# Two Sigma Connect: Rental Listing Inquiries

### 1. Introduction

Renting houses is an issue for people who travel to new place and plan to settle down. In our team, every teammate has experience of travelling to different places and searching for an ideal house. A lot of information needs to be considered so that the decision is made, but information bursting from Google, Craigslist increases the difficulties of decision-making. The decision making process takes people a lot of time to do background research of house, to communicate with different agencies and to negotiate price with house owners. The whole process of finding a house is time and energy consuming.

The project aims to simplify the decision making process and to increase the time and money efficiency so that users can find a high performance house based on their interest level. In short, the objective of the project is to optimize the user's’ searching experience in house searching area. From another perspective, the project also aim to make resource allocation sensible. In house market, because of lacking enough information, renter can’t find an ideal house and house owner can’t find an affordable renter either. Therefore, the object of the project is to build up a communication bridge between house owner and renters so that resources can be found out. Moreover, in the process of the project, features like bathrooms, address, latitude will be taken into account so that complete the standard of users. In other words, the objective is to enrich the searching standard by adding detail features. Furthermore, different background of our teammates contributes to the solution of project and interpersonal communication will be improved during the process of learning by doing. The objective of the project can be improved through communication between different background and practice the skills learned from the class.

To conclusion, the objectives of the project are the following:

1. Simplify the decision making process and to increase the time and money efficiency;

2. Build up a communication bridge between house owner and renters;

### 2. Dataset

The dataset has 49352 samples, and 15 features for each sample. The features are as following:

• bathrooms: number of bathrooms.

• bedrooms: number of bathrooms.

• building id

• created

• description

• display address

• features: a list of features of the room.

• latitude

• listing id

• longitude

• manager id

• photos: a list of photo links.

• price: in USD.

• street address

• interest level: this is the target variable. It has 3 categories: ’high’, ’medium’, ’low’.

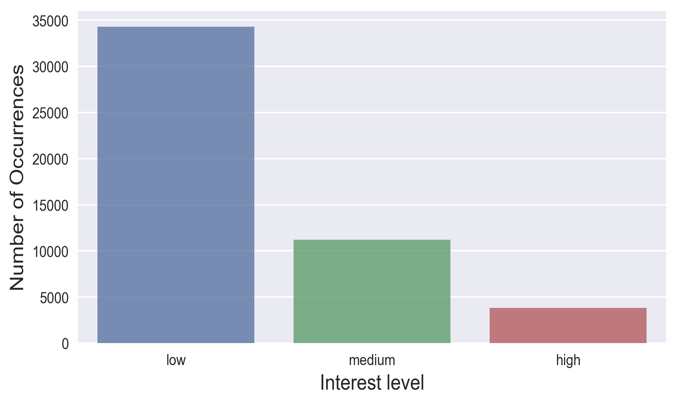
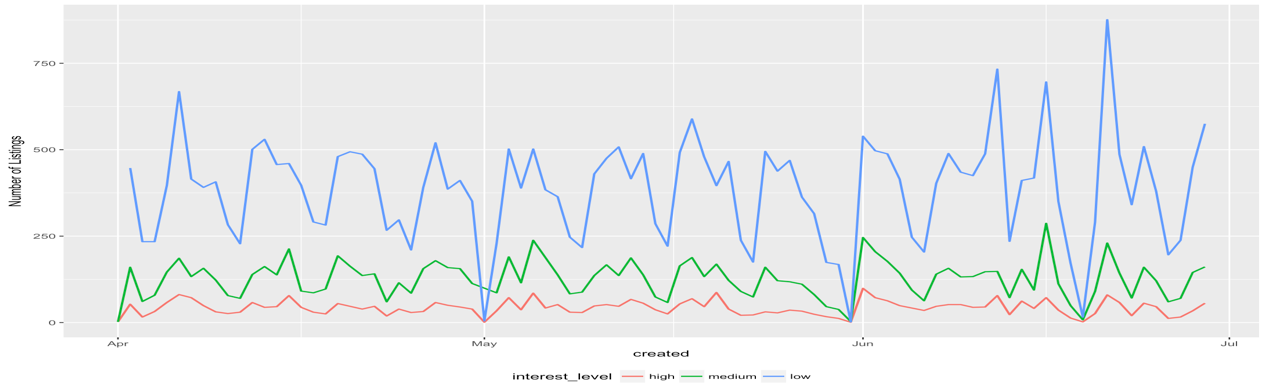


Figure 1: Distribution and time series of each interest levels

First, we plot the distribution of each interest levels. We found that 69 percent of the samples have low interest level, 23 percent of the samples have medium interest level, and only 8 percent of the them have high interest level. And then we plot the time series relation of time created and listing numbers of each interest level.

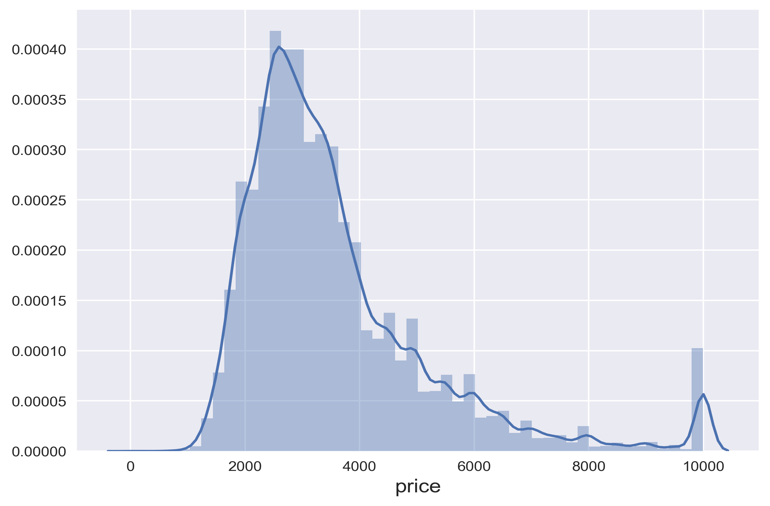


Figure 2: Distribution of prices.

Second we plot the distribution of prices and put prices over 10000 to be 10000.

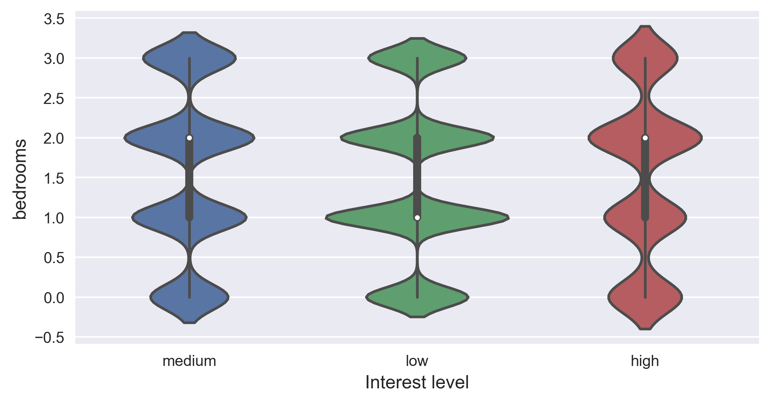
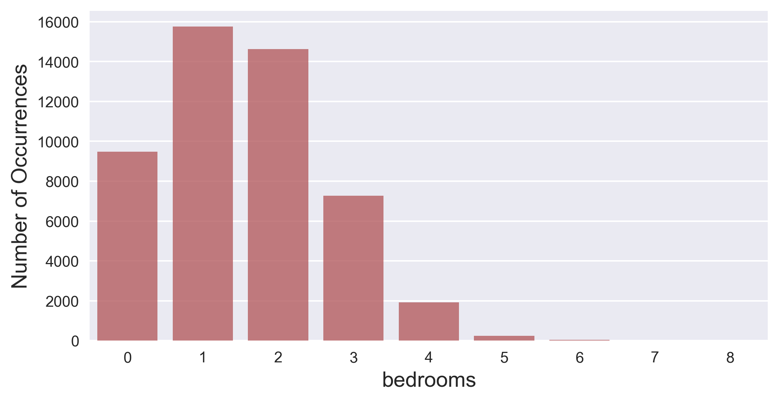


Figure 3: Distribution (left) and relationship (right) between number of bedroom and the interest level.

We could see from figure 3 left that most listings have 1 and 2 bedrooms. From right, we could see that for one bedroom listings, they are most likely to be low interest level and for high and medium interest level listing, it’s most likely to have 2 bedrooms.

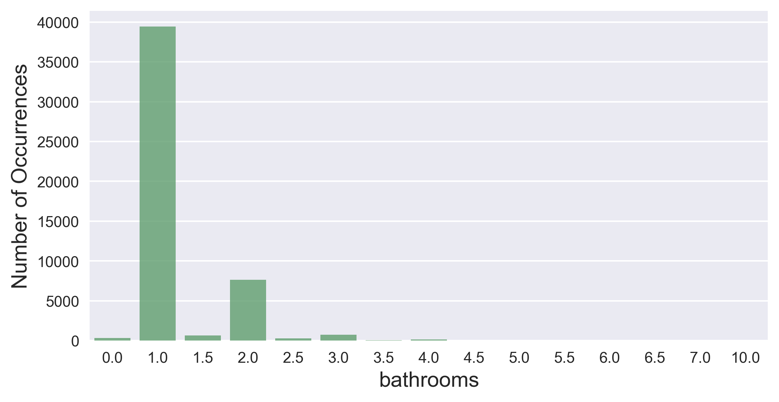
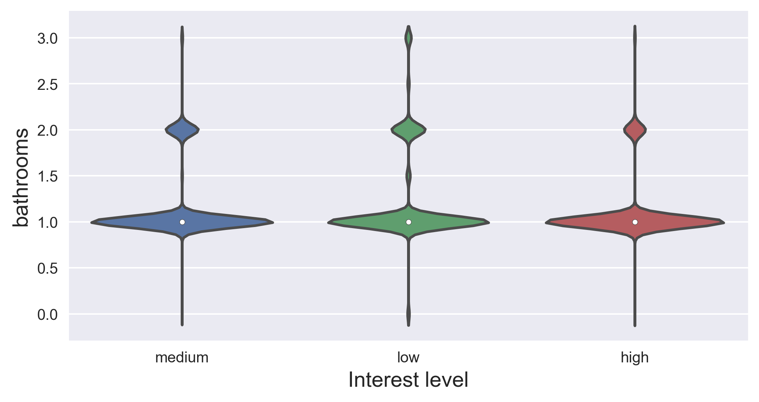


Figure 4: Distribution (left) and relationship (right) between number of bathroom and the interest level.

From figure 4, we could see that most listings have 1 bathroom among low, medium and high interest level.

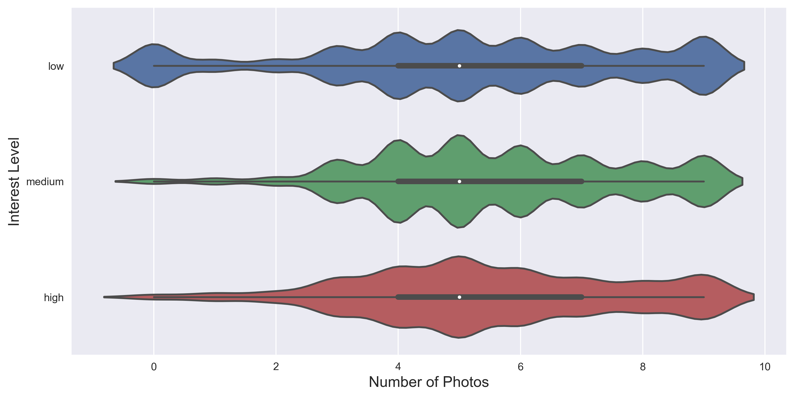
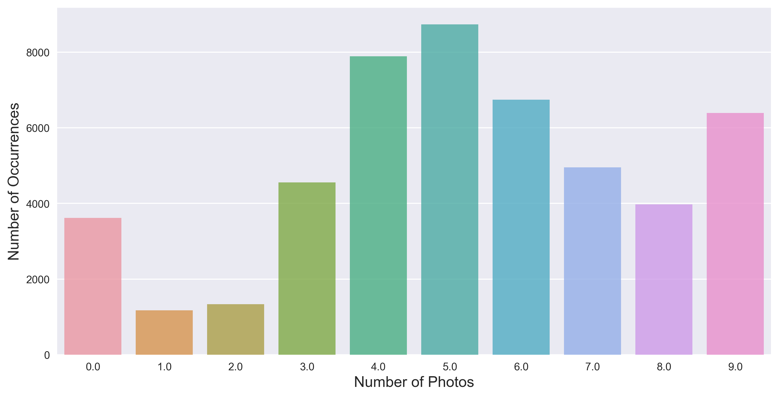


Figure 5: Distribution (left) and relationship (right) between number of photos and the interest level.

Figure 5 left shows that most listing have photos numbers around 5. And right shows that if a listing has no photos, it’s mostly has low interest level.

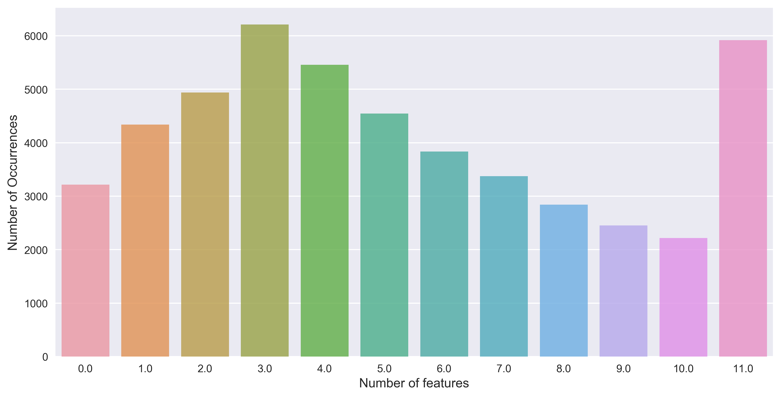
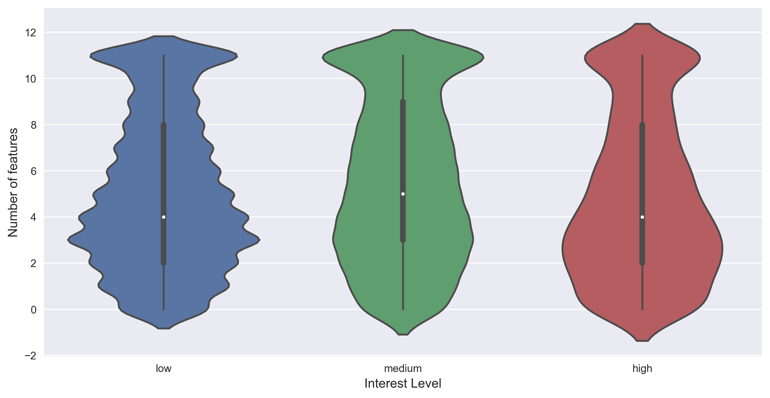


Figure 6: Distribution (left) and relationship (right) between number of features and the interest level.

Figure 6 indicates that most listing have number of features of 3.

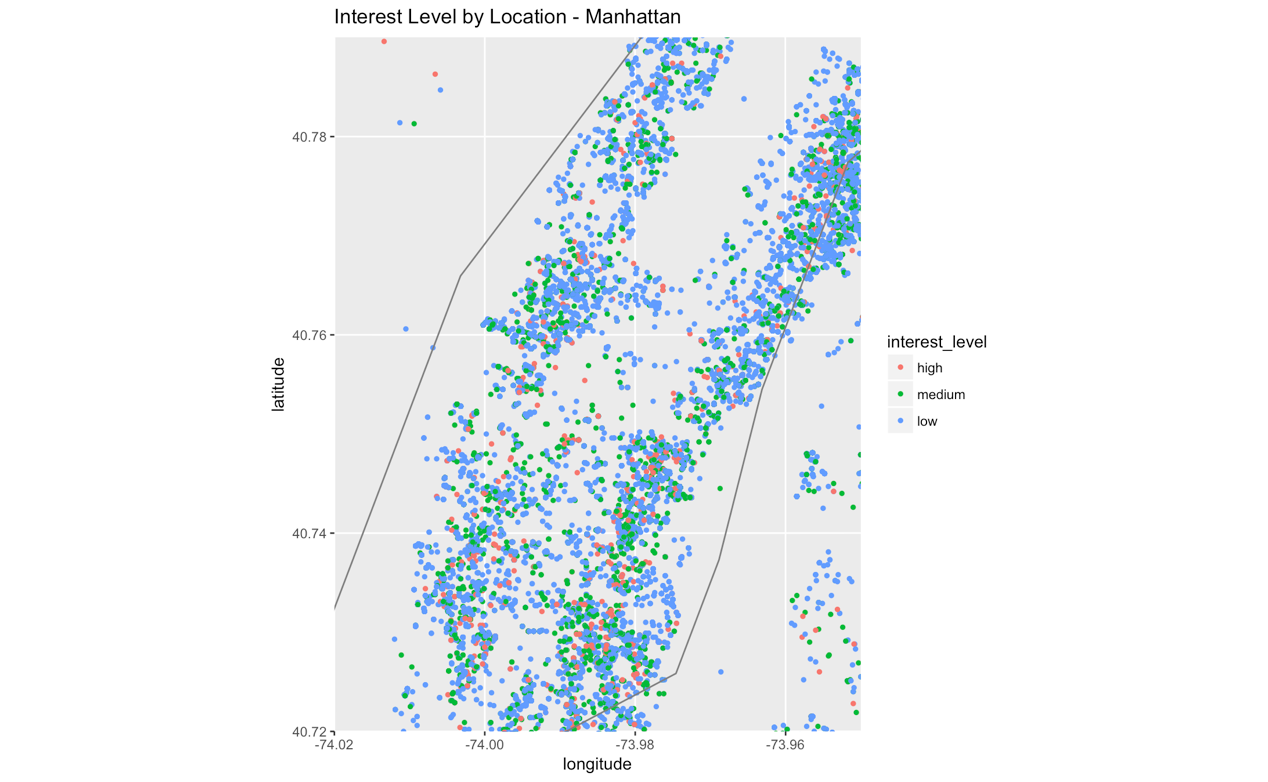
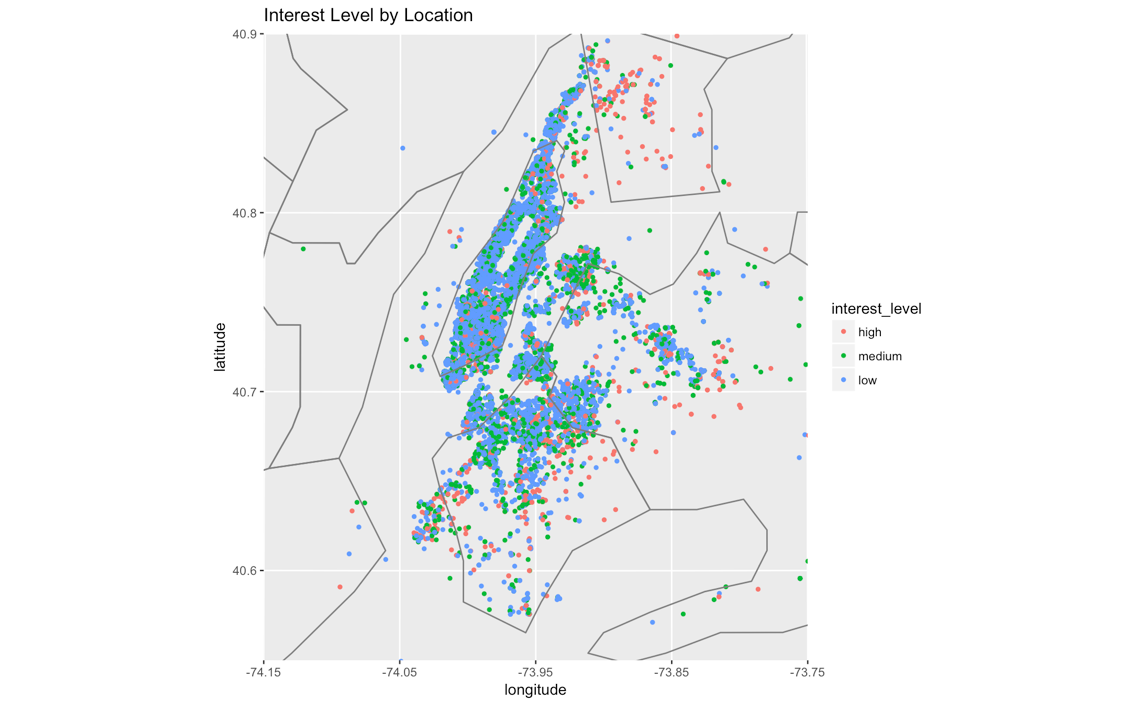


Figure 7: Locations of listings.

From figure 7 we could see that most low interest level listings exit on Manhattan and north Brooklyn. High interest level listings are in more suburban area.