

Group 13 Proposal: Data Analytics of Best Airbnb's in San Diego

Problem:

Looking for the ideal Airbnb may be daunting for travelers and hosts, especially not knowing what factors come into play. By understanding the host-customer relationship one may determine the important factors that contribute to the decision-making process for the housing.

Dataset:

Airbnb datasets: <http://insideairbnb.com/get-the-data.html>

- Price (per night) & monthly/annual revenues
- Region (city of listings as well as latitude and longitude)
- Dates reviews received for 10-year span
- Number of listings in area
- Room types
- Occupancy rate for hosts & availability throughout the year

Background:

Airbnb Datasets are created by Airbnb, an online marketplace company that connects people who want to rent out their homes with people who are looking for accommodation in that locale. Airbnb Datasets contain a lot of different csv files during different time periods. They record the price, room type, occupancy rate, host rating, location of room, etc. By analyzing Airbnb Datasets, we can find some relationship between those factors.

Proposed Solution and Real World Application:

Analyze the correlation between price and various external factors (i.e. reviews, room type, region, time of season, host rating, occupancy rate). Analyze the profitability for hosts through review metrics, price, number of listings per host, estimated occupancy, etc.). Cross-correlate to ultimately determine the best Airbnb in San Diego. From the host side, they'll have an example for how to improve their listings while the travelers find their ideal vacation spot.

Project steps

Steps	Estimated completion time	Person(s) in charge (among the group of 5)
1. Extracting and cleaning up data	One week	Jesi Miranda-Santos
2. Data analysis & Model Future Prediction	Two weeks	Zhang Zhizuo & Zibo Zhang
3. Data visualization (to obtain data statistics and user behavior)	One week	Ahnaf Ahmed & Kevin Hau