Questions :

1. Réaliser une courte analyse exploratoire.

- De l’univariante sur les colonnes considérées + une ACP suffirait ? Je peux aussi faire une bivariante avec un boxplot, fais gaffe aux colonnes String, mapping first ? Addresse ?

Questions :

-

2. StringTransformers

- Avec ou sans l-ajout des strings, ou tu fais simplement 2 pipelines

- Un qui definit les colonnes que tu ajoutes ?

- EncodingMode : [OneHotEncoder](https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.OneHotEncoder.html" \l "sklearn.preprocessing.OneHotEncoder) ou [TargetEncoder](https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.TargetEncoder.html" \l "sklearn.preprocessing.TargetEncoder) ou [OrdinalEncoder](https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.OrdinalEncoder.html" \l "sklearn.preprocessing.OrdinalEncoder)

Check this : [https://scikit-learn.org/stable/auto\_examples/preprocessing/plot\_target\_encoder.html#sphx-glr-auto-examples-preprocessing-plot-target-encoder-py](https://scikit-learn.org/stable/auto_examples/preprocessing/plot_target_encoder.html" \l "sphx-glr-auto-examples-preprocessing-plot-target-encoder-py)

C’est tout. Puis tu cr ees tes pipelines, ils vont servir a preparer les donnes.

Suis le livre ensuite, tu es prêt à essayer plusieurs modèles.

Commence simple avec peu de colonnes que tu ajoutes, tu buildes ensuite par-dessus. C’est bien d’avoir un pipeline qui marche avec un modèle et plus pour mercredi comme ça tu peux poser les questions.