

Airbnb Listing Recommendations

Davis Vance

Motivation



Motivation

- ❖ Recommenders as a product reduce the burden of time spent on research
- ❖ Recreate Airbnb similar listings from scratch utilizing Natural Language Processing

Similar listings

The image shows three separate screenshots of an Airbnb search interface, each displaying a different listing under the heading "Similar listings". Each listing includes a thumbnail image, the room type, the listing name, price per night, and the number of reviews.

- PRIVATE ROOM · 1 BED**
The Justice Tower 2
\$125 per night
★★★★★ 11
- PRIVATE ROOM · 1 BED**
Enjoy Brooklyn & Proximity to Train
\$60 per night
★★★★★ 10
- PRIVATE ROOM · 1 BED**
Large Room w/ Wi-Fi in Brooklyn
\$46 per night
★★★★★ 25

Data

- ❖ InsideAirbnb.com
- ❖ Focus on New York 2018
- ❖ 100k Airbnb listings
- ❖ 100 data points per listing
 - Descriptions
 - Amenities
 - Lat/ Long
 - Pricing

Counts of Cities in Dataset



Natural Language Processing

Recommendations based on text data to distinguish
users' interests they didn't know existed

Natural Language Processing

- ❖ Bucket text into 2 categories & counting words and word pairs
 - Listing information
 - Host information
 - uni/bigrams
 - 182k terms
 - 72k terms

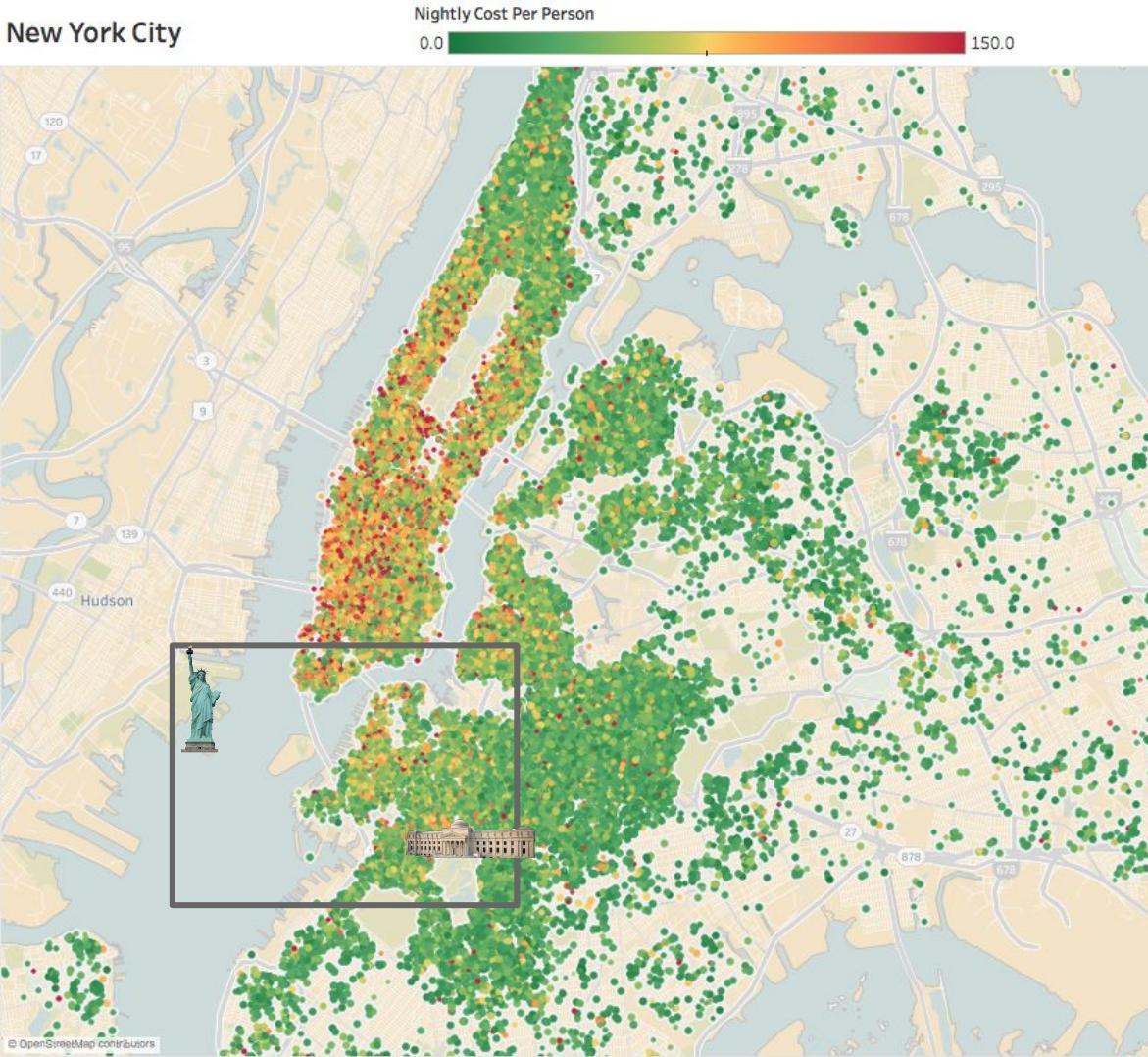
- ❖ Condense listings to calculate similarity
 - Latent Semantic Indexing

The System

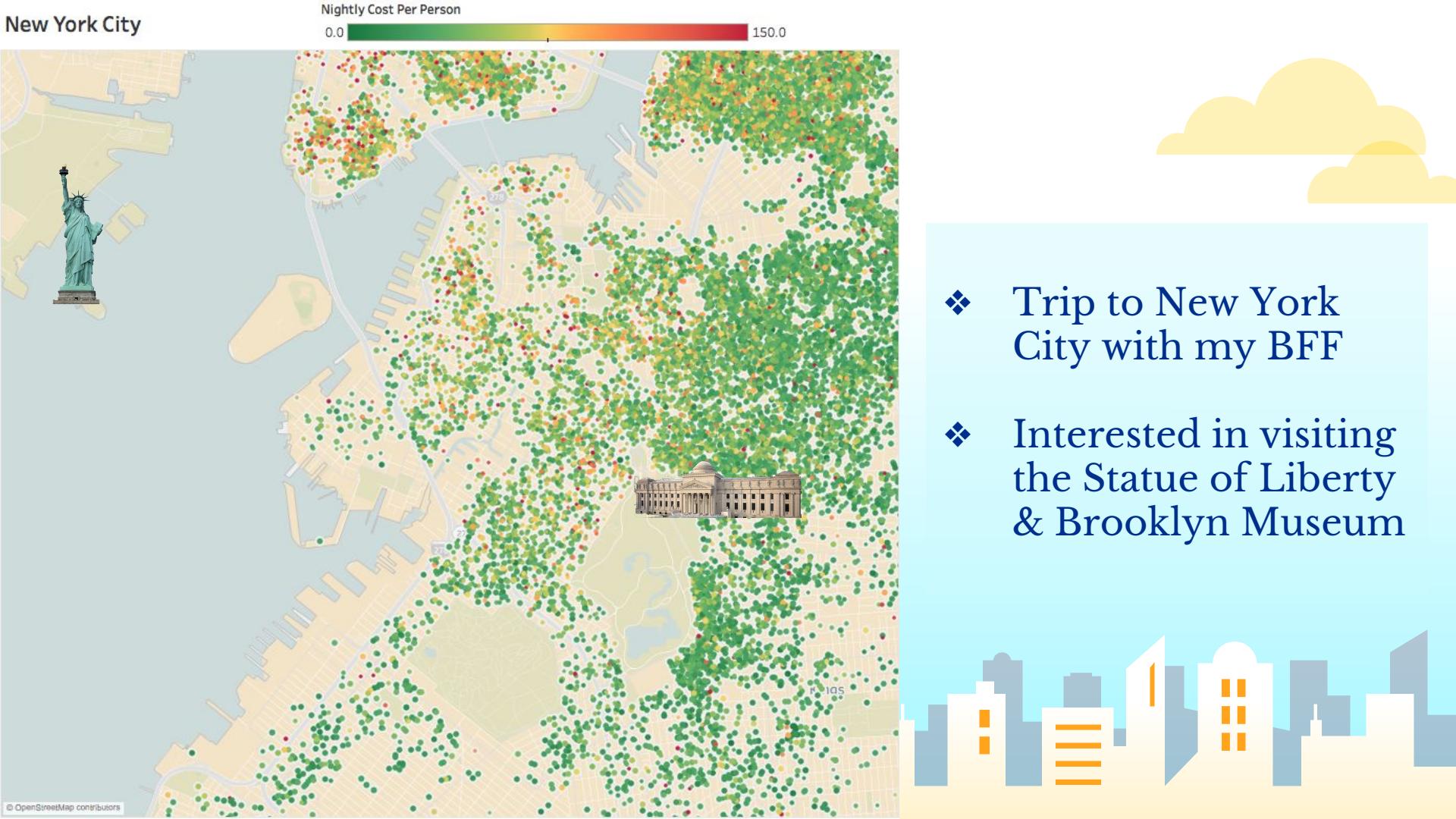
Combining the engineered NLP features with selected
features for calculating an overall similarity score

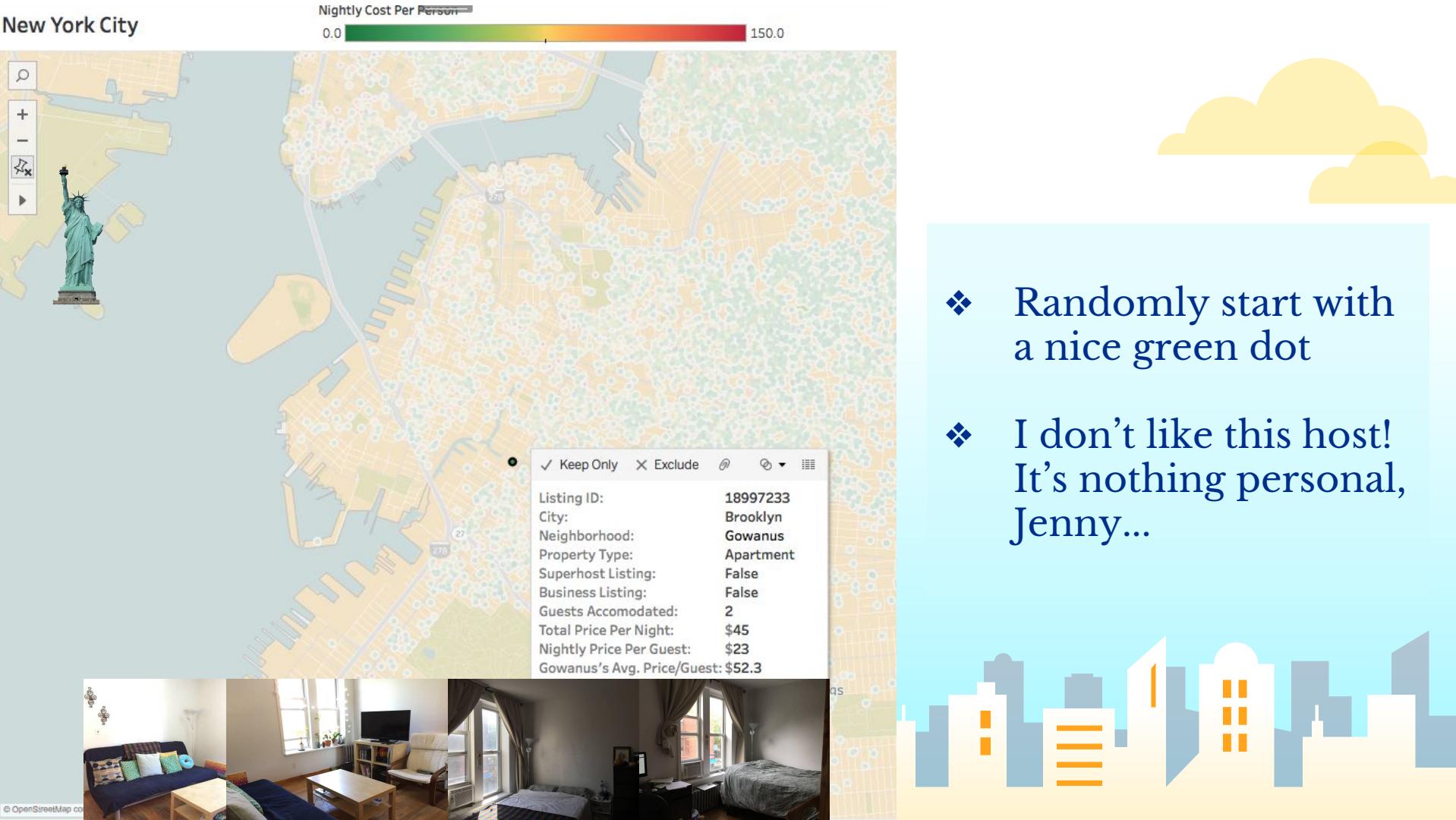
Selected Features

- ❖ 8 selected features power the recommendations
 - Text similarities
 - Price & location
 - Helper features for extra stability
- ❖ Rescale all features for equidistant comparison
 - Robust quantile transformer
- ❖ Ask user for preferences → Re-weight accordingly

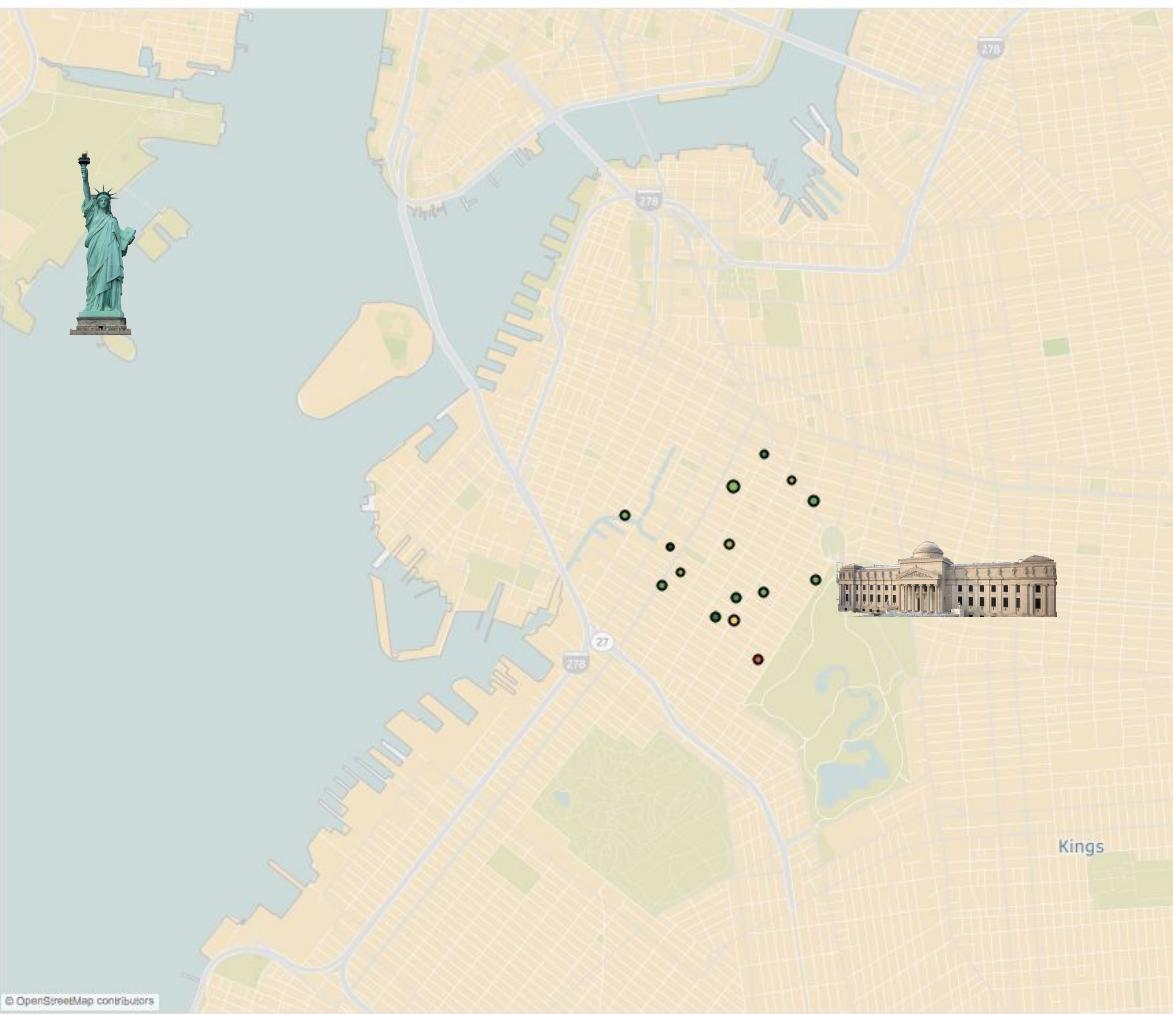


- ❖ Trip to New York City with my BFF
- ❖ Interested in visiting the Statue of Liberty & Brooklyn Museum





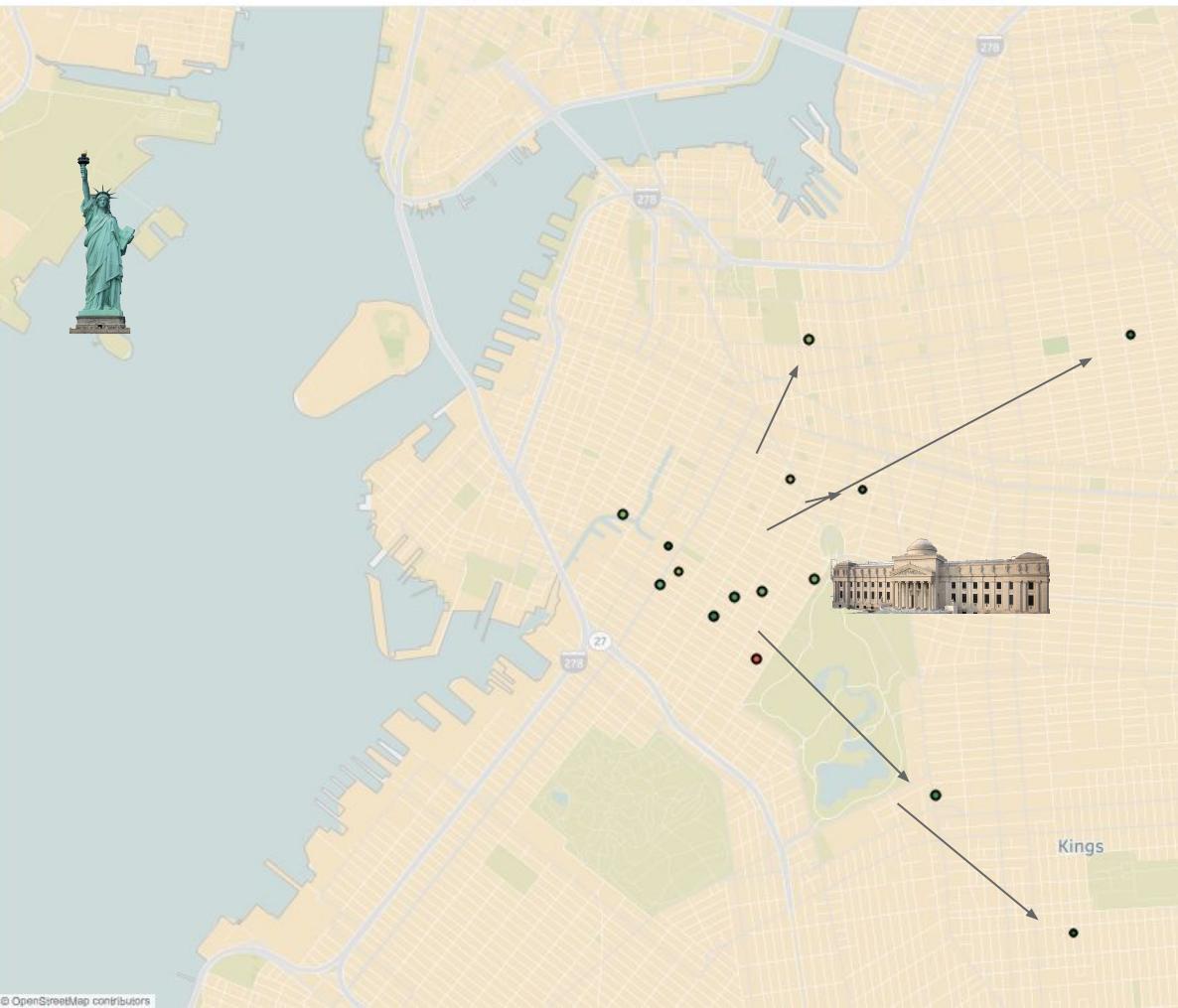
New York City



User Prompt:

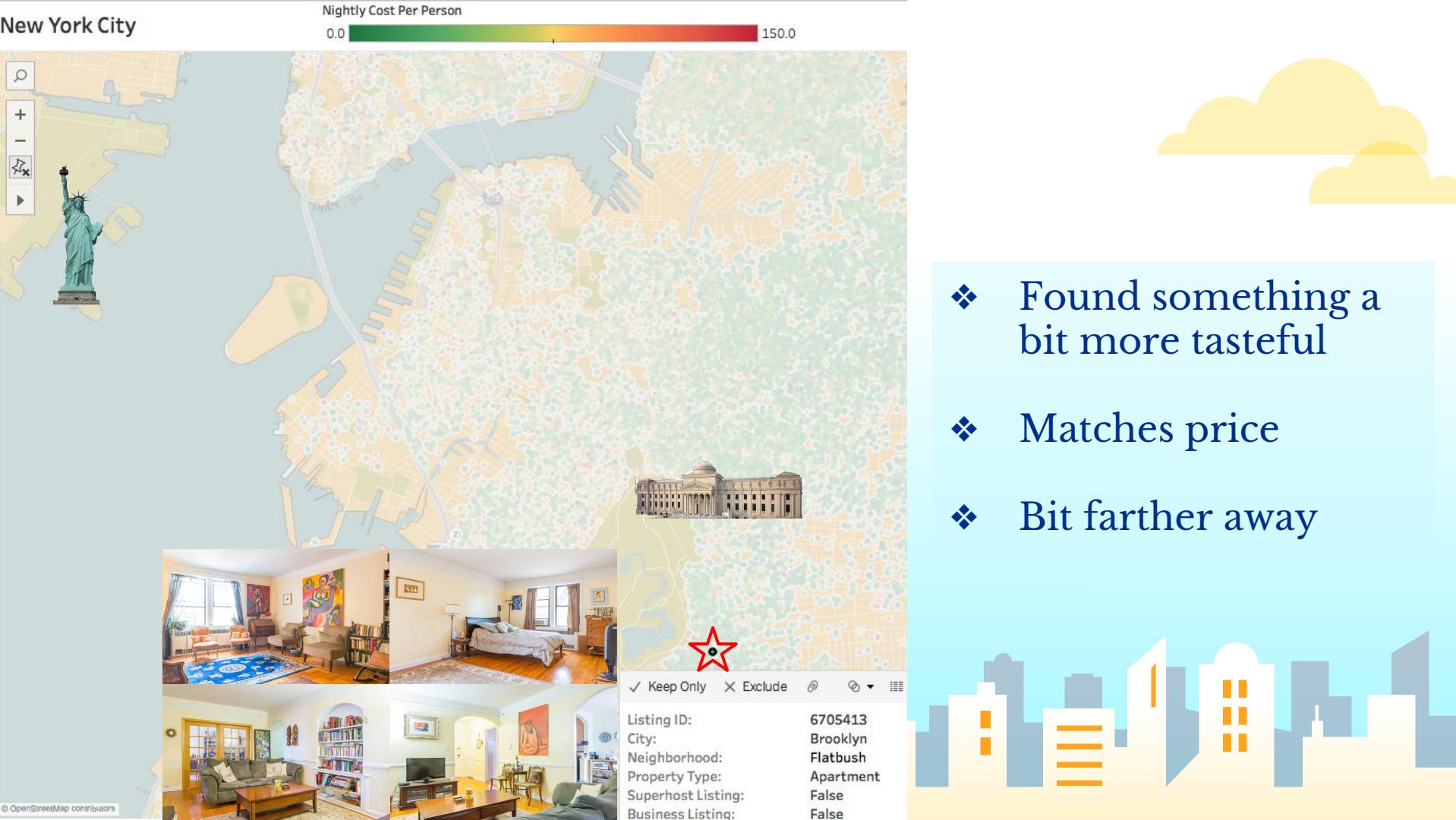
- ❖ Same listing description? - Yes
- ❖ Same host description? - No
- ❖ Business trip? - No
- ❖ Similar price? - 1/5
- ❖ Similar area? - 5/5
➤ 5 = close proximity

New York City



User Prompt:

- ❖ Same listing description? - Yes
- ❖ Same host description? - No
- ❖ Business trip? - No
- ❖ Similar price? - 5/5
- ❖ Similar area? - 1/5
➤ 1 = exploration



What Else?

- ❖ Add more features, like using these landmarks as signals
- ❖ Track page viewing behavior to understand what a user cares about and learn weights **implicitly**
- ❖ Use image data to extract more details with CNN's

“Search Ranking and Similar Listings drive 99% of [Airbnb’s] booking conversions.”

Mihajlo Grbovic

Senior Machine Learning Scientist - Airbnb





me

Thank You!



Daviscvance@gmail.com



Linkedin.com/in/daviscvance



Github.com/daviscvance



Davis Vance
B.A. in Economics
San Diego State Univ.

Credits

Special thanks to all the people who made and released these awesome resources for free:

- ❖ Murray Cox for all the data at InsideAirbnb.com
- ❖ Presentation template by [SlidesCarnival](https://SlidesCarnival.com)
- ❖ House Photos on [Unsplash](https://Unsplash.com)
 - Thanks [Jeremy Kovac](https://Unsplash.com/photos/jeremykovac), [Bailey Hall](https://Unsplash.com/photos/baileylh), [Abbie Bernet](https://Unsplash.com/photos/abbiebernet)