

Raport Final — Analyse des facteur influencant le prix des logement en Californie

### 1. Introduction

Dans ce projet, j'ai travaillé comme Data Scientiste (lol) pour comprendre qu'est ce qui influence le prix des maisons en Californie. Le marcher immobilier est compliquer et dépend de plusieurs choses comme la surface, l'age des maisons, le revenue des habitants et aussi ou elle se trouve. Le but c'était de trouver quelles caractéristique ont le plus d'impact sur le prix, pour aider les agences immobilières à mieux viser leur client.

### 2. Methodologie

J'ai utilisé le jeu de données "California Housing Prices" (pris de Kaggle). On a chargé les données avec la fonction pandas `read_csv("housing.csv")` puis affiché les premières lignes avec `df.head()`. Pour voir des infos rapides, j'ai utilisé `df.describe()` qui donne la moyenne, médiane, écart-type etc. Pour visualiser j'ai fait : - `sns.histplot()` pour voir la distribution des prix - `sns.heatmap()` pour voir la corrélation entre les colonnes - `sns.boxplot()` pour comparer le prix par âge médian - Un `st.slider()` dans Streamlit pour filtrer par plage de prix - `st.dataframe()` pour montrer les données filtrées

### 3. Resultat principaux

Après analyse, on a trouvé que : - Le revenue médian est très lié au prix (plus le revenue est haut plus les prix sont haut) - Les grandes maisons avec plus de chambres coûtent plus cher - Les maisons récentes coûtent plus cher que les vieilles - La localisation joue beaucoup (certains endroits sont toujours plus chers)

### 4. Recommendations

- Se concentrer sur les quartiers avec revenue haut
- Mettre en valeur les logements grands avec plusieurs chambres
- Proposer des logements récents ou bien rénovés
- Adapter le marketing selon les zones géographiques
- Utiliser des outils comme notre app Streamlit pour suivre les tendances

### 5. Conclusion

Ce projet m'a appris à manipuler pandas, seaborn et streamlit ensemble pour analyser un jeu de données. L'application finale est simple mais utile pour voir les relations entre variables et prendre de bonnes décisions.