أكاديمية سدايا
SDAIA Academy

Agoda linear regression

1. Abstract

- the data relates to room prices in agoda website , the goal of this project is to implement machine learning (regression) algorithms, we will predict the room prices based on their features.

2. Design

- does the bedroom count affect the price ?

-does the bed count affect the price?

3. Data

The data to be tested in this project are scraped from Agoda .com
Website link https://www.agoda.com/ar-ae/ -

- Dataset shape [2247raws-6columns]

4. Algorithms

1. Problem understanding
2. Data collection
3. Data preparation
4-Build the model.
5-Train the model
6-Evaluate the model
7-choose the right model(conclusion)

5. Tools

• Technologies

- Python
- Jupiter Notebook

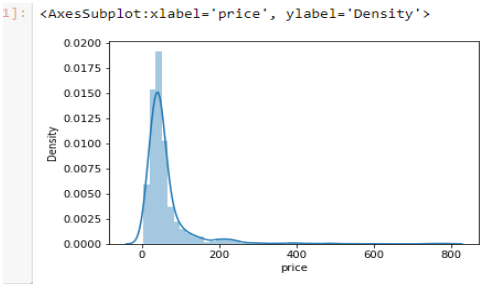
• Libraries

- NumPy
- Pandas
- Matplotlib
- Seaborn
- Sklearn
- Math Libraries

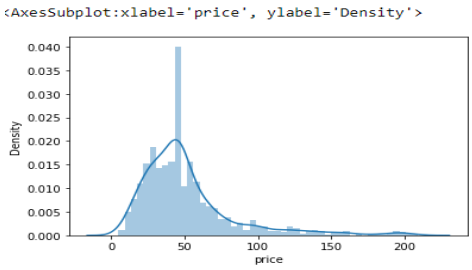
6. Communication

- there were a lot of outliers in the data

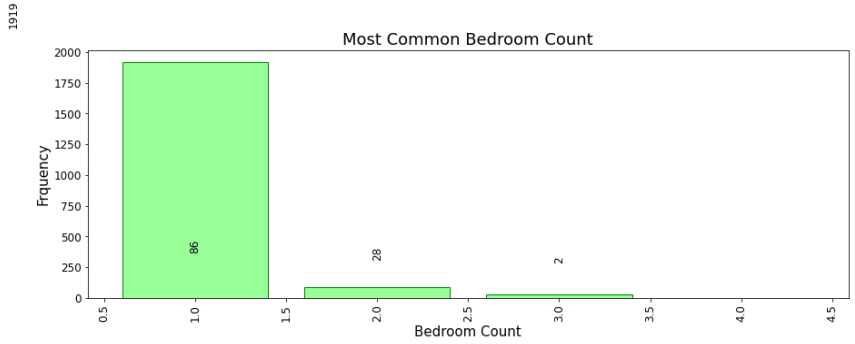
Before cleaning



after cleaning

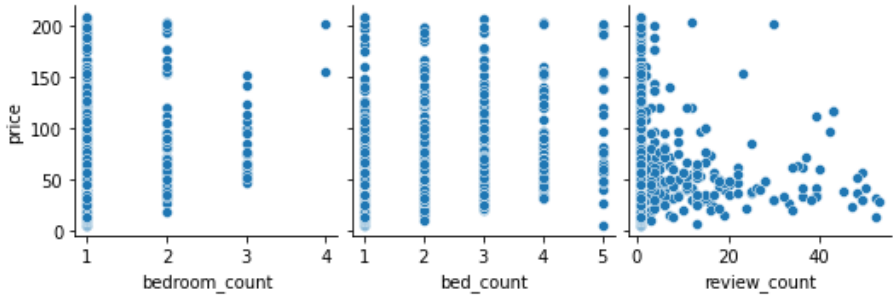


- Most common bedroom count and frequency



- The relation between all the features and the target

]: <seaborn.axisgrid.PairGrid at 0x7f8921f7a5e0>



6. Conclusion

The best model is Decision Tree then Polynomial.

REGRESSION MODELS

Model Type	Cost Function training	Cost Function predict
Polynomial	0.18	0.13
Ridge	0.16	0.15
Lasso	0.18	0.17
Decision Tree	0.21	0.16