

# IMMOBILIER



## Analyse data-centric

---

LOU, MATHIEU, DJESSIM & RYAN

Une société immobilière souhaiterait récupérer et stocker des données Open Data des ventes immobilières afin de mener une étude lui permettant d'avoir une meilleure connaissance du marché en utilisant l'IA dans un futur proche. En amont de cela, elle souhaiterait donc pouvoir créer un flux permettant le stockage de données mais aussi des analyses de l'historique.



# Lou - Nettoyage de données

J'ai pris le parti de nettoyer les données pour n'avoir qu'une ligne par Transaction. En effet, en analysant, je me suis rendu compte qu'il y avait des doublons et que les dépendances, caves, garages, etc. qui possédaient la même Valeur foncière. Dans un premier temps, l'objectif était de les linker ensemble. Finalement, j'ai préféré garder que les transactions Appartement, Maison et Local Commercial qui représentent les lots principaux des ventes.

Dans un premier temps, pour obtenir ce résultat, j'ai contrôlé les colonnes et leur type ainsi que la proportion de NaN au sein des colonnes. Avec l'équipe, nous avons décidé de ne conserver que les colonnes avec moins de 60 % de NaN. Les colonnes Surface carrez n'étaient pas concernées et nous les avons gardées pour les analyses futures.

Par la suite, j'ai concaténé les colonnes N° voie, type de voie et Voie afin d'obtenir une colonne Adresse. J'ai fait la même chose avec les colonnes Code département et Code commune afin d'obtenir un Code postal propre. Au préalable, il aura fallu préparer ces dites colonnes. Enfin, une fois terminé, j'ai décidé de supprimer les doublons.

J'ai finalement trouvé un dataset des régions et départements que j'ai merge avec notre dataframe afin de pousser les analyses un peu plus loin. Au final, j'ai traité le dataframe obtenu pour n'avoir qu'une seule transaction passant de plus de 4 millions de lignes à 1 471 825 lignes. J'en ai profité pour remplacer tous nos NaN par des 0.

Enfin, j'ai renommé et réorganisé toutes les colonnes afin d'obtenir un dataframe propre et prêt à l'emploi. Je l'ai exporté en CSV pour que mon équipe puisse travailler dessus. À partir de là, nous avons pu préparer les dataframe correspondant à nos tables qui ont été établies dans notre MCD. Cela nous a permis d'importer sans souci nos tables dans notre base de données.

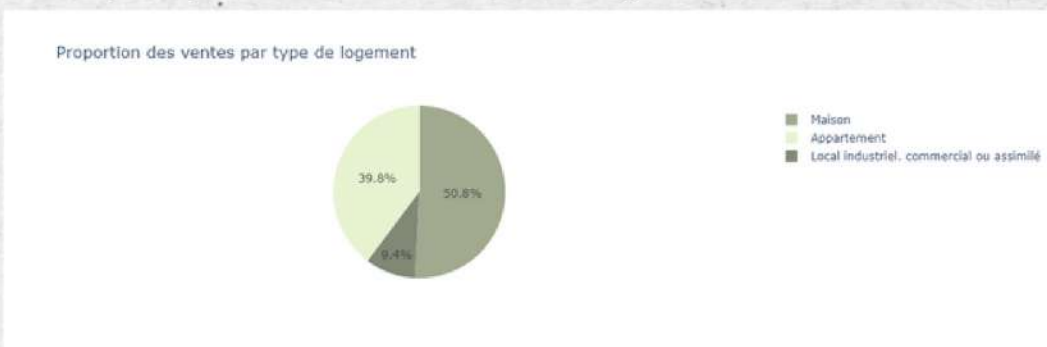




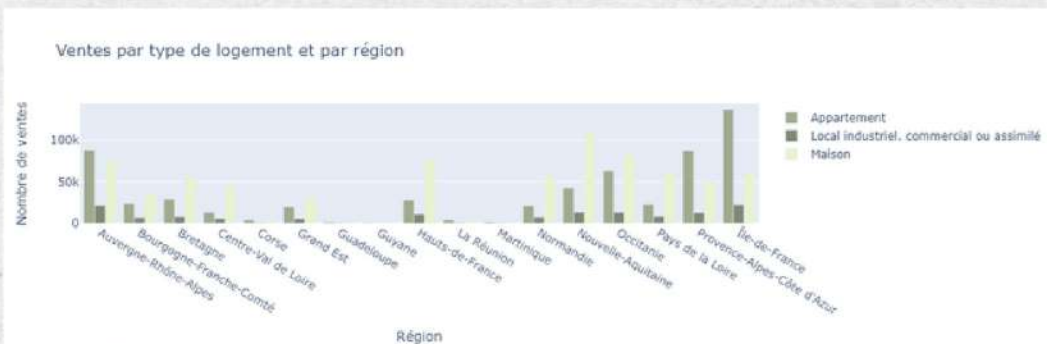
# Viz Lou



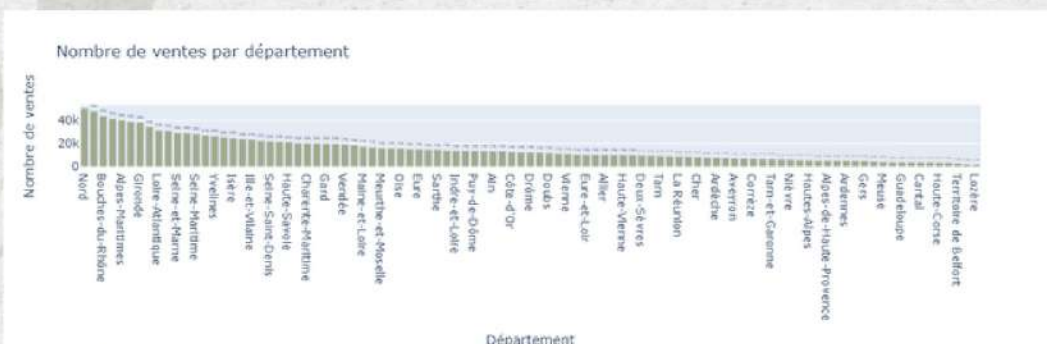
On constate que le 1er et le 2eme lot concerne surtout les ventes de ces trois types. N'ayant pas gardé les lots annexes, on en déduit que les 3 autres lots doivent principalement concerner les autres types.



Il est clairement visible que les maisons est le type qui se vend le plus ensuite les appartements puis les locaux commerciaux. Les maisons se vendent autant que les appartements et locaux réunis.



On voit qu'en fonction des régions, la tendance n'est pas la même. Cela peut se justifier par le type de région si elle est constituée de ville en majorité ou de village.



Ce graphique est intéressant pour visualiser les ventes par département. On constate que les ventes les plus importantes se situent majoritairement sur les côtes littorales.



# Viz Lou



Ces trois graphiques traitent des moyennes en fonction du type de logement. Nous avons donc :

- Moyenne du prix
- Moyenne des m²
- Moyenne du Prix au m²

On constate alors que les locaux commerciaux possèdent à chaque fois la plus grande part du donut. Cela se justifie par le fait qu'ils soient plus grands et donc plus chers que les autres types.

## Conclusion

Avec toutes les analyses effectuées, les maisons se démarquent clairement du type de logement. C'est celui qui se vend le mieux et ce dans la majorité des régions. On constate que les départements où les régions influent énormément sur les achats. Les maisons font partie des valeurs foncières les plus élevées.

Pour autant, ces derniers ne sont pas soumis à la saisonnalité puisque les ventes restent constantes sur les 4 trimestres 2022. On peut à la rigueur soulignée les faits que les ventes augmentent légèrement au printemps et à l'été, le soleil donne envie aux acheteurs d'investir. Cependant, il y a une grosse chute des ventes au mois d'Août et au mois de Novembre.

Si l'on s'attarde sur les appartements, on rapporte que ce sont les 2 pièces puis les 3 pièces qui vont se vendre globalement mieux.

Concernant les locaux commerciaux, eux, se démarquent vraiment lorsqu'on effectue les moyennes par prix, surfaces et prix au m², sinon on ne les voit pas tant que ça.

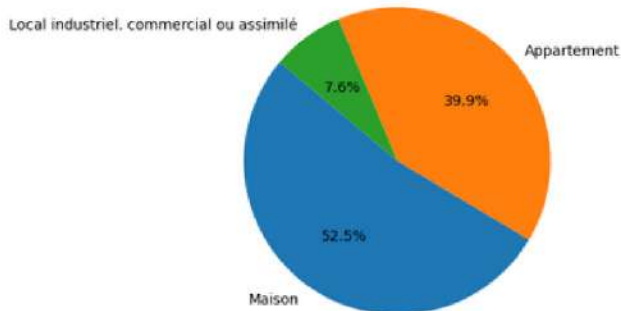
---



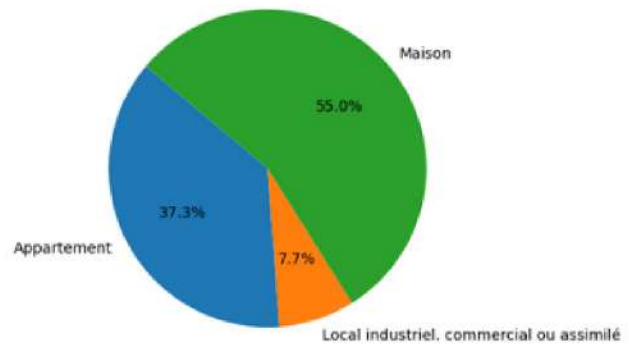


# Viz Mathieu

Proportion de Ventes par Type de Local



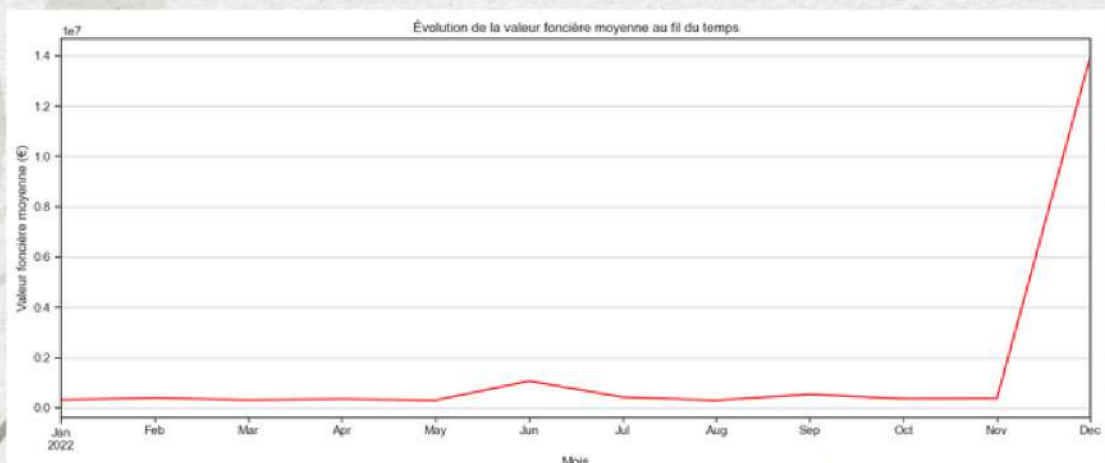
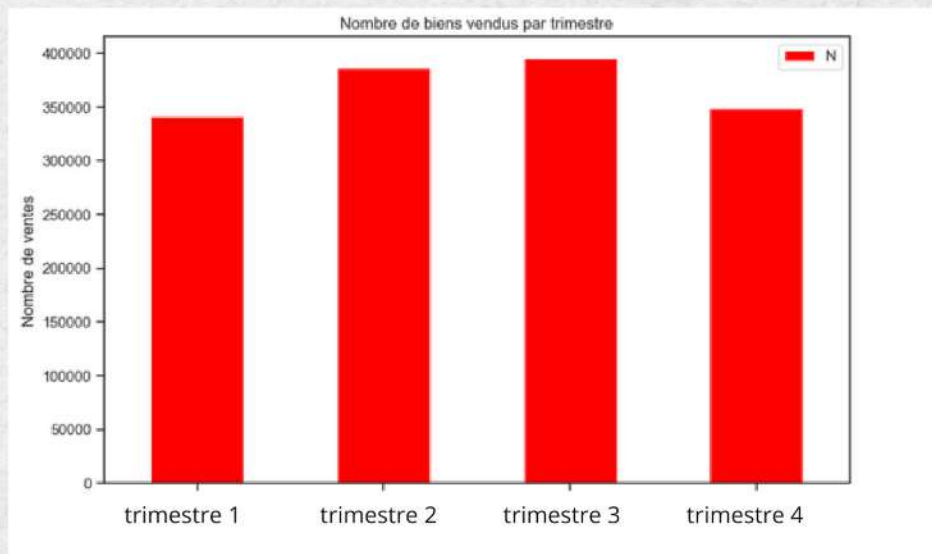
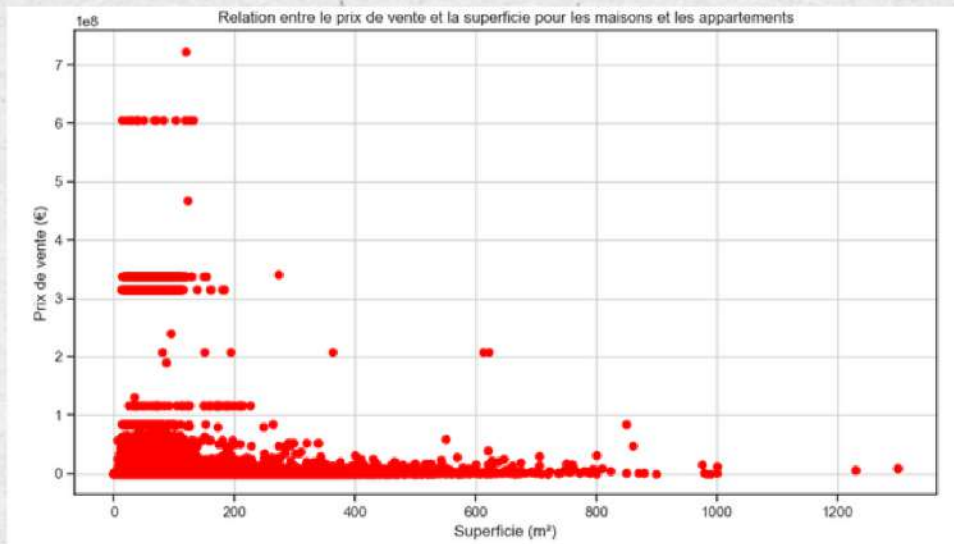
Proportion de la somme des Valeurs Foncières par Type de Local



## Conclusion

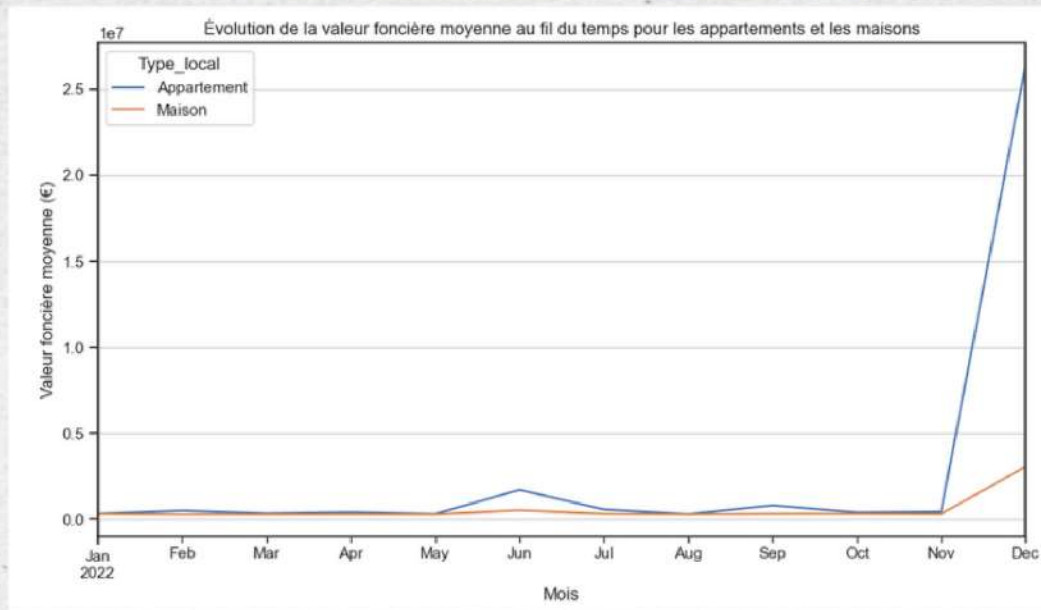
- proportion par type de local : **préférence nette pour les maisons** dans les transactions immobilières.
- proportion de la somme des valeurs foncières : bien que les maisons constituent la majorité des ventes, leurs proportions en valeurs est plus faible. Cela pourrait suggérer que les **ventes d'appartement sont plus élevé en valeur**. contre-intuitif.
- somme des valeurs foncières par régions : grande disparité, **Ile de France important**, qui se confirme par la présence de Paris.
- Surface réelle du Bâtiment par Régions : Ile de France a une surface bâtit relativement basse, ce qui suggère que les prix au m2 sont élevé
- évolutions des ventes par mois : **augmentation en juin et juillet** probablement du à la volonté des acquéreurs de validé leur achat avant les vacances d'août.

# Viz Djessim





# Viz Djessim



## Conclusion



# Requêtes Ryan

1/ Nombre d'appartements et Maisons vendus en 2022

	nb_maisons_appart	Type_Local
0	744942	Maison
1	585515	Appartement

2/ Nombre de biens vendus par trimestre

	Trimestre	Nombre_de_biens_vendus
0	1	341413
1	2	386302
2	3	395501
3	4	348609

3/ Proportion des ventes de biens par trimestre

	Trimestre	Proportions_de_biens_vendus
0	1	23.1966
1	2	26.2465
2	3	26.8715
3	4	23.6855

4/ Proportion d'appartements vendus par nombre de pièces

	Proportions	Nombre_pieces_principales
0	3.9359	5.0
1	29.9309	3.0
2	15.5190	4.0
3	18.6461	1.0
4	30.6310	2.0
5	0.1944	7.0

5/ Les 10 départements où il y a eu le plus de ventes immobilières.

	biens_vendus	Département
0	50645	Nord
1	48071	Paris
2	43911	Bouches-du-Rhône
3	41670	Var
4	40215	Alpes-Maritimes
5	39022	Rhône
6	38380	Gironde
7	34456	Hérault
8	31419	Loire-Atlantique
9	31033	Haute-Garonne

6/ Les 10 départements où il y en a eu le moins

	Biens_vendus	Département
0	1748	Lozère
1	1962	Guyane
2	2841	Territoire de Belfort
3	3421	Martinique
4	3452	Haute-Corse
5	3639	Corse-du-Sud
6	3693	Cantal
7	3709	Creuse
8	3763	Guadeloupe
9	3811	Haute-Marne

7/ Prix moyen du mètre carré en IDF

	prix_moyen
0	4407.342152

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

ID_Bien	Type_local	Surface_reelle_batiment	Nombre_pieces_principales	Surface_terrain	Nombre_lots	ID_Bienification	Adresse	Code_postal	Commune	Département
A_472303	Appartement	65.0	3.0	772.0	0	1233533	586.0 AV VICTOR HUGO	83061	FREJUS	V
A_472308	Appartement	80.0	4.0	772.0	0	1233534	586.0 AV VICTOR HUGO	83061	FREJUS	V
A_383002	Appartement	75.0	3.0	3244.0	0	1405799	1.0 RUE DES TISSERANDS	9740116	SAINT PIERRE	La Réunion
A_382997	Appartement	28.0	1.0	3244.0	0	1405793	1.0 RUE DES TISSERANDS	9740116	SAINT PIERRE	La Réunion
A_383003	Appartement	62.0	3.0	3244.0	0	1405801	1.0 RUE DES TISSERANDS	9740116	SAINT PIERRE	La Réunion
A_383001	Appartement	29.0	1.0	3244.0	0	1405798	1.0 RUE DES TISSERANDS	9740116	SAINT PIERRE	La Réunion
A_382995	Appartement	43.0	2.0	3244.0	0	1405791	1.0 RUE DES TISSERANDS	9740116	SAINT PIERRE	La Réunion
A_382998	Appartement	71.0	4.0	3244.0	0	1405794	1.0 RUE DES TISSERANDS	9740116	SAINT PIERRE	La Réunion

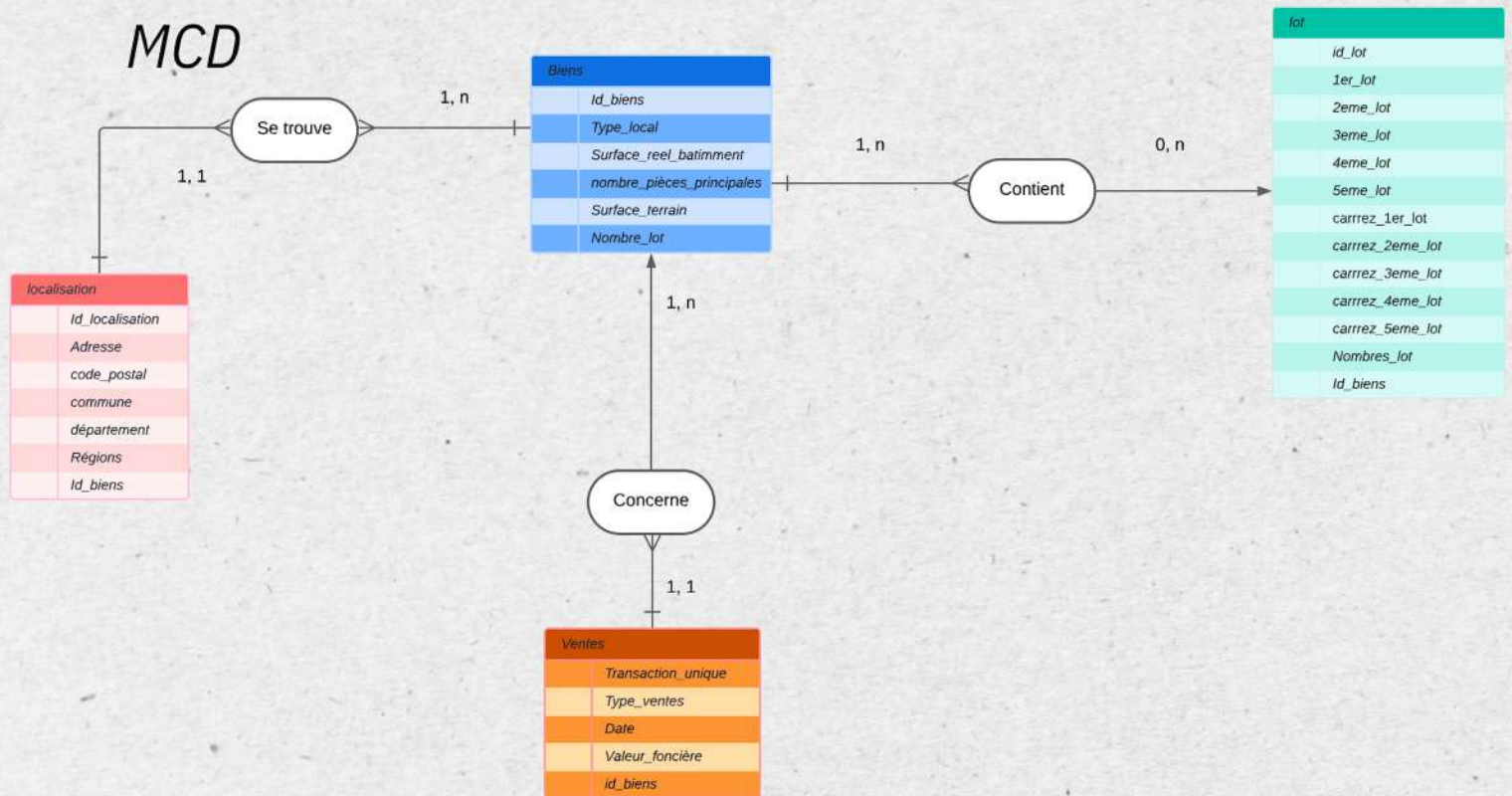
9/ Taux d'évolution du nombre de ventes entre le premier et le second trimestre de 2020.

	Taux_evolution
0	0.1315

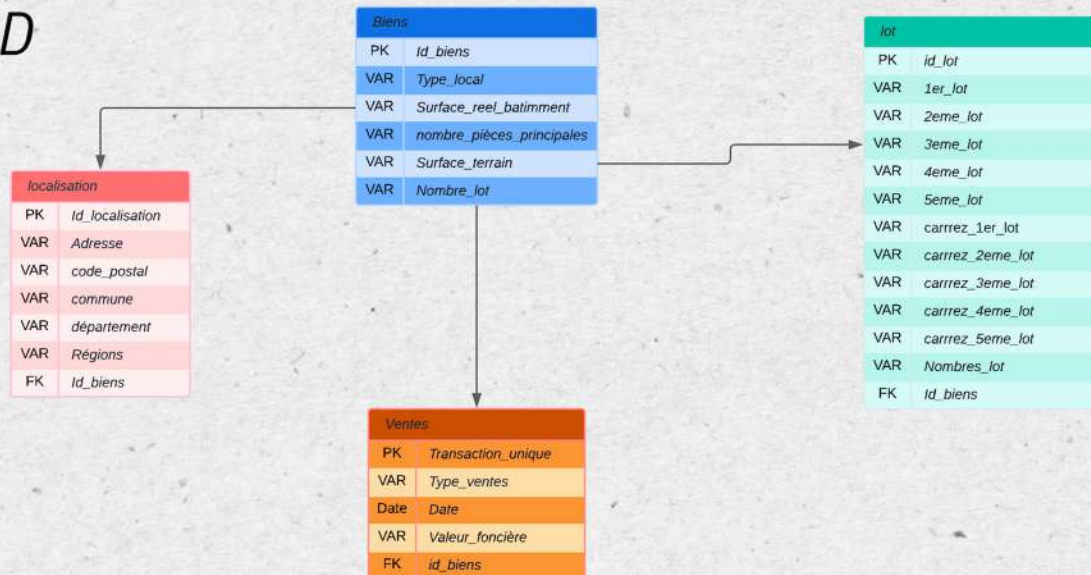


# MCD/MPD Ryan

## MCD



## MPD



## Conclusion

- Nous avons faits notre MPD autour de la table 'Biens', cependant nous aurions pu le faire autour de 'Ventes'.
- La valeur foncière global en moyenne a augmenté de 10% du 1er au 4e trimestre.
- Le territoire Français ayant des territoires d'outre mer, il faut parfois séparer les différentes zone pour avoir des visualisations et analyses correct. . Notamment la valeur foncière en moyenne qui a de gros écart entre les territoires.
- Le nombre de vente par mois chute soudainement des le mois d'août.