

# BedVal: A Visualization of Airbnb Pricing Factors

CSE 6242 Data Visualization and Analytics, Spring 2019

Team 17

Team Members: Usman Ashraf, Azlan Shah Bin Abdul Jalil,  
Jiun-Yu Lee, Jing Shi, He Zhang



## Motivation/Introduction

Understanding the pricing and demand rationale of the sharing economy will be important.



Do the ease of access to public transportation, sentiments in the listing description and nearby attractions contributes to the listing price or the occupancy rate of a listed property?

The ultimate goal of this project is to help the hosts to determine their strategies of listing, and to help the guests to better understand the rationale behind the pricing mechanism of Airbnb.

## Methodology & Approach

1, Sentimental analysis by using the Opinion Lexicon in the NLTK module.  
2, Linear regression and T-test to see whether those factors are significant.

3, Experiment with DNN, Random Forest and Linear Regression predictions and build up the backend function for the prediction interface.  
4, Build up the prediction interface with Leaflet in JavaScript, HTML and Python.

5, Build up the user interface by using density map and choropleth map with D3 and Leaflet.

## Data

1, The public transportation and the attraction location data are scraped with API(600MB)  
2, The listing property data is downloaded(10MB)

## Analytics and Results

- 1, More positive words in the listing title and more neutral words in the listing description do increase the demand. Positive Descriptive word such as beautiful, quiet, comfortable help get more bookings for property.
- 2, The more stations within the walkable distance there are, the higher the demand will be.
- 3, The more attractions within walkable distance there are, the higher the demand will be.
- 4, No matter which model we looked at, our model are better than the original model built by previous researchers. The MSE is reduced by 3%-12%.
- 5, For the price prediction, linear regression will be the best model. For the demand prediction, random forest regression is the best.

## User Interface for Hosts

10  
Anticipated available days in a month.

Accommodates  
4  
How many people can be accommodated?

Beds  
4  
How many beds are available?

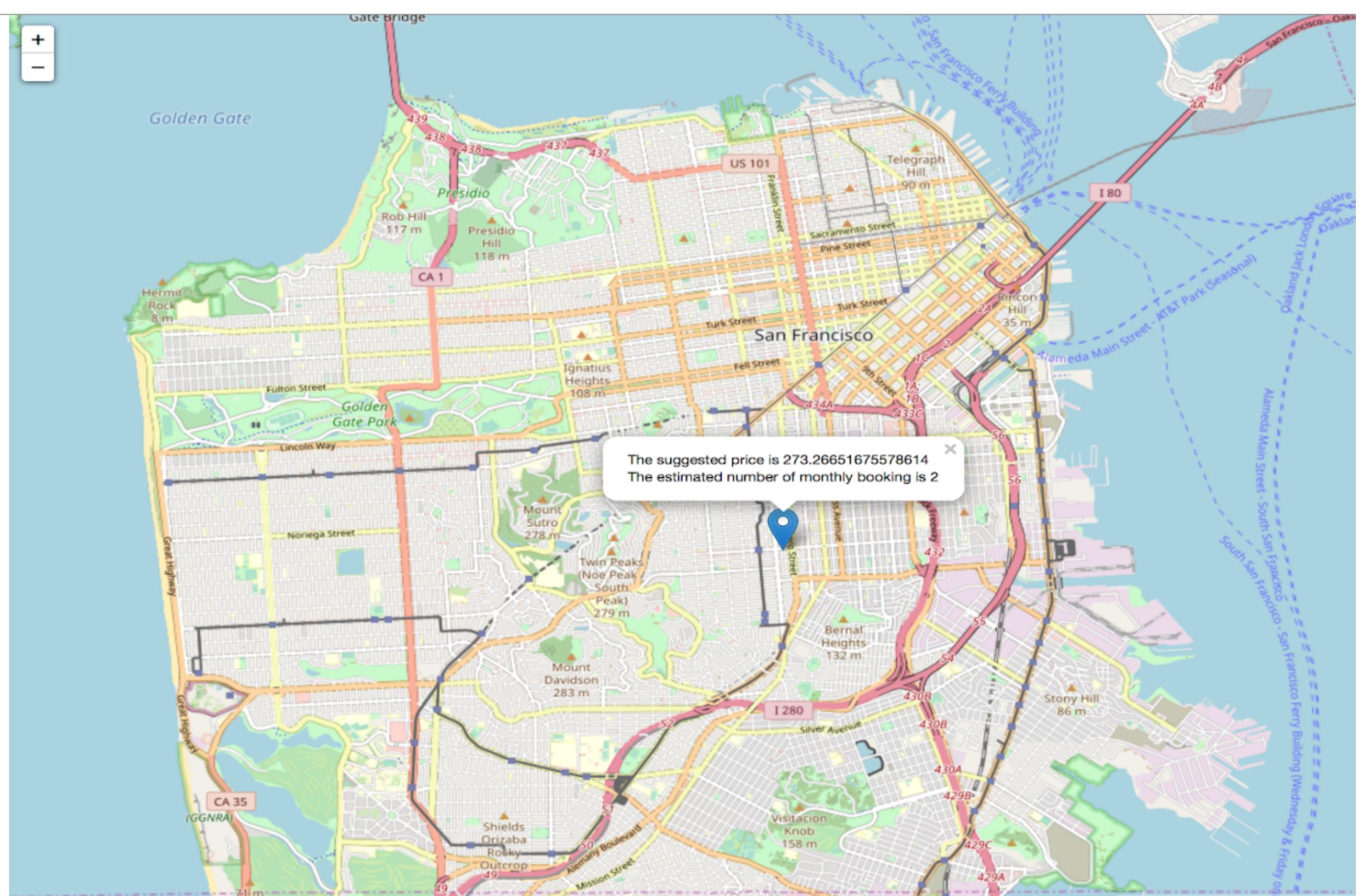
Review Score  
90  
Current review score.

Is the property an apartment?  
 Is the property a house?  
 Are you listing the whole property?  
 Are the beds real beds or futon/sofa?  
 Is the property Instantly Bookable?  
 Do you require guest's phone verification?  
 Is WiFi available?

Host Information  
Host Listing Counts  
2  
How many properties listed by you.

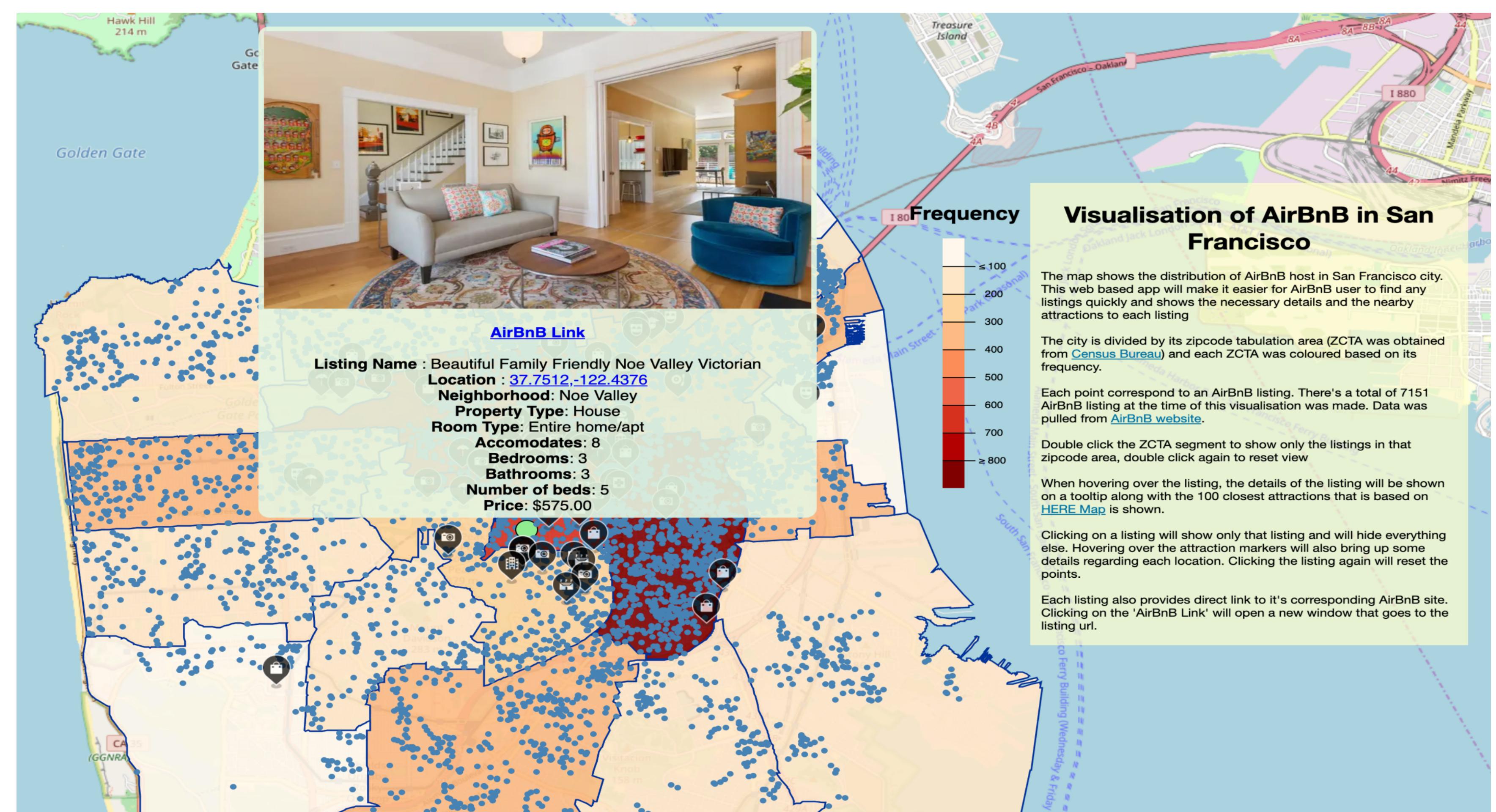
Are you a super host? (4.8+ rating, 10+ stays, 0 cancellation, 90%+ response rate)  
 Is your Identity Verified?  
 Do you have a picture in your profile?

Submit



By pinning down the location of the property on the map and choosing the options in the userform, this application will automatically generate the suggested price for listing and the anticipated monthly bookings.

## User Interface for Guests



Users can choose property listing id to explore the near by attractions(restaurants, cinemas, museums etc.) Users can also look at the listing density and average price in that area.