



L'IA au service des agents immobiliers - Silicon Valley

Sommaire

- Contexte
- Résultat
- EDA
- Modèle
- Inférence
- Itérations
- Organisations

Contexte



Résultats

Linear Regression

```
base r2 : 0.6348249131619182
base_rmse : -69612.35723615656
0.7013336656722939
-0.3109894876297067
difference entre base r2 et nouveau r2 : 0.06650875251037569
difference entre base rmse et nouveau rmse : 69612.04624666894

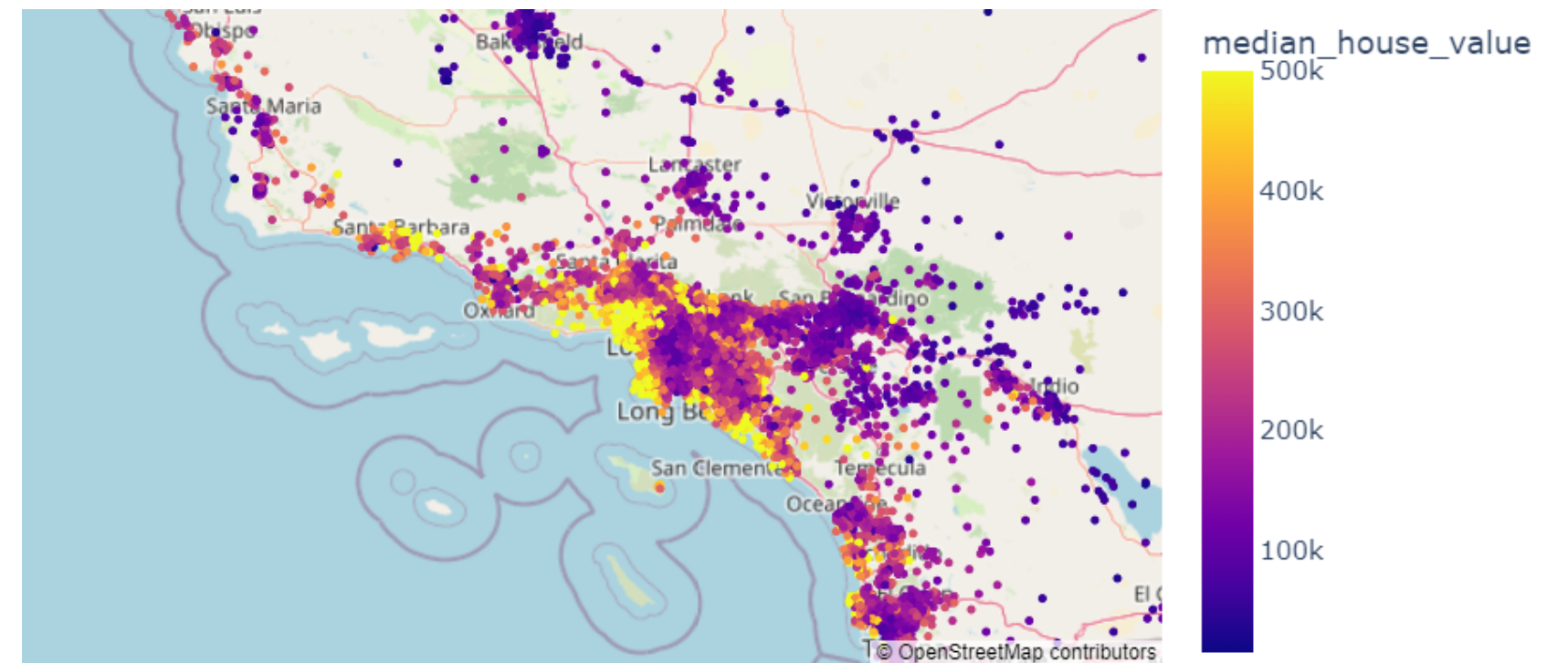
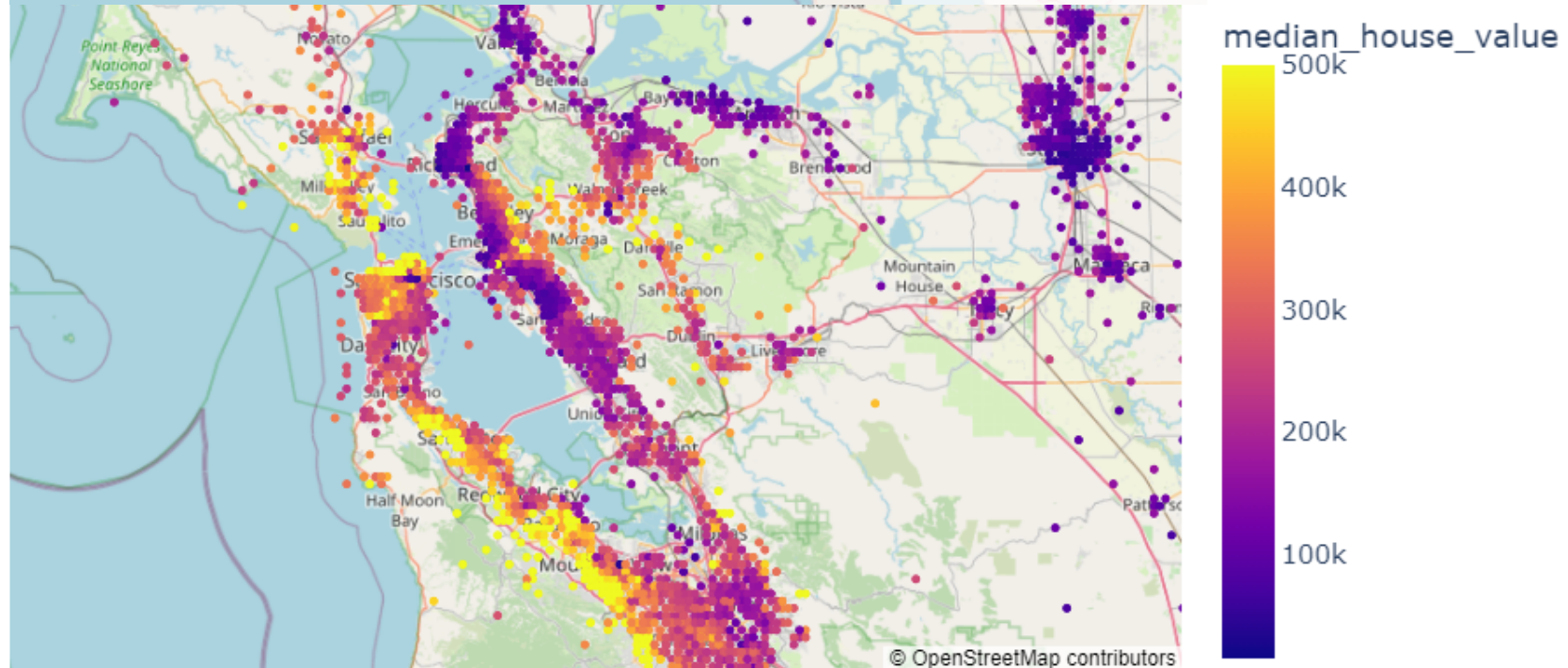
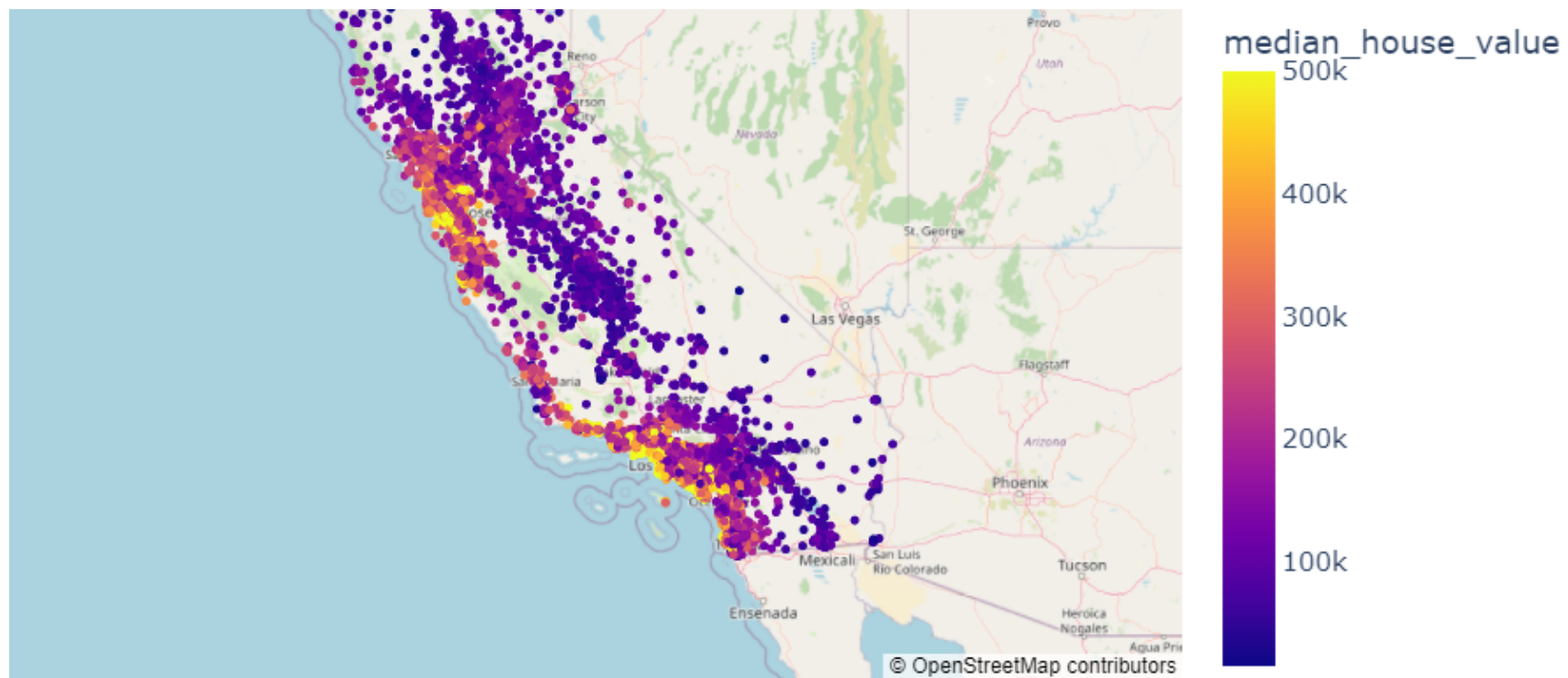
69657.48879271372
```

Random Forest

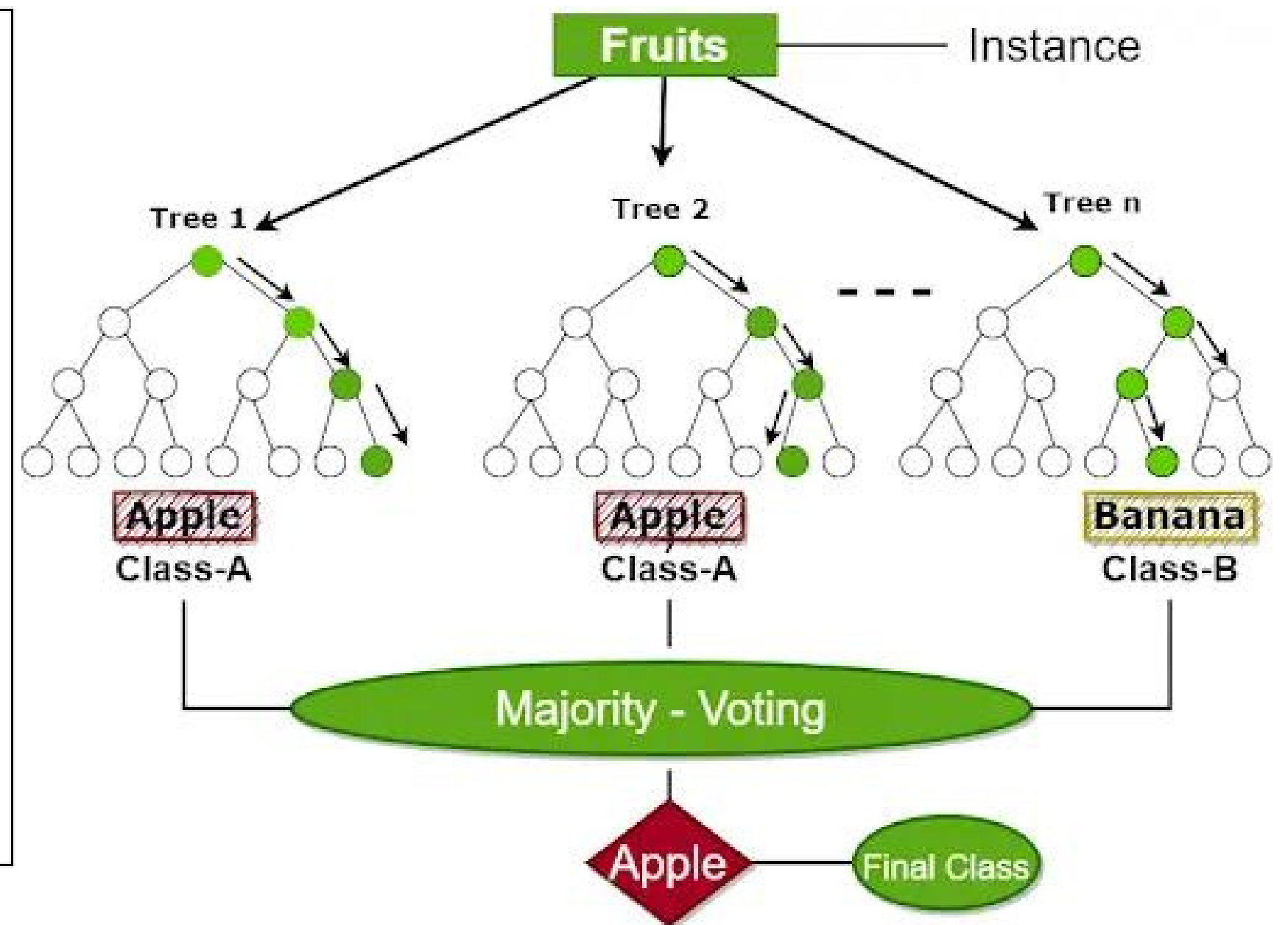
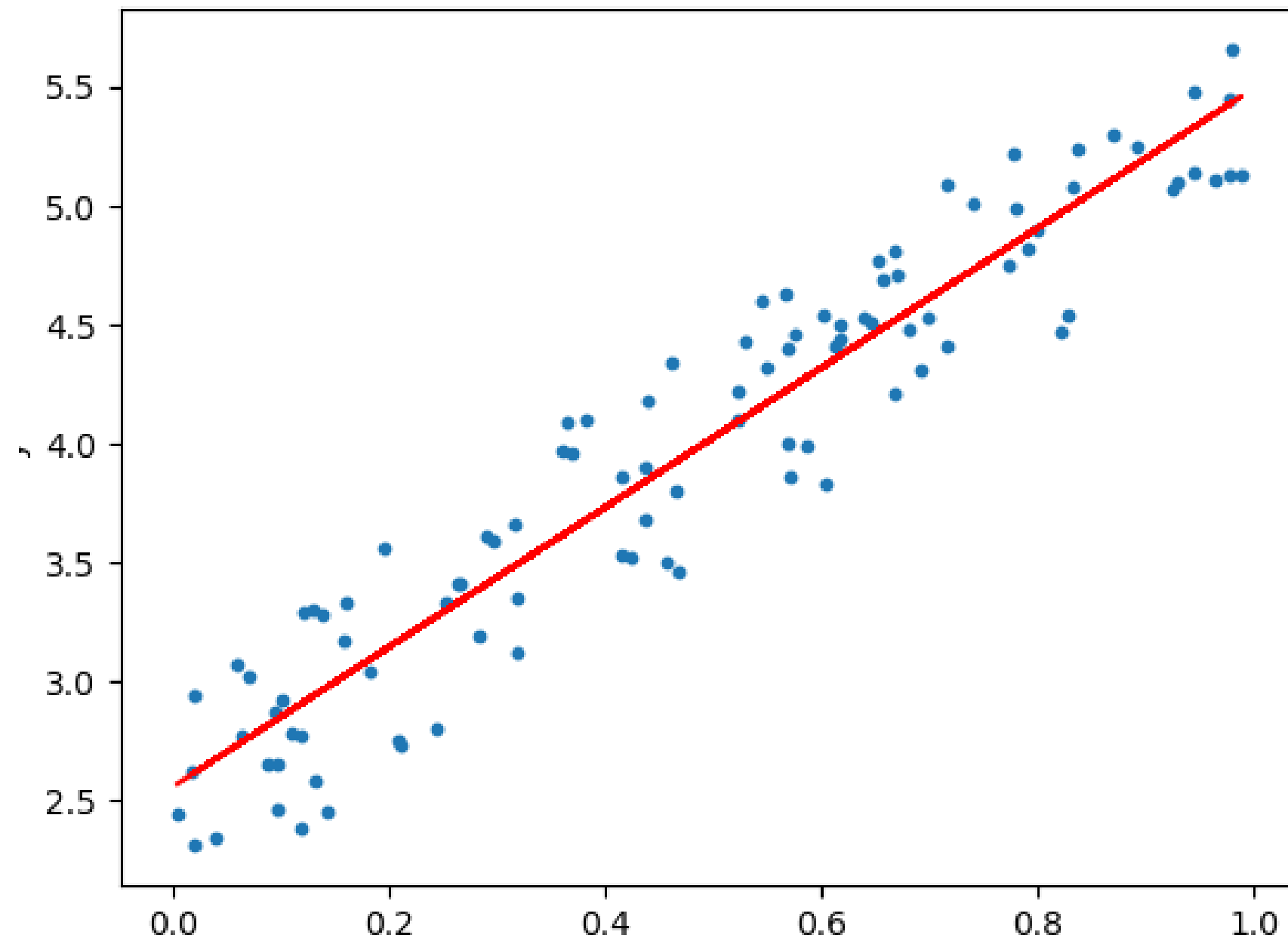
```
base r2 : 0.6348249131619182
base_rmse : -69612.35723615656
0.8497168493745553
-0.2204545221012949
difference entre base r2 et nouveau r2 : 0.21489193621263702
difference entre base rmse et nouveau rmse : 69612.13678163446

47377.257958645474
```


Exploration de données



Modèle



Inférence

Pearson

l'hypothèse de base est qu'il n'y a pas de corrélation
entre une feature et la target

```
group_house_age_0,10  
(-0.01452489003499134, 0.06198662722408942)  
group_house_age_20,30  
(0.004096274817808571, 0.5986591587949732)  
group_house_age_30,40  
(7.176077340653036e-05, 0.9926432164308154)  
group_house_age_40,50  
(-0.0021133445385515615, 0.7859738461289058)
```

Nous avons ici 4 features qui ne sont pas corrélé a la
target et statistiquement insignifiant

Itérations

Missing value, duplicates, imputation

Recherche des valeurs manquantes et des duplicats.

Remplacement des Nan par la moyenne

Outliers avec zscore

Recherche des outliers avec un Zscore de 3 pour garder

99.7% des données


Encoding, scaling

Encoder la variable qualitative.

Normalisation et standardisation des données


Feature engineering et feature selection


Selections de certaines features et enrichissement de la données





Projet immo ML

L'objectif est de créer un modèle avec vos données (train) pour prédire la valeur du prix médian des maisons par district / bloc (medianHouseValue). A la fin du projet, vous devez évaluer ce modèle avec les données (validation) que seul votre client dispose (le prof).

 Board View

 Galerie

 Table



<div>To Do1</div> <div><div>12/13 janvier</div><div>25 janvier</div><div>1/2 février</div><div>présentation le 3</div><div><div><div>✓</div>intro-machine-learning</div><div><div>✓</div>intermediaire-machine-learning</div></div></div> <div><div><div>📄</div>Jours avec Arturo</div><div>To Do</div></div>

+ Nouvelle page