How to make money in LA with Airbnb

Erdos Data Science Bootcamp Project
Team 39

Zheyu Ni, Ricky Oropeza, Reza Averly, Shubhrika Ahuja, Praveen Shahani

Goal

- To predict the rental price of a new airbnb listing for a host in Los Angeles.
- Recommend amenities to make homestay popular

Exploratory Data Analysis

- **Data gathering:** We took the Inside Airbnb Listings dataset for Los Angeles which has information all listings in Los Angeles till 7th March, 2023.

- **Data cleaning**: Our modified data set has over **29,000** rows. We narrowed down from **75 features to 68 features** which are classified in six categories: **identification, host, location, property, reviews and number of listings by host.**

Price Distribution by room type

Entire home/apartment are listed at higher prices compared to other room types.

Top 10 Neighbourhoods

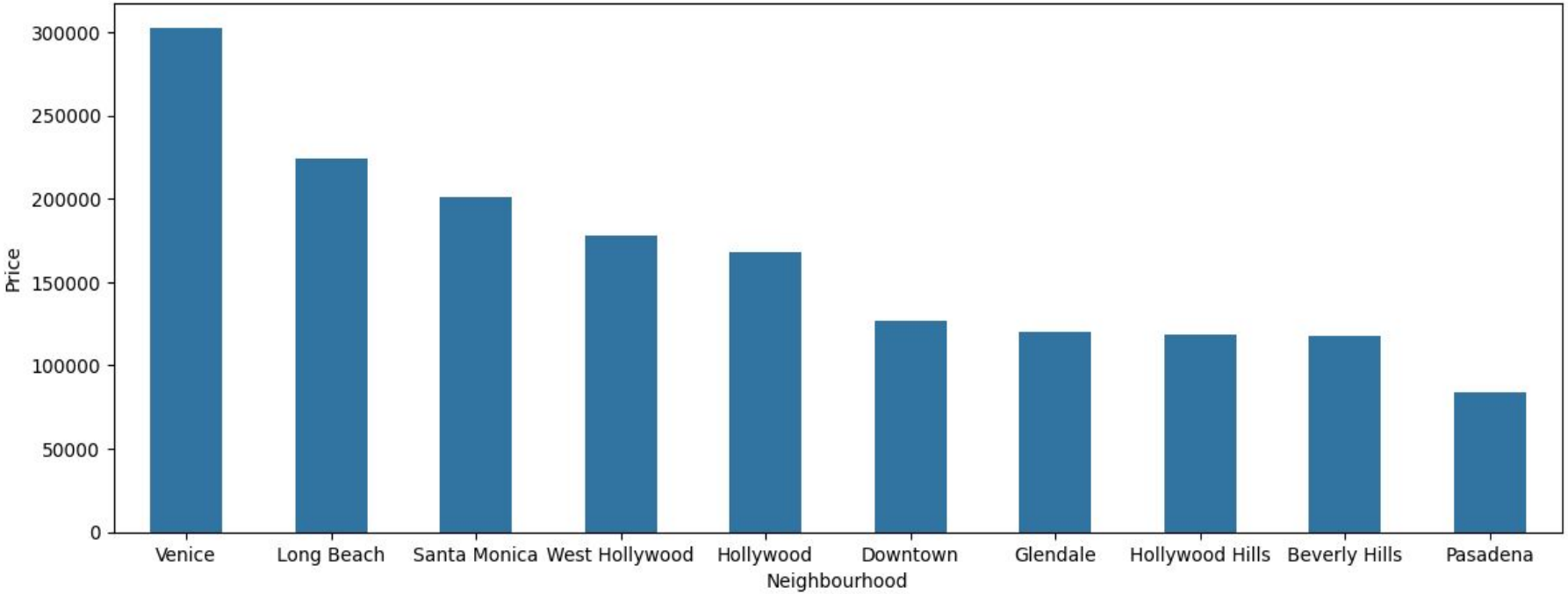
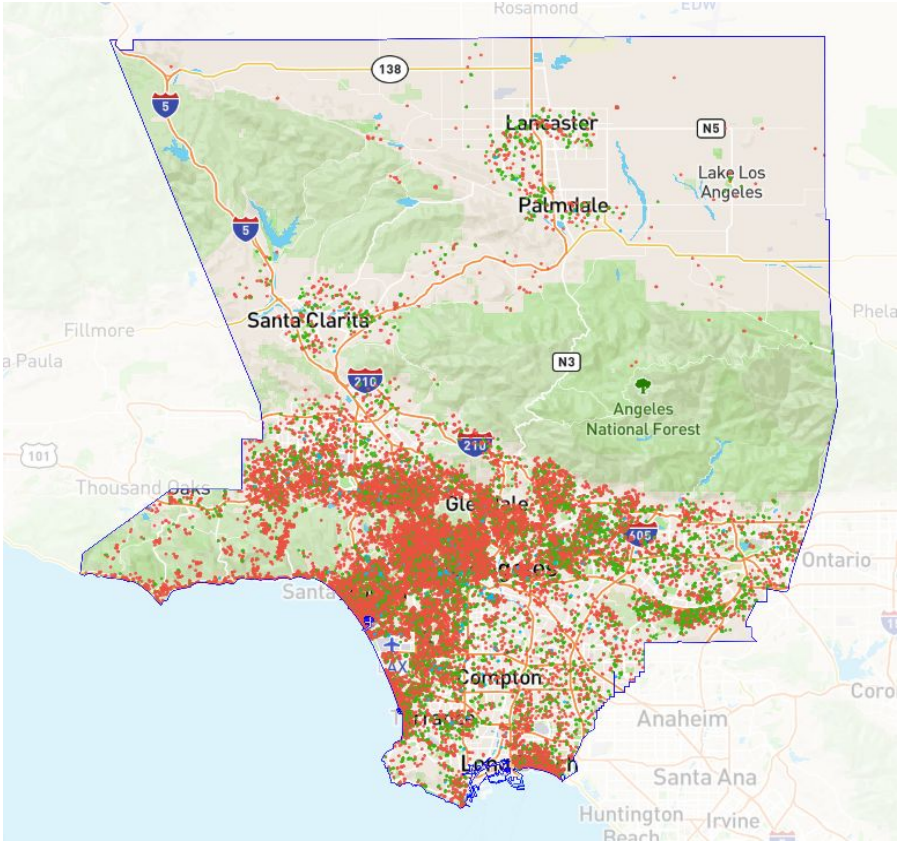
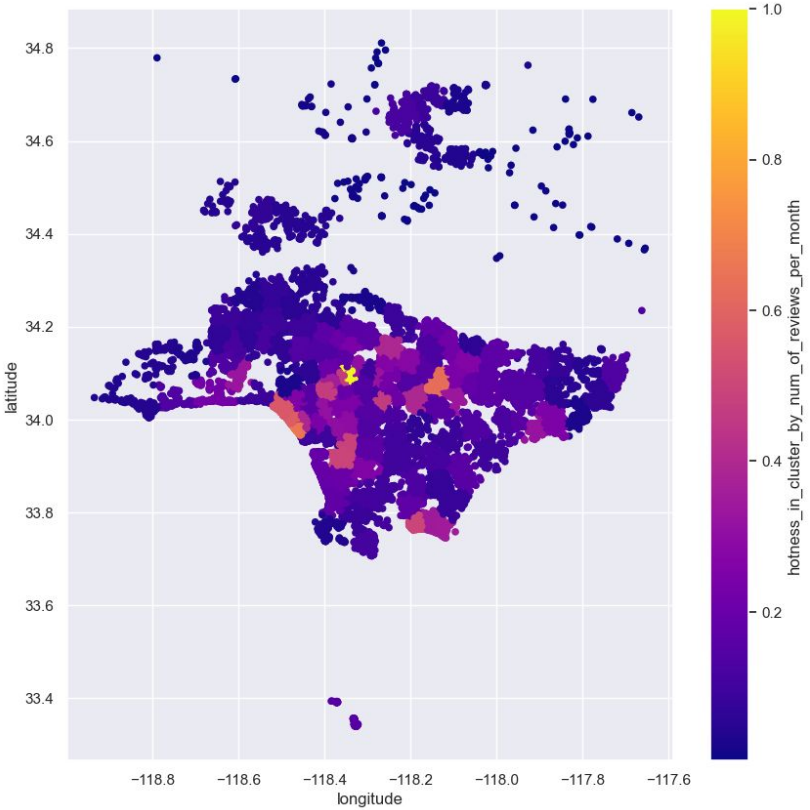


Image of all the data we received from Insider Airbnb



Clustering/Hotspot

- K-means clustering is utilize to find neighborhoods using latitude and longitude
- A hotspot can be a small cluster with many listings, many reviews, and/or many reviews per month
    - Image on the right is of number of reviews per month within a cluster



What Does Structural Modeling Tell?

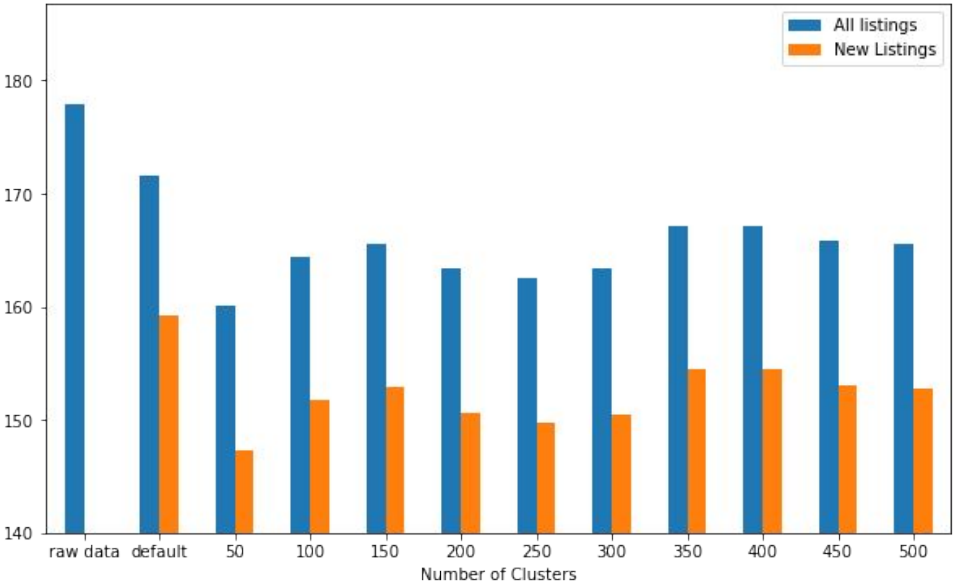
- Consumers' choices reveal their preference.
- The variations in prices, features and combined with sales among listings tell how consumer value different features.
- Consumers are willing to
  - Pay $2/night more for superhost.
  - Pay $9.3/night more for 1pt higher review scores
  - Pay $71/night more upgrading from private room to entire home.

	coef	std err	t	P>|t|	[0.025	0.975]
price	-0.1038	0.016	-6.577	0.000	-0.135	-0.073
beds	4.2210	0.634	6.662	0.000	2.979	5.463
license	4.5189	0.677	6.679	0.000	3.193	5.845
Entire home/apt	-8.5537	1.171	-7.305	0.000	-10.849	-6.259
Hotel room	-10.4800	1.825	-5.742	0.000	-14.057	-6.903
Private room	-15.9352	1.561	-10.207	0.000	-18.995	-12.875
Shared room	-20.2491	2.235	-9.061	0.000	-24.630	-15.869
host_is_superhost	0.1972	0.194	1.017	0.309	-0.183	0.577
review_scores_rating	0.9700	0.240	4.040	0.000	0.499	1.441

Footnotes: superhost: 0.1972/0.1038; review score: 0.9700/0.1039; private room→ entire home (15.9352-8.5537)/0.1039

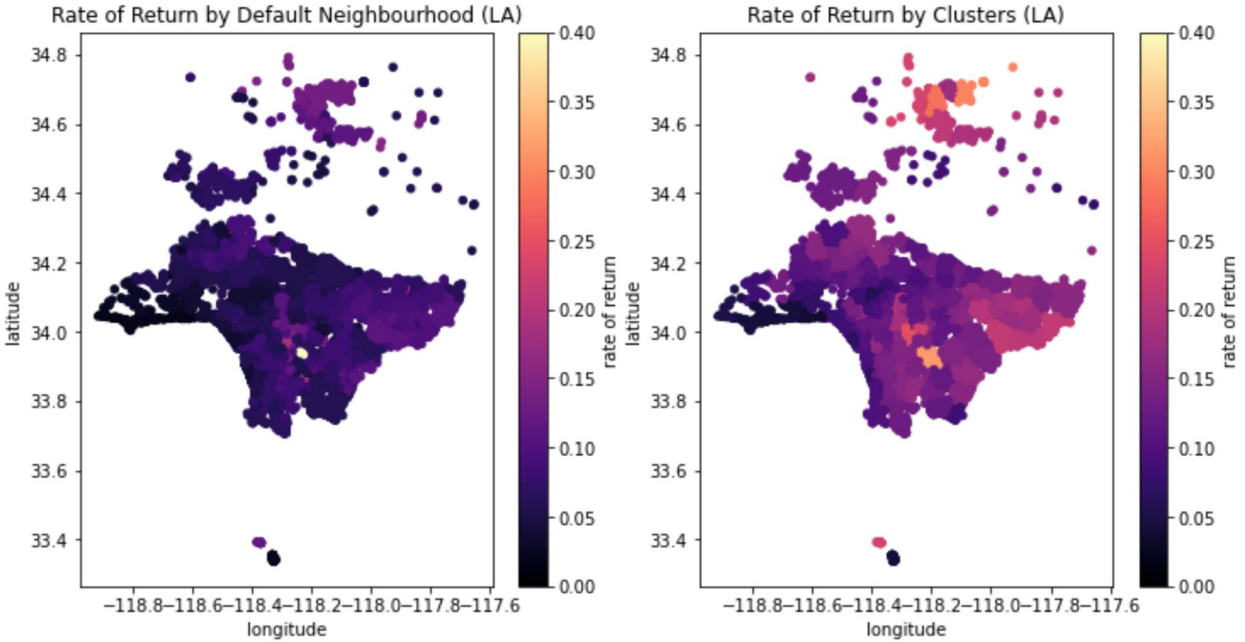
Optimal Pricing with Structural Model

- Customize the price given property locations, features (amenities), host ownership, etc.
  - maximize host's profit
  - market demand and substitution between other listings.
- Shows competition effects coming from added new listings.
- Price (clusters) << the price (w/o clusters) due to strong competition within hotspots.



Comparison: Default Neighbourhood V.S. Clustered

- Rates of Return differ significantly
- Recommend an area with higher return for potential hosts.



Smart Pricing

HOW: use machine learning to find the best rental price:

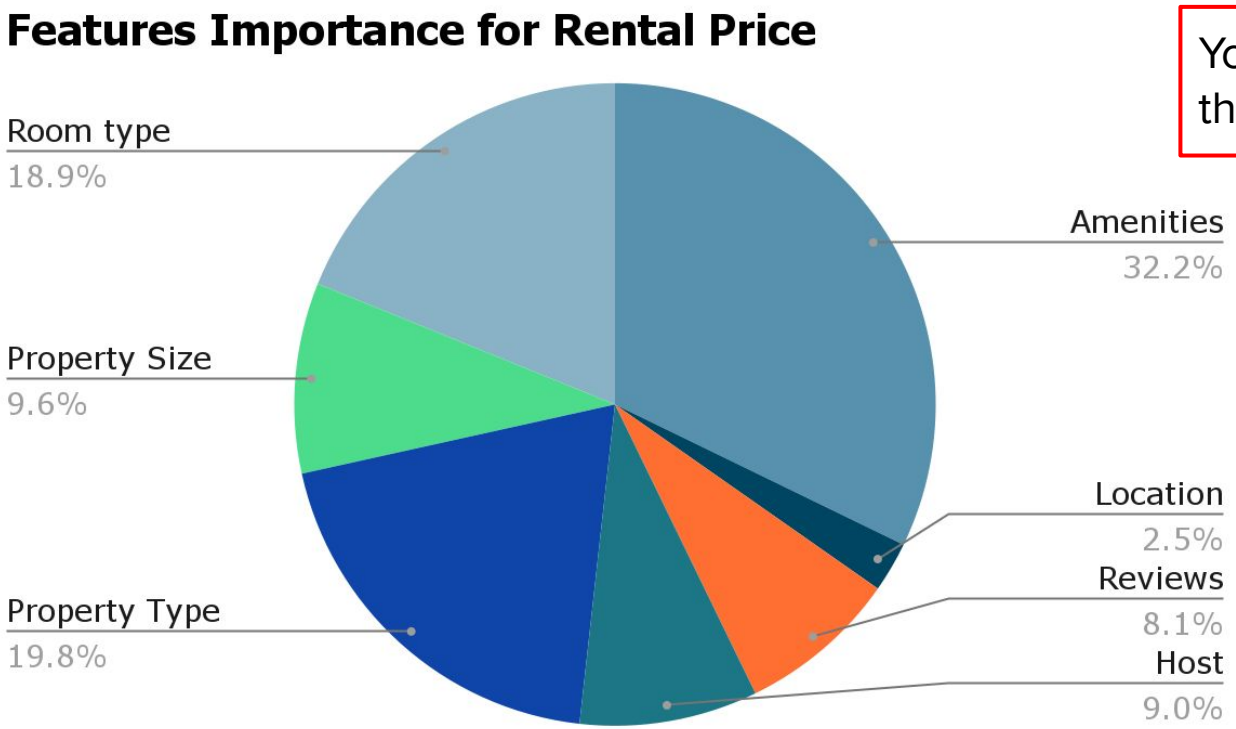
Host characteristics, reviews, location, properties, amenities, etc.

EXPERIMENT:

- Data cleaning and preprocessing
- 5 splits cross validation
- Hyperparameter tuning

MODEL	**R^2-score**
Linear Regression	0.563
K Nearest Neighbors (10 neighbors)	0.467
Random Forest (500 estimators, depth=5)	0.604
XGBoost (100 estimators, depth=5)	**0.709**

Smart Pricing



Your amenities makes the biggest difference!

Attracting Customers

Amenities that attract most people:

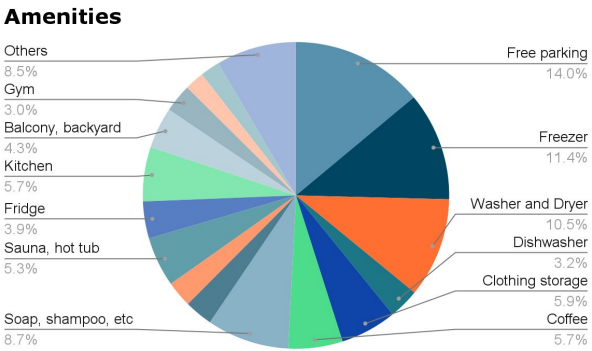
1. **Free parking!**
2. Freezer
3. Washer and dryer

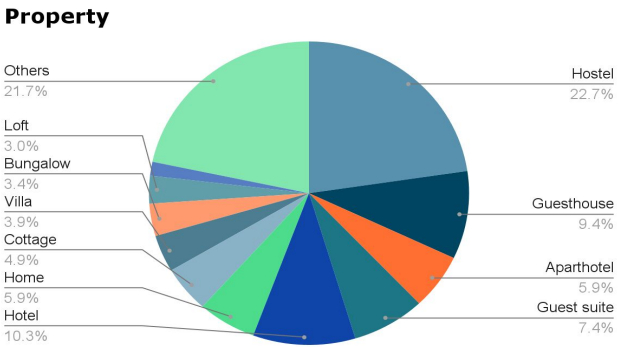
Property that attracts most people:

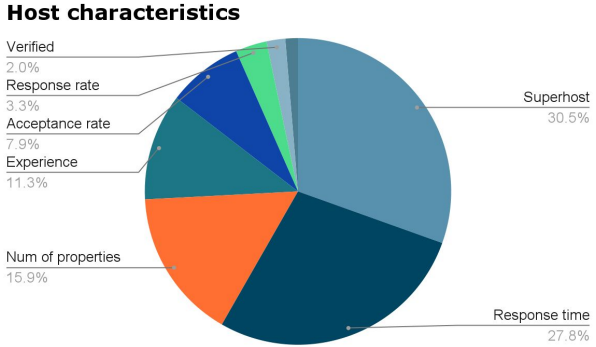
1. **Hotel**
2. Guest House
3. Apartment Hotel

Host that attract most people:

1. **Superhost**
2. Response within an hour
3. Number of listings the host owns







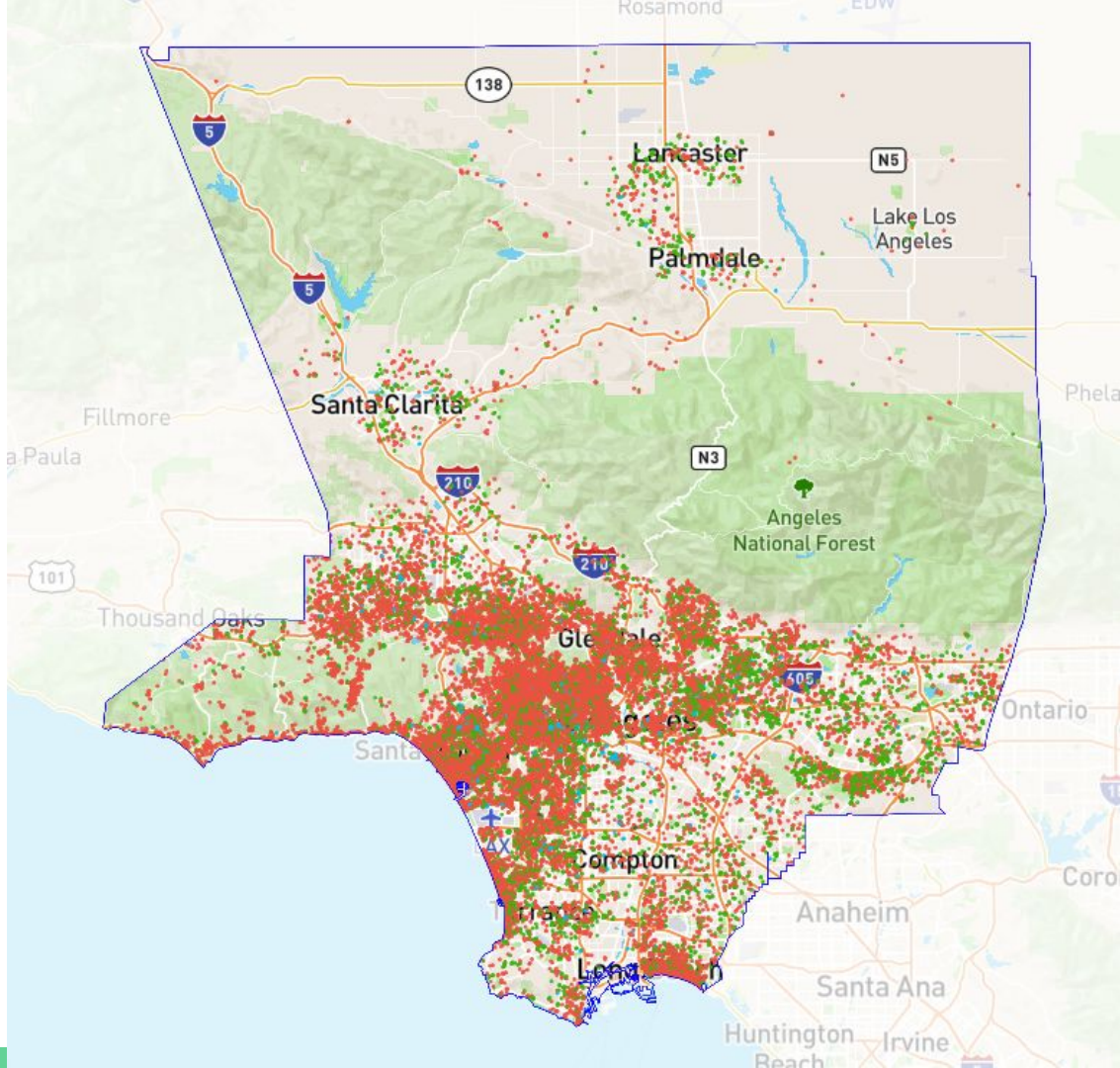
How to make money in LA with Airbnb

1. Choose an area that has the best rate of return from structure modeling
2. Provide amenities and property that attract most people
3. Become a superhost and respond fast to messages from customers

Thank you

Extra Slide

Plot image of all the data we received from Insider AirBnb



Hotspot by total reviews in cluster