



airbnb hosts in NYC

Modeling Social Data Group Project

Tejas Dharamsi, Vedant Dharnidharka + Molly Hanson



Agenda

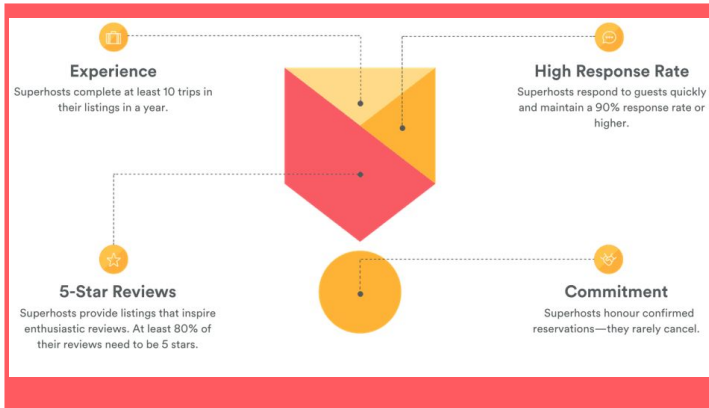
- **Introduction: Airbnb Hosts**
- **Motivating Questions + The Data**
- **Modeling Task: Can we predict Superhost status?**
- **Multi-Listers: Are hosts violating the “One Host, One Home” policy?**
- **Key Takeaways**



Introduction

- Online marketplace: host rent out vacant living space to short-term renters
- Disruptive technology
 - Provides consumers with a service that did not previously exist
 - Shifting of power within hotel industry
 - Opposition from multiple parties

How can hosts benefit?



What are hosts at risk of?

Under NY state law, it is illegal to rent out your apartment for less than 30 days, but not yet illegal to advertise it.

“Airbnb is dominated by commercial operators with **multiple listings** who are stealing our supply of **affordable housing**” - Neal Kwart

Motivating Questions + The Data

- Are we able to predict a host's Superhost status?
 - Goal: Can we predict superhost status using only the 4 traits Airbnb lists as benchmarks? Or are there other factors that are at play?
 - Importance: Is the Superhost program requirements as transparent as it seems? How can I increase the likelihood of becoming a Superhost?
- What is the concentration of hosts with multiple listings?
 - Goal: Assess the concentration and characteristics of multi-listing hosts
 - Importance: Are hosts violating the "One Host, One Home" policy? Are multi-listers affecting the affordable housing issue in NYC?

Where did we get the data?

- Retrieved from www.insideairbnb.com
 - Sources the data from publicly available information from the Airbnb site
 - Data pre-cleaned and aggregated (where appropriate) to facilitate public discussion
- New York City listings, compiled April 2, 2017



Modelling Task

- **Goal: Can we predict Superhost status based on...**
 - 22 features (e.g. toom_type, accommodates, availability_30)
 - 4 main features: host_response_time, host_response_rate, number_of_reviews, review_scores_rating
- **Random Forest Classifier (Ensemble Method)**
- **~22,891 listings**
- **Unbalanced classes: 3,320 Superhost, and 19,571 non-Superhost**
 - Balanced Weight Representation, Under Sampling, Over Sampling
- **Assessed how different features weigh on super-host tag**



Modeling Result : 22 features, RF, UnderSampling

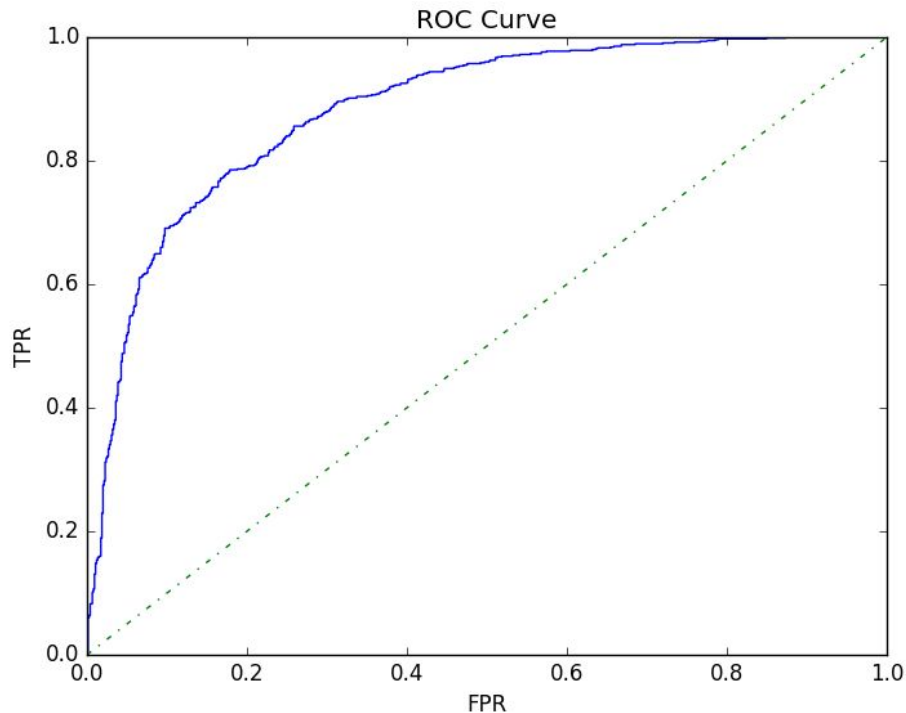
Test Accuracy: 79.4%,

AUC: 0.883

	Precision	Recall	F1-score
0	.79	.80	.80
1	.80	.79	.79
Avg	.79	.79	.79

Confusion Matrix:

	0	1
0	600	150
1	159	591



Modeling Result: 4 features, RF, UnderSampling

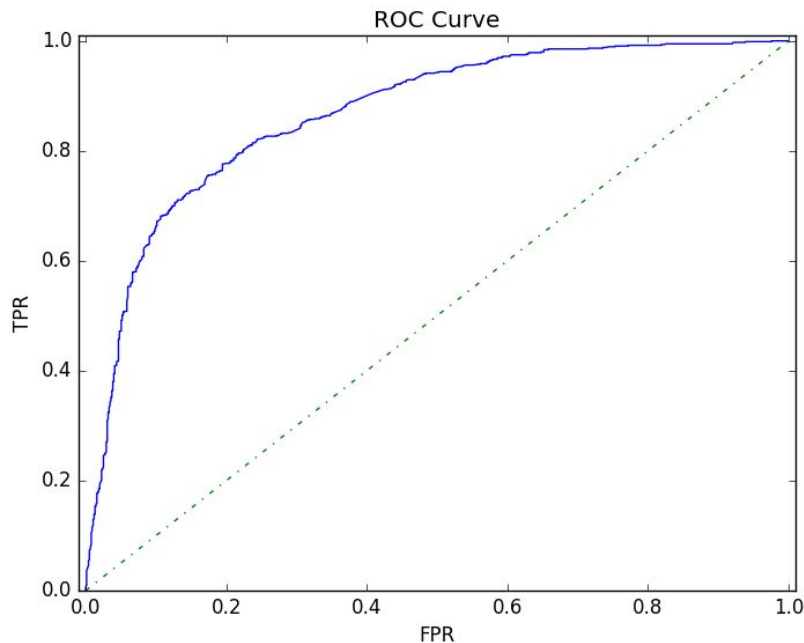
Test Accuracy: 78.86%,

AUC: .8665

	Precision	Recall	F1-score
0	0.77	.83	.80
1	.81	.75	.78
Avg	.79	.79	.79

Confusion Matrix:

	0	1
0	622	128
1	189	561

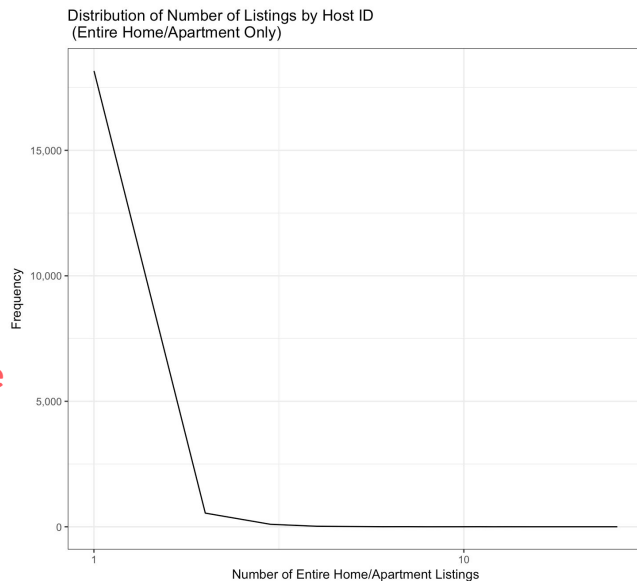


Key Modeling Takeaways:

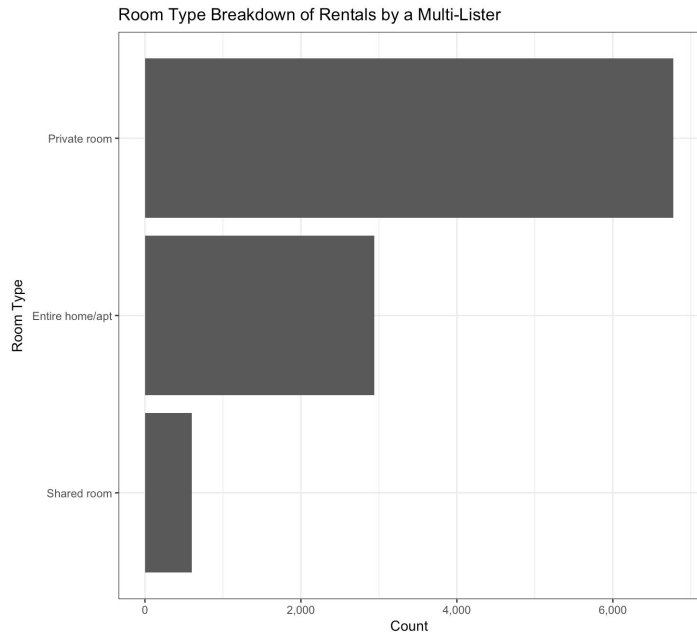
- Additional 18 features barely influence the Superhost status
- Program is transparent

Concentration of Multi-Listers

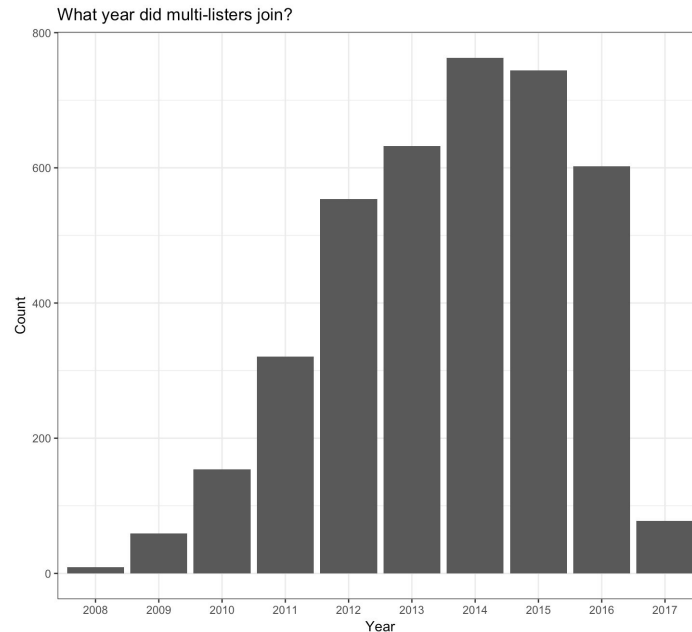
- 34,223 unique Host IDs
- 11.4% of hosts have multiple listings (all room types)
- Are private or shared rooms of concern to affordable housing advocates?
 - Of the 18,856 hosts that list entire apartments, 3.6% have multiple entire home listings
- Findings consistent with public statements
 - “96% of New Yorkers renting out an entire unit were offering their own home”



Who are Multi-Listers?



Entire home/apartment represent 30% of multi-lister listings

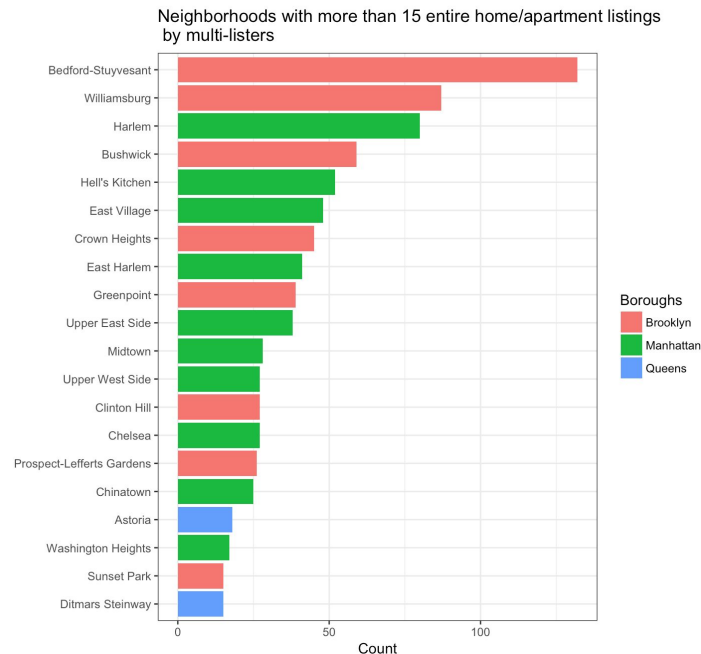


Fewer current NYC multi-listers joined 2015 or 2016, compared to 2014

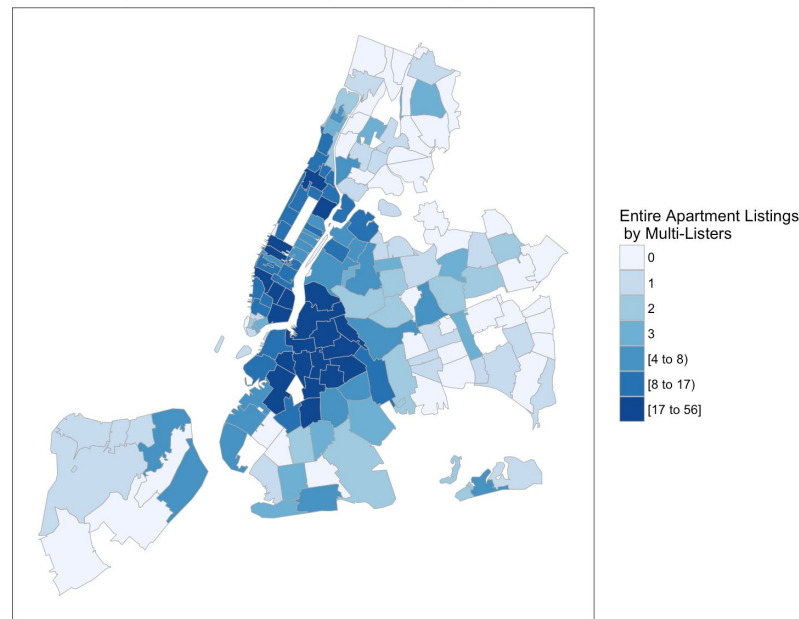


Where are apartments by Multi-Listers located?

Affordable housing advocates are concerned with short-term rentals of entire apartments



New York City Airbnb Entire Home/Apartment Listings by Multi-Listers



Key Takeaways

- Based on our decision tree model on the given dataset, we affirm that the features stated by Airbnb regarding Superhost's are good predictor and the Superhost program is quite transparent
- “One Host, One Home” policy clearly violated by some, but implementing laws to ban advertising short-term listings hurts all of hosts
 - Update the tech: hosts should not be able to create multiple listings within the platform
 - Crackdown on bad actors: focus on commercial listers abusing platform, rather than eliminating the service all together



Questions?

