

PROJET DATA IMMO

ROMAIN RAMAT

JANVIER 2024

LAPLACE IMMO

CONTEXTE

- Mise à disposition d'une base de données permettant d'identifier l'évolution du marché de l'immobilier en prévision des prix de vente des biens immobiliers.

SOURCES :

- ☐ Des données extraites du site open data des **Demandes de valeurs foncières (DVF)**
- ☐ **Des données de l'INSEE** avec les résultats des recensements de la population
- ☐ **Des données de data.gouv sur les régions**, avec le référentiel géographique français, communes, unités urbaines, aires urbaines, départements, académies, régions

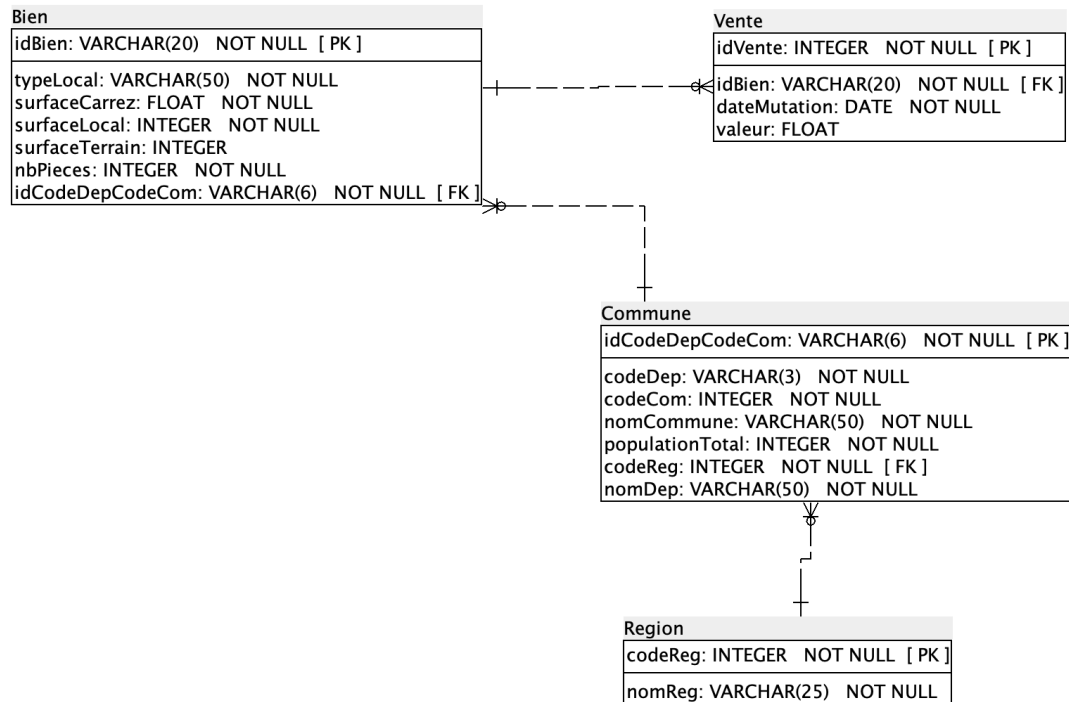


DICTIONNAIRE DES DONNÉES SANS DONNÉES PERSONNELLES

CODE	SIGNIFICATION	TYPE	LONGUEUR	NATURE	REGLE DE GESTION	REGLE DE CALCUL
idBien	Identifiant unique de la table BIEN	VARCHAR	20	CLE PRIMAIRE	NON NULL	Concaténation de : CodeDépartement, CodeCommune, Section, N°Plan, Code TypeLocal, nb_lot}
typeLocal	Maison, Appartement, dépendance, autres...	VARCHAR	50	Elémentaire	NON NULL	
surfaceCarrez	Surface selon loi Carrez	FLOAT		Elémentaire	NON NULL	Convertir les "," en "." pour interprétation dans la base de donnée
surfaceLocal	Surface de l'habitation	INTERGER		Elémentaire	NON NULL	
surfaceTerrain	Surface du terrain	INTERGER		Elémentaire		
nbPieces	Nombre de pièce dans l'habitation	INTERGER		Elémentaire	NON NULL	
idCodeDepCodeCom	Clé étrangère regroupant le code département et code commune	VARCHAR	6	CLE ETRANGERE	NON NULL	Concatenation du code dep et code commune
idVente	Identifiant unique de la table Vente	INTERGER		CLE PRIMAIRE	NON NULL	Auto incrémentation nouvelle vente
idBien	Identifiant unique de la table BIEN	VARCHAR	20	CLE ETRANGERE	NON NULL	
dateMutation	Date de signature de l'acte du bien	DATE		Elémentaire	NON NULL	Format AAAA/MM/JJ
valeur	Valeur foncière du bien immobilisé (montant ou évaluation)	FLOAT		Elémentaire		Convertir les "," en "." pour interprétation dans la base de donnée
idCodeDepCodeCom	Identifiant composé du numéro de département et code de la commune CADASTRALE	VARCHAR	6	CLE PRIMAIRE	NON NULL	Concatenation du code dep et code commune
codeDep	Code Département	VARCHAR	3	Elémentaire	NON NULL	
codeCom	Code Commune	INTERGER	3	Elémentaire	NON NULL	
nomCommune	Nom de la commune	VARCHAR	50	Elémentaire	NON NULL	
populationTotal	Nombre total d'habitant dans la commune	INTERGER		Elémentaire		
codeReg	Code de la région	INTERGER	2	CLE ETRANGERE	NON NULL	
nomDep	Libellé du département	VARCHAR	50	Elémentaire	NON NULL	
codeReg	Code de la région	INTERGER	2	CLE PRIMAIRE	NON NULL	
nomReg	Nom de la région	VARCHAR	25	Elémentaire	NON NULL	

Pas d'utilisation des données personnelles permettant d'identifier une personne physique

ETABLISSEMENT D'UN MODÈLE CONCEPTUEL DES DONNÉES



```
Preview SQL Script

Your Target Database is not configured.

CREATE TABLE Region (
  codeReg INTEGER NOT NULL,
  nomReg VARCHAR(25) NOT NULL,
  CONSTRAINT id_region_pk PRIMARY KEY (codeReg)
);

CREATE TABLE Commune (
  idCodeDepCodeCom VARCHAR(6) NOT NULL,
  codeDep VARCHAR(3) NOT NULL,
  codeCom INTEGER NOT NULL,
  nomCommune VARCHAR(50) NOT NULL,
  populationTotal INTEGER NOT NULL,
  codeReg INTEGER NOT NULL,
  nomDep VARCHAR(50) NOT NULL,
  CONSTRAINT id_codedep_codecommune PRIMARY KEY
(idCodeDepCodeCom)
);

CREATE TABLE Bien (
  idBien VARCHAR(20) NOT NULL,
  typeLocal VARCHAR(50) NOT NULL,
  surfaceCarrez FLOAT NOT NULL,
  surfaceLocal INTEGER NOT NULL,
  surfaceTerrain INTEGER,
  nbPieces INTEGER NOT NULL,
  idCodeDepCodeCom VARCHAR(6) NOT NULL,
  CONSTRAINT id_Bien_pk PRIMARY KEY (idBien)
);
```

FORMES NORMALES :

- 1NF** Chaque table a des colonnes atomiques
- 2NF** Les colonnes non-clé dépendent entièrement de la clé primaire
- 3NF** Aucune dépendance transitive entre colonnes non-clé

PRÉPARATION ET IMPORTATION DES DONNÉES DANS SQLITE STUDIO



Bien.csv



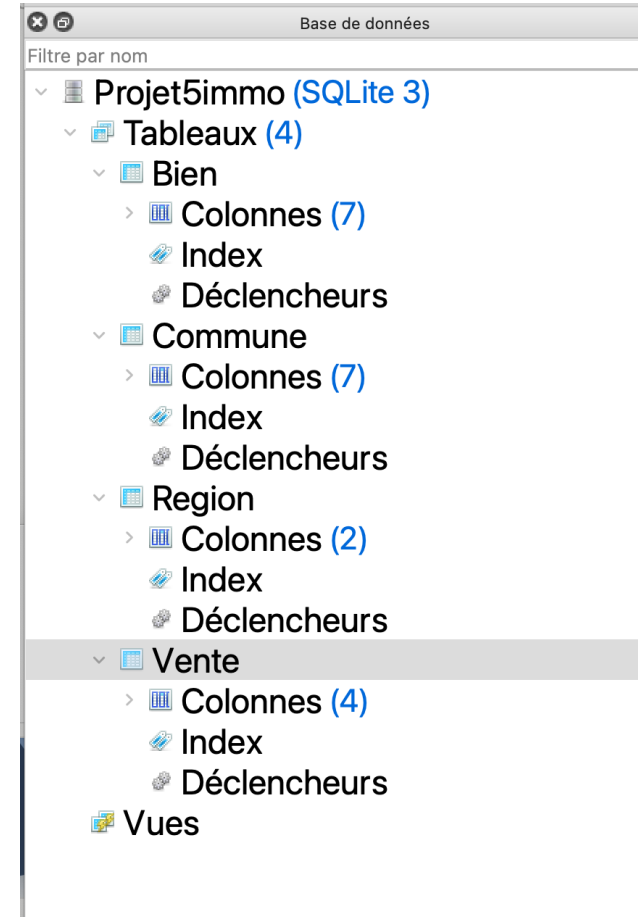
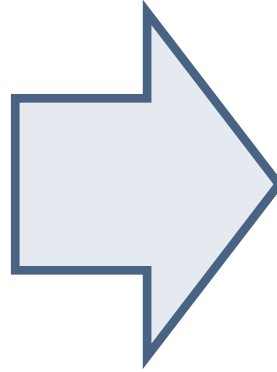
Vente.csv



Commune.csv



Region.csv



VÉRIFICATION DES DONNÉES

SELECT COUNT(*) FROM
Bien → 34169



SELECT COUNT(*) FROM
Vente → 34169



SELECT COUNT(*) FROM
Commune → 34991



SELECT COUNT(*) FROM
Region → 19

VÉRIFICATION DES DONNÉES

SELECT * FROM Bien

JOIN Vente ON Bien.idBien = Vente.idBien

JOIN Commune ON Bien.idCodeDepCodeCom = Commune.idCodeDepCodeCom

JOIN Region ON Commune.CodeReg = Region.CodeReg

LIMIT 10

Nombre de lignes chargées : 10																				
	idBien	typeLocal	surface	surface	surface	nbPiec	idCode	idVente	idBien:1	dateMutatio	valeur	idCode	codeDe	codeCr	nomCommune	populatio	codeRe	nomDep	codeRe	nomReg
1	01103A3022122	Appartement	48.22	48		3	01103	1	01103A3022122	2020/01/02	165000	01103	01	103	Chevry	2196	84	Ain (01)	84	Auvergne-Rhône-Alpes
2	06004CP18621322	Appartement	39.11	40		1	06004	2	06004CP18621322	2020/01/02	355680	06004	06	4	Antibes	74523	93	Alpes-Maritimes (06)	93	Provence-Alpes-Côte d'Azur
3	06088LS1692991	Appartement	80.25	82		3	06088	3	06088LS1692991	2020/01/02	229500	06088	06	88	Nice	345528	93	Alpes-Maritimes (06)	93	Provence-Alpes-Côte d'Azur
4	06123AO34822421	Appartement	27.51	27		1	06123	4	06123AO34822421	2020/01/02	125000	06123	06	123	Saint-Laurent-du-Var	29352	93	Alpes-Maritimes (06)	93	Provence-Alpes-Côte d'Azur
5	13005AW22422181	Appartement	47.33	47		2	13005	5	13005AW22422181	2020/01/02	90000	13005	13	5	Aubagne	47986	93	Bouches-du-Rhône (13)	93	Provence-Alpes-Côte d'Azur
6	13028AB69221	Appartement	25.31	24		1	13028	6	13028AB69221	2020/01/02	93000	13028	13	28	La Ciotat	36374	93	Bouches-du-Rhône (13)	93	Provence-Alpes-Côte d'Azur
7	13208C2121122	Appartement	70.84	70		3	13208	7	13208C2121122	2020/01/02	298100	13208	13	208	Marseille 8e Arrondissement	82690	93	Bouches-du-Rhône (13)	93	Provence-Alpes-Côte d'Azur
8	13212A97210802	Appartement	67.19	66		3	13212	8	13212A97210802	2020/01/02	163500	13212	13	212	Marseille 12e Arrondissement	61958	93	Bouches-du-Rhône (13)	93	Provence-Alpes-Côte d'Azur
9	14338AC18121552	Appartement	18.89	19		1	14338	9	14338AC18121552	2020/01/02	53000	14338	14	338	Houlgate	1748	28	Calvados (14)	28	Normandie
10	14366BH931322	Maison	105.37	99		4	14366	10	14366BH931322	2020/01/02	136000	14366	14	366	Lisieux	20692	28	Calvados (14)	28	Normandie

RÉSULTATS DES REQUÊTES 1/4

Nombre d'appartements vendus 1 Semestre 2020:

Local	Nombre	DateMin	DateMax
1 Appartement	31378	2020/01/02	2020/06/30

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

Nombre de lignes chargées : 1	
Region	Prix_M2
1 Ile-de-France	3675.79 Euros

Liste des 10 Appartements les plus chers avec la région et le nombre de mètre carrés :

	Bien	Prix_De_Vente	Region	Surface_M2
1	75116DD128531	9000000.0	Ile-de-France	9.1
2	91174AX1623221	8600000	Ile-de-France	64
3	75107AK3421005	8577713	Ile-de-France	20.55
4	75117CO1002152	7620000	Ile-de-France	42.77
5	75106AU1421045	7600000	Ile-de-France	253.3
6	75101AZ862792	7535000	Ile-de-France	139.9
7	75116DV46251	7420000	Ile-de-France	360.95
8	75116CN242112	7200000	Ile-de-France	595
9	75101BC67228	7050000	Ile-de-France	122.56
10	75101AT421811	6600000	Ile-de-France	79.38

Nombre de ventes d'appartements par région pour le 1er Semestre 2020.

	Region	Nb_vente
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

RÉSULTATS DES REQUÊTES 2/4

Taux d'évolution du nombre de vente entre le 1er et le second trimestre 2020

	VenteT1	VenteT2	Variation_Pourcentage
1	16776	17393	3.68 %

Proportion de vente d'appartements par le nombre de pièces

	NB_Pieces	NB_Appart	Proportion
1	2	9783	31.18
2	3	8966	28.57
3	1	6739	21.48
4	4	4460	14.21
5	5	1114	3.55
6	6	204	0.65
7	7	54	0.17
8	0	30	0.1
9	8	17	0.05
10	9	8	0.03
11	10	2	0.01
12	11	1	0

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

	Departement	Prix_M2
1	Paris (75)	11898.72
2	Hauts-de-Seine (92)	7238.19
3	Val-de-Marne (94)	4831.37
4	Alpes-Maritimes (06)	4587.83
5	Haute-Savoie (74)	4159.22
6	Seine-Saint-Denis (93)	4065.39
7	Yvelines (78)	4002.16
8	Rhône (69)	3893.5
9	Corse-du-Sud (2A)	3752.76
10	Gironde (33)	3567.93

RÉSULTATS DES REQUÊTES 3/4

Le classement des régions par rapport aux prix au mètre carré un appartement de 2 pièces et un appartement de 3 pièces.

	Region	Prix_VenteT2	Prix_VenteT3	Prix_Moyen
1	Ile-de-France	7308.34	6241.35	6774.85
2	Corse	4350.39	3548.97	3949.68
3	Provence-Alpes-Côte d'Azur	3749.88	3389.15	3569.52
4	Auvergne-Rhône-Alpes	3490.27	3015.28	3252.78
5	Nouvelle-Aquitaine	3215.4	2883.54	3049.47
6	Pays de la Loire	2897.55	2636.07	2766.81
7	Martinique	2950.28	2530.77	2740.53
8	Normandie	2823.07	2333.6	2578.34
9	Occitanie	2639.35	2465.07	2552.21
10	Bretagne	2654.99	2441.64	2548.31
11	Guyane	2609.7	2404.64	2507.17
12	Hauts-de-France	2632.55	2236.04	2434.3
13	La Réunion	2308.44	2497.87	2403.15
14	Centre-Val de Loire	2151.05	1700.11	1925.58
15	Bourgogne-Franche-Comté	1715.31	1728.04	1721.67
16	Grand Est	1783.63	1594.7	1689.16

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

	PrixM2T2	PrixM2T3	Différence
1	4903.56	4299.88	-12.31%

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

Nombre de lignes chargées : 47		
	NB_Ventes	Nom_commune
1	228	Paris 17e Arrondissement
2	215	Paris 15e Arrondissement
3	209	Paris 18e Arrondissement
4	173	Nice
5	169	Paris 11e Arrondissement
6	165	Paris 16e Arrondissement
7	157	Bordeaux
8	146	Paris 14e Arrondissement
9	127	Paris 20e Arrondissement
10	119	Nantes
11	116	Paris 19e Arrondissement
12	110	Paris 12e Arrondissement
13	109	Paris 10e Arrondissement
14	106	Paris 9e Arrondissement
15	106	Grenoble
16	99	Boulogne-Billancourt
17	94	Paris 13e Arrondissement
18	87	Paris 7e Arrondissement
19	86	Paris 6e Arrondissement
20	81	Marseille 8e Arrondissement

RÉSULTATS DES REQUÊTES 4/4

Les moyennes de valeurs foncières pour le top3 des communes des départements 6, 13, 33, 59 et 69

	Departement	Commune	Rang	Valeur_Moyenne
1	Alpes-Maritimes (06)	Saint-Jean-Cap-Ferrat	1	968750.0 Euros
2	Alpes-Maritimes (06)	Eze	2	655000.0 Euros
3	Alpes-Maritimes (06)	Mouans-Sartoux	3	476898.1 Euros
4	Bouches-du-Rhône (13)	Gignac-la-Nerthe	1	330000.0 Euros
5	Bouches-du-Rhône (13)	Saint-Savournin	2	314425.0 Euros
6	Bouches-du-Rhône (13)	Cassis	3	313416.88 Euros
7	Gironde (33)	Lège-Cap-Ferret	1	549500.64 Euros
8	Gironde (33)	Vayres	2	335000.0 Euros
9	Gironde (33)	Arcachon	3	307435.93 Euros
10	Nord (59)	Bersée	1	433202.0 Euros
11	Nord (59)	Cysoing	2	408550.0 Euros
12	Nord (59)	Halluin	3	322250.0 Euros
13	Rhône (69)	Ville-sur-Jarnioux	1	485300.0 Euros
14	Rhône (69)	Lyon 2e Arrondissement	2	455217.27 Euros
15	Rhône (69)	Lyon 6e Arrondissement	3	426968.25 Euros

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

	Commune	NB_Vente	Population	Ratio_1000
1	Paris 2e ...	127	21735	5.84
2	Paris 1er ...	79	16055	4.92
3	Paris 3e ...	161	34306	4.69
4	Arcachon	55	11898	4.62
5	La Baule-Escoublac	77	16797	4.58
6	Paris 4e ...	120	29390	4.08
7	Roquebrune-Cap-...	52	13041	3.99
8	Paris 8e ...	139	36250	3.83
9	Sanary-sur-Mer	60	17160	3.5
10	Paris 9e ...	208	60563	3.43
11	La Londe-les-Maures	37	10776	3.43
12	Paris 6e ...	139	41171	3.38
13	Saint-Cyr-sur-Mer	38	11725	3.24
14	Chantilly	35	11178	3.13
15	Saint-Mandé	69	22576	3.06
16	Pornichet	35	11440	3.06
17	Paris 10e ...	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

LES POINTS D'ATTENTION

- Remplacer les « , » par des « . » dans les tables, sinon les nombres n'étaient pas correctement classés malgré le « Float »
- Importation en Latin1 pour la gestion des accents
- Usage des différentes jointures
- Découverte de fonctions avancées :
 - WITH
 - COALESCE
 - RANK () OVER() PARTITION BY

