



#### Scénario

Le but est de construire un modèle de prédiction de biens immobiliers.

Télécharger les données foncières de l'année 2021 sur cette page : <https://www.data.gouv.fr/fr/datasets/demandes-de-valeurs-foncieres/>

Une agence immobilière de Toulouse souhaite faciliter le travail d'estimation de ses agents, en fournissant un outil de prédiction de prix de vente.

En discutant avec lui, vous apprenez que le responsable est déçu de sa dernière recrue. En effet, il a réalisé de nombreuses erreurs.

Par exemple, il avait conseillé un prix de vente à 200,000 euros alors que le bien valait au moins 250,000. Cette erreur n'est pas acceptable. Il cite ensuite un autre exemple dans lequel un appartement a été estimé à 150,000 alors que le concurrent l'a correctement estimé à 175,000. C'est dommage, car le responsable estime que même en l'estimant à 170,000, ils auraient pu gagner le contrat. Le contact avec le client avait été excellent.

#### Instructions

En vous appuyant sur la méthodologie abordée dans la formation, construisez le meilleur modèle afin de déterminer le prix de vente d'une maison ou d'un appartement dans la ville de Toulouse.

Chacun de vos choix devra être justifié. Les éléments suivants devront être présents :

- justification de la métrique d'évaluation et définition de seuils d'acceptabilité (excellent, bon, moyen, mauvais)
- exploration et description des données
- nettoyage et préparation des données selon leur type
- mise en place d'un modèle de référence
- comparaison d'au moins 3 algorithmes
- hypertuning
- sauvegarde du modèle en vue d'un déploiement



### Introduction au Machine Learning

Une conclusion est également à fournir sur le client présentant votre modèle, ses points forts, ses faiblesses et comment il pourrait être amélioré lors d'une prochaine mission.

Conseils :

- Pour cet exercice, n'hésitez pas à supprimer les colonnes et les lignes qui ne vous semblent pas pertinentes (en tout cas, dans le cadre de cet exercice au temps limité)
- Pour débiter, conserver une dizaine de colonnes, celles qui vous semblent, a priori, les plus pertinentes
- N'ayez pas peur de faire des hypothèses simplificatrices pour cet exercice.

