

Airbnb provides great accommodation services and its search engine provides plenty of options for users. However, there still exists some flaws that can be fixed to improve the service.

Problems:

1. The search input doesn't support free text query. Users are only expected to type location names, but queries like "Best scenic place in Chicago" are not processed well enough to give relevant search results.
2. The search results do not include socioeconomic information, like crime rate and population ratio, etc. Travelers are usually new to their visiting places, so this kind of information is important to consider when travelers are making decisions. For example, consumers may want to stay in a safe neighborhood (in cities like Chicago) or a specific population neighborhood (like the Sweden neighborhood in Chicago) to have better cultural experience in the cities new to them.

Solution:

1. To deal with the first issue, we need to design a new algorithm to analyze user input. To process the query "Best scenic place in Chicago", we can use the review dataset from [insideairbnb.com](https://www.insideairbnb.com) to provide relevant results. By searching in the consumer review data, we can surely provide better answer.
2. To provide socioeconomic information for the user queries, we need support from other data source. Take Chicago as an example, we can get the public crime and population data based on neighborhood from Chicago Police Department (CPD) and other public data centers, and show the information on additional maps to let users consider. Here we would need tools from geospatial data science.
3. To build index and process queries, we would use Elasticsearch to help with that. Relevant datasets can be obtained from [insideairbnb.com](https://www.insideairbnb.com) and CPD. To make the project possible to be completed, we would constrain our focus to the city of Chicago, and use this to demonstrate my solution to the issues mentioned above.