



Compagnon Immobilier

Cursus concerné : Data Scientist

Niveau de difficulté : 7/10

Description du projet :

Développer une solution permettant aux acheteurs de logements d'explorer et de comparer différents territoires en termes de prix de l'immobilier, démographie, transports, services, éducation, criminalité, et économie.

L'application doit offrir une Data Visualization, permettant aux utilisateurs d'établir des classements et de visualiser les forces et faiblesses relatives des territoires.

Deux objectifs principaux sont à considérer :

1. La prédiction de l'évolution du prix des logements selon les territoires.
2. L'estimation du prix au m² d'un logement donné. Une première prédiction peut être effectuée à l'aide des données tabulaires disponibles, et peut être approfondie avec des données relatives à l'annonce du bien, comme le texte descriptif ou les photos du logement.

L'objectif global est d'aider les acheteurs à prendre des décisions éclairées en traduisant des données complexes et nombreuses en informations utiles et accessibles.

Ressources à consulter :

- **Données :**
- Annonces de vente/location :
 - [https://raw.githubusercontent.com/klopstock-dviz/immo_vis/master/
data/ech annonces ventes_68.csv](https://raw.githubusercontent.com/klopstock-dviz/immo_vis/master/data/ech annonces ventes_68.csv)
 - [https://raw.githubusercontent.com/klopstock-dviz/immo_vis/master/
data/ech annonces locations_68.csv](https://raw.githubusercontent.com/klopstock-dviz/immo_vis/master/data/ech annonces locations_68.csv)
 - Prix, surface, nb de pièces, type, position géographique (France des 5 dernières années):



<https://www.data.gouv.fr/fr/datasets/demandes-de-valeurs-foncieres-geolocalisees/>

- Autres données open source (Ventes DVF, Insee, Data gouv et diverses API publiques) à scraper

- **Ressource additionnelle :**

Une application en lien avec le sujet :

<https://comparateur-communes.fr/>

Conditions de validation du projet :

- un **rapport** d'exploration, de data visualisation et de pre-processing des données ;
- un **rapport** de modélisation ;
- un **rapport** final et le **GitHub** associé.