

Jedha
dse – od – 06

Projet DVF

Sommaire

- 01 Introduction et présentation
- 02 Choix techniques
- 03 Traitement des données
- 04 Résultats
- 05 Conclusion et ouverture

Introduction et présentation

2024 → Amiens



Trouver un bien immobilier pour acheter



Comment est définie la valeur foncière d'un bien immobilier dans la ville d'Amiens ?

Data : *DVF_80* soit le fichier des Demandes de Valeur Foncière dans le département de la Somme. Base provenant de *data.gouv.fr*

Choix techniques



SQL



Traitement des données

```
1 ALTER TABLE DVF_80
2 DROP COLUMN date_mutation,
3 DROP COLUMN numero_disposition,
4 DROP COLUMN adresse_numero,
5 DROP COLUMN adresse_suffixe,
6 DROP COLUMN adresse_nom_voie,
7 DROP COLUMN adresse_code_voie,
8 DROP COLUMN code_commune,
9 DROP COLUMN nom_commune,
10 DROP COLUMN code_departement,
11 DROP COLUMN ancien_code_commune,
12 DROP COLUMN ancien_nom_commune,
13 DROP COLUMN id_parcelle,
14 DROP COLUMN ancien_id_parcelle,
15 DROP COLUMN numero_volume,
16 DROP COLUMN lot1_numero,
17 DROP COLUMN lot1_surface_carrez,
```

```
1 DELETE FROM DVF_80
2 WHERE nature_mutation != "Vente"
```

```
1 DELETE FROM DVF_80
2 WHERE type_local LIKE "Dépendance"
3
4 DELETE FROM DVF_80
5 WHERE type_local LIKE ""
6
7 DELETE FROM DVF_80
8 WHERE valeur_fonciere ISNULL
9
10 DELETE FROM DVF_80
11 WHERE surface_reelle_bati ISNULL
```

DROP des colonnes non pertinentes

Je ne regarde que les ventes

**Suppression des NULL et des
"Dépendance"**

Résultats après retraitements de données et manipulations

1ère hypothèse

Relation directe entre la
valeur_fonciere et la
surface_reelle_bati

Train_set : 0.01

Test_set : -0.04

2ème hypothèse

Relation directe entre la
valeur_fonciere et la
surface_reelle_bati + le
type_bien

Train_set : 0.1

Test_set : 0.07

3ème hypothèse

Relation directe entre la
valeur_fonciere et la
surface_reelle_bati + le
type_bien et en
supprimant les
valeur_fonciere extrêmes

Train_set : 0.01

Test_set : 0.006

Conclusion et ouverture

- Remise en cause du jeu de données : manque d'exhaustivité des données
- Remise en cause du modèle : régression linéaire pas pas adaptée
- Remise en cause du contexte : volatilité du marché

Ouverture

Pensant qu'il était possible de comprendre le fonctionnement d'un marché, la réalité est beaucoup plus complexe. Des études seraient à refaire avec d'autres sortes de modèles et plus de données (un traitement moins strict de mes colonnes de bases)

Jedha
dse – od – 06

Merci de votre attention
