Problem Statement

Once there lived an atrocious King with the finest sword a man could bear at that time.

Alzar, the record keeper, lost papers that had prices for houses in the kingdom.

As he trembled with mortal fear, he went to Elric the sorcerer seeking for help.

"King is very specific and rather precise with numbers!" exclaimed Elric seeing the records.

Fortunately some records were still present, but they were too scattered!

King has commanded Alzar to present to him the complete record with price of each house against its unique ID.

Now Elric invites you through time travel to help poor Alzar lest he should lose his life to sword.

Alzar will present to you the information that he has.

1) Each paper is specific to one builder family with details of houses that they built.

2) Alzar has sorted for you the house details with builder family name and 'Not Known' where builder's information was lost. "But certainly there are only ten builder families" he remarks.

"Careful! Black Magic has scraped off some more data from the records" says Elric as you begin to think upon…

<u>FORMALLY:</u> Files contain information about 20k houses in total however 16.5k houses have price information against them (house_prices.csv) and 3.5k houses are without price information (missing.csv).

You need to use the former .csv file for training and testing your model. Finally, predict the prices for 3.5 k houses and generate a .csv file similar as house_prices.csv

Problem set by Shubham Pandey