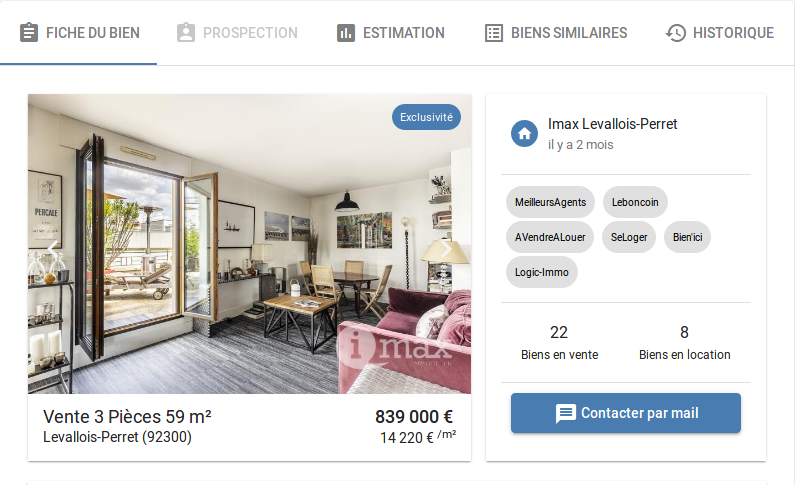
# Test Technique - Data Engineer/Scientist

Afin de valider vos compétences techniques, voici un fichier contenant l’ensemble des annonces immobilières à la vente en cours de publication sur la ville de Levallois-Perret :

<https://docs.google.com/spreadsheets/d/1XUjqeXVgjZJ8jVNAn9MeIb9zHLhVj4mwRs-hk1P050o/edit?usp=sharing>

Nous vous proposons de regrouper les annonces immobilières portant sur un même bien immobilier. En effet, plusieurs professionnels (voire particuliers) publient sur plusieurs portails immobiliers des annonces portant sur le même bien immobilier et il y a un intérêt fort de les dédoublonner pour rendre la recherche plus efficace mais aussi pour avoir des statistiques plus pertinentes.

Voici une capture d’écran d’annonces déjà dédoublonnées depuis notre application :



Nous attendons de votre part à la fois le code source, sous forme de notebook Python de préférence, décrivant votre méthodologie, votre algorithme et les pistes d’améliorations. Envoyez votre notebook en répondant au mail reçu.

Le code n’est pas une finalité, ce qui nous intéresse est aussi votre méthode de travail, votre analyse des données, ainsi que la restitution des résultats.

Vous pouvez aussi regarder la documentation de notre [API](https://www.yanport.com/data/api/documentation) qui permet de mieux comprendre les modèles de données (le modèle ‘Ad’).