

Abstract

This project aims to analyze the types of rooms in the site Airbnb and find out the average prices for them in each city, guests and hosts have uses Airbnb to expand on traveling possibilities and present more unique, personalized way of experiencing the world. The used data in this project are provided by Kaggle.

Design

This project is one of the T5 Data Science BootCamp requirements. Data provided by Kaggle has been used in this project.

Data

The dataset is provided in .csv format. It contains 48895 rows, each row has 16 columns. The most features of this project are city, room type, price and availability. This dataset describes the listing activity and metrics in NYC, NY for 2019.

Algorithms

Feature Engineering

- Cleaning the dataset by dropping some columns are id, host_id, latitude, longitude, last_review and rename name and neighbourhood_group to name_host and city.
- Cleaning the missing values by drops rows.
- Cleaning the duplicated by dropping duplicated.

Tools

- Numpy and Pandas for data manipulation
- Matplotlib and Seaborn for plotting

Communication

The slides are provided [here](#) .