Predicting the Price for Airbnb Listings

Presented by Nicholas Reise

Roadmap

1. What is Airbnb
2. Problem/Solution
3. Data
4. Model Description
5. Findings
6. Risks/Limitations
7. Next Steps

What is Airbnb?

Total Guests
200,000,000+

Castles
1,400+

Cities
65,000+

Countries
191+

Problem

Airbnb does not offer substantial guidelines for hosts listings their homes.

"The price you charge for your listing is completely up to you. . .”

- Airbnb

Problem

Inaccurate listing can cause:

1. Inefficient allocation of marketing budget
2. Mislead guests, causing host to incur fines
3. Underbooked listings due to over pricing

Solution

Create a model that calculates the listing price for a given Airbnb

Customer Lifetime Value

Allocate their marketing budget more efficiently

Calculate prices for online marketing

Solution

Provides clear guidelines to hosts and guests

Prevent penalties incurred by hosts

Encourage prospective hosts to join Airbnb

Reassure guests that the listing price is fair

Airbnb → Goals → **Data** → Model → Findings → Limits → Next Steps

USA
Los Angeles
is cinematic cityscapes

Top sights in Los Angeles

1 **Huntington Library, Art Collections & Botanical Gardens**

PASADENA & THE SAN GABRIEL VALLEY

2 **Broad**

DOWNTOWN LOS ANGELES & BOYLE HEIGHTS

3 **Getty Center**

BEVERLY HILLS, BEL AIR, BRENTWOOD & WESTWOOD

4 **Griffith Observatory**

LOS FELIZ & GRIFFITH PARK

5 **Los Angeles County Museum of Art**

WEST HOLLYWOOD & MID-CITY

6 **City Hall**

DOWNTOWN LOS ANGELES & BOYLE HEIGHTS

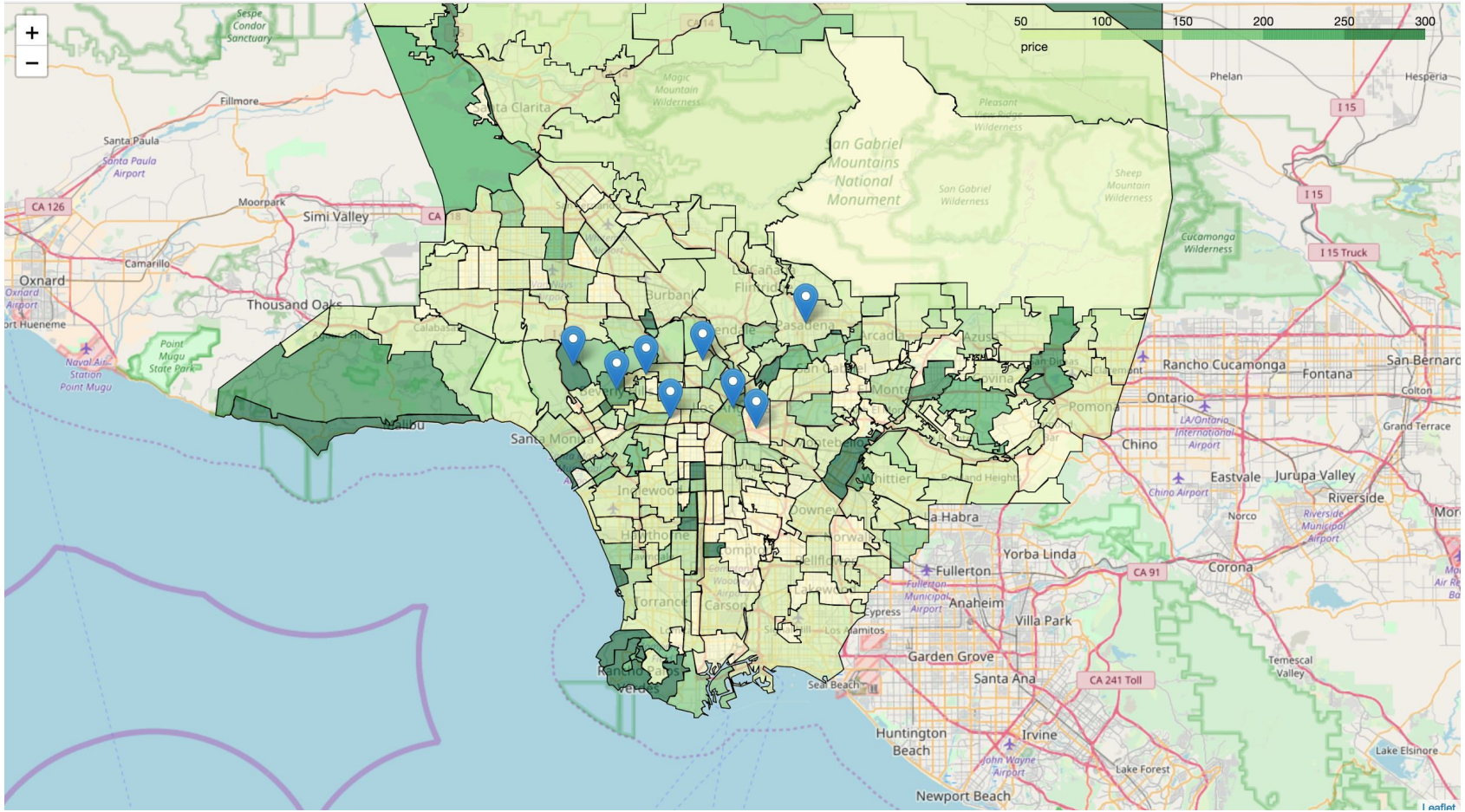
7 **Museum of Tolerance**

BEVERLY HILLS, BEL AIR, BRENTWOOD & WESTWOOD

8 **Petersen Automotive Museum**

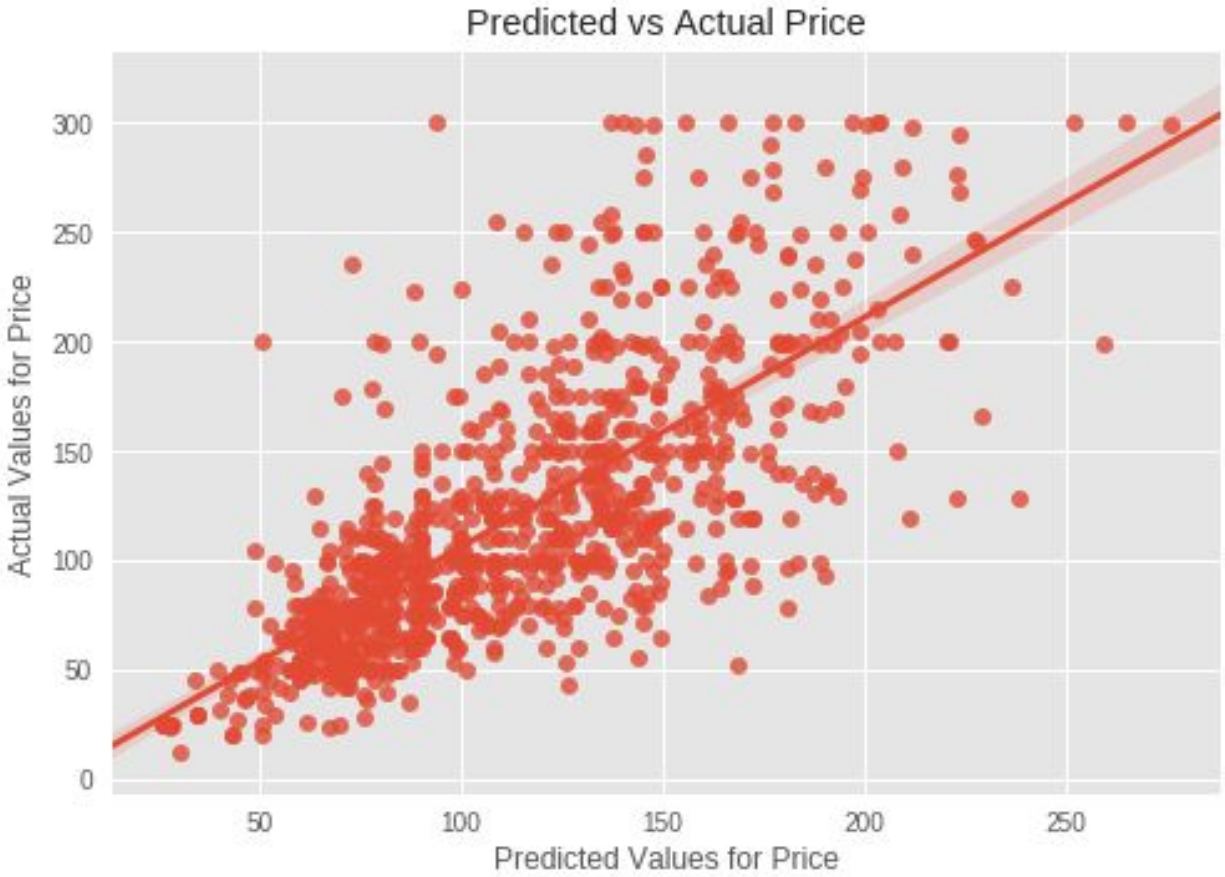
WEST HOLLYWOOD & MID-CITY

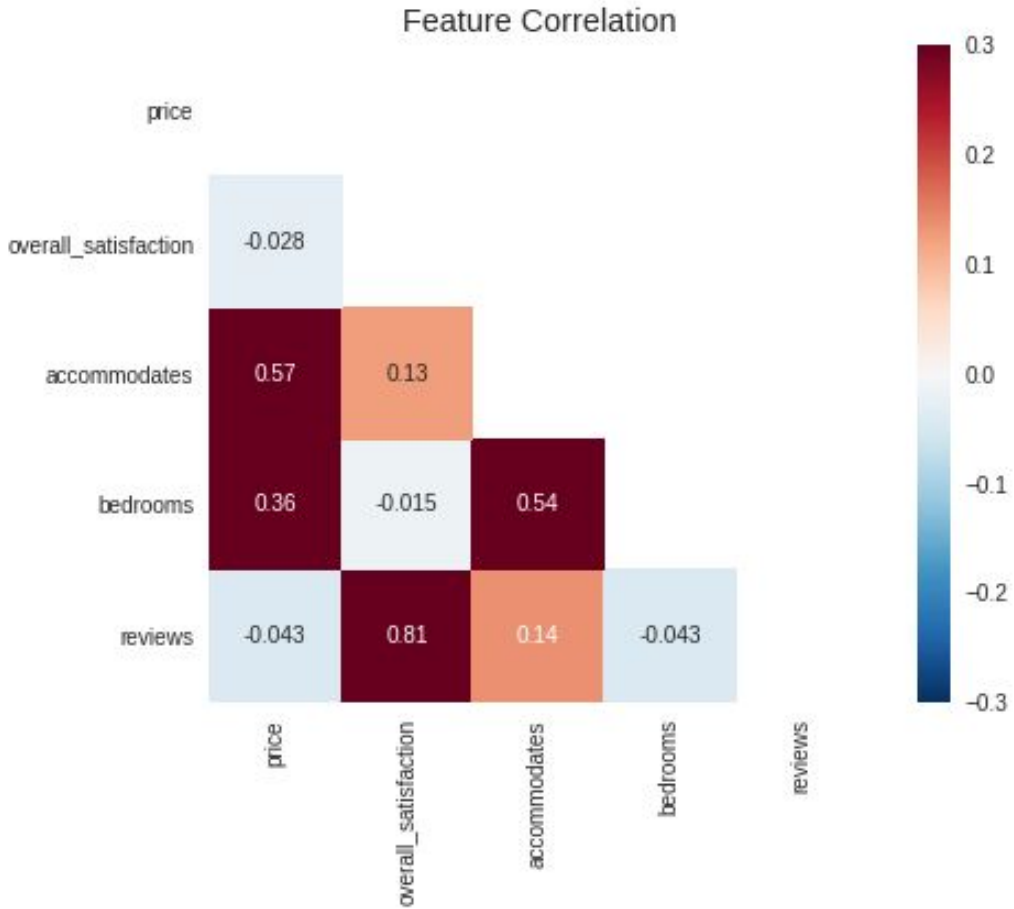
Price Distribution - All Neighborhoods < 1000
Mean
Number of Listings
12000
10000
8000
6000
4000
2000
0
0
200
400
600
800
1000
Price



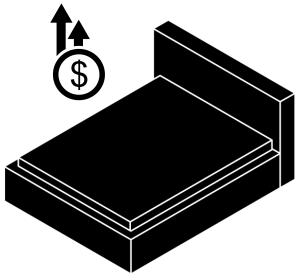
Airbnb → Goals → **Data** → Model → Findings → Limits → Next Steps

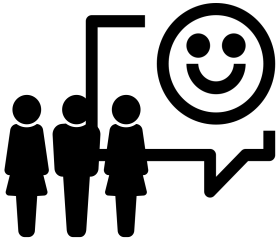
Price Distribution by Neighborhood
300
250
200
150
100
50
0
Price
Los Feliz
West Hollywood
Beverly Hills
Mid-City
Pasadena
Boyle Heights
East Los Angeles
Neighborhoods

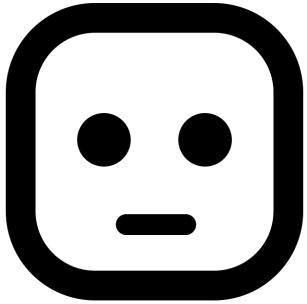

Predicted vs Actual Price
Actual Values for Price
Predicted Values for Price
0
50
100
150
200
250
300
50
100
150
200
250


Feature Correlation
price
overall_satisfaction
accommodates
bedrooms
reviews
-0.028
0.57
0.13
0.36
-0.015
0.54
-0.043
0.81
0.14
-0.043
price
overall_satisfaction
accommodates
bedrooms
reviews
0.3
0.2
0.1
0.0
−0.1
−0.2
−0.3

Findings







Risks/Limitations

1. Model is limited to neighborhoods
2. 60% of the variability in price is explained by the model
3. Future iterations will requires additional features

Next Steps

1. Incorporate more data
   a. National and International
   b. Archived data
2. Neighborhoods and marketing
3. Build a web application

Questions?

References & Resources

- https://www.airbnb.com/about/about-us
- https://www.airbnb.com/help/article/52/how-should-i-choose-my-listing-s-price?topic=212
- https://www.lonelyplanet.com/usa/los-angeles
- https://medium.com/airbnb-engineering/automated-machine-learning-a-paradigm-shift-that-accelerates-data-scientist-productivity-airbnb-f1f8a10d61f8
- https://medium.com/airbnb-engineering/using-machine-learning-to-predict-value-of-homes-on-airbnb-9272d3d4739d
- https://thenounproject.com/grega.cresnar/
- https://thenounproject.com/aneeque/
- https://thenounproject.com/muhammadfaisal40/

References & Resources

- https://thenounproject.com/homan.chris/collection/search/
- https://thenounproject.com/vectormarket01/
- https://thenounproject.com/marialuisa.iborra/
- https://thenounproject.com/coquet_adrien/
- https://thenounproject.com/seymur/collections/
- https://thenounproject.com/hoangloi/collections/