

BIA-660D Final Project Proposal - RS

Data Analyst / Scientist Listing Crawling & Analytics

Members: Rui Song

Project initiative

People considering taking trips as an essential way of recreation, it is pleasant to go to new place, to see new scenery, to meet new people, but the planning of a trip is not always easy. Housing for example, is particularly hard to decide, due to the constraints of transportation, budget, space quality. Airbnb, provides a new route to solve this problem and has greatly enriched people's option. Yet it gets mundane and time consuming to find a satisfying one. With programming and analytic skills explored in this course, I hope to improve the user experience.

Goal

To improve user experience can be too general, therefore to be more precisely, I am targeting shortening the overall time to find the right list as the project objective. The measure of which will involve experimental design.

To achieve this broad goal, the project will involve the following steps:

1. Limit geographical locations to selected ones, each of which should be a travel attraction (Miami, Orelana, NYC, LA, SF, BOSTON, etc.) Scrape listings from these cities.
2. Build filter and model to narrow down choice for the user.
3. provide key info screen shot of these listing accordingly and render it to the user.

Reference link: <http://phantomjs.org/screen-capture.html>

4. Provide pipeline that allows the user to go back to the Airbnb website and make his final call.

Data source

Airbnb <https://www.airbnb.com/>

Through web crawling or API

Analytics to be involved

1. Experimental design to measure time difference.
2. Simple presentation of summary statistics using visualization
3. NLP to extract key information from review
4. Predictive model based on user preference and listing key metrics.