Deep Trip to California

Runze Liu, Wei-Fan Chen, and Yen-Chen Chou

[1] University of Southern California, Los Angeles CA 90007, USA
[2] INF 554 Data Visualization Final Project Report
[3] Deep Trip to California
[4] {runzeliu,weifanch,yenchenc}@usc.edu

Abstract. The paper presents an analysis of Airbnb housing availability in California from different aspects of location, room type and price using advanced visualization created by D3.js.

Keywords: Airbnb · Travel· Visualization · Analysis · California.

1 Introduction

When you come to California for traveling or business, finding a place to stay during the trip may be the biggest problem for you, especially if you are on a limited budget or you have some other preferences. With intention to deal with this issue, we collect open source data released by Airbnb and show the analysis result by visualizing data in different ways by D3.js. In order to enhance user experience, we build a responsive web application involving VUE.js and Bootstrap to introduce some ways of interaction between users and our app

2 Data

We obtained our data directly from Inside Airbnb, the open data released from Airbnb. There are tons of dataset available in different cities around the world. In this project, we focus on the data in few cities of California. There are different dataset available, such as listing, reviews calendar. Originally, there were about XX features, we extracted about XX features then integrated into one csv files. To fit in our map format, we converted the file into geojson.

3 Approach

3.1 Design consideration

In most situations, price is the most important deciding factor when we book a room. Therefore, comparison of price with other features is the main purpose of our project.

A usual approach is to show statistical graphs reflecting all data. But we aim to highlight the users' selection to make an obvious comparison. Therefore, we simply invent a new way of interaction. When a user selects one house, the part of it on statistical graphs will be highlighted.

3.2   Technical consideration

Since we have over 70,000 Airbnbs in our dataset, it will take some time for Vue to read and process them. Therefore, to improve the user experience, we optimize the web by loading the data once instead of reading separately and avoid the implementation of some functions. For example, the number of top of the barplot is fixed because I just want to speed up calculation.

We choose VUE.js as Web framework since it has detailed documentation and is very easy to understand and implement a project. In order to achieve responsive web application, we use a Bootstrap table and design all map and graphs responsively.

3.3   Development and Evaluation

We first prepossessed our data into the format we need and use python to analyze them, trying to find some patterns. Then we drew a bubble using MapBox, designed four statistical graphs by D3.js and finally integrated them together to form a web application involving VUE.js.

We evaluate our system from few aspects. Firstly, our web has a high speed of response. Except for web initialization, all other interactions will be done fast. What's more, our app provide concrete housing information. For example, when you hover over one point on the map, you can see the information about that Airbnb house and when you mouse over some parts of statistical graphs, you can know the detailed statistics result as well. Finally, the outlook of our app is pretty and layout is reasonable.

4   System

4.1   Layout

The layout of our system is shown below: Left part is a bubble map and we focus on California area. On the right hand side, there are two areas. The one above is the filter and the one below is the area where you can see the statistics graphs.
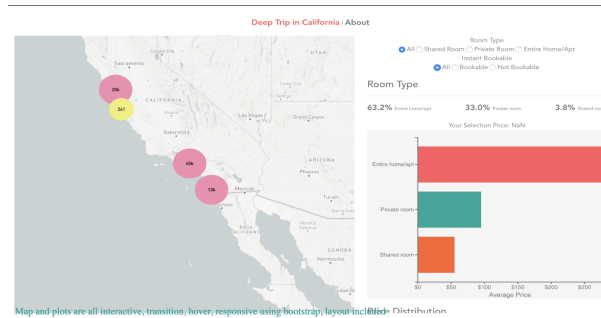

Fig. 1. Layout

4.2   California Map

We drew a map with all the hotels spot using MapBox. Also, the spots were summarized by the bubble-like clusters that can easily show the hotel availability around the area. User can click to zoom in, the information will be shown on the single spot while hover. Also, user can click on the single spot to see further comparison on bar plot and histogram.

4.3   Bar plot

The bar plot shows the average price of the houses with respect to different room types. As for the function part, when you click a point on the bubble map, a line corresponding to that point will occur in Barplot, so that you can easily tell the distribution of your selected point over the whole dataset.

4.4   Histogram

The histogram shows the average price distribution of the houses. As for the function part, when you click a point on the bubble map, the bar corresponding to that point will be highlighted in histogram, so that you can easily tell the distribution of your selected point over the whole dataset.
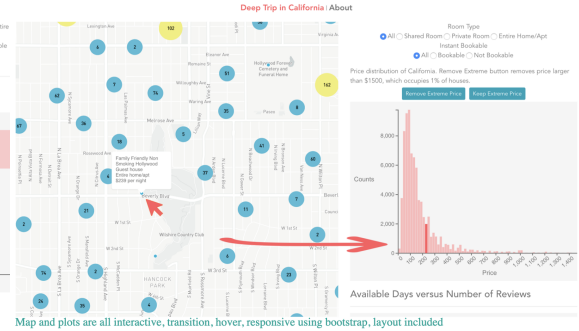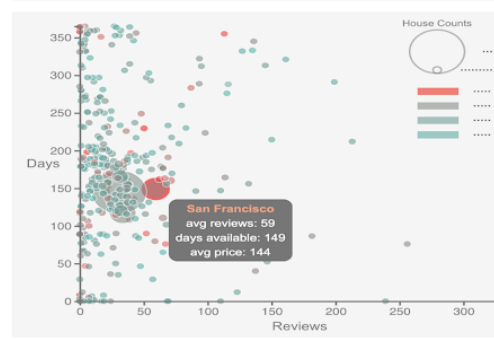

Fig. 2. Bar Plot   **Fig. 3.** Histogram

4.5   Bubble plot

The bubble plot is about showing how popular places are by looking at the available days and the number of reviews. As for the legend, the the bubble size declares the amount of house and the color bars show different price range. As for the function, we can see a larger circle in the below chart. It is because you pick up that point on the map.

4.6   Circular bar

This is the plot we extract words that described hotels and see how hotel descriptions are going to affect the price. The distribution of house types and their corresponding average price. The bar represents the number of houses in that type of house style.
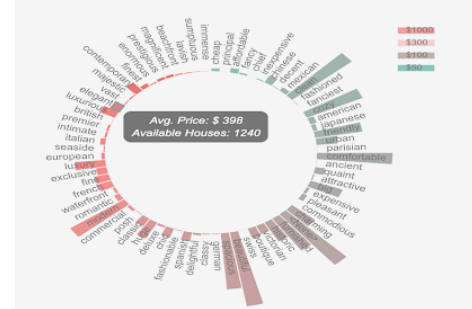

Fig. 4. Bubble Plot

Fig. 5. Circular Bar

4.7   Map

We draw a map using D3.js to show the density of house distribution.

5   Related work

We learned from a demo below about its layout and some functionalities. Obviously, the way it works is that you can choose the features you care about and the web can help give information about them, like this map and this statistics graph. We have this filter like the demo as well. But what we have done is a little different. We introduce a new way of interaction between users and our app. You pick up a point which is an Airbnb house on the map and the statistics graphs on the right will change depending on what you choose.

6   Future work

Due to the time limitation, we have a lot of ideas haven't fulfilled. For example, we mainly look from the view of customers, therefore, we compare price and
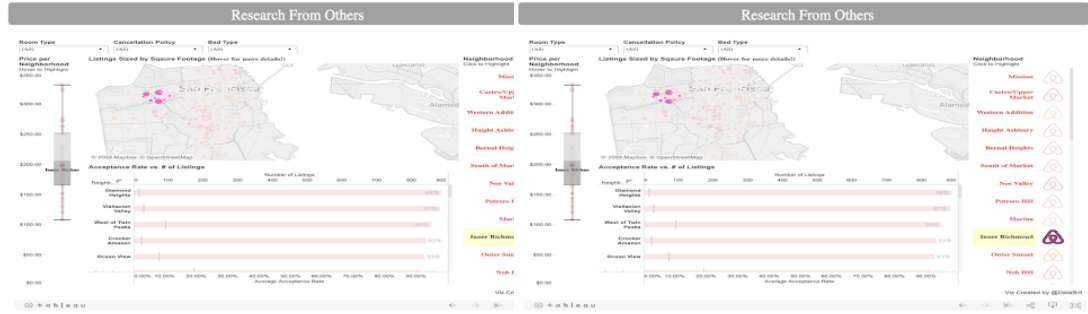
Fig. 6. Bubble Plot

Fig. 7. Other Demo

room type to provide customer to make their decision. On the other hand, we really want to analyze the dataset from the perspective of hosts or anyone who are interested in running Airbnb as a business. We could provide a big picture of the market. To be more specific, they would like to know the key to running a successful Airbnb, such as what kind of description are more attractive, easier to get an eye from the customers.

7   Conclusion

Through this project, we discover the basic environment of house renting market in California area. We learn a lot of Web technologies and data visualization techniques as well. But in our project, all processing data is done in client side. The speed of web response will be slow sometimes. Maybe we could improve it by developing a separate backend server involving like Springboot. I think by backend server we can easily manipulate our data in backend to speed up web response.

References