

The Promise and Ethical Concerns of Using Geotagged Social Media Data in Spatial Research



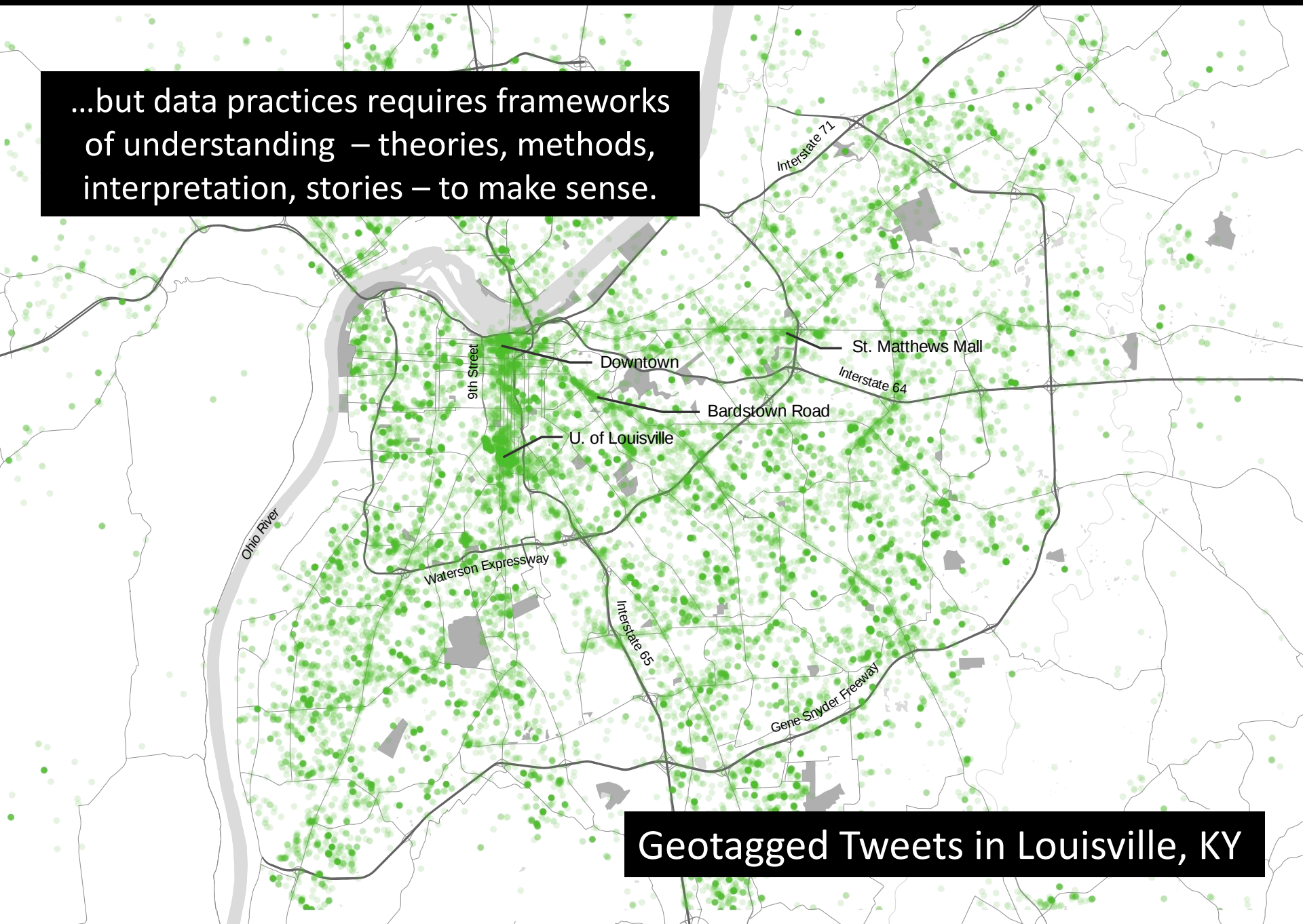
Prof. Matthew Zook
University of Kentucky
May 15, 2018

National Academies of Sciences, Engineering,
and Medicine, Geographical Sciences
Committee (GSC).

The promise of micro-spatial data for human geographic research is **not solely its size** but the **wealth of social processes** that are **encoded** in such data...



...but data practices requires frameworks of understanding – theories, methods, interpretation, stories – to make sense.



Geotagged Tweets in Louisville, KY

Approach

- Important to think “beyond the geotag”, *i.e.*, the relatively simple spatial ontology of latitude and longitude coordinates.
- Need holistic methodological and theoretical approaches to space (see Jessop *et al*, 2008; Crampton *et al* 2013) to consider:
 1. unevenness across scales and time;
 2. spatially-distanciated, relational networks which “jump” space; and
 3. mixed method approaches to these data.

Opportunities for Human Geography

Event based identification

Metropolitan mobility patterns and
activity spaces

Mapping the spaces and networks of
economic knowledge

Hurricane Sandy
October 24-31, 2012
4227 sites

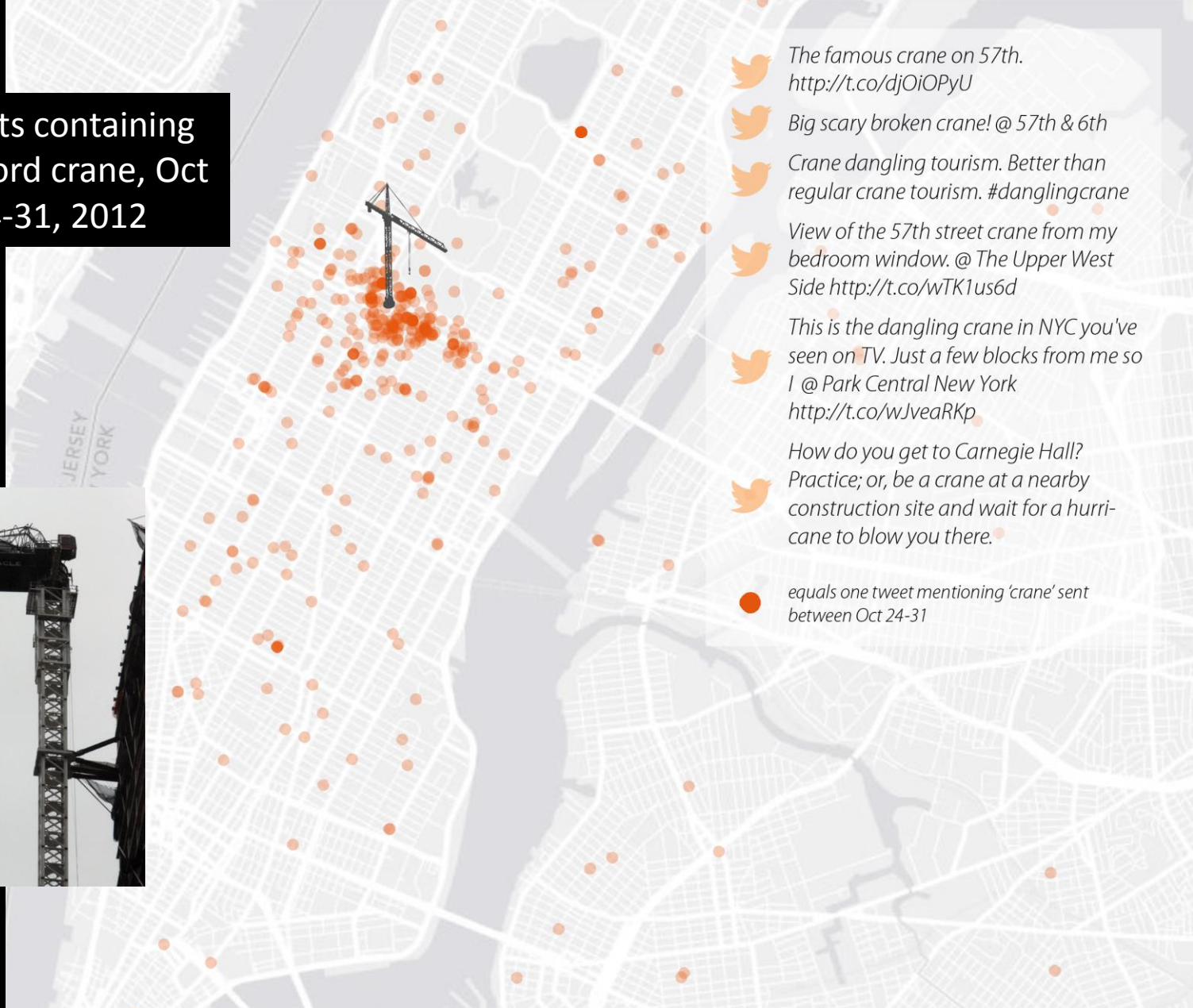
Track

1
3
5
7
10

Maximum: 12.83"
Bellevue 0.3 NNE, MD



Tweets containing
the word crane, Oct
24-31, 2012



The famous crane on 57th.
<http://t.co/djOiOPyU>

Big scary broken crane! @ 57th & 6th

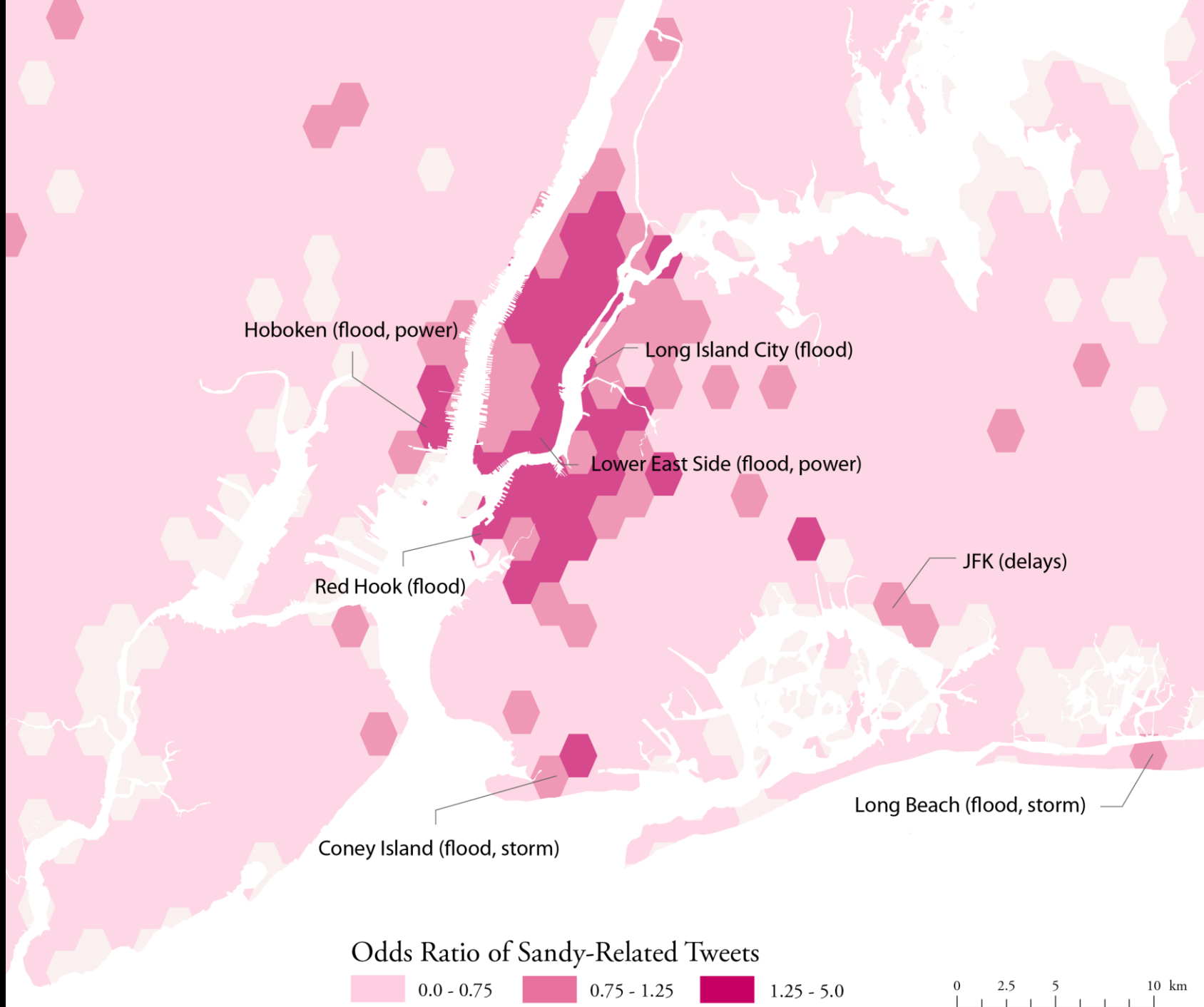
Crane dangling tourism. Better than
regular crane tourism. #danglingcrane

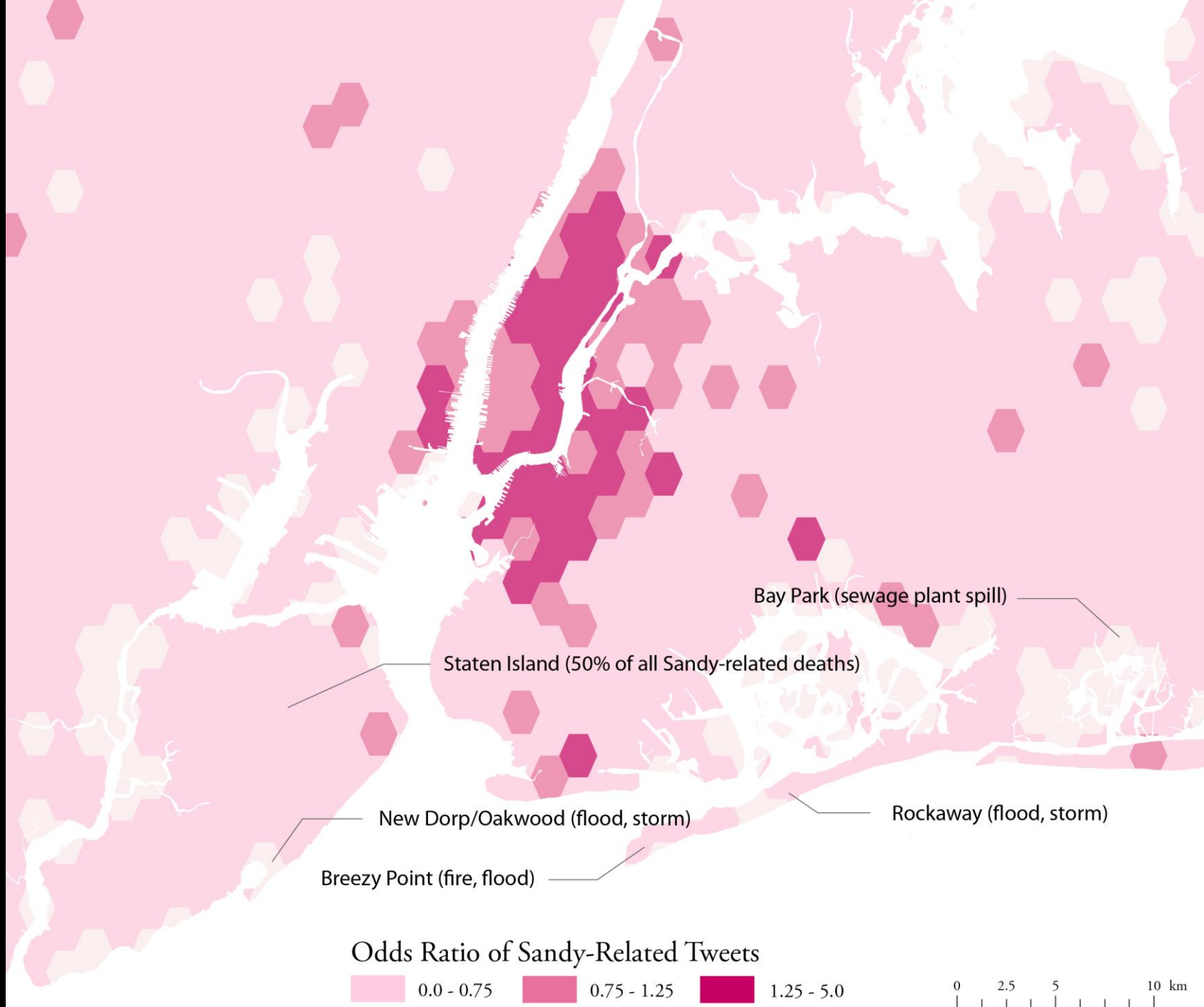
View of the 57th street crane from my
bedroom window. @ The Upper West
Side <http://t.co/wTK1us6d>

This is the dangling crane in NYC you've
seen on TV. Just a few blocks from me so
I @ Park Central New York
<http://t.co/wJveaRKp>

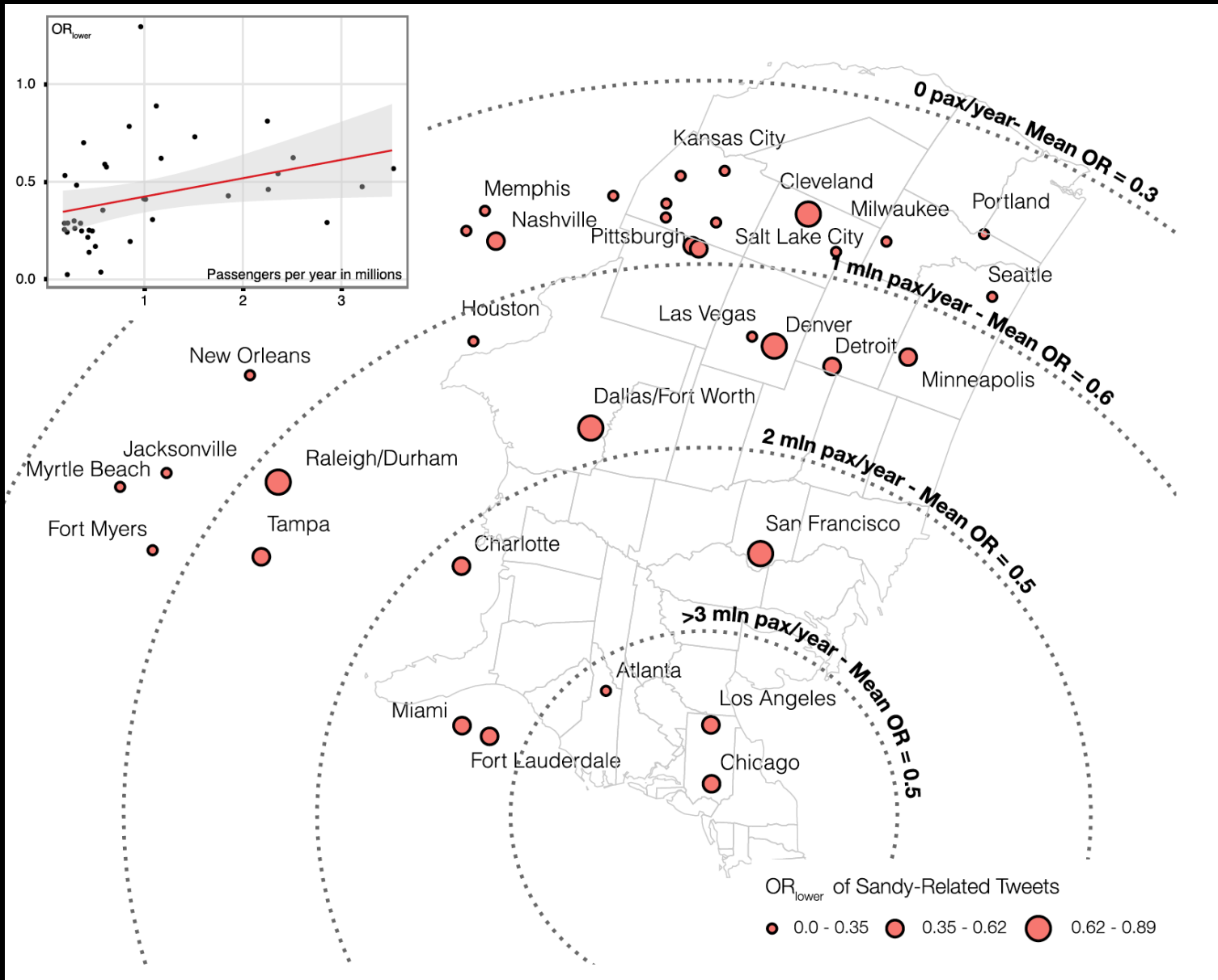
How do you get to Carnegie Hall?
Practice; or, be a crane at a nearby
construction site and wait for a hurri-
cane to blow you there.

● equals one tweet mentioning 'crane' sent
between Oct 24-31





Relational connections to NY



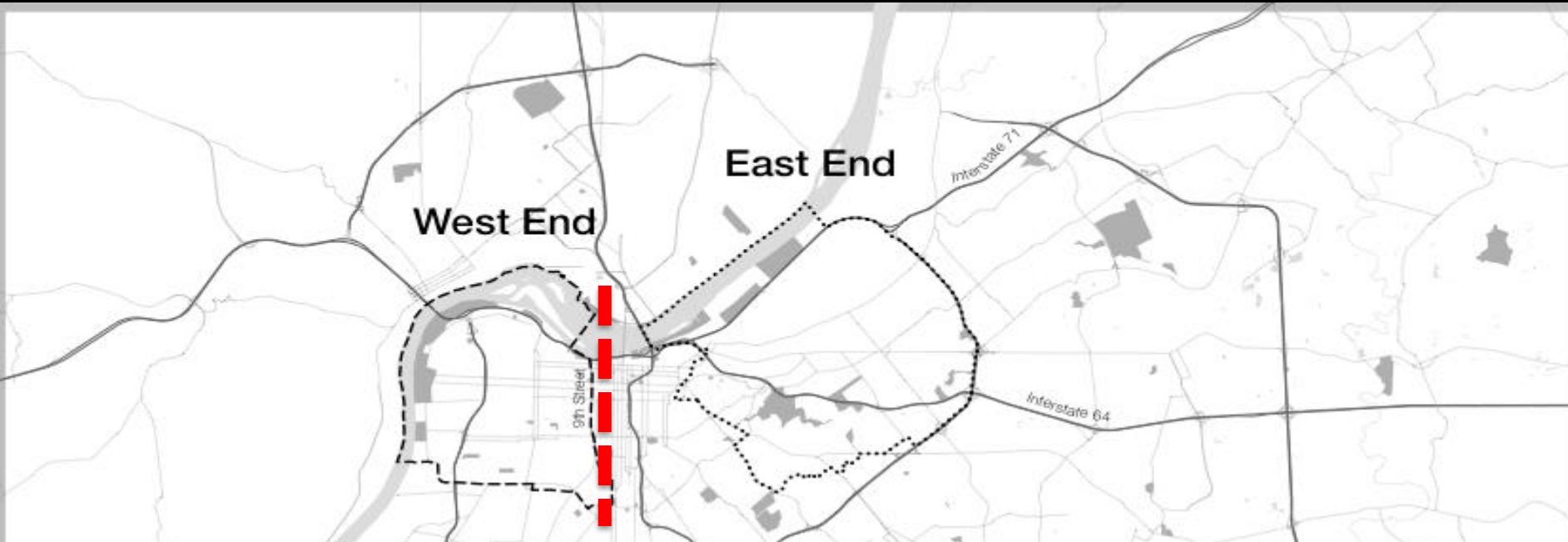
Opportunities for Human Geography

Event based identification

Metropolitan mobility patterns and
activity spaces

Mapping the spaces and networks of
economic knowledge

Louisville and the 9th Street Divide

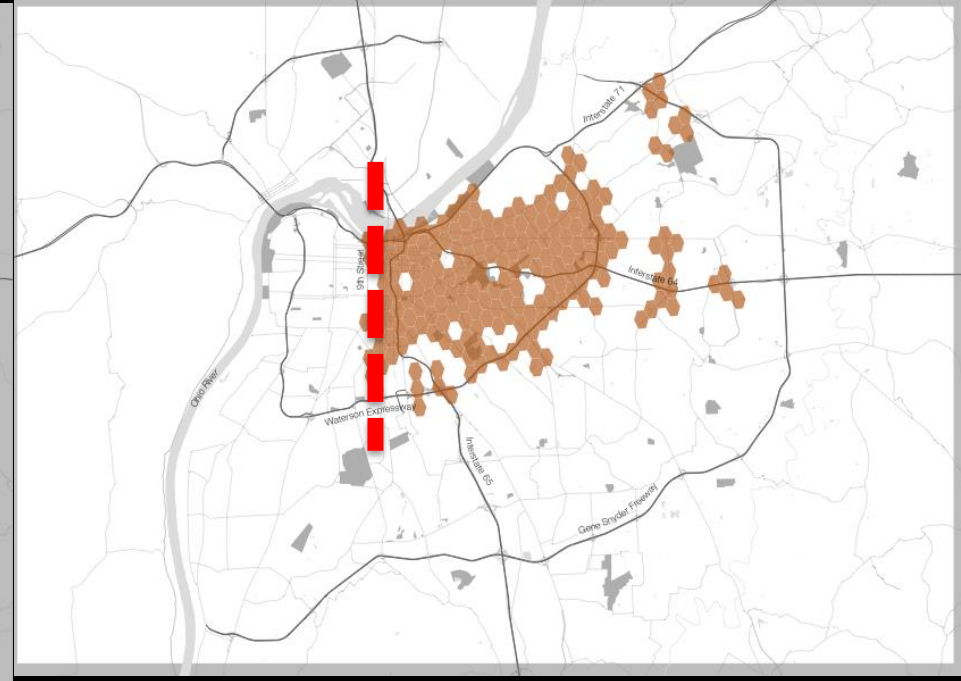
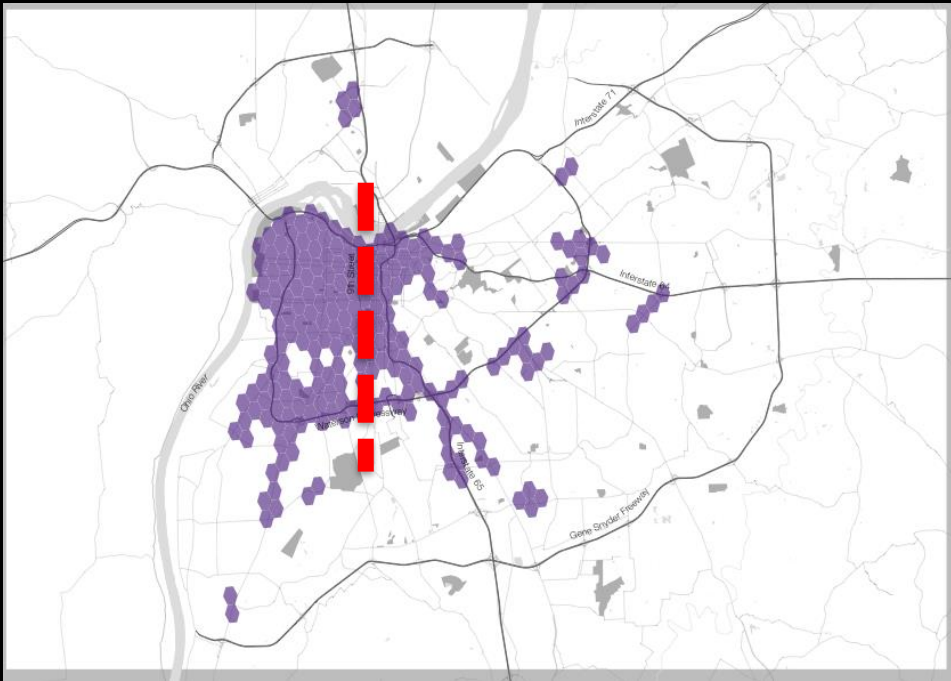


“Any Louisvillian who has lived here for more than a few years knows, almost instinctively, the boundary line between west Louisville and the rest of Louisville: Ninth Street. Most white Louisvillians know it because they’ve heard some variation of the warning, ‘Don’t go west of Ninth Street.’...Although the notion that west Louisville is a dangerous and even foreign place is embedded in the mental map that many of us – even the most bleeding-heartedly liberal and racially tolerant, if we will admit it – carry around in our heads, it is rarely talked about in public” (Crutcher 2013: 25)

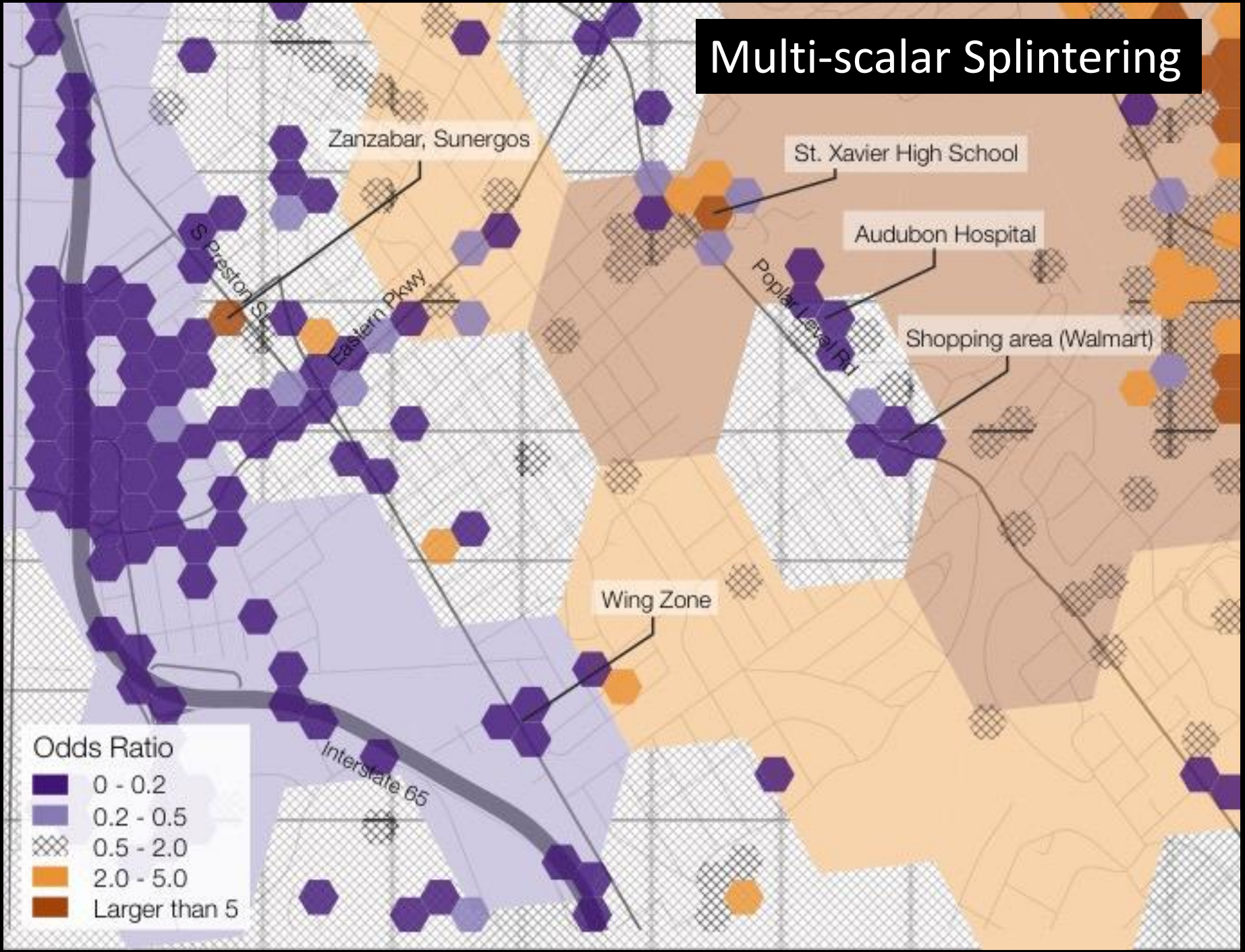
- Defined west and east ends neighborhoods.
- Identified all users with 50+ geotagged tweets in either neighborhood.
- Users with 51 percent of their tweets in either the west/east end were included.
- Analyzed the spatial distribution of these users' tweeting activity.



Redefining the Boundaries of the West End and the East End



Multi-scalar Splintering

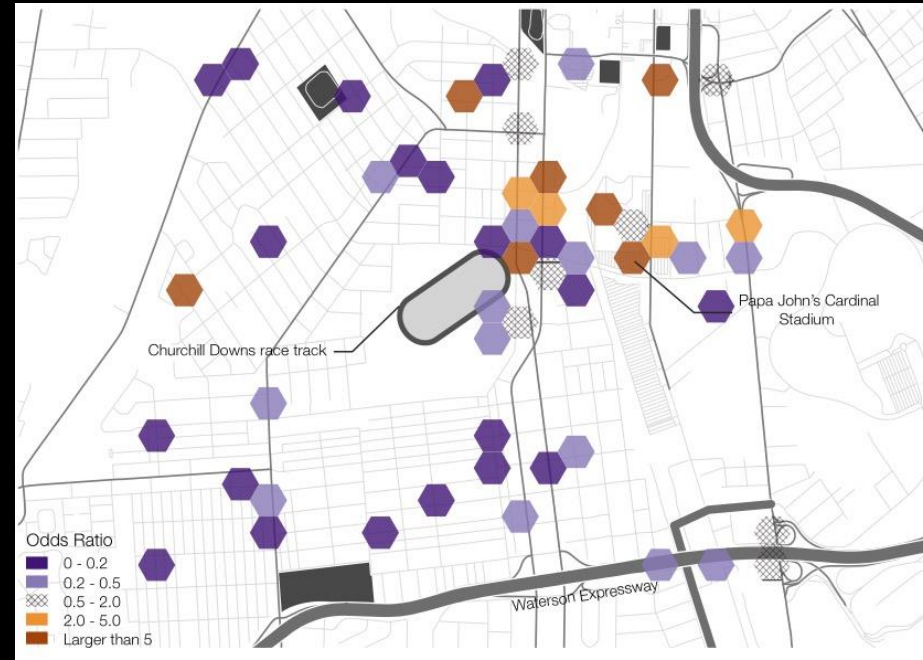


Temporal Differences

Racing Season



Non-Racing Season



“most often this area is a spatial extension of the West End, though it temporarily transforms into a space of (relatively) elite consumption and a site of global attention, despite still being surrounded by a neighborhood from which many visitors have little connection whatsoever”

Opportunities for Human Geography

Event based identification

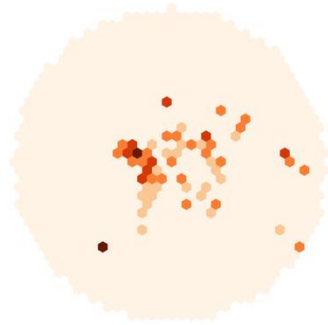
Metropolitan mobility patterns and
activity spaces

Mapping the spaces and networks of
economic knowledge

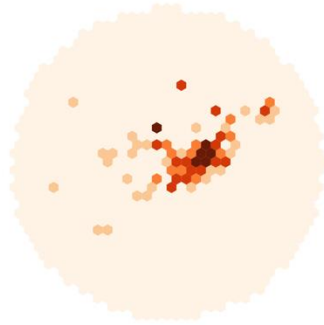
Measuring Knowledge and the Attention Economy

- The key economic input of knowledge is hard to measure, particularly within cultural industries.
- Attracting, creating and processing attention is a form of knowledge work (Walker, 2007) and is an inherent part of the cultural economy.
- Within the fashion industry firms (and other actors) compete intensely to curate attention for their particular aesthetic/brand.

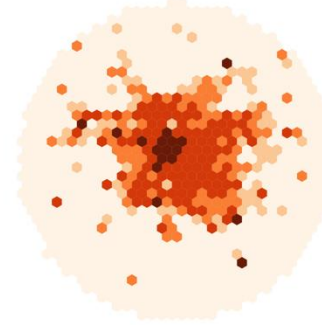
Capetown



Sydney



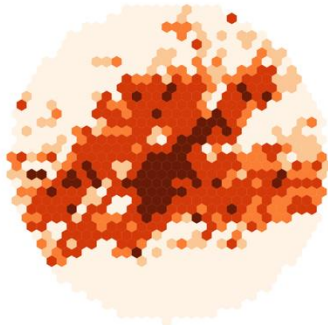
Sao Paulo



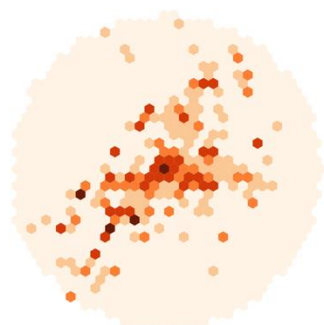
Singapore



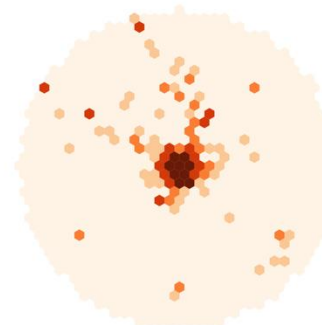
New York



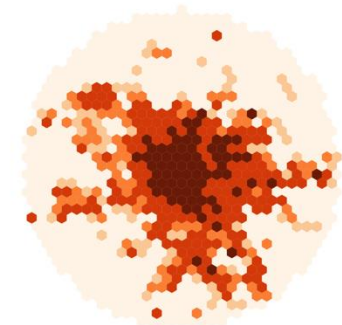
Minneapolis



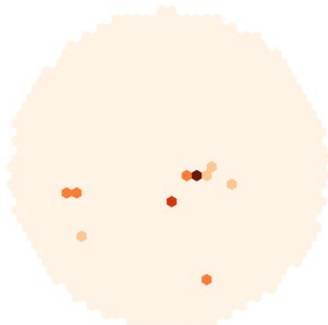
Milan



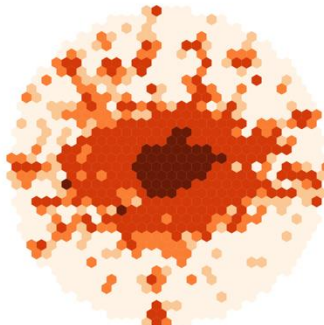
Paris



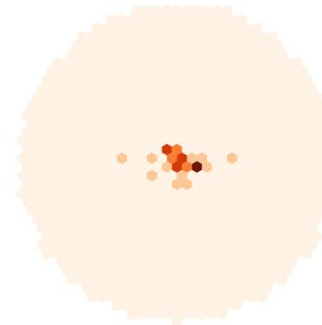
Frankfurt



London



Warsaw



Count (absolute intensity)

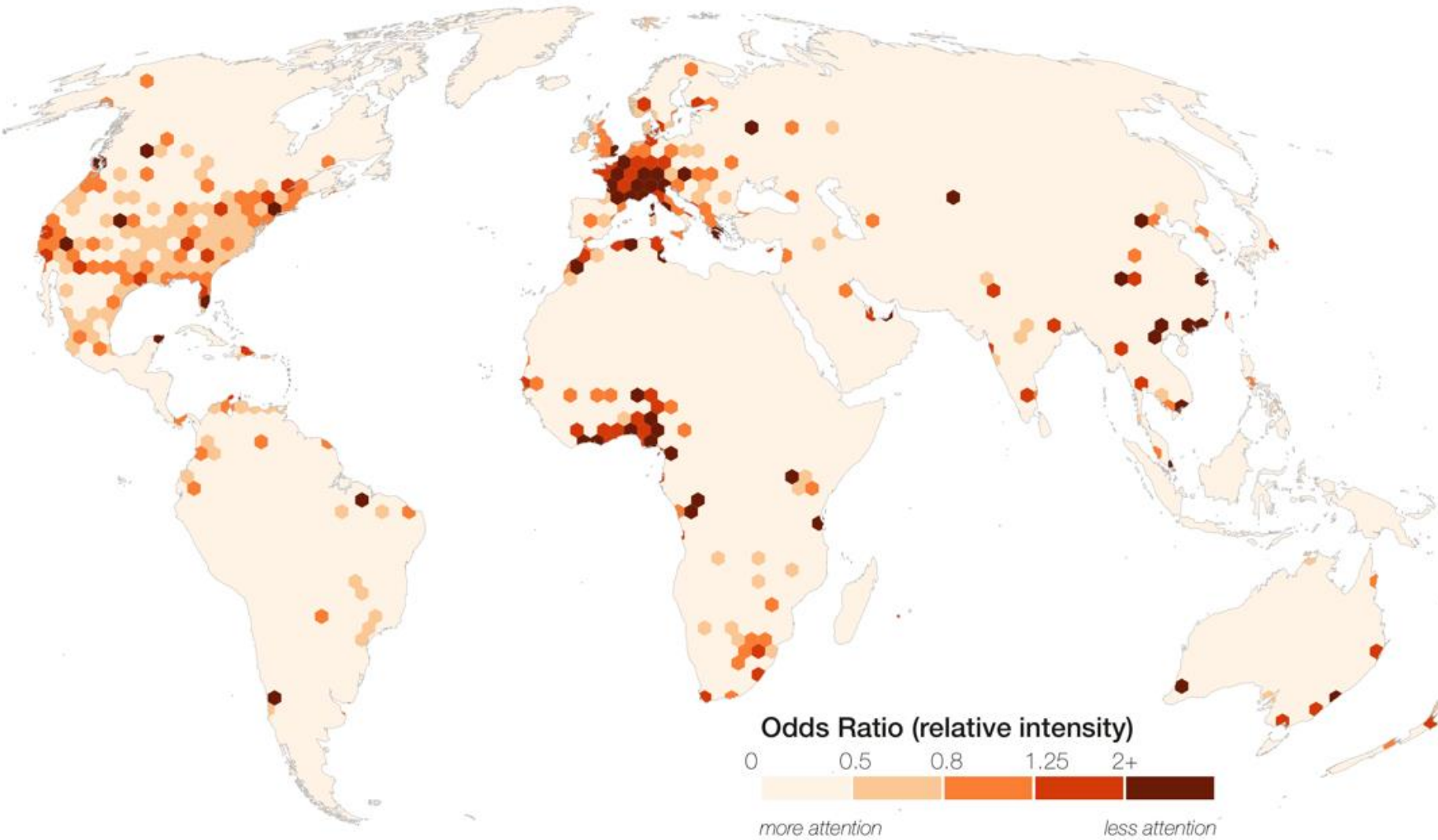


more attention

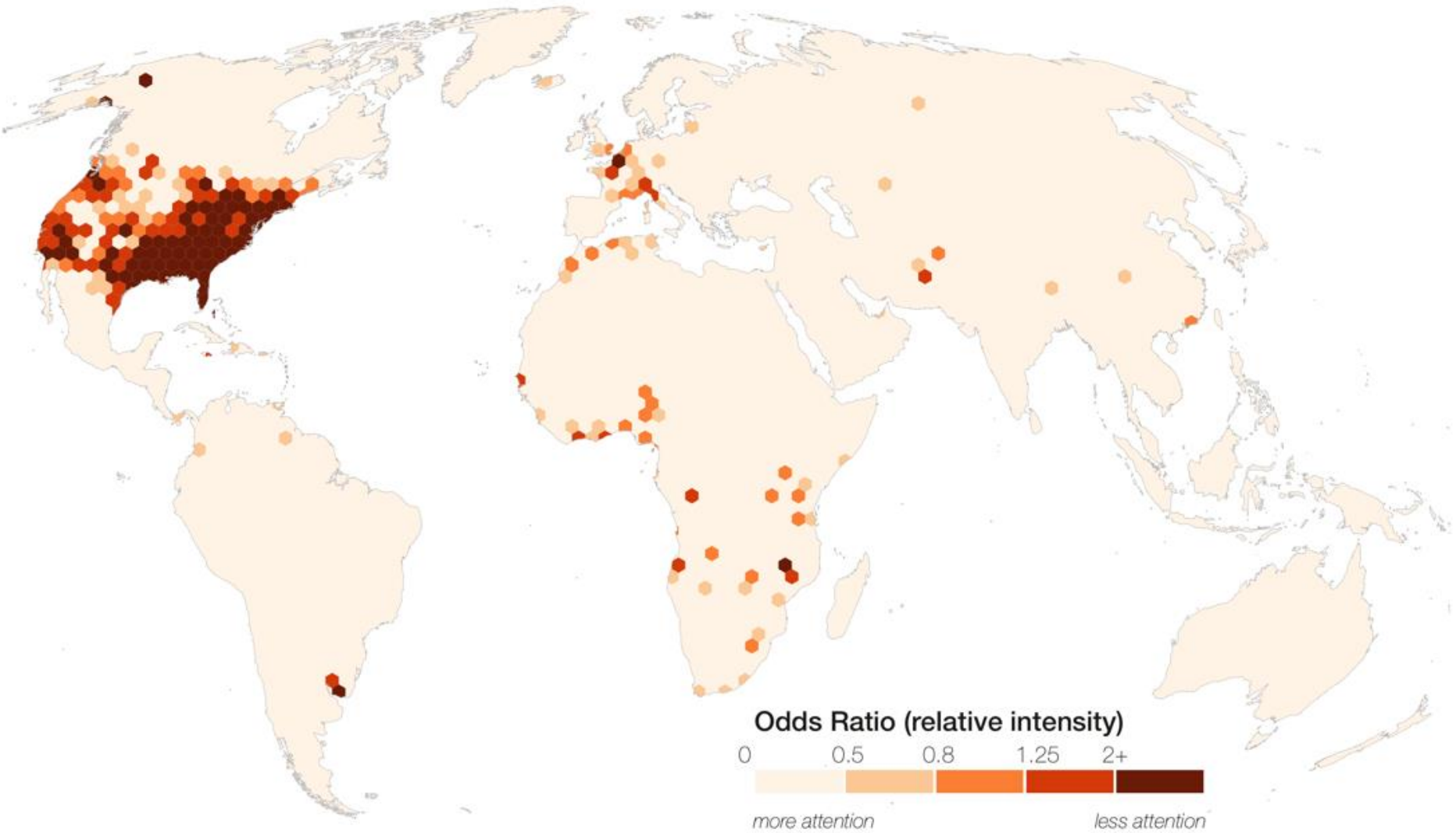
less attention

Used the Business of Finance 500 index to identify keywords (names, firms, brand) for selecting social media postings (tweets)

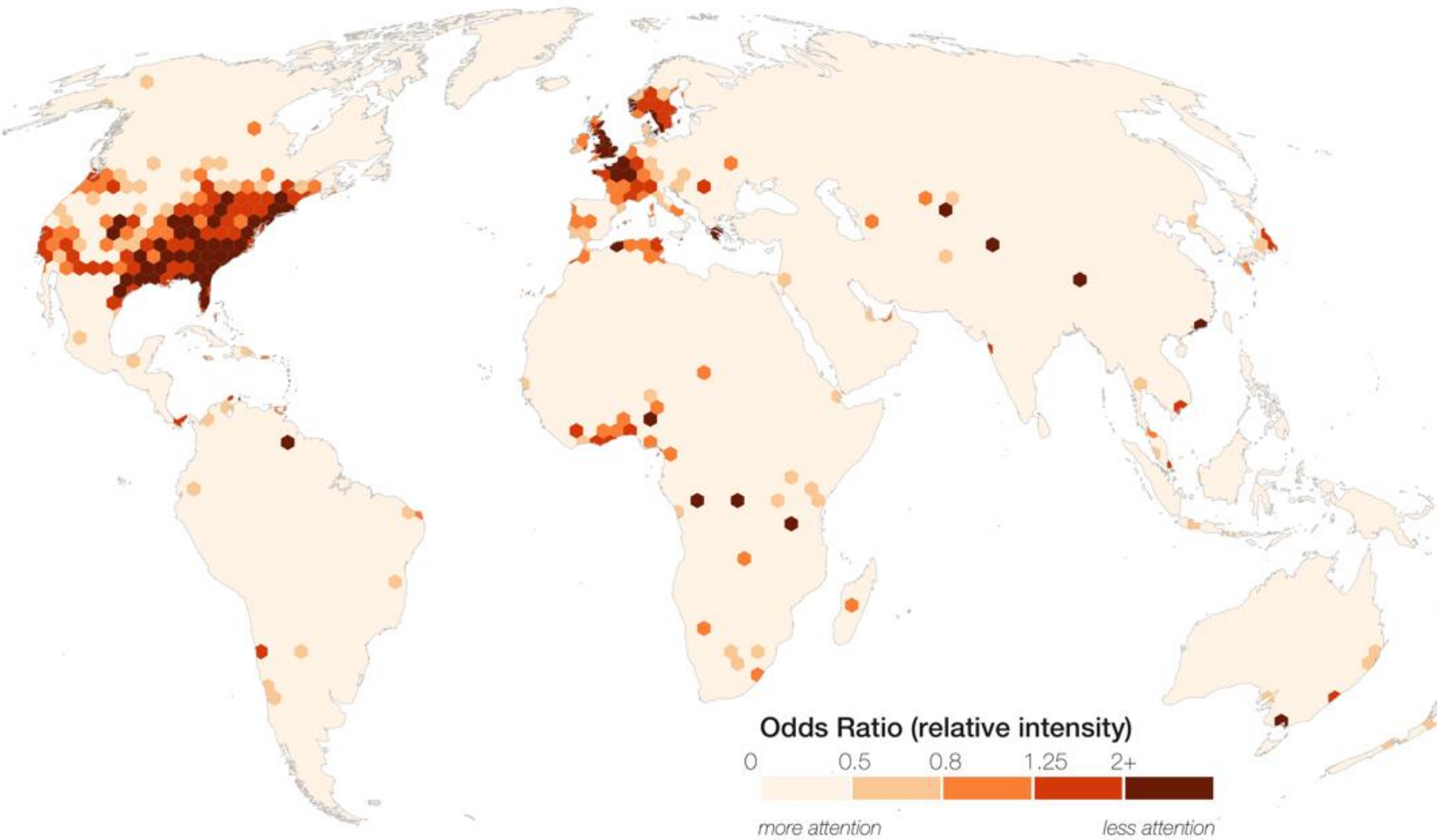
Total Attention to Louis Vuitton



Total Attention to Gucci



Total Attention to Ralph Lauren



Concerns: Privacy and Consent

- Standards for using and protecting social media data are ad hoc with uncertain protocols.
- “Publically shared” does not mean any subsequent use is unproblematic.
 - A single Instagram photo versus all posts.
- Location is uniquely sensitive and revealing.
 - 87 percent of the US population of the US could be uniquely identified based only on 5-digit ZIP, gender, date of birth (Sweeny 2000).

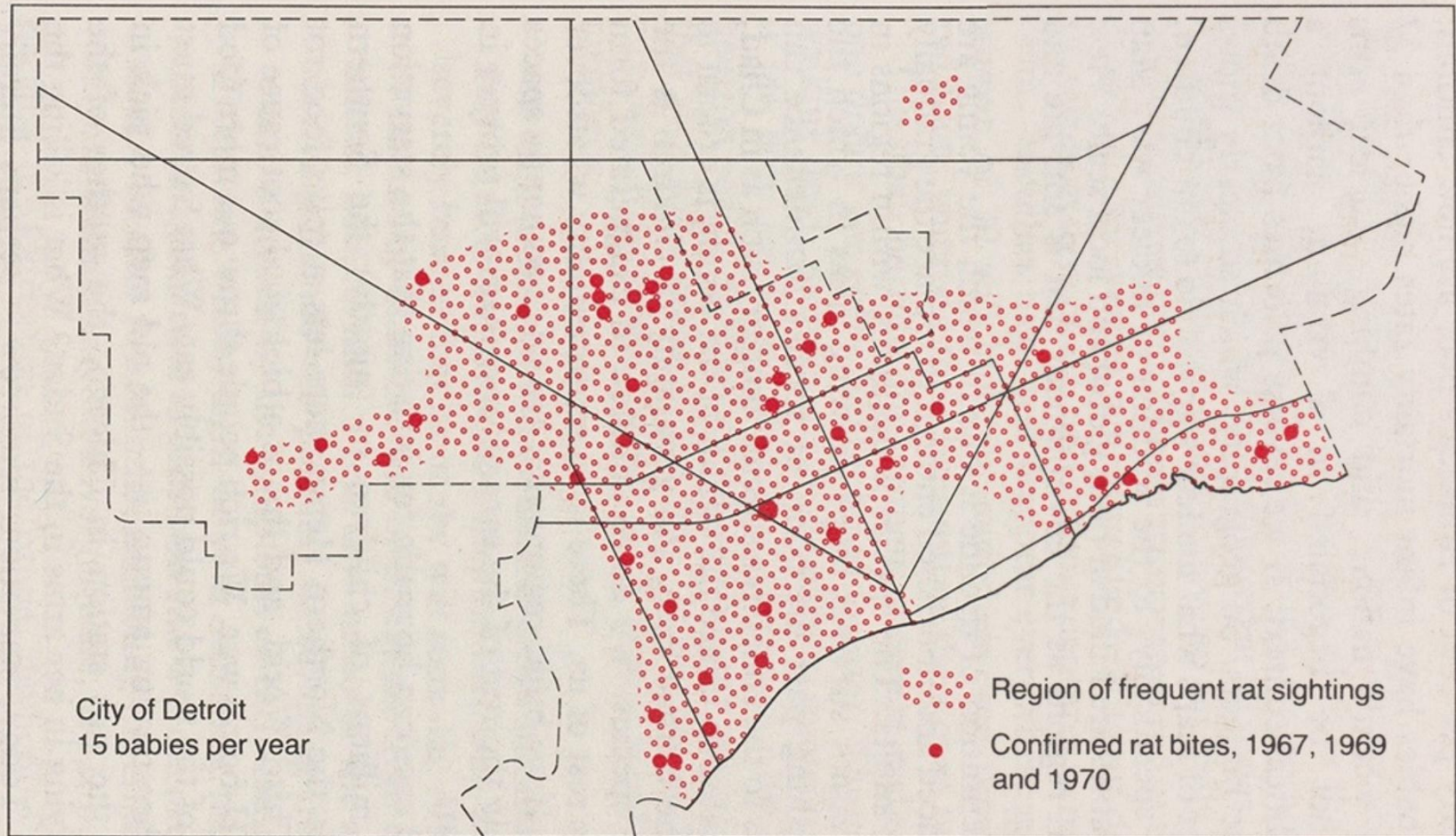
Concerns: Sharing Data

- Anonymization of data is difficult, lots of failures (Barbaro et al, 2006; Cox, 2016; Panduragnan, 2014).
- Data that is not directly personal - such as location and movement, call records and even high resolution satellite images of the earth - can be used to re-identify people (Kloumann and Kleinberg, 2014).
- Hard to know vulnerable points a priori and “harmless data” may very well prove to be a significant vector of identification (e.g., battery usage varies by location).

Concerns: Producing Social Inequality

- Use of data goes beyond single individuals and extends to groups such as disproportionately impacting specific demographics.
 - Social networks shaping credit-access
 - Recidivism metrics that shape parole decisions in a racially disparate manner (Angwin et al, 2016),
 - Zip code categorization resulting in less access to Amazon Prime for African-Americans in US cities (Ingold and Soper, 2016).
- What gets measured and valourized?

Region of Rat-Bitten Babies, Detroit



Map 3.16 Region of rat-bitten babies

If a middle-class suburbanite sees a rat in a slum ghetto he may feel disgust, but a slum mother feels horror. The rats – predacious carnivores – are looking for any food to eat, be it garbage or her infant's toes, fingers and nose. In New York City there are an estimated 4 million rats, almost one to every person. But in slum ghettos like Harlem there are 20 rats per baby, hunting, hunting.

Source: Detroit Department of Health, 1967, 1969 and 1970.

Bunge, 1971

Thanks

zook@uky.edu
@mattzook