VISUALIZING URBAN HOSPITALITY

YUNHAN XU
HARVARD COLLEGE
B.A. CANDIDATE
STATISTICS
APRIL 2016

We use data from Airbnb, an online accommodations market, to interpret the characteristics of local peer-to-peer exchange between hosts and visitors in New York City. Using the variables of annual availability, proximity to mass transit, proximity to other listings, pricing, and guest ratings and reviews, we can construct a hospitality index that assigns a composite score to each neighborhood, thus locating the optimal regions for visitors.

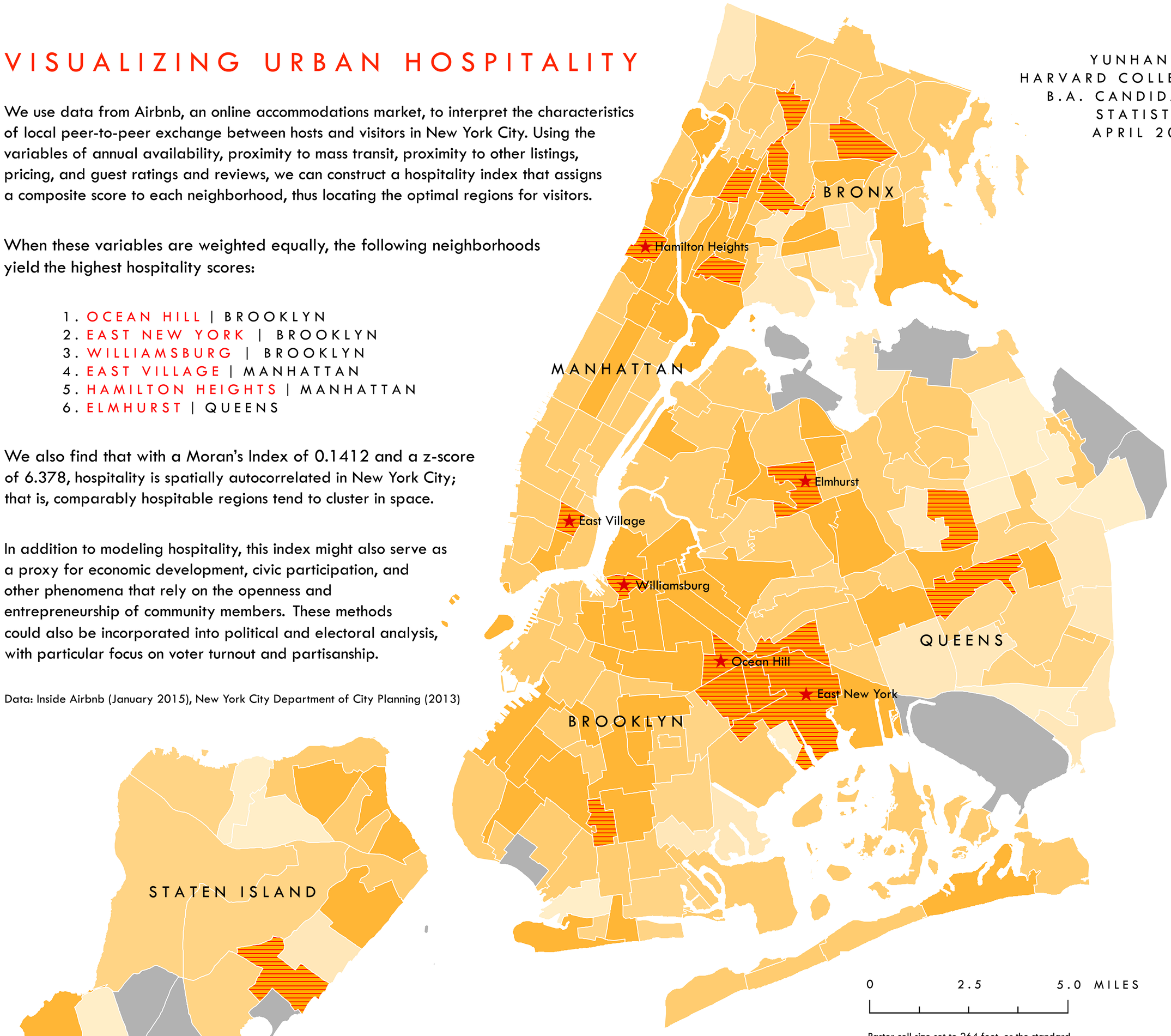
When these variables are weighted equally, the following neighborhoods yield the highest hospitality scores:

1. OCEAN HILL | BROOKLYN
2. EAST NEW YORK | BROOKLYN
3. WILLIAMSBURG | BROOKLYN
4. EAST VILLAGE | MANHATTAN
5. HAMILTON HEIGHTS | MANHATTAN
6. ELMHURST | QUEENS

We also find that with a Moran's Index of 0.1412 and a z-score of 6.378, hospitality is spatially autocorrelated in New York City; that is, comparably hospitable regions tend to cluster in space.

In addition to modeling hospitality, this index might also serve as a proxy for economic development, civic participation, and other phenomena that rely on the openness and entrepreneurship of community members. These methods could also be incorporated into political and electoral analysis, with particular focus on voter turnout and partisanship.

Data: Inside Airbnb (January 2015), New York City Department of City Planning (2013)

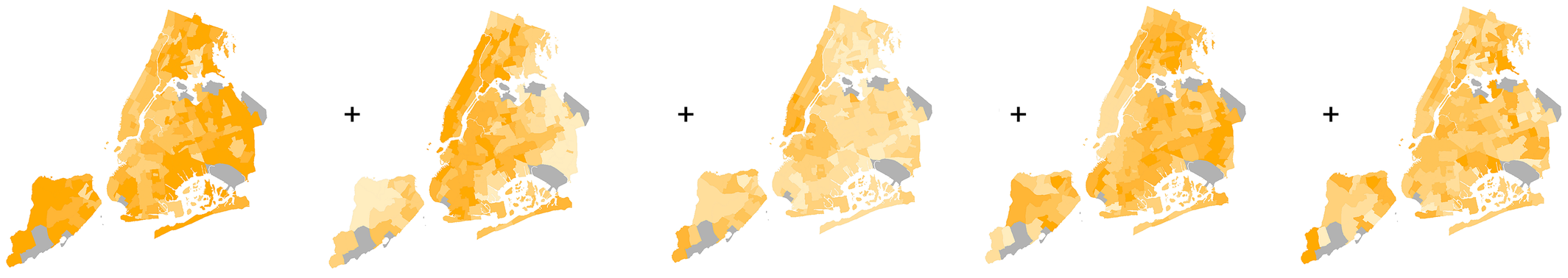


Raster cell size set to 264 feet, or the standard length of a city block in Manhattan.

HOSPITALITY INDEX

≤4.0 | 4.8 | 5.2 | 5.6 | 6.0 | 6.4 | ≥7.0

AVAILABILITY + ACCESSIBILITY + RENTAL DENSITY + AFFORDABILITY + QUALITY



Days per year that listing is available for rental

Proximity of listing to subway stations

Proximity of listing to other listings + rental density per neighborhood

Price of listing per night

Average rating of listing + calculated score of parsed reviews

1. Rasterization 2. Reclassification 3. Weighted overlay 4. Zonal statistics — ArcMap

5. Text analysis — R, Python