

CASE STUDY 019

[Python]

Kaggle House Prices: Advanced Regression Techniques Competition



Here are some clues in case you are stuck with the case study:

- 1- Remember, you need to combine two datasets before doing anything. Pandas has a function it.
- 2- When dealing with missing values, it is up to you, you can keep whatever columns you want, but make sure to have reasoning behind it
- 3- Filling the missing values should make sense. Again, it is up to you, you can choose how to fill them and which ones to drop or fill. To drop columns, you can use `pd.dropna()`.
- 4- To normalize a column, you can use log function
- 5- For plotting, you may consider using seaborn. It is really powerful package. For data cleaning, please use pandas

Good Luck!