Test task for Data Scientist position

We hope it will help us to build further communication and think it will be interesting for you. Please write your response in English. Looking forward to your answer!

Goal:

The goal is to build a model for accommodation price prediction. The challenge is determining the optimal nightly rent price. In many areas, would-be renters are presented with a good selection of listings and can filter by criteria like price, number of bedrooms, room type, and more.

Dataset:

This data file includes all needed information to find out more about hosts, geographical availability, necessary metrics to make predictions and draw conclusions.

Dataset contains 45 columns. 44 - independent variables and 1 dependent. The dependent variable is in the **Price** column. There are different types of data: numeric, text, datetime. Feel free to use any of them for the best result.

Train dataset

Train data **Train.csv** consists of 3013 records for data analysis and model training.

Submission dataset

Submission dataset **Submission.csv** consists of 618 records for your prediction only. The independent columns are presented. You need to add an additional column - **price_prediction** based on your model results.

Key points:

1. Provide with useful insights from the data
2. Find the top impact factors on the price in the data
3. Select the best metric for the model evaluation and justify the choice
4. Prepare Jupyter Notebook with clear research and descriptive comments
5. Describe the business value of your results