



Immobilier Ellana

IA au service des agents immobiliers

Un guide pour acheter votre maison de La Silicon Valley





SOMMAIRE

- Planning
- EDA
- Modèle de régression linéaire
- Inférence statistique
- Modèle de prédiction

Planning



1

4

2

3

🔧 Avancées 12,13,25 jan et 1, 2 fev

Par statut Épopées Calendrier des épopées Sprints 5 de plus... Filt

Pas commencé

8

...

En cours

7

Terminé

16

<https://www.kaggle.com/learn/feature-engineering>

faire hypothèse de départ

présenter un planning de travail

faire un analyse d'inférence statistique sur les données (notebook_inference)

présenter la procédure suivie pour trouver un modèle adapté (notebook_model)

proposer un protocole d'utilisation du modèle pour faire des prédictions y compris les étapes de data cleaning (notebook_predict)

faire des push

sélectionner les données (features) pertinents pour le modèle (dans le notebook model)

<https://makina-corpus.com/data-science/initiation-au-machine-learning-avec-python-la-pratique>

<https://mrmint.fr/multivariate-regression>

finir le readme du github

créer un modèle d'IA entraîné et adapté au problème (format pickle)

créer un repo git et un espace sur github pour le projet

créer canva

[Eclaircissement](#)

10

préparer dossier dans ellanaca avec données de départ et notebook

- techniques de data cleaning : savoir traiter duplicates, missing data, outliers, scaling, balancing, encoding, discretizing, feature engineering

notebook eda notebook inference et notebook model

EDA

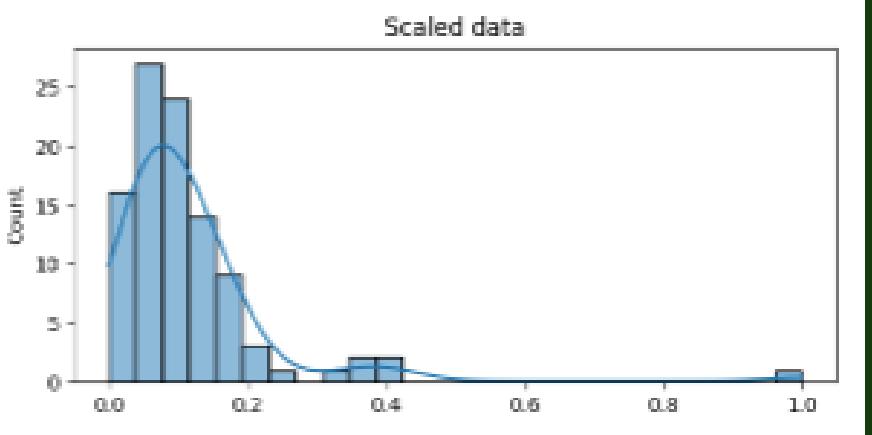
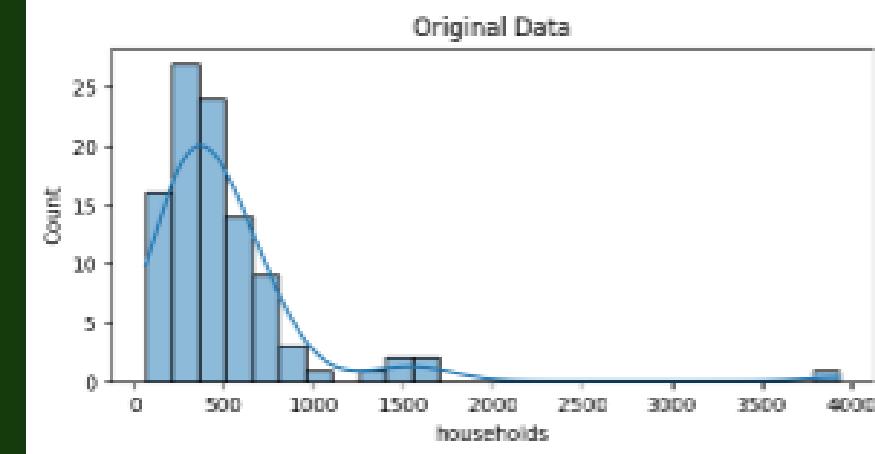
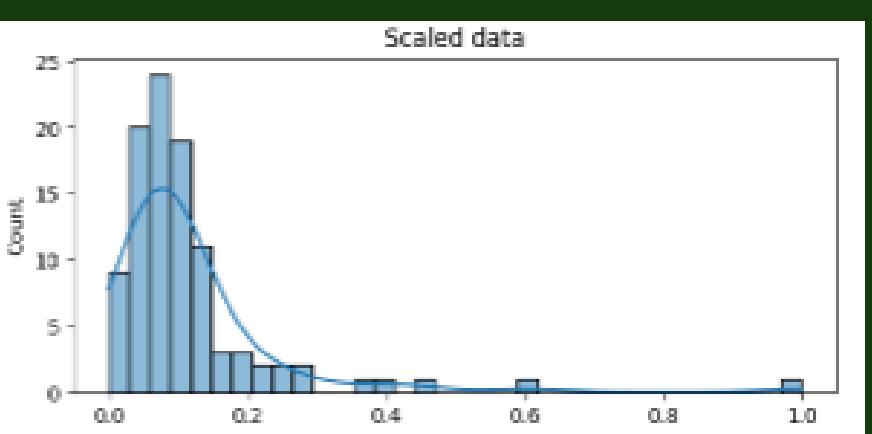
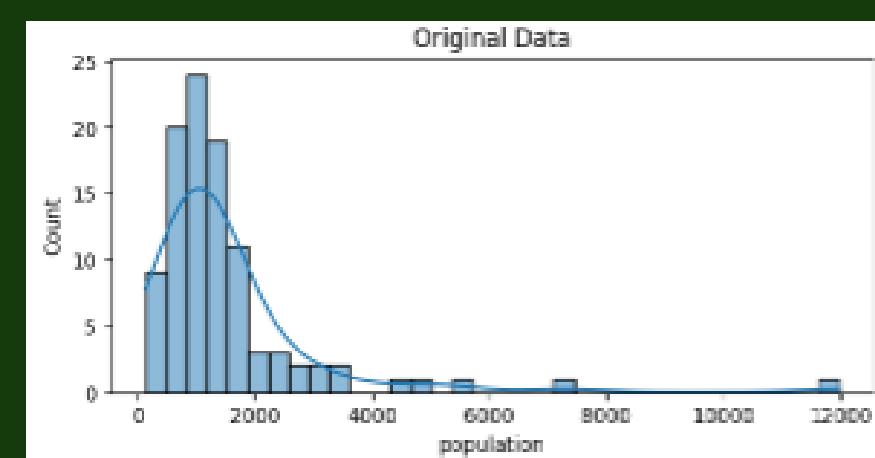
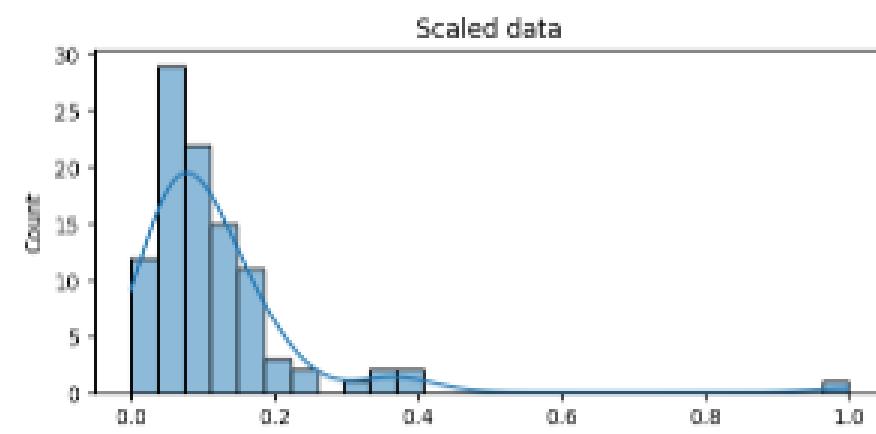
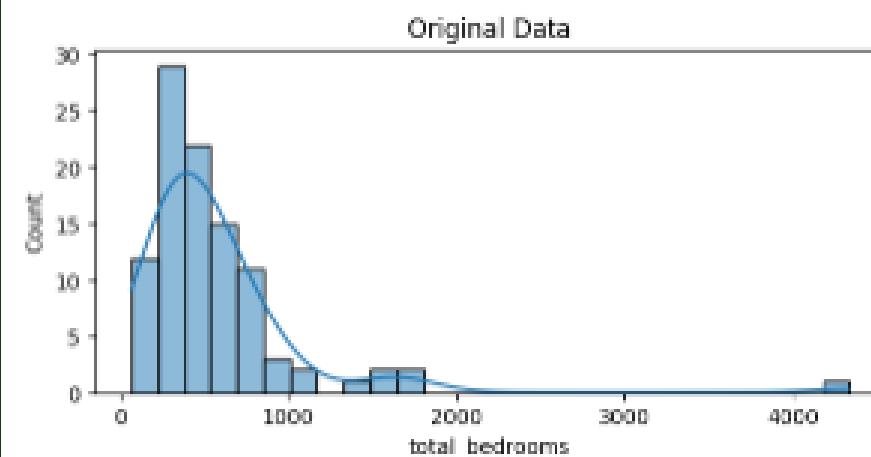
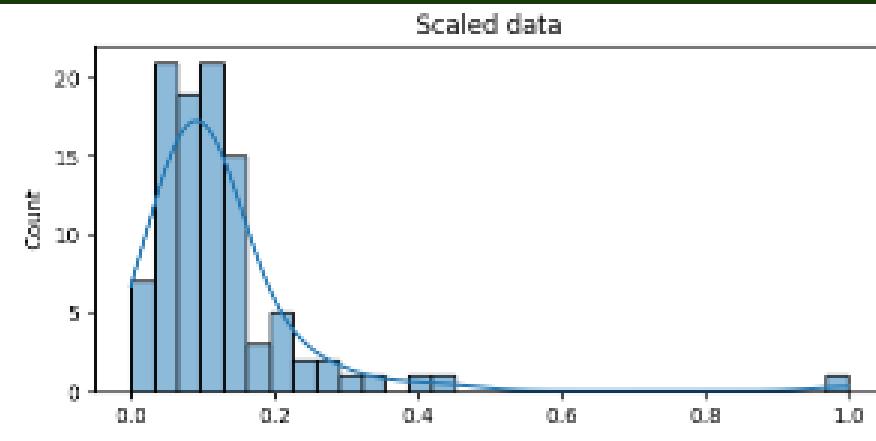
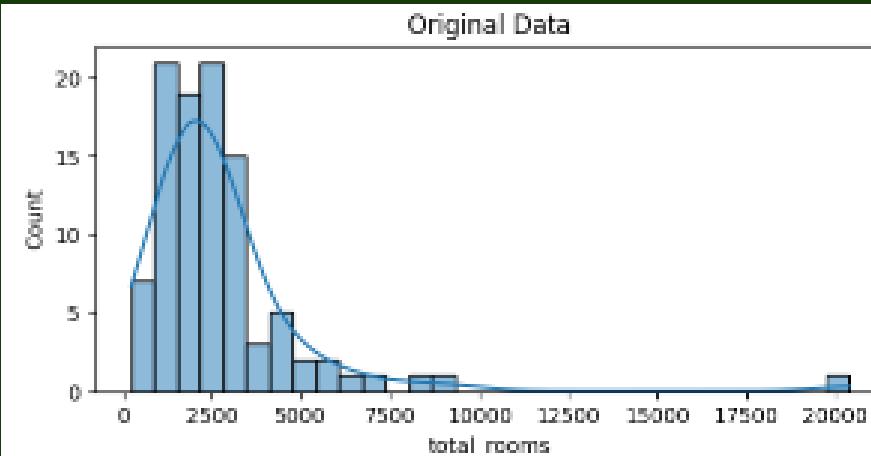


1

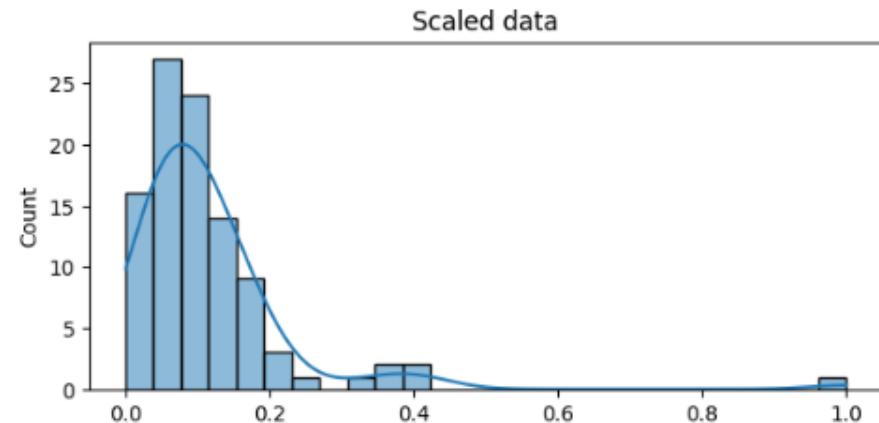
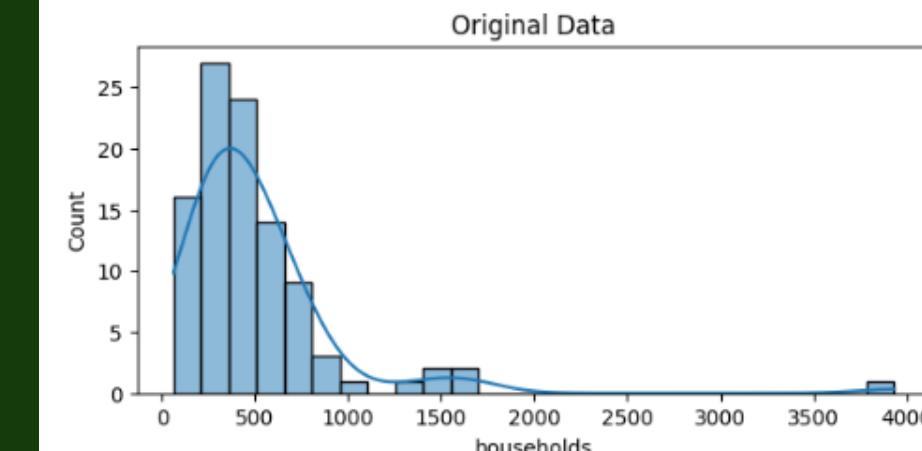
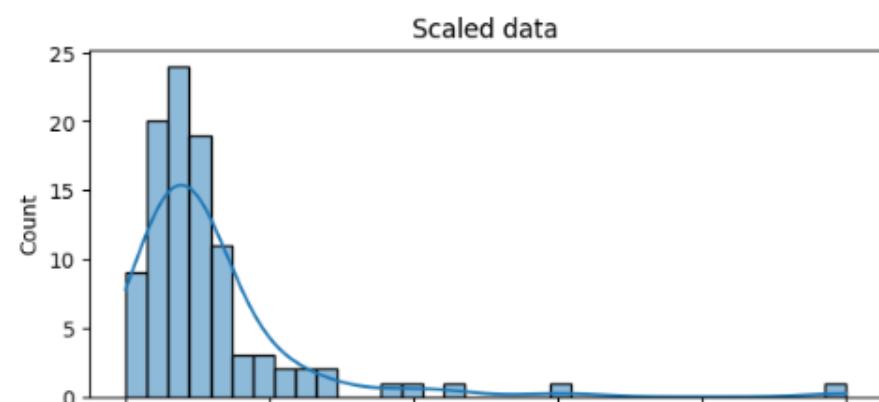
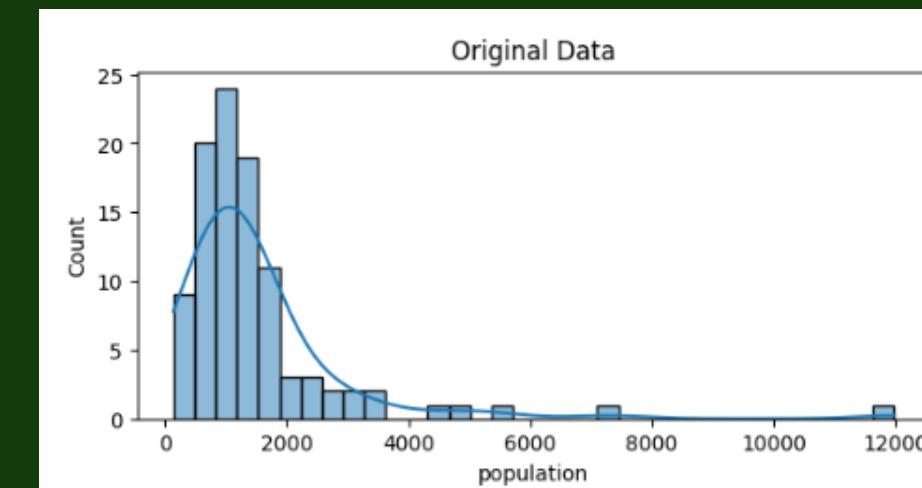
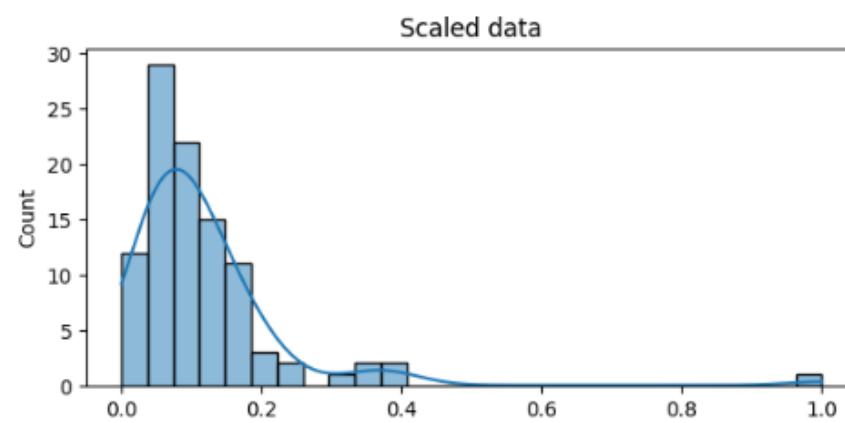
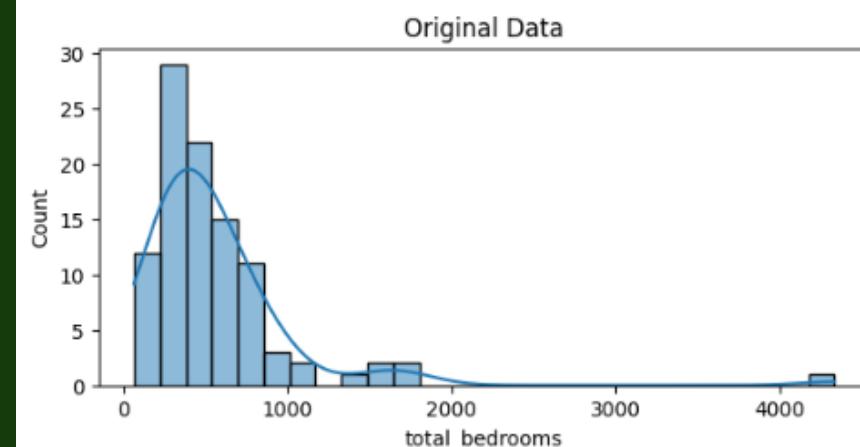
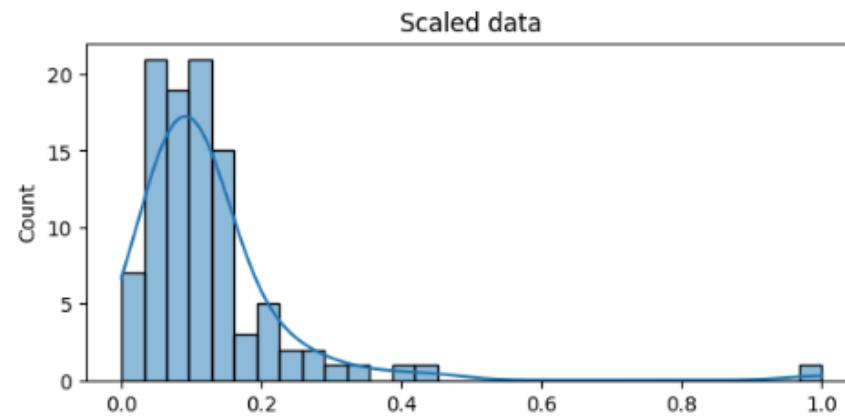
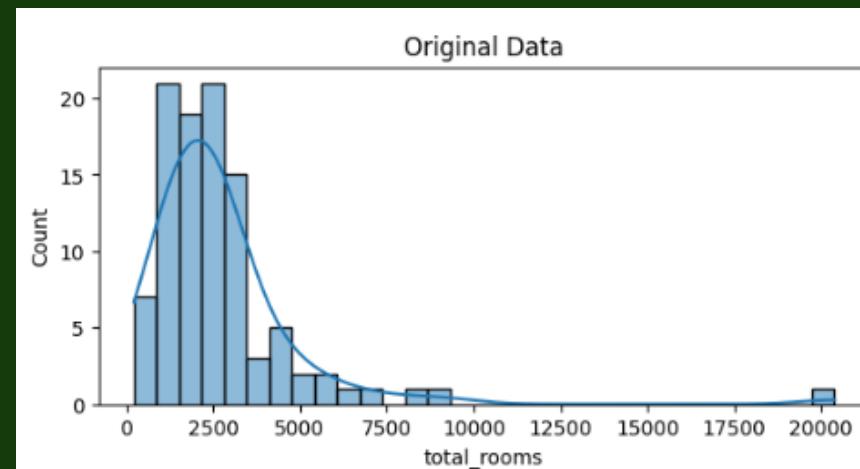
4



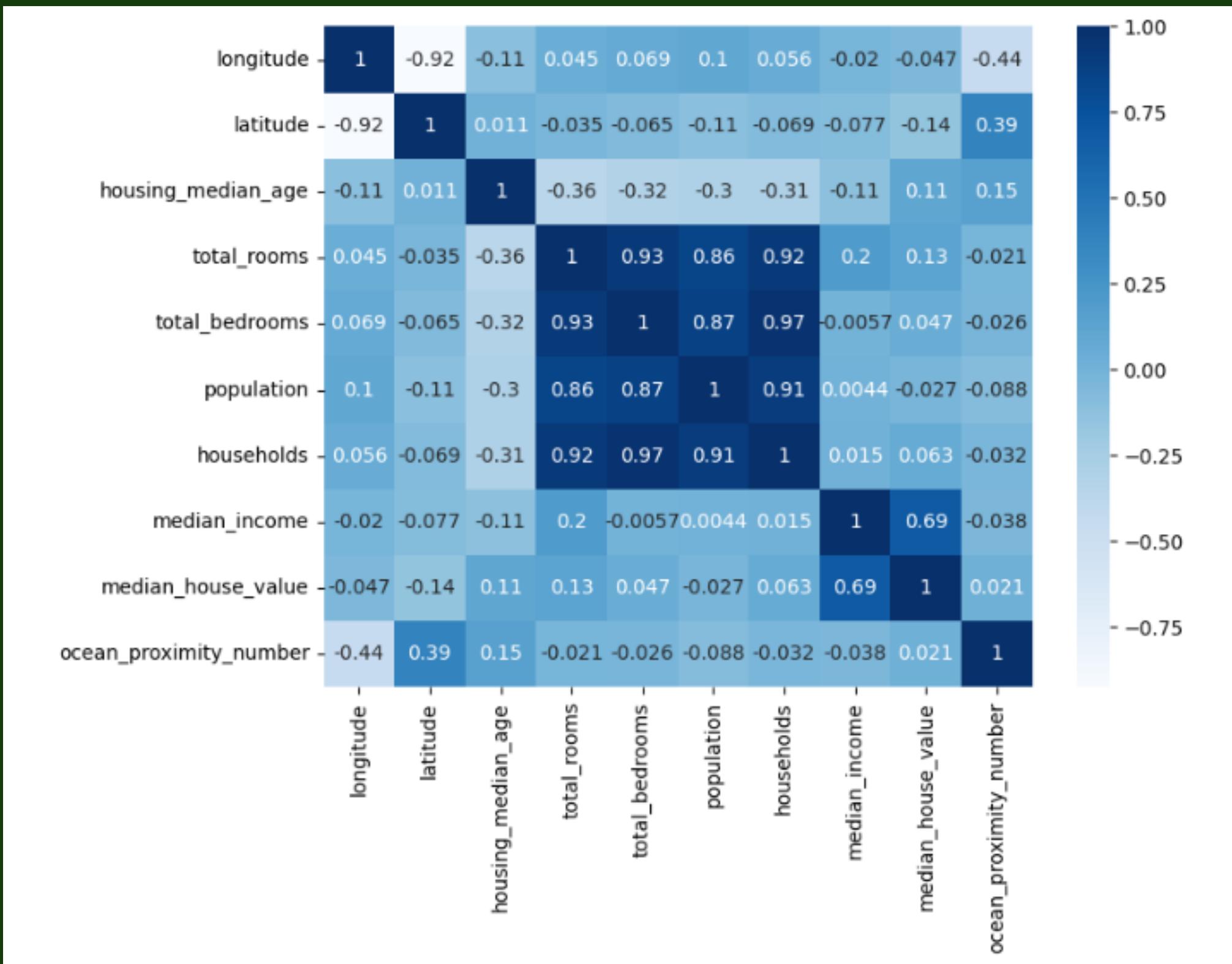
SCALING VISUELS



NORMALIZATION VISUELS



HITMAP



NETTOYAGE

Imputation des données manquantes:

- par la médiane
- par la moyenne

Modèle de régression linéaire

4

1

ITERATIONS

- Données de base pour baseline
- Délétions des outliers
- Imputation des données: mean, median
- Encoding avec le One Hot Encoder

Inférence statistique

4

1

INFERENCE

- Pearson

Modèle de prédiction

4

1

MODELE DE PREDICTION

- 2 fonctions que j'ai réuni pour faire un pipeline
- Puis export dans un fichier pickle

Ouverture



1

4

OUVERTURE

- Log
- Balancing
- Utiliser un pipeline Scikitlearn
- Utiliser Scikitlearn pour les imputations

Merci !