CSC478 - Final Project Proposal

Airbnb House Analysis and Recommender Application

Team Member: Jie Bao, Lanny Xu, Kuangyi Zhang

## Overview

Airbnb provides the short-term leasing and renting service for people during tourism. Users could register and create their own account, then book different holiday cottages, apartments, homestays and hostel beds according to price, location, or review. In our project, by using the original dataset of Chicago’s airbnb house listings and user reviews, we provide host an evaluation of targeted residential housing market (data mining), as long as personalized recommendations to lodgers based on their selections or history records (recommender system).

## Dataset

### 1. Source

The data is published on the Inside Airbnb site, which is sourced from publicly available information from the Airbnb site. In our project, we use the **Listings** and **Reviews** dataset for Chicago, which was compiled on May 10th, 2017 and can be accessed through url below:

http://insideairbnb.com/get-the-data.html

### 2. Data Attributes

**Listings**

The dataset of Listings includes full descriptions of each listing and average review score.

Listings dataset: 5207 rows and 95 attributes

**#Location and environment#**

Id

Listing\_url

Scrape\_id Last\_scraped

Name

Summary

Space

Description

Experiences\_offered

Neighborhood\_overview

Notes

Transit

Access

Interaction

House\_rules

Thumbnail\_url Medium\_url picture\_url xl\_picture\_url host\_id host\_url host\_name

host\_since host\_location

host\_about host\_response\_time host\_response\_rate host\_acceptance\_rate

host\_is\_superhost host\_thumbnail\_url host\_picture\_url host\_neighbourhood

Host\_listings\_count

host\_total\_listings\_count host\_verifications host\_has\_profile\_pic host\_identity\_verified street

neighbourhood

neighbourhood\_cleansed neighbourhood\_group\_cleansed

**# Space and rent detail#**

City

State

Zipcode

Market

Smart\_location Country\_code Country

Latitude Longitude

Is\_location\_exact

Property\_type Room\_type

Accommodates Bathrooms bedrooms

Beds Bed\_type Amenities Square\_feet

Price

Weekly\_price Monthly\_price

Security\_deposit Cleaning\_fee Guests\_included

extra\_people minimum\_nights maximum\_nights

Calendar\_updated

Has\_availability

Availability\_30 Availability\_60 Availability\_90 Availability\_365 Calendar\_last\_scraped

**#Review#**

Number\_of\_reviews

First\_reviewLast\_review

Review\_scores\_rating Review\_scores\_accuracy Review\_scores\_cleanliness

Review\_scores\_checkin review\_scores\_communication review\_scores\_location

Review\_scores\_value

Requires\_license License Jurisdiction\_names

Instant\_bookable Cancellation\_policy

Require\_guest\_profile\_picture require\_guest\_phone\_verification

calculated\_host\_listings\_count reviews\_per\_month

**Reviews**

The dataset of Reviews includes unique id for each reviewer and detailed comments.

Review dataset: 13605649 rows and 6 attributes

Listing\_id

Id

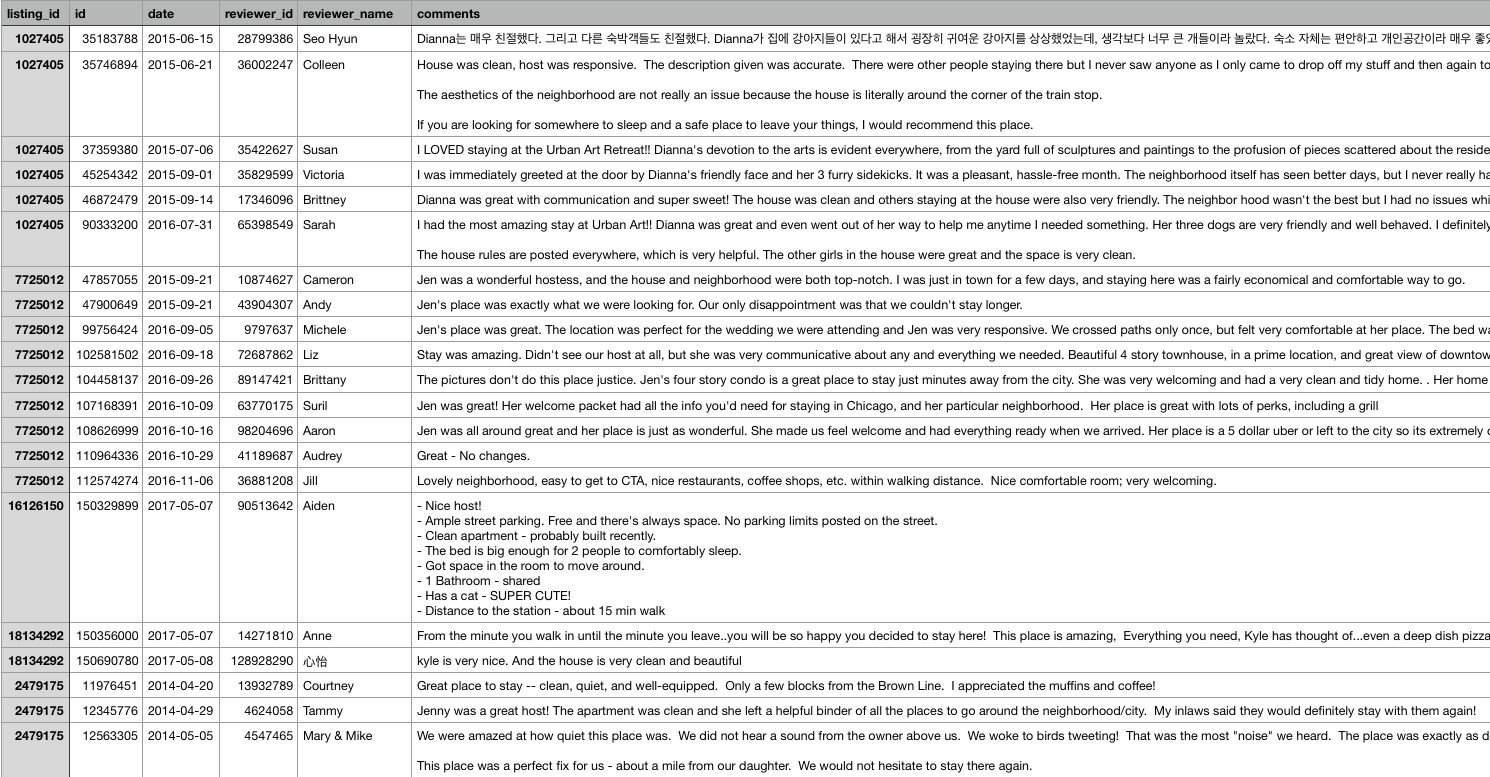
Date

Reviewer\_id

Reviewer\_name

Comment

*Reviews data illustration:*

****

### Application Features

1. Data analysis for hosts

The application provides an insight of residential housing market to home hosts who plan to offer their accommodation service through Airbnb platform.

To do that, we first conduct data preprocessing and preparation on the original data, then apply data explorations, data visualization or other techniques on the data (e.g., classification, estimation, clustering, association rule discovery, etc.).

Home hosts can obtain basic information about the Airbnb listings in their neighborhoods, such as their locations, prices, type of homes and departments, or major customer types. Besides, our application can provide hosts with reasonable pricing range based on the home profile, or important home amenities features as an guidance to improve their popularity on Airbnb.

2. Recommender system for lodgers

On the other hand, the application help users to make their selection on Airbnb platform. On the basis of learning user’s rental history in Listings and Reviews data, it provides personalized recommendations through below two ways:

**To existing users:**

The application processes text mining based on user reviews history. By analyzing their comments and ratings on particular home, it discovers the users’ favourite features of home, and recommend similar homes to them.

**To new users:**

The application lists couples of most popular homes for user to choose at the very beginning. After user picks up the one he is interested in, the system will also recommend other homes with similar features to the one they have picked.