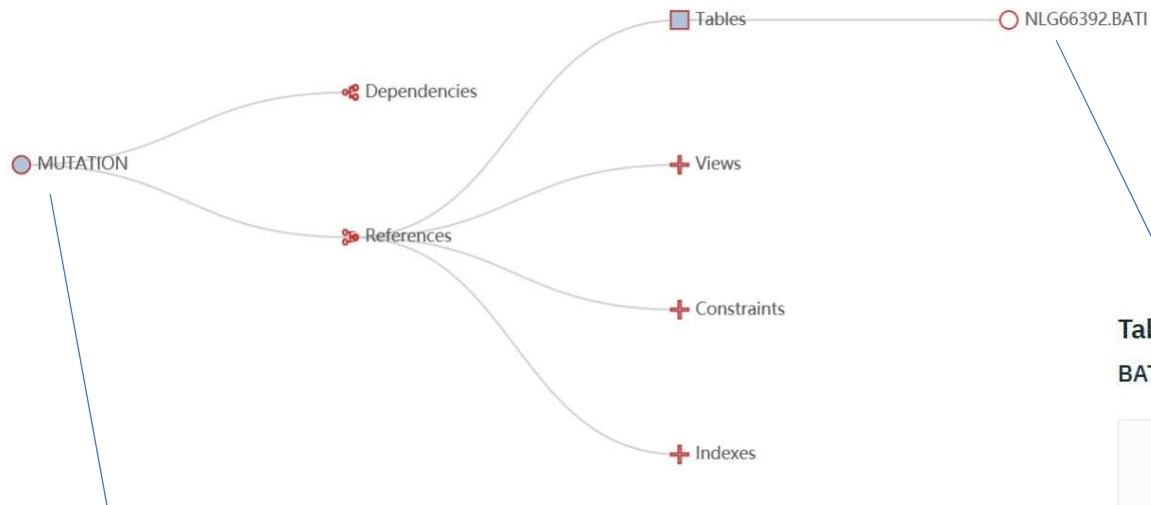


Table Definition

MUTATION

Approximate 34147 rows (1376 KB)

Updated on 2021-06-01 13:34:34

COLUMN NAME ▲	DATA TYPE	NULLABLE	LENGTH	SCALE
MUTATION_ID	INTEGER	N		0
NUMERO_DISPOSIT...	INTEGER	N		0
DATE	DATE	N	4	0
NATURE_MUTATION	VARCHAR	N	34	0
VALEUR_FONCIERE	DECIMAL	N	10	2
BATI_ID	INTEGER	N		0

Table Definition

BATI

Approximate 34150 rows (1408 KB)

Updated on 2021-05-31 11:14:35

COLUMN NAME	DATA TYPE	NULLABLE	LENGTH	SCALE
BATI_ID	INTEGER	N		0
NUMERO_LOT	VARCHAR	N	10	0
SURFACE_CARREZ_LOT	DECIMAL	N	7	2
TYPE_LOCAL	SMALLINT	N		0
SURFACE_REELLE	INTEGER	N		0
NOMBRE_PIECES	INTEGER	N		0
ADRESSE_ID	INTEGER	N		0
REF_CADASTRALE_ID	INTEGER	N		0

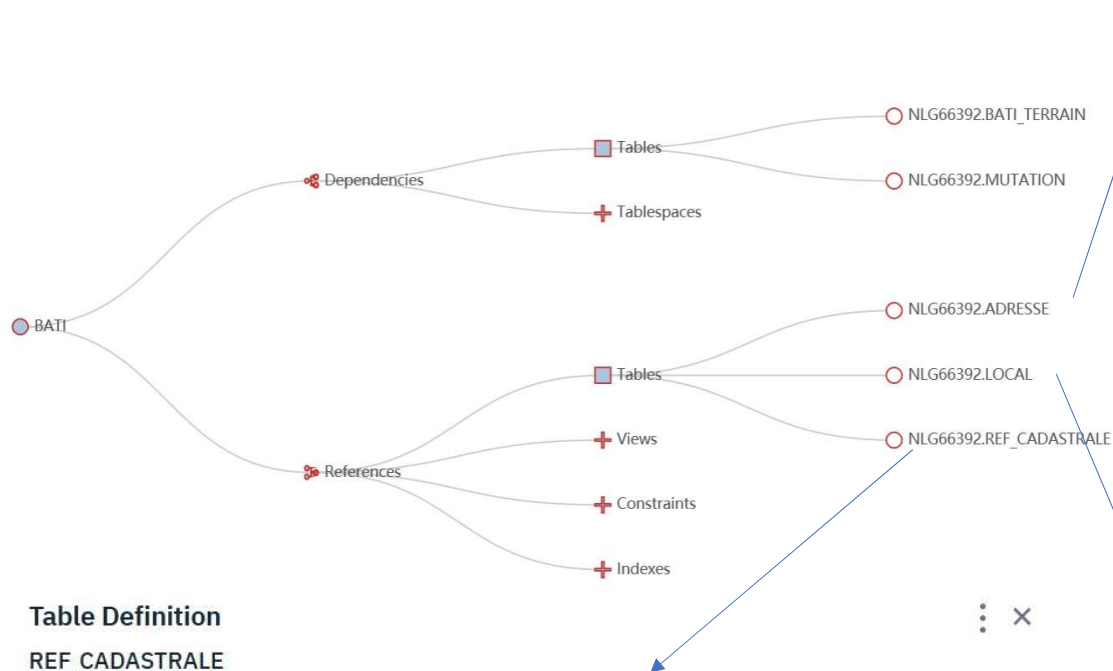


Table Definition

REF_CADASTRALE

No statistics available.

COLUMN NAME	DATA TYPE	NULLABLE	LENGTH	SCALE
REF_CADASTRALE_ID	INTEGER	N		0
CODE_DEPARTEMENT	VARCHAR	N	3	0
CODE_COMMUNE	INTEGER	N		0
PREFIXE_SECTION	INTEGER	Y		0
SECTION	VARCHAR	Y	2	0
NUMERO_PLAN	INTEGER	N		0

Table Definition

ADRESSE

Approximate 30784 rows (1600 KB)

Updated on 2021-05-31 09:34:35

COLUMN NAME	DATA TYPE	NULLABLE	LENGTH	SCALE
ADRESSE_ID	INTEGER	N		0
NUMERO	INTEGER	Y		0
INDICE_REPETITION	VARCHAR	Y	1	0
TYPE_VOIE	VARCHAR	Y	4	0
NOM_VOIE	VARCHAR	N	30	0
ID_COMMUNE	INTEGER	N		0

Table Definition

LOCAL

Approximate 2 rows (32 KB)

Updated on 2021-05-31 12:06:12

COLUMN NAME	DATA TYPE	NULLABLE	LENGTH	SCALE
TYPE_LOCAL	SMALLINT	N		0
NOM_LOCAL	VARCHAR	N	43	0

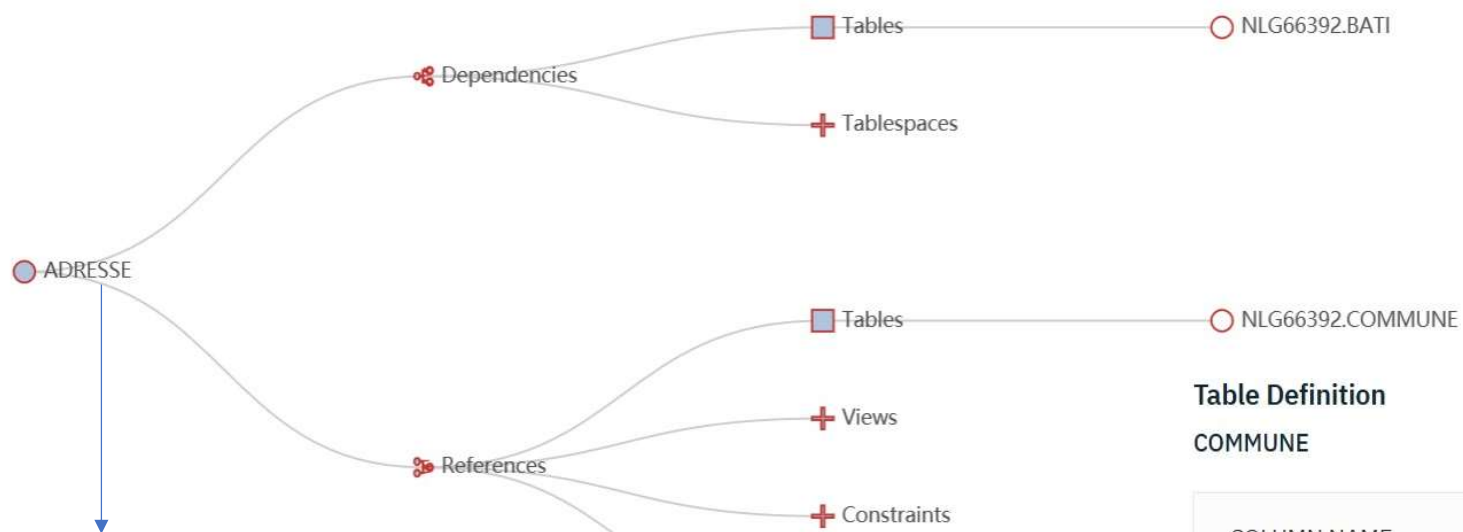


Table Definition

ADRESSE

Approximate 30784 rows (1600 KB)

Updated on 2021-05-31 09:34:35

COLUMN NAME	DATA TYPE	NULLABLE	LENGTH	SCALE
ADRESSE_ID	INTEGER	N		0
NUMERO	INTEGER	Y		0
INDICE_REPETITION	VARCHAR	Y	1	0
TYPE_VOIE	VARCHAR	Y	4	0
NOM_VOIE	VARCHAR	N	30	0
ID_COMMUNE	INTEGER	N		0

Table Definition

COMMUNE

Approximate 3214 rows (128 KB)

Updated on 2021-05-31 10:01:08

COLUMN NAME	DATA TYPE	NULLABLE	LENGTH	SCALE
ID_COMMUNE	INTEGER	N		0
CODE_POSTAL	VARCHAR	N	5	0
NOM_COMMUNE	VARCHAR	N	50	0

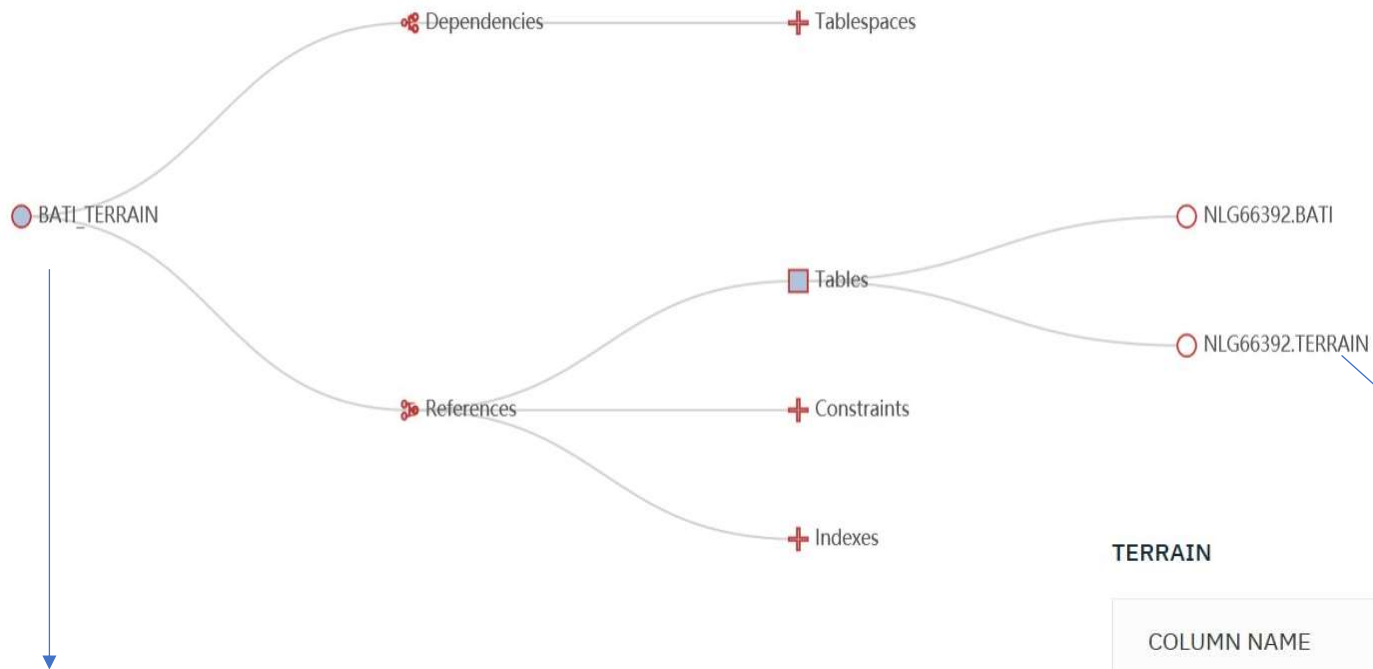


Table Definition

BATI_TERRAIN

Approximate 253 rows (32 KB)
Updated on 2021-05-31 10:01:09

COLUMN NAME	DATA TYPE	NULLABLE	LENGTH	SCALE
BATI_ID	INTEGER	N		0
TERRAIN_ID	INTEGER	N		0

TERRAIN

Approximate 208 rows (32 KB)
Updated on 2021-05-31 09:54:34

COLUMN NAME	DATA TYPE	NULLABLE	LENGTH	SCALE
TERRAIN_ID	INTEGER	N		0
NATURE_CULTURE	VARCHAR	N	6	0
NATURE_CULTURE_S...	VARCHAR	Y	5	0
SURFACE_TERRAIN	INTEGER	N		0

Interrogation de la base de données



Requêtes SQL et
résultats

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

Requête SQL

```
SELECT COUNT(*) AS "Nombre total d'appartements vendus au 1er semestre 2020."  
FROM mutation  
WHERE (nature_mutation = 'Vente' AND bati_id IN  
(SELECT bati_id FROM bati WHERE type_local = '2'));
```

Résultat

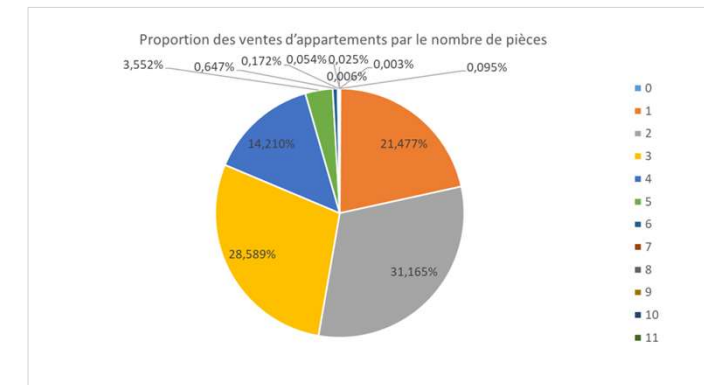
Nombre total d'appartements vendus au 1er semestre 2020.

31358

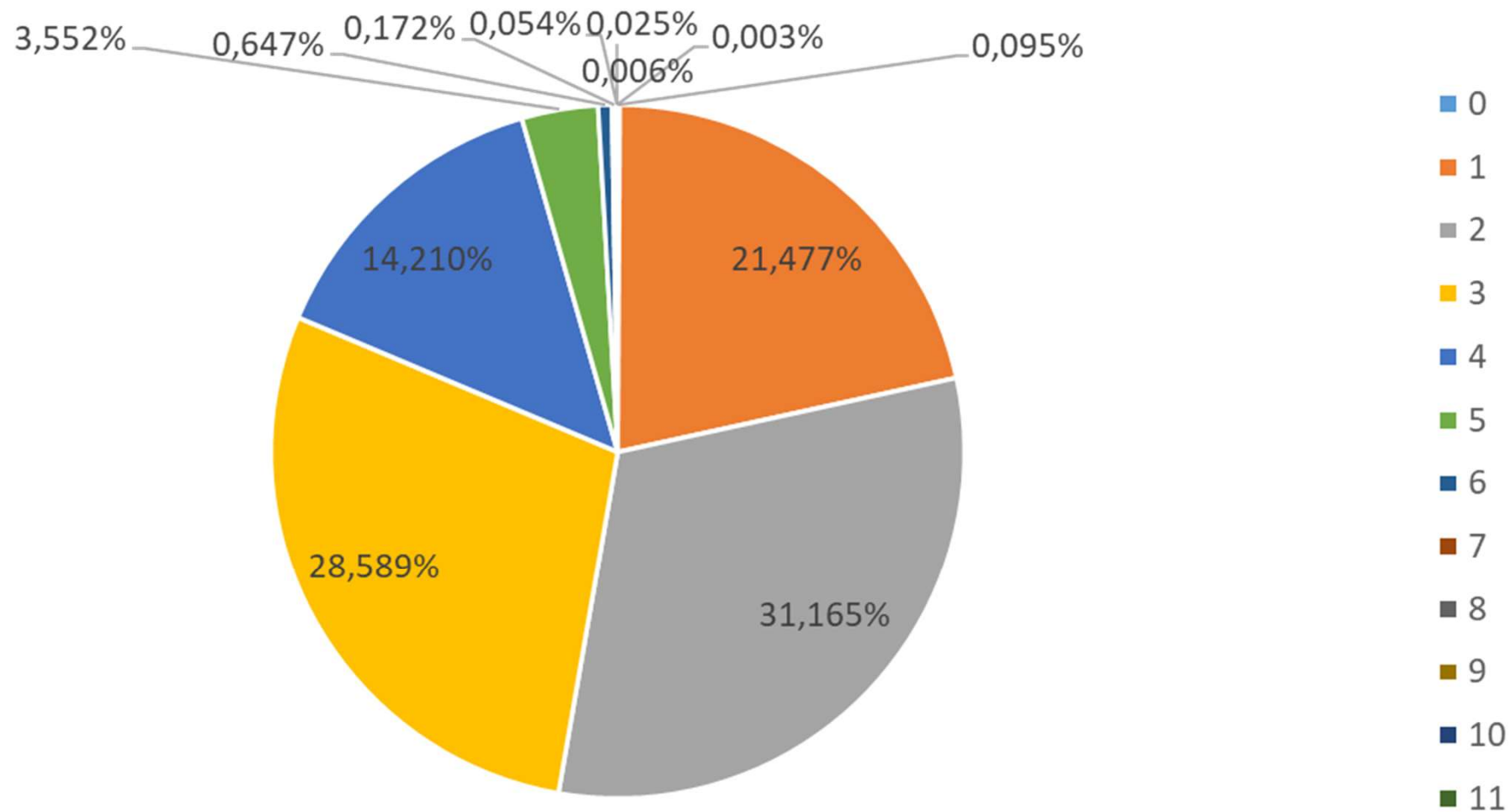
2. Proportion des ventes d'appartements par le nombre de pièces

Requête SQL

```
SELECT nombre_pieces AS "Nombre de pièces principales",  
CONCAT(DECIMAL((total_par_nb_pieces/nb_total_appart)*100, 6,3), ' %')  
AS "Pourcentage sur tous les appartements"  
FROM  
(SELECT nombre_pieces, COUNT(nombre_pieces) AS total_par_nb_pieces  
FROM bati  
INNER JOIN mutation ON mutation.bati_id = bati.bati_id  
WHERE nature_mutation= 'Vente' AND bati.type_local = '2'  
GROUP BY nombre_pieces) ,  
(SELECT COUNT(*) AS nb_total_appart  
FROM mutation  
WHERE (nature_mutation = 'Vente' AND bati_id IN  
(SELECT bati_id FROM bati WHERE type_local = '2')));
```



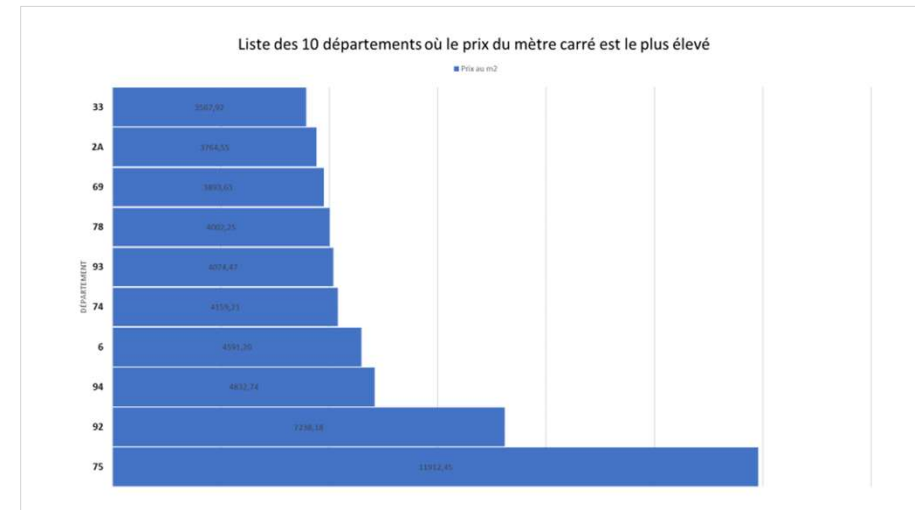
Proportion des ventes d'appartements par le nombre de pièces



3. Liste des 10 départements où le prix du mètre carré est le plus élevé (surface loi Carrez)

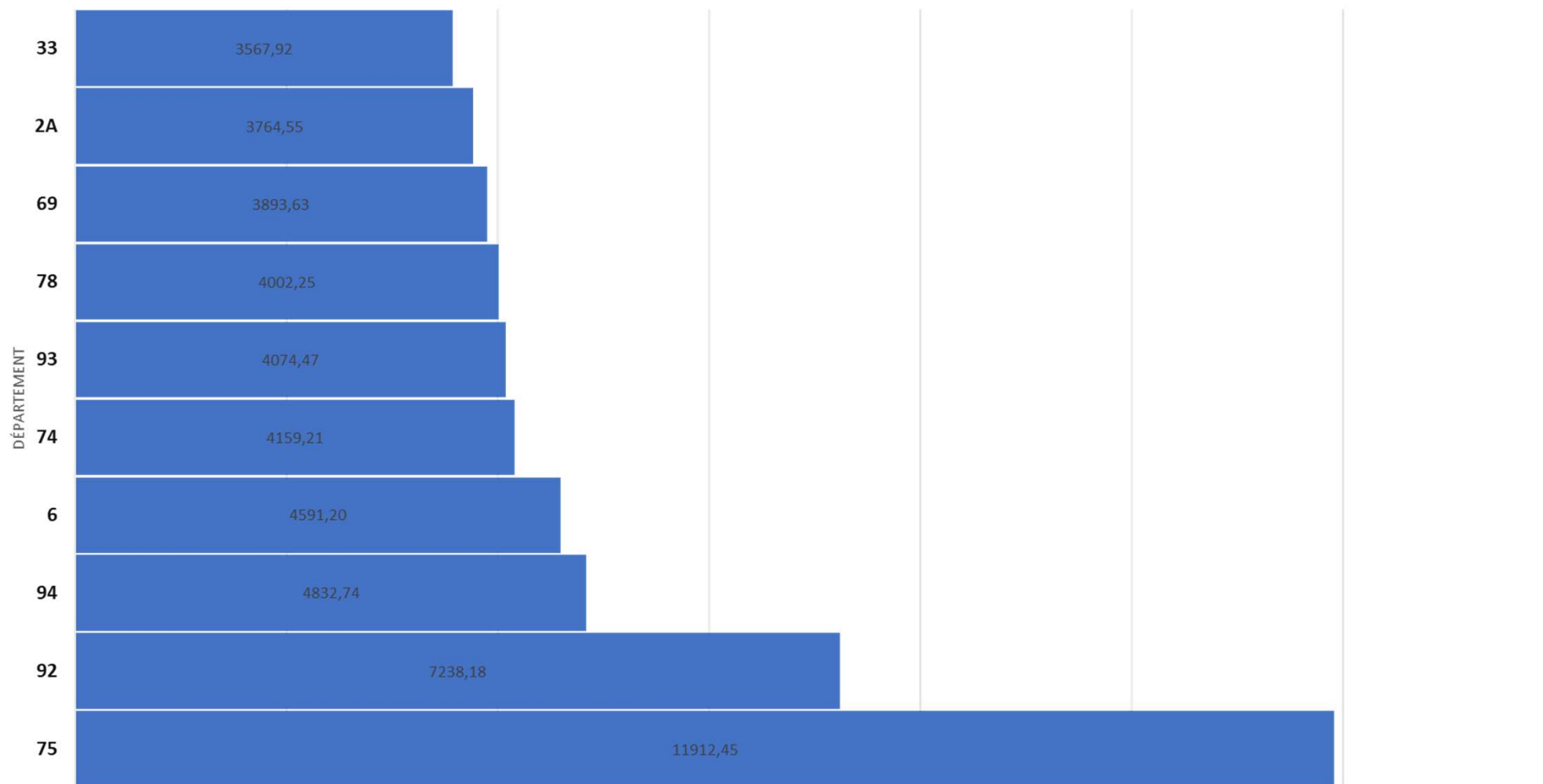
Requête SQL

```
SELECT ref_cadastrale.code_departement AS  
"Département",  
DECIMAL((SUM(valeur_fonciere)/SUM(surface_carrez_lot)),  
8, 2) AS "Prix au m2"  
FROM bati, ref_cadastrale  
INNER JOIN mutation ON mutation.bati_id = bati.bati_id  
WHERE bati.ref_cadastrale_id =  
ref_cadastrale.ref_cadastrale_id  
GROUP BY ref_cadastrale.code_departement  
ORDER BY "Prix au m2" DESC LIMIT 10
```



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

■ Prix au m2



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

Requête SQL

```
SELECT DECIMAL(SUM(mutation.valeur_fonciere/bati.surface_carrez_lot)/COUNT(*),6,2) AS "Prix  
moyen du m2"  
FROM bati, ref_cadastrale  
INNER JOIN mutation ON bati.bati_id = mutation.bati_id  
WHERE (bati.ref_cadastrale_id = ref_cadastrale.ref_cadastrale_id  
AND ref_cadastrale.code_departement IN ('75','77','78','91','92','93','94','95')  
AND bati.type_local = '1')
```

Résultat

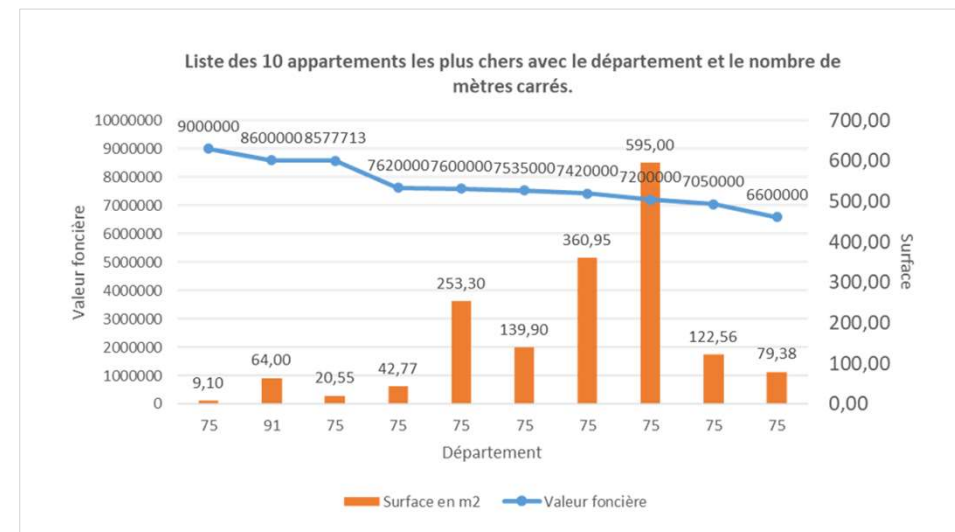
Prix moyen du m2

3745,01

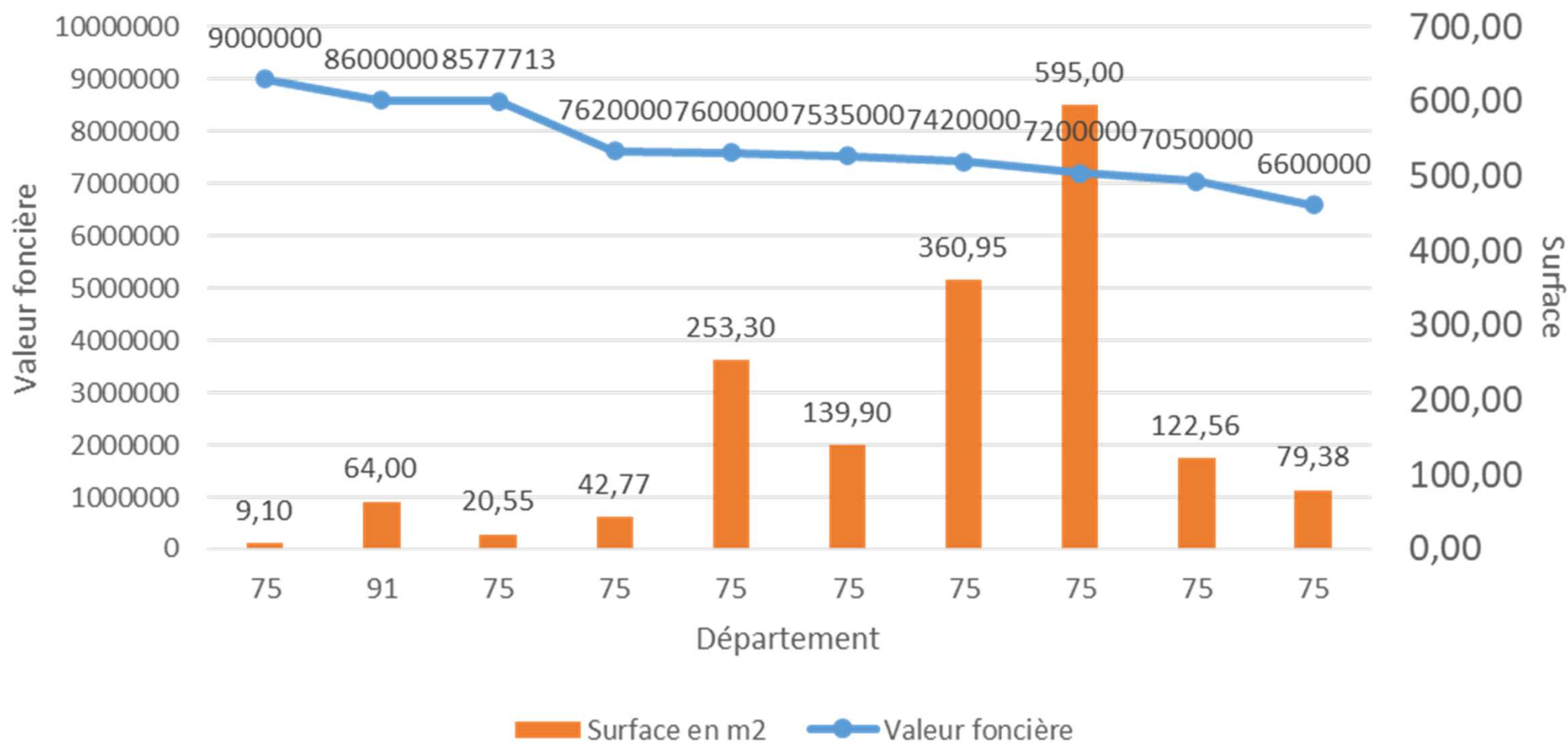
5. Liste des 10 appartements les plus chers avec le département et le nombre de mètres carrés

Requête SQL

```
SELECT * FROM  
(SELECT mutation.valeur_fonciere AS "Valeur  
foncière",  
bati.surface_carrez_lot AS "Surface en m2",  
ref_cadastrale.code_departement AS "Département"  
FROM bati,ref_cadastrale  
INNER JOIN mutation ON bati.bati_id =  
mutation.bati_id  
WHERE bati.ref_cadastrale_id =  
ref_cadastrale.ref_cadastrale_id  
AND bati.type_local = '2')  
ORDER BY "Valeur foncière" DESC LIMIT 10
```



Liste des 10 appartements les plus chers avec le département et le nombre de mètres carrés.



6. Taux d'évolution du nombre de ventes entre le premier et le deuxième trimestres de 2020

Requête SQL

```
SELECT CONCAT(DECIMAL (((ventes_2eme_trim-ventes_1er_trim)/ventes_1er_trim) * 100, 4,3),' %')  
AS "Taux d'évolution du nombre de ventes entre le premier et le deuxième trimestres de 2020."  
FROM (SELECT COUNT(*) AS ventes_1er_trim FROM mutation  
WHERE MONTH(DATE) IN (1,2,3)),  
(SELECT COUNT(*) AS ventes_2eme_trim FROM mutation  
WHERE MONTH(DATE) IN (4,5,6));
```

Résultat

Taux d'évolution du nombre de ventes entre le premier et le deuxième trimestres de 2020.
3.655 %

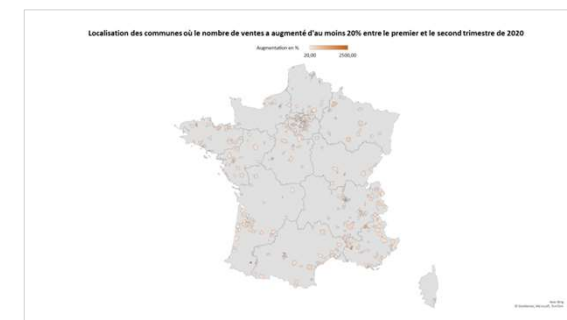
7. Liste des communes où le nombre de ventes a augmenté d'au moins 20% entre le premier et le second trimestre de 2020

Requête SQL (optimisée)

```
WITH ventes_com_1(id_commune, nb_ventes_1er_trim) AS
(SELECT adresse.id_commune, COUNT(mutation.mutation_id) AS nb_ventes_1er_trim FROM adresse
  INNER JOIN bati ON bati.adresse_id = adresse.adresse_id
  INNER JOIN mutation ON bati.bati_id = mutation.bati_id
 WHERE MONTH(DATE) IN (1,2,3)
 GROUP BY id_commune),
ventes_com_2(id_commune, nb_ventes_2eme_trim) AS
(SELECT adresse.id_commune, COUNT(mutation.mutation_id) AS nb_ventes_2eme_trim FROM adresse
  INNER JOIN bati ON bati.adresse_id = adresse.adresse_id
  INNER JOIN mutation ON bati.bati_id = mutation.bati_id
 WHERE MONTH(DATE) IN (4,5,6)
 GROUP BY id_commune),
ventes_com(id_commune, augmentation) AS
(SELECT id_commune, augmentation FROM
  (SELECT ventes_com_1.id_commune, DECIMAL(((ventes_com_2.nb_ventes_2eme_trim -
ventes_com_1.nb_ventes_1er_trim)/ventes_com_1.nb_ventes_1er_trim)*100,6,2) AS Augmentation
 FROM ventes_com_1
  INNER JOIN ventes_com_2 ON ventes_com_1.id_commune = ventes_com_2.id_commune)
 WHERE Augmentation >= 20)
SELECT commune.nom_commune AS "Commune", commune.code_postal AS "Code postal",
ventes_com.augmentation AS "Augmentation en %"
FROM commune
  INNER JOIN ventes_com ON commune.id_commune = ventes_com.id_commune
ORDER BY augmentation DESC;
```

Le choix arrêté il y a 48h était d'utiliser des vues, avec 'CREATE VIEW', toutefois je propose ici l'utilisation de 'WITH' qui permet de n'écrire qu'une unique requête.

[Revenir au fichier Excel des résultats](#)

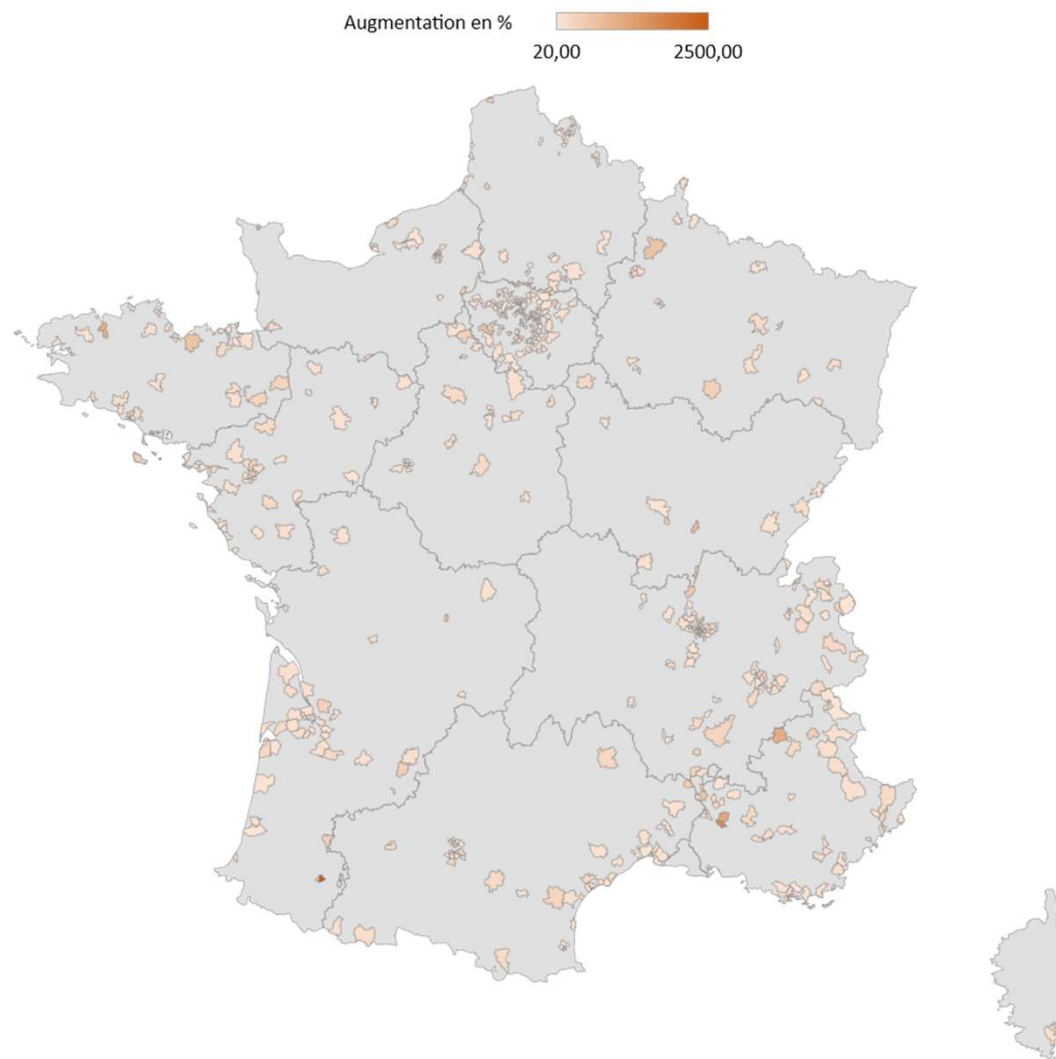


Diapositive 25

SP1

Sophie Piekarec; 16/06/2021

Localisation des communes où le nombre de ventes a augmenté d'au moins 20% entre le premier et le second trimestre de 2020



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

Requête SQL

```
SELECT CONCAT(DECIMAL(((prix_moyen_T3 - prix_moyen_T2)/prix_moyen_T2)*100,5,2), ' %')
  AS "Différence de prix moyen T3 T2"
FROM
  (SELECT SUM(mutation.valeur_fonciere/bati.surface_carrez_lot)/COUNT(*) AS prix_moyen_T2
FROM bati
INNER JOIN mutation ON bati.bati_id = mutation.bati_id
WHERE bati.type_local = '2' AND bati.nombre_pieces = '2'),
  (SELECT SUM(mutation.valeur_fonciere/bati.surface_carrez_lot)/COUNT(*) AS prix_moyen_T3
FROM bati
INNER JOIN mutation ON bati.bati_id = mutation.bati_id
WHERE bati.type_local = '2' AND bati.nombre_pieces = '3');
```

Résultat

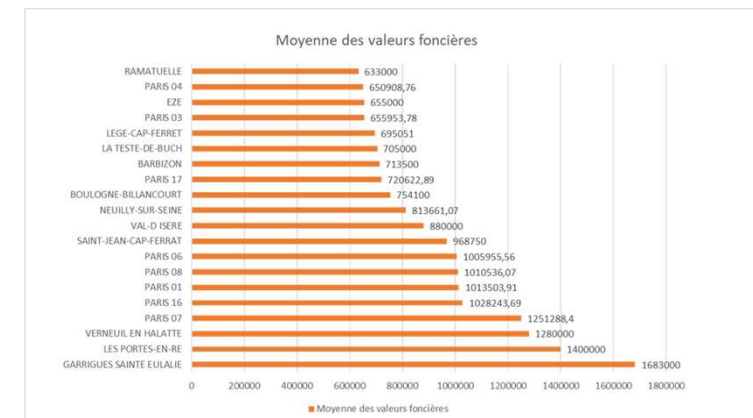
Différence de prix moyen T3 T2

-12.39 %

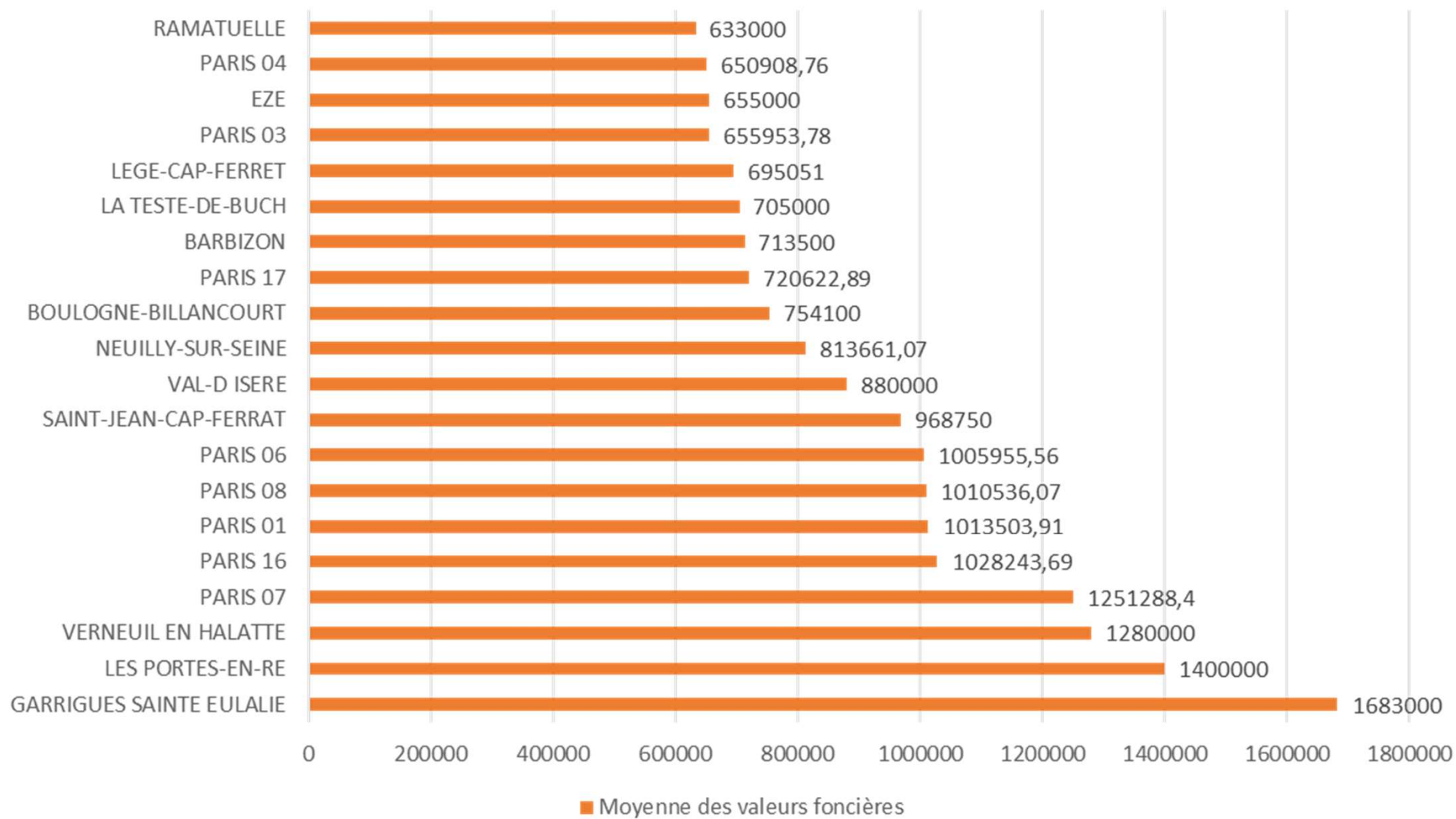
9. Moyennes de valeurs foncières pour le top 20 des communes

Requête SQL

```
SELECT commune.nom_commune AS "Commune",  
commune.code_postal AS "Code postal",  
DECIMAL(AVG(mutation.valeur_fonciere), 12,2) AS "Moyenne des  
valeurs foncières"  
FROM commune  
INNER JOIN adresse ON commune.id_commune =  
adresse.id_commune  
INNER JOIN bati ON bati.adresse_id = adresse.adresse_id  
INNER JOIN mutation ON bati.bati_id = mutation.bati_id  
GROUP BY commune.nom_commune, commune.code_postal  
ORDER BY "Moyenne des valeurs foncières" DESC LIMIT 20 ;
```



Moyenne des valeurs foncières



Merci de votre attention 😊