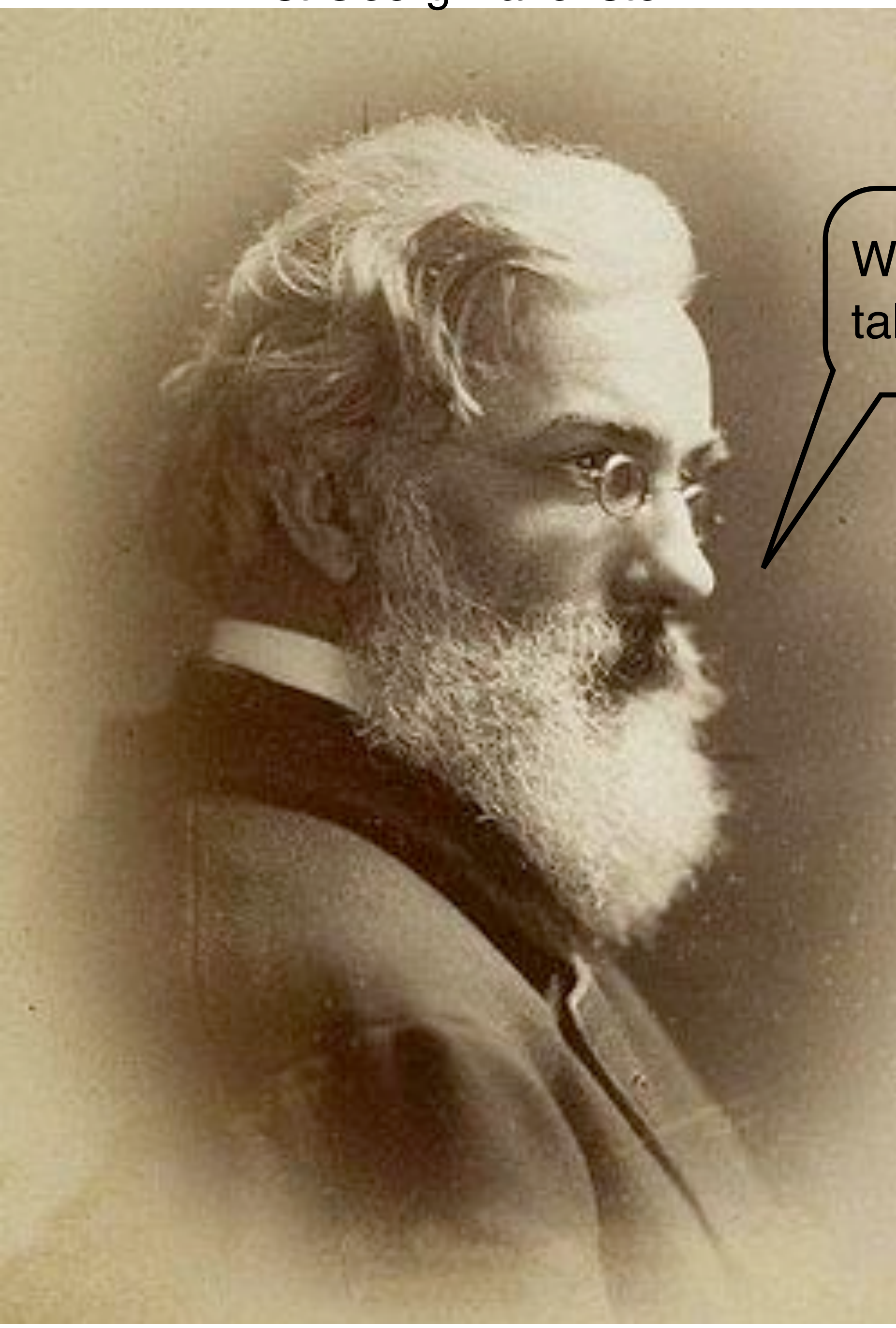
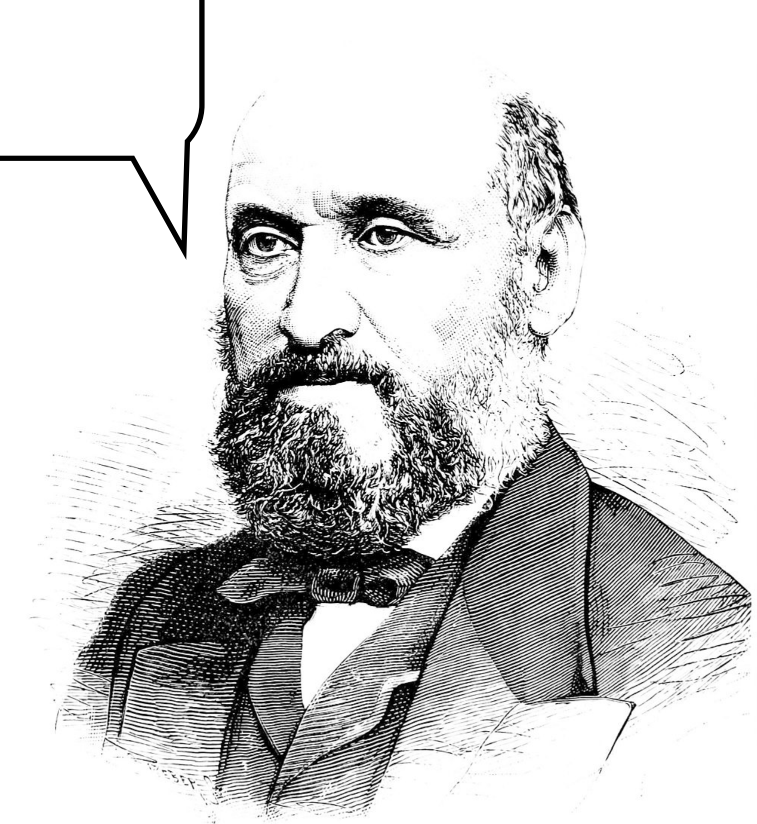
Social and Technological Network Analysis: Spatial Networks, Mobility and Applications

Anastasios Noulas
Data Science Institute
School of Computing & Communications
Lancaster University

History of modern human mobility studies

Ernst Georg Ravenstein





William Farr ... or Dark Vader

the main man ...

Lancaster University

The laws of human migration

Lancaster University

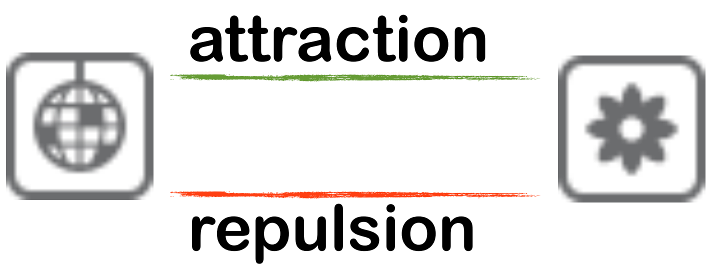
The following was a standard list after Ravenstein's (1834-1913) proposal in the 1880s. The theories are as follows:

1. *every migration flow generates a return or countermigration.*
2. *the majority of migrants move a short distance.*
3. *migrants who move longer distances tend to choose big-city destinations.*
4. *urban residents are often less migratory than inhabitants of rural areas.*
5. *families are less likely to make international moves than young adults.*
6. *most migrants are adults.*
7. *large towns grow by migration rather than natural increase.*

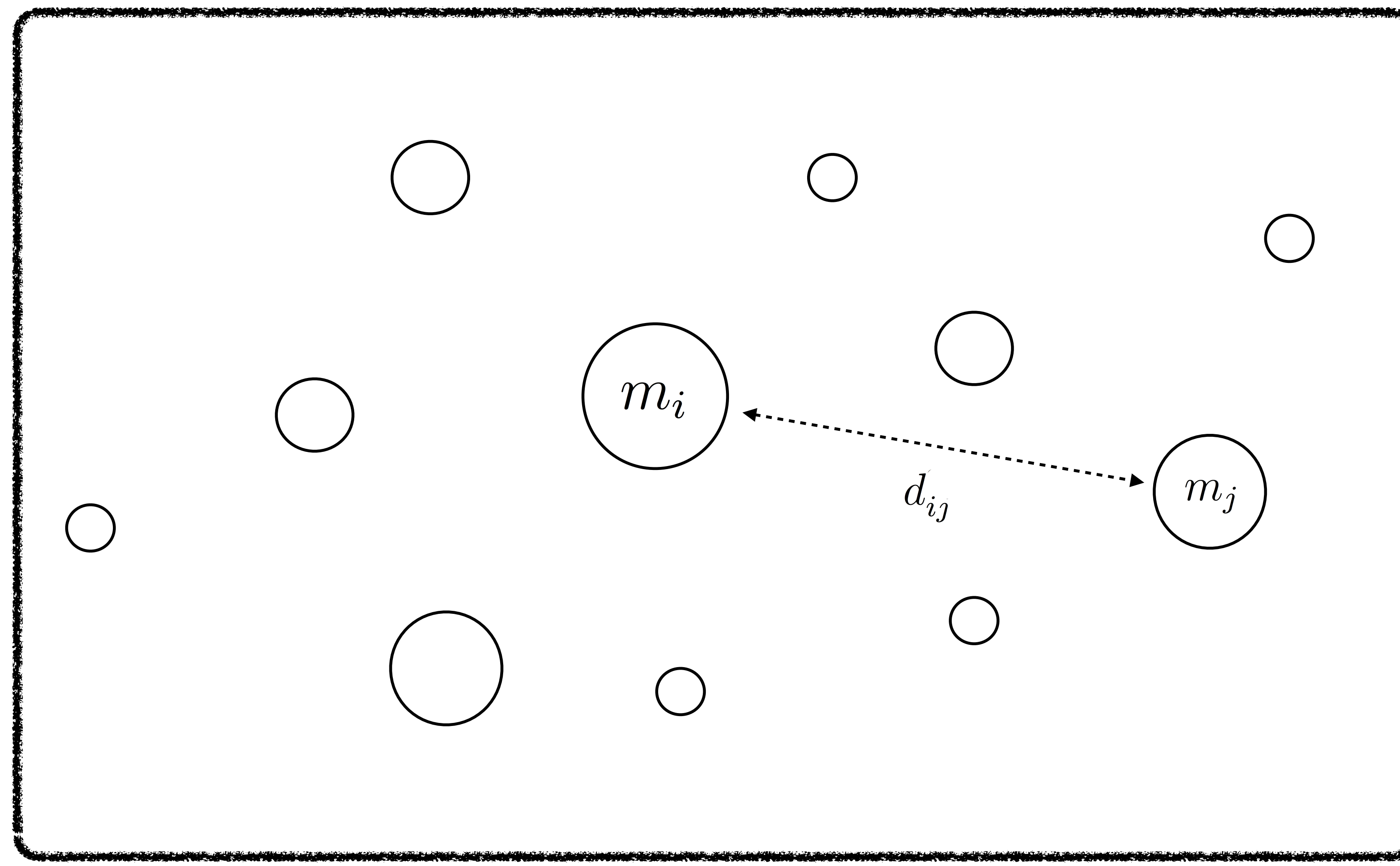
Ravenstein exploited census data from the United Kingdom to support empirically his findings ...

E. G. Ravenstein. The laws of migration. Journal of the Royal Statistical Society, 1885.

Gravity Models



Inspired by Newtonian physics, gravity models suggest that two places attraction is proportional to their mass and inversely proportional to their geographic distance.



$$F_{ij} = \gamma \frac{m_i \; m_j}{d_{ij}^2}$$

Urban Transport Modeling

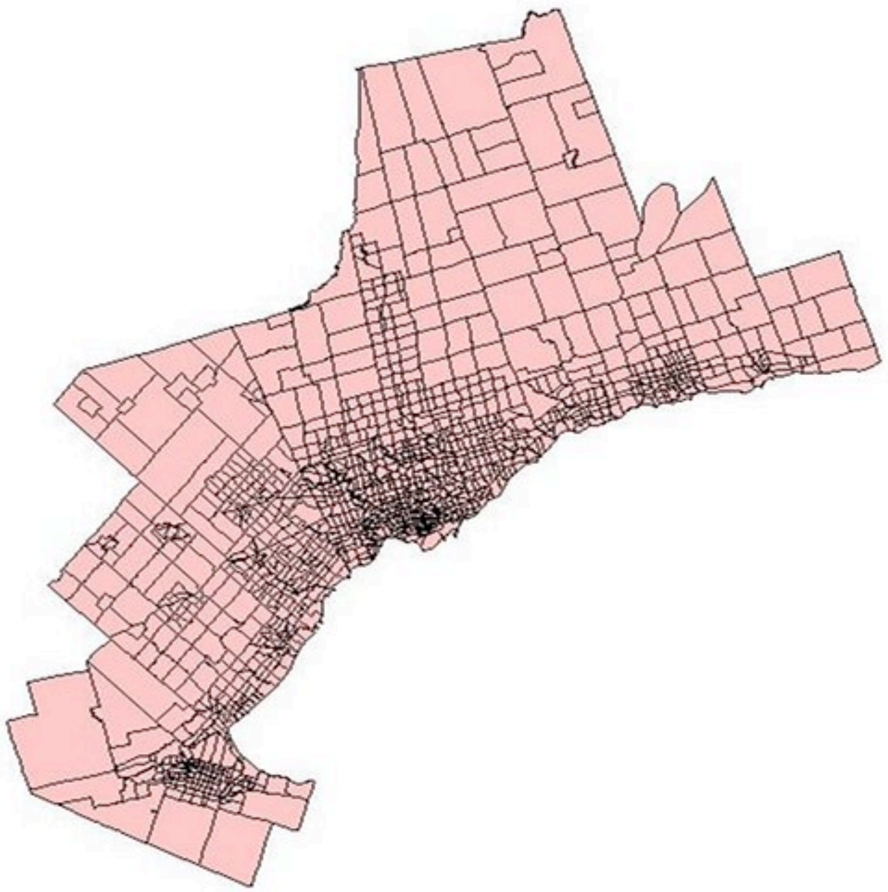


Table: Illustrative trip table

Origin \ Destination	1	2	3	Z
1	T_{11}	T_{12}	T_{13}	T_{1Z}
2	T_{21}			
3	T_{31}			
Z	T_{Z1}			T_{ZZ}

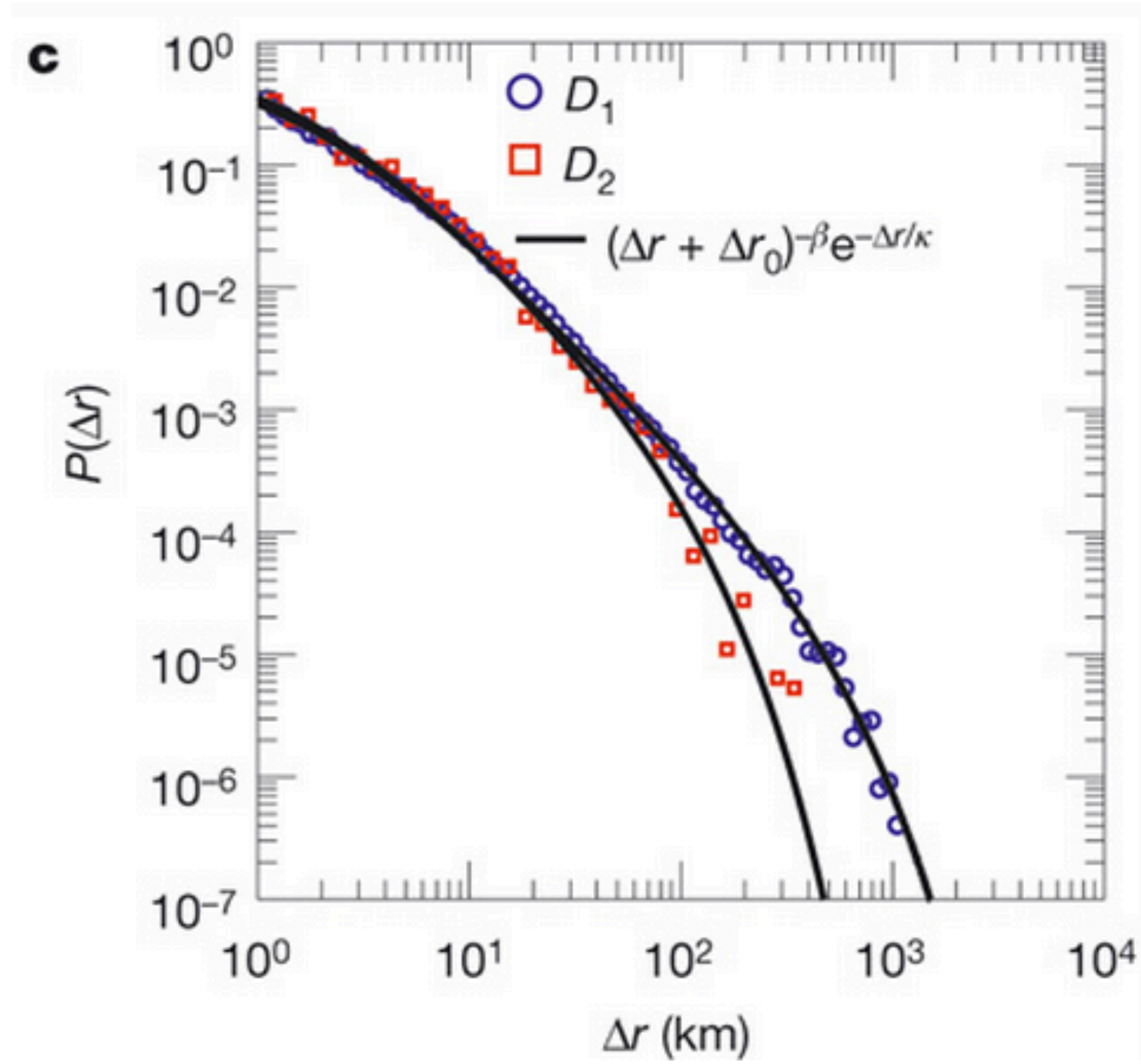
$$T_{ij} = k\frac{O_i D_j}{d_{ij}^2}$$

Data in urban transport modeling has been based primarily on surveys...

Ortuzar, J. de, and Luis G. Willumsen. *Modelling transport*. 2011.

Cellular Datasets

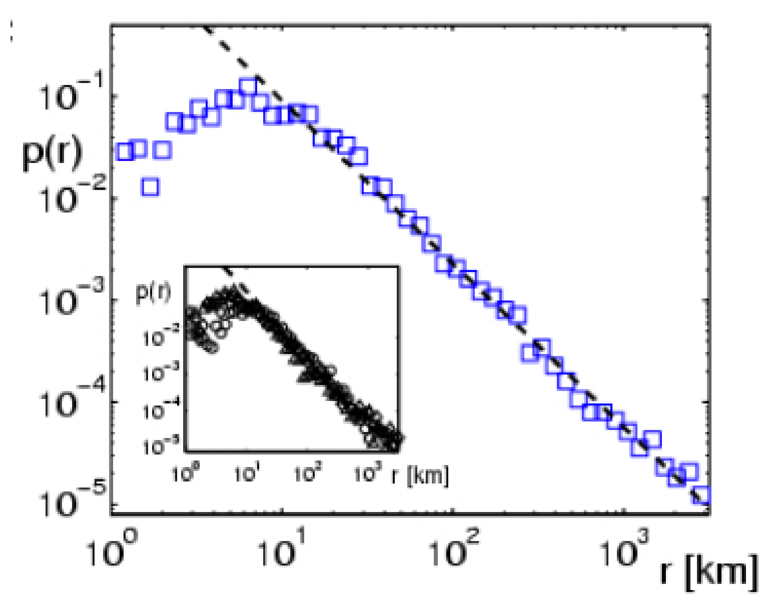
One of the first large scale studies of human movement with modern mobile datasets...

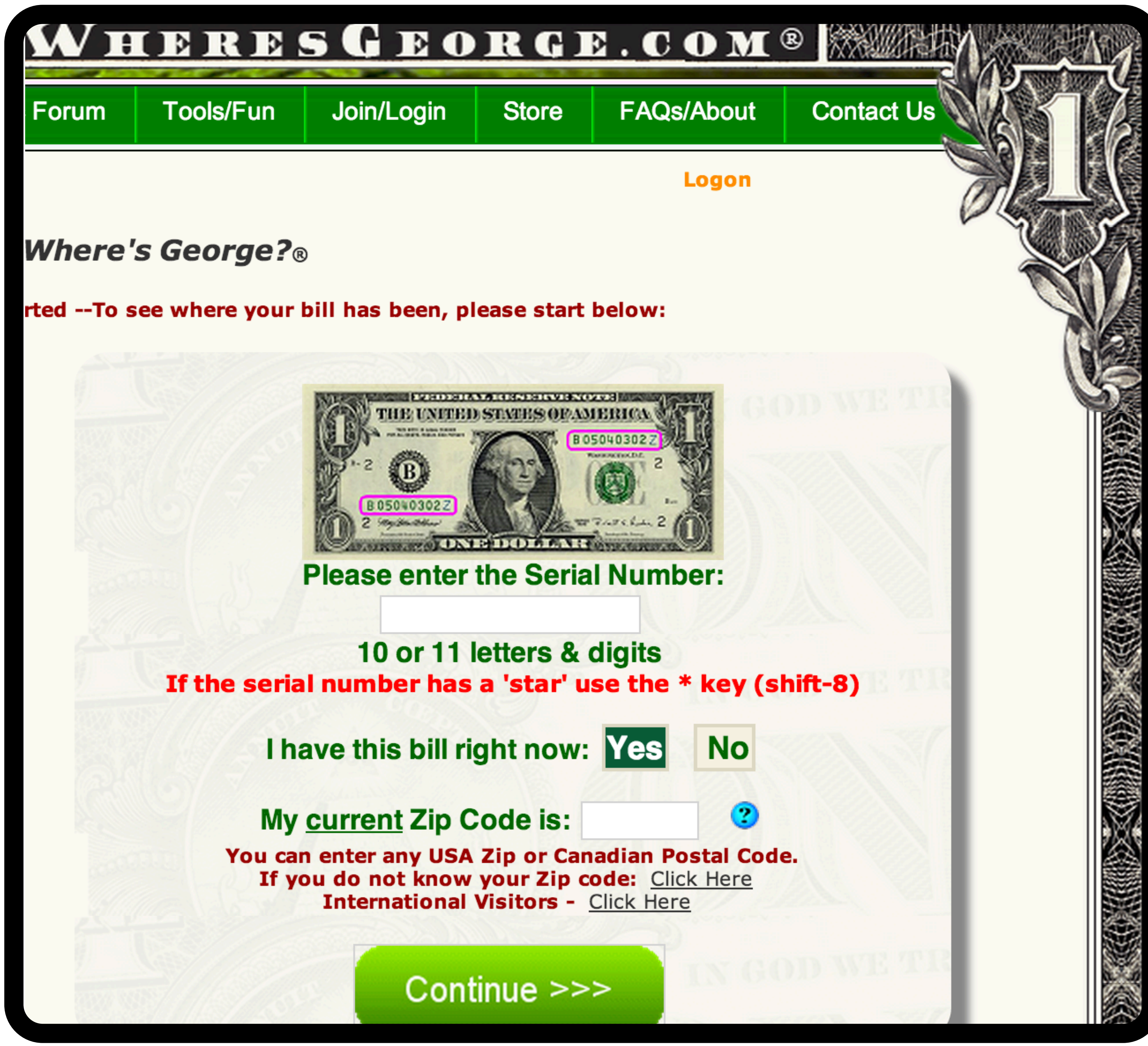


Gonzalez, Marta C., Cesar A. Hidalgo, and Albert-Laszlo Barabasi. "Understanding individual human mobility patterns." *Nature* 453.7196 (2008): 779-782.

Where's George ?

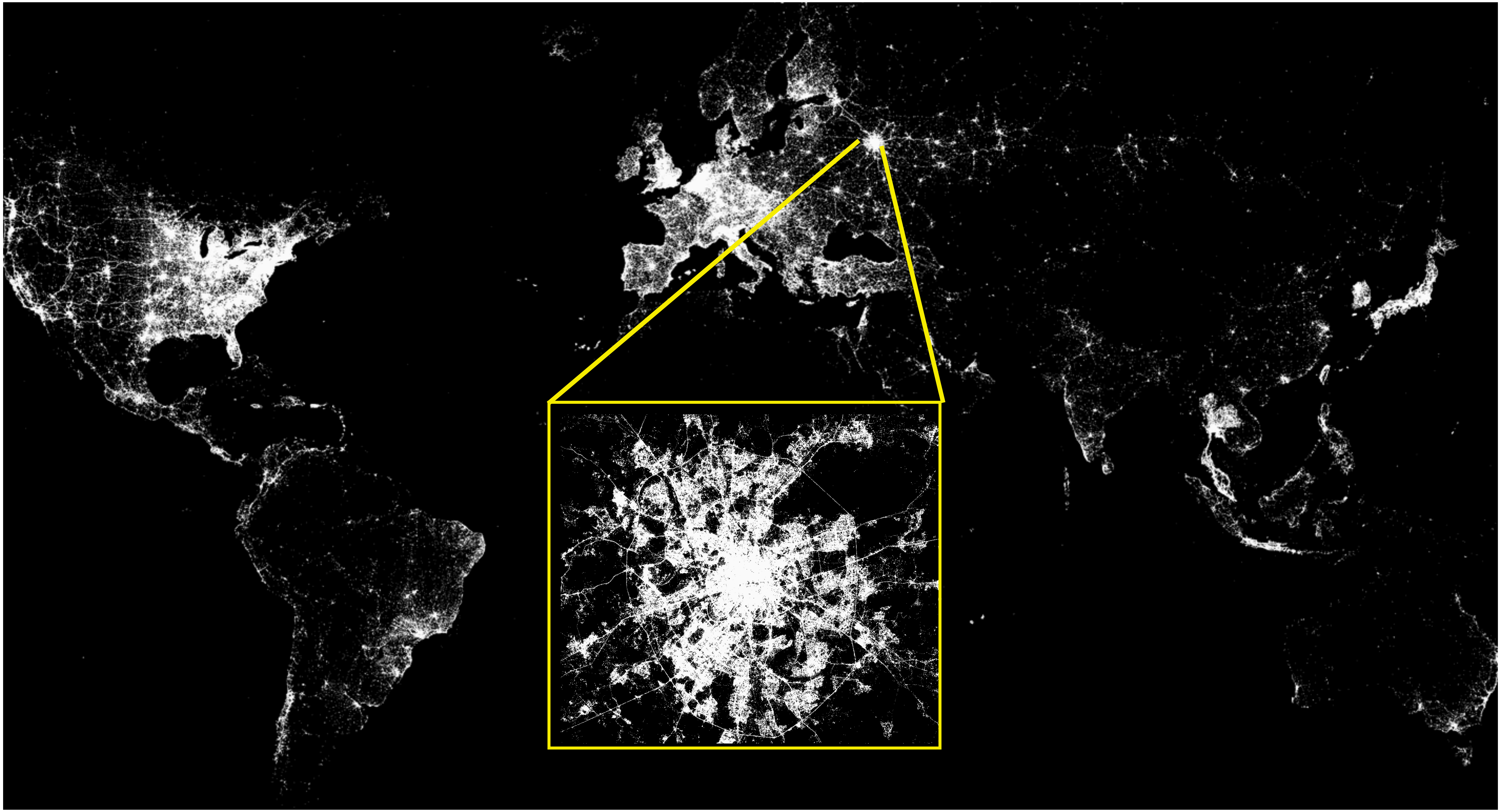
One of the most creative ways to study human movement that has used the displacement of dollar bills as a proxy to human mobility...





Brockmann, Dirk, Lars Hufnagel, and Theo Geisel. "The scaling laws of human travel." *Nature* 439.7075 (2006): 462-465.

Mobile users are the stars



https://foursquare.com/infographics/500million

Dataset Statistics

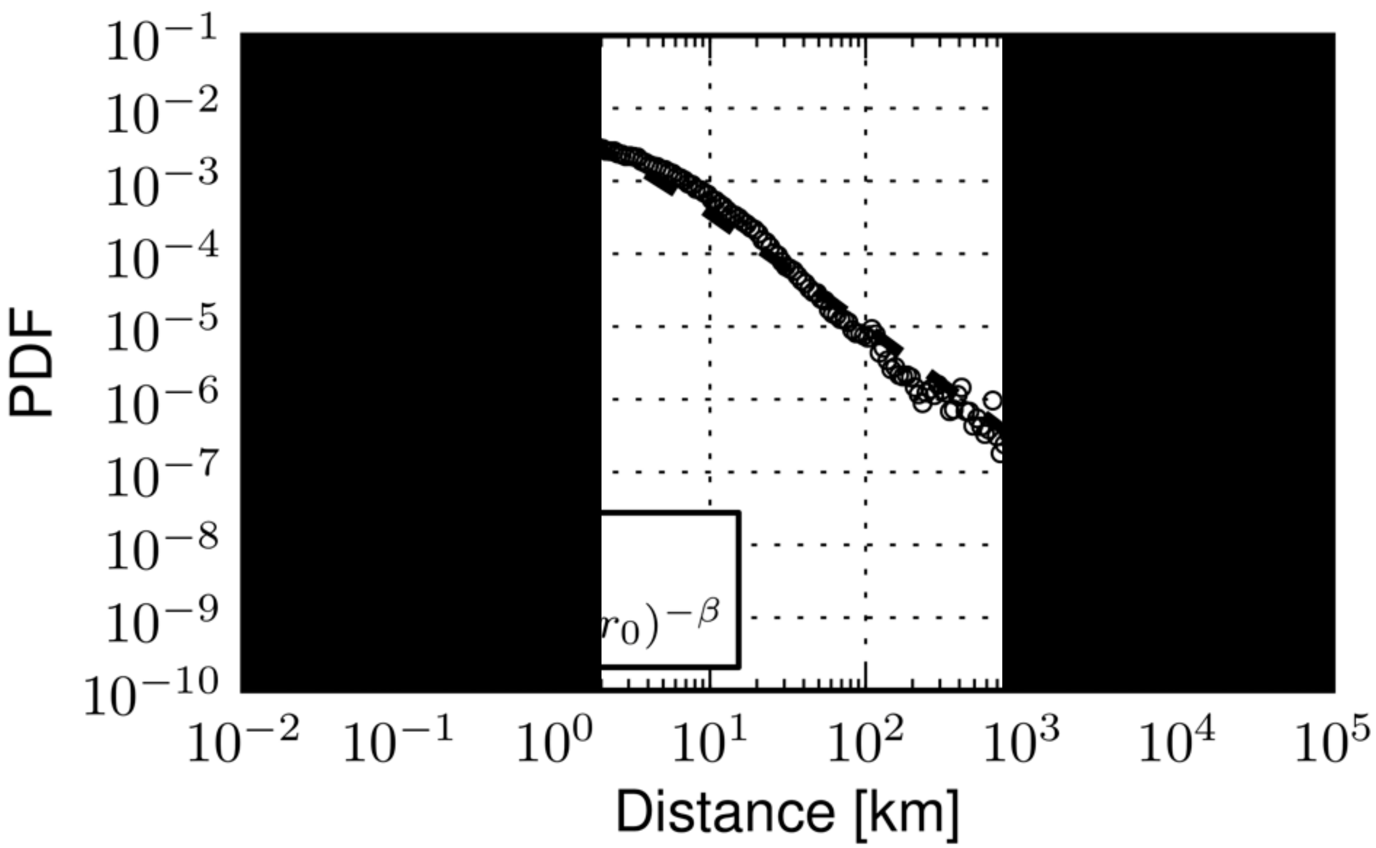
925,030 users around the globe over a period of 6 months in 2010.

34 Cities that span 4 continents and 11 countries.

For the first time human mobility is analyzed in light of 5 million recorded settlements (places).

Power-law tales ...

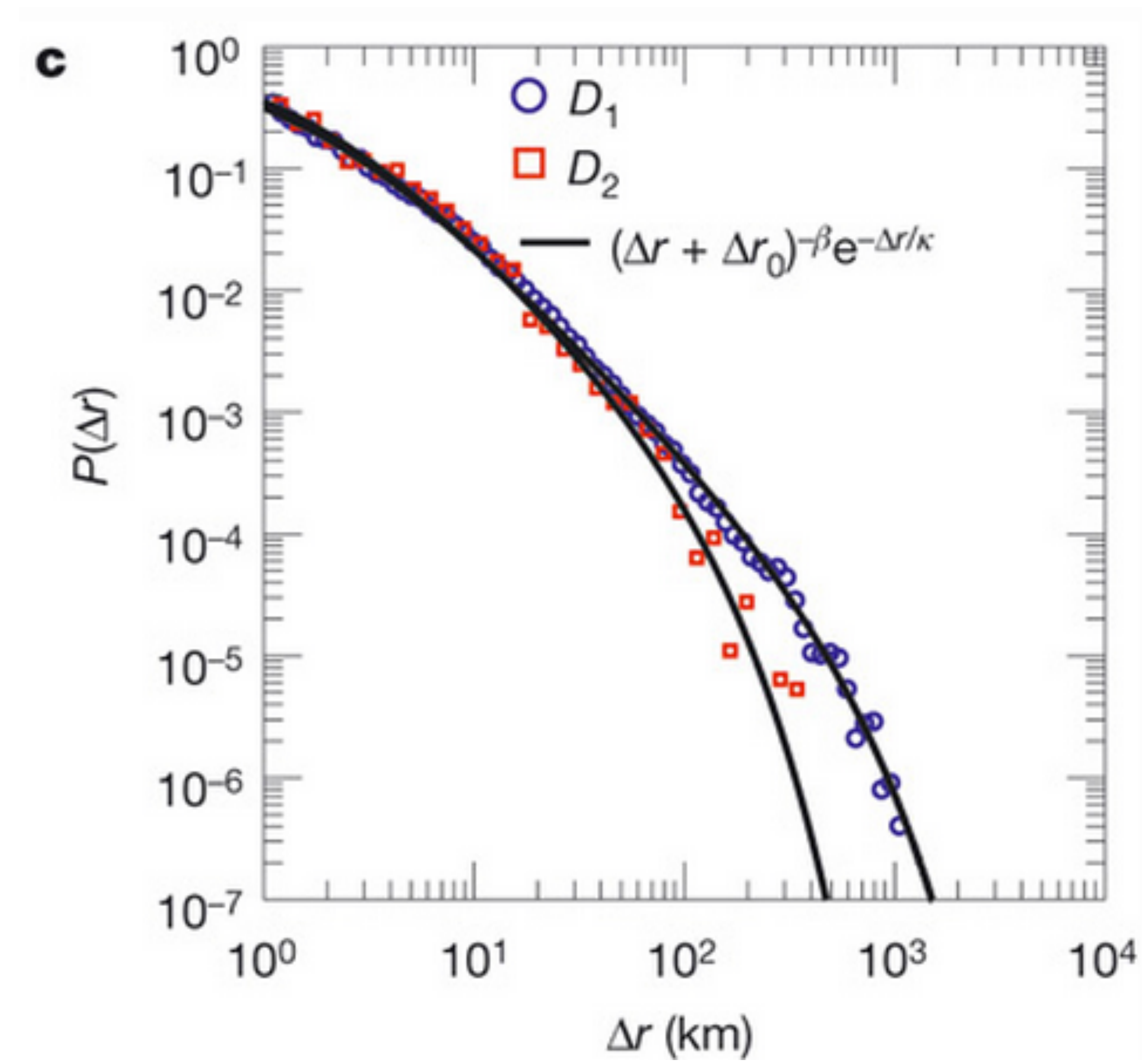
Mobile Social Network Data



$(\Delta r + \Delta r0)^{-\beta}$

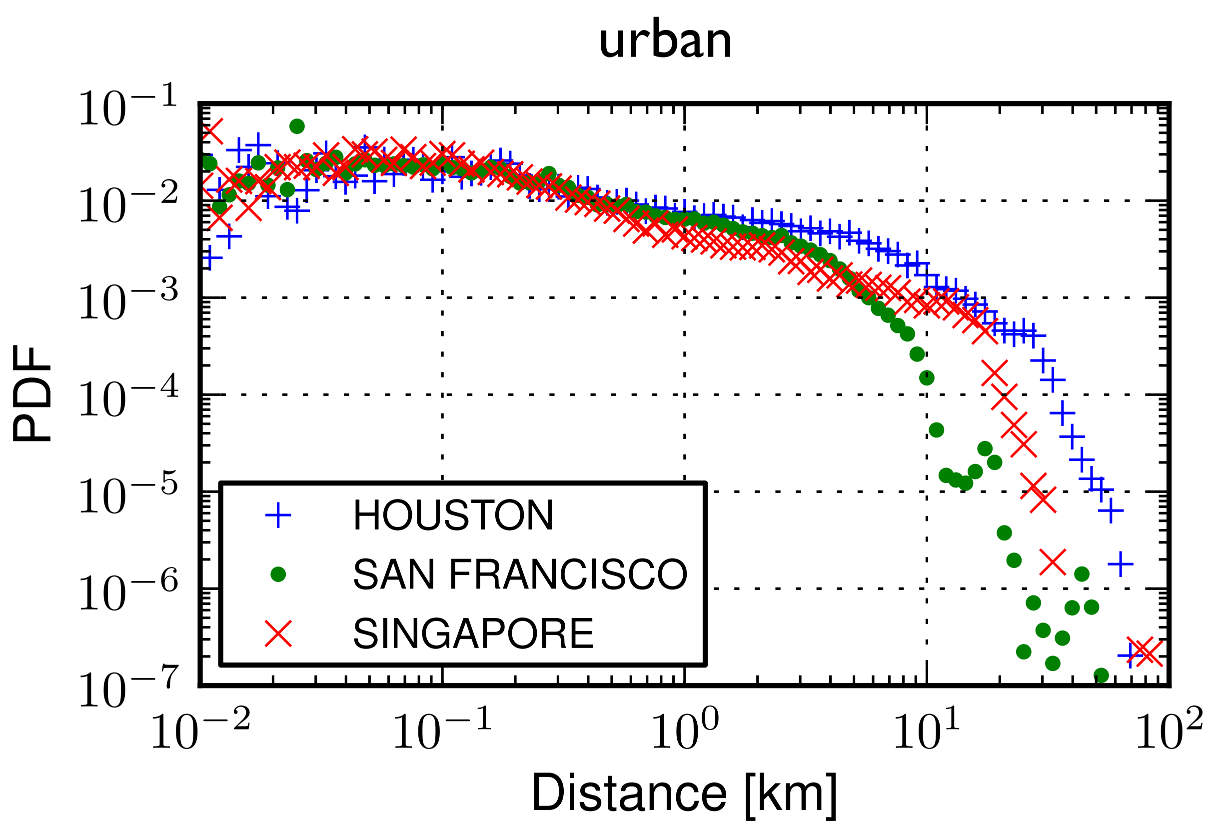
exponent $\beta = 1.50$

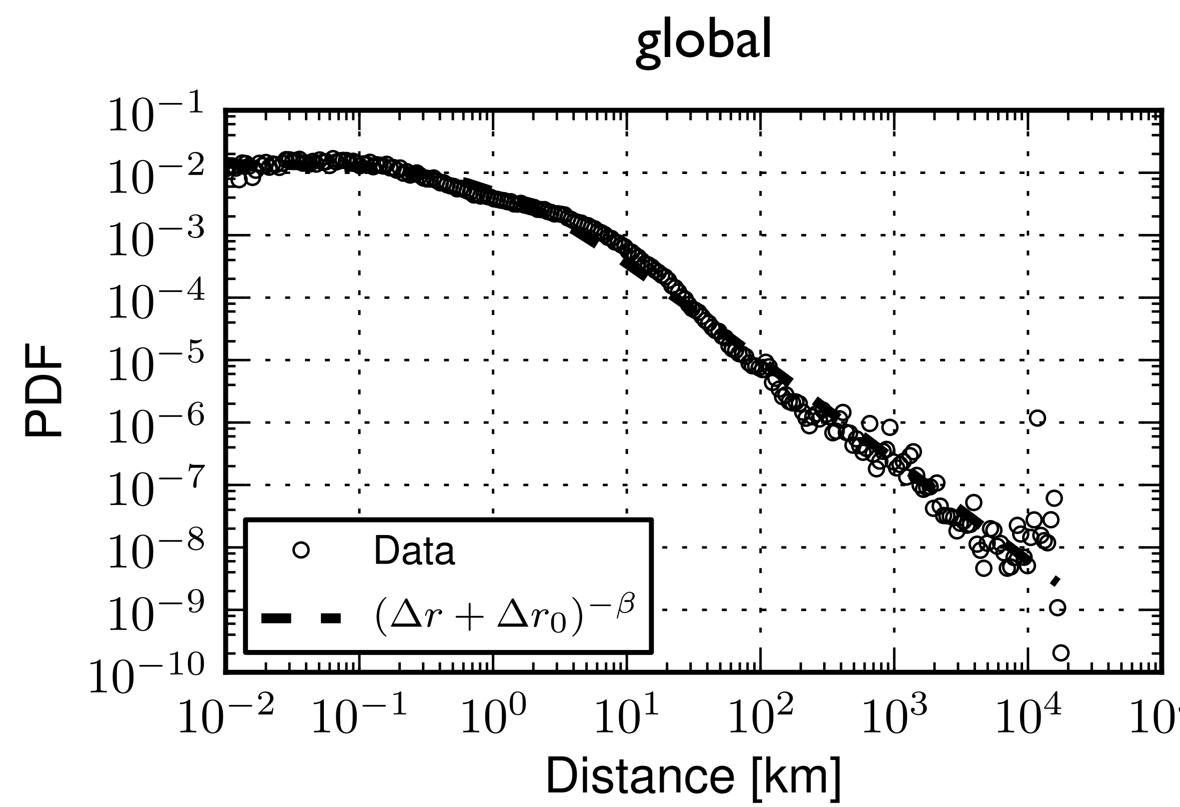
Nature **453**, 779-782(5 June 2008)



exponent $\beta = 1.75$

Urban vs Global mobility

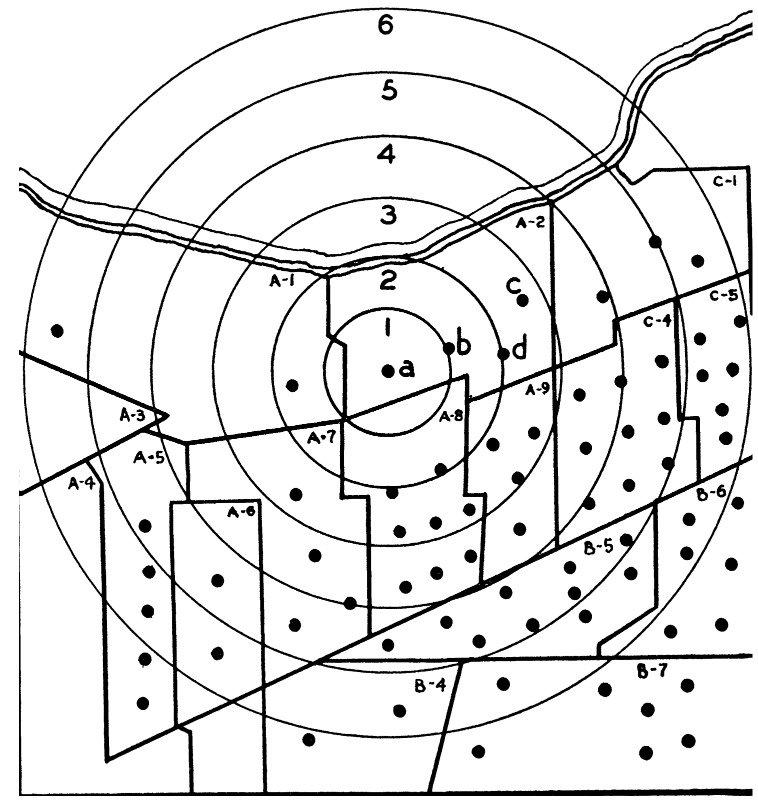




Power law kicks in at 18.42km!!!

Samuel A. Stouffer

Stouffer's **law of intervening opportunities** states, *"The number of persons going a given distance is directly proportional to the number of opportunities at that distance and inversely proportional to the number of intervening opportunities."* *

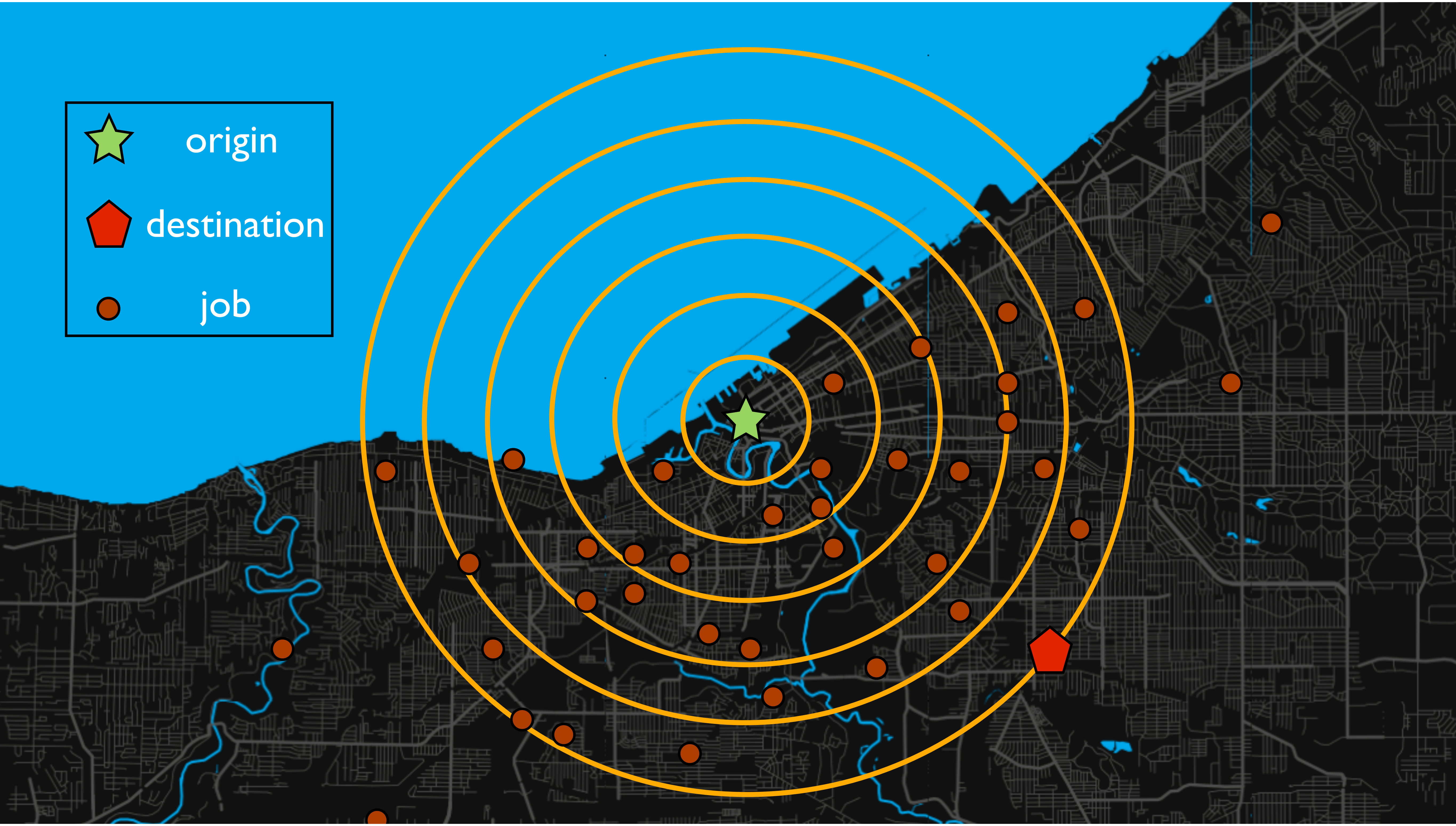


- Empirically proven using data for migrating families in the city of Cleveland.

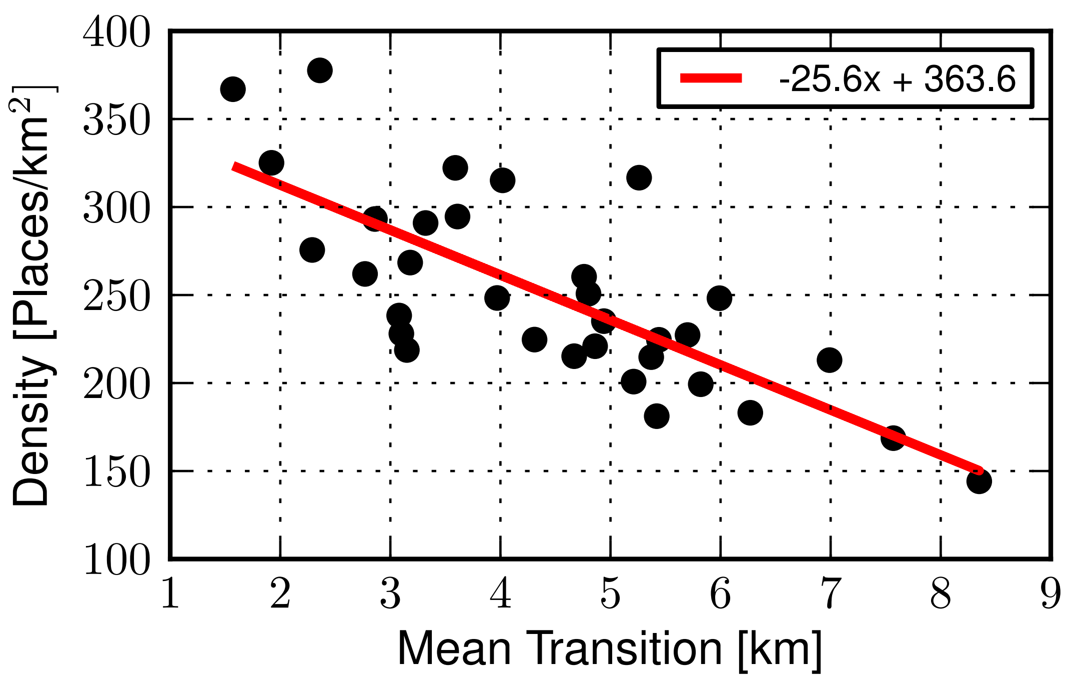
- We investigate the plausibility of the theory for urban movements in Foursquare.

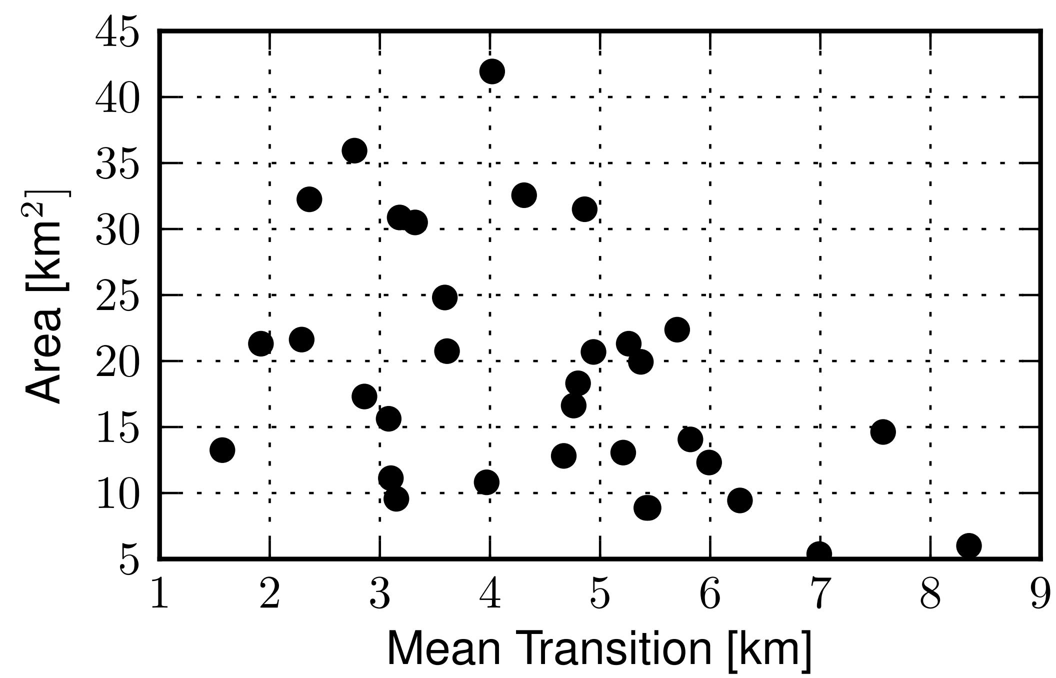
* S. Stouffer (1940) Intervening opportunities: A theory relating mobility and distance, American Sociological Review 5, 845-867

Samuel A. Stouffer was a big data pioneer!



The importance of density

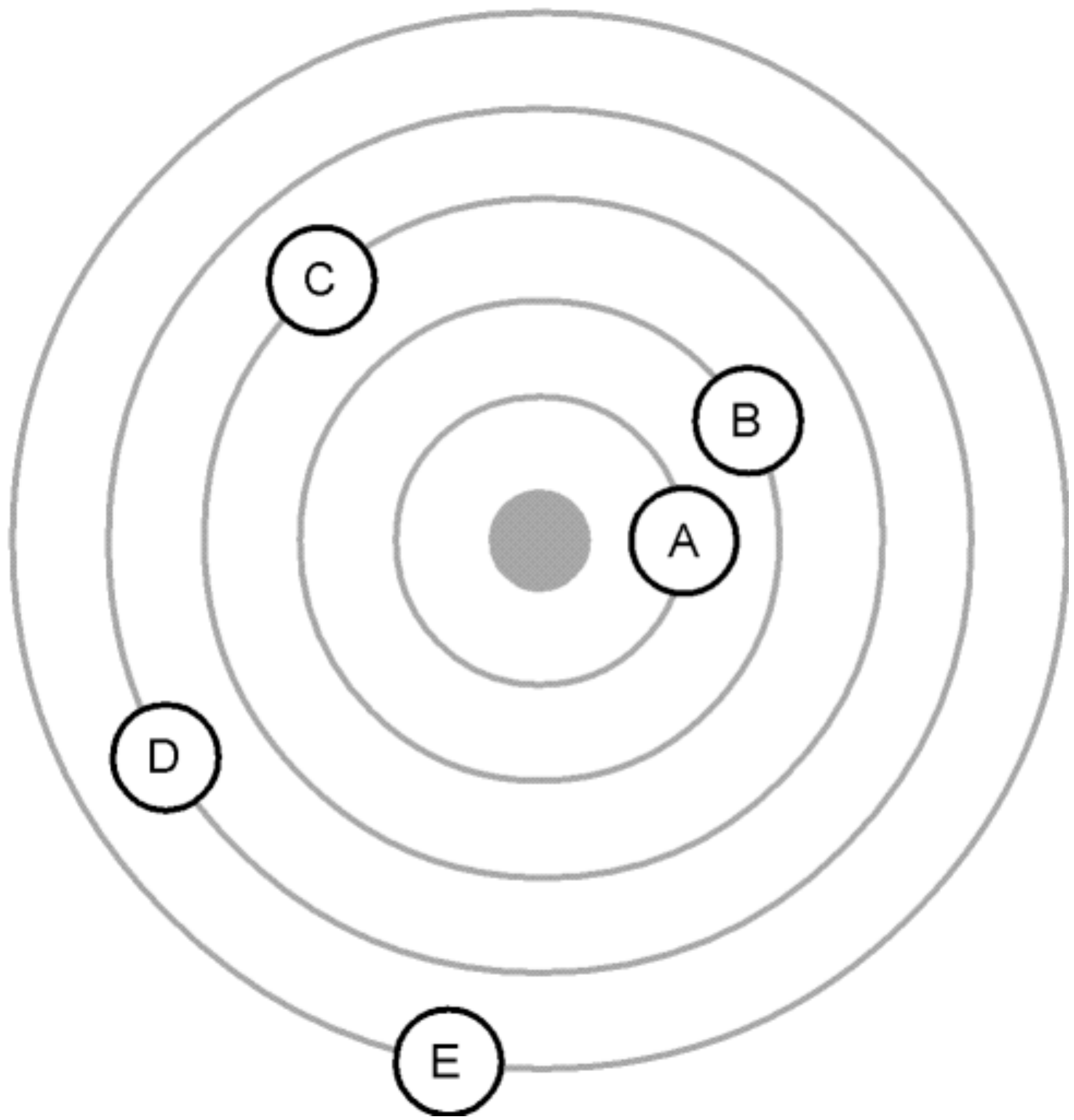




- Stouffer's Theory of Intervening Opportunities motivated us to inspect the impact of places(=opportunities) in human mobility.

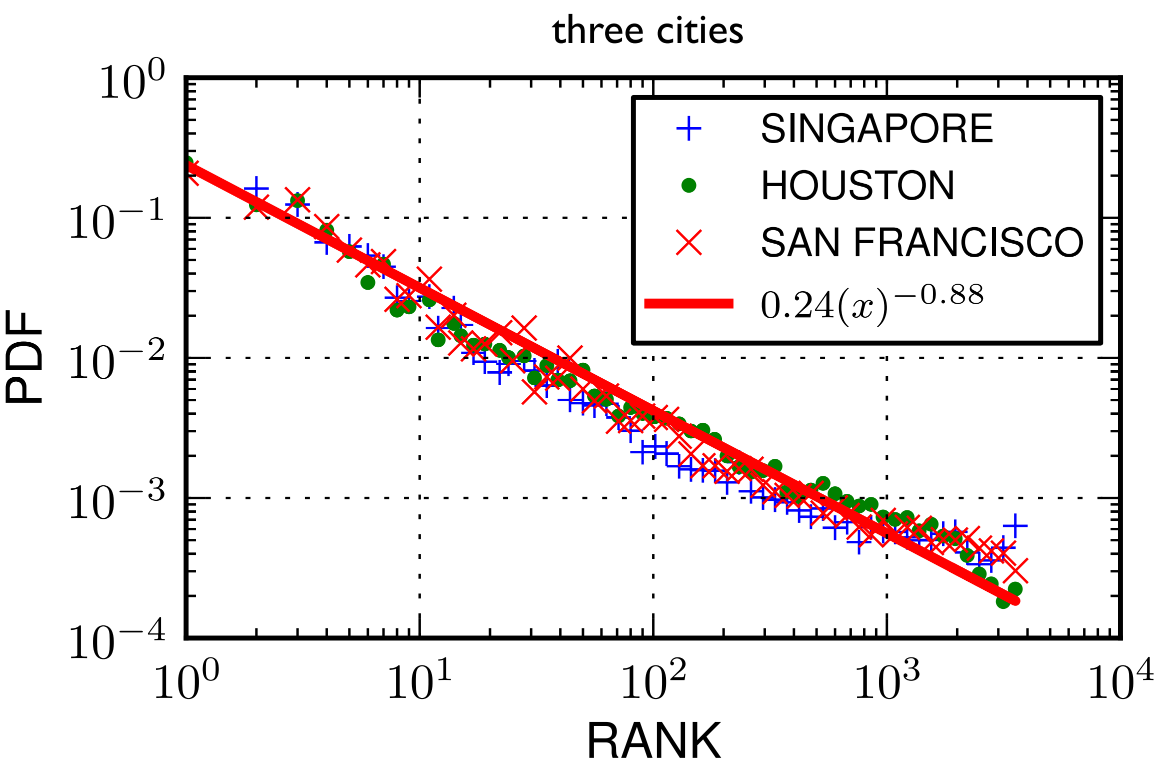
- Place density by far more important than the city area size with respect to mean length of human movements (R^2 = 0.59 and 0.19 respectively).

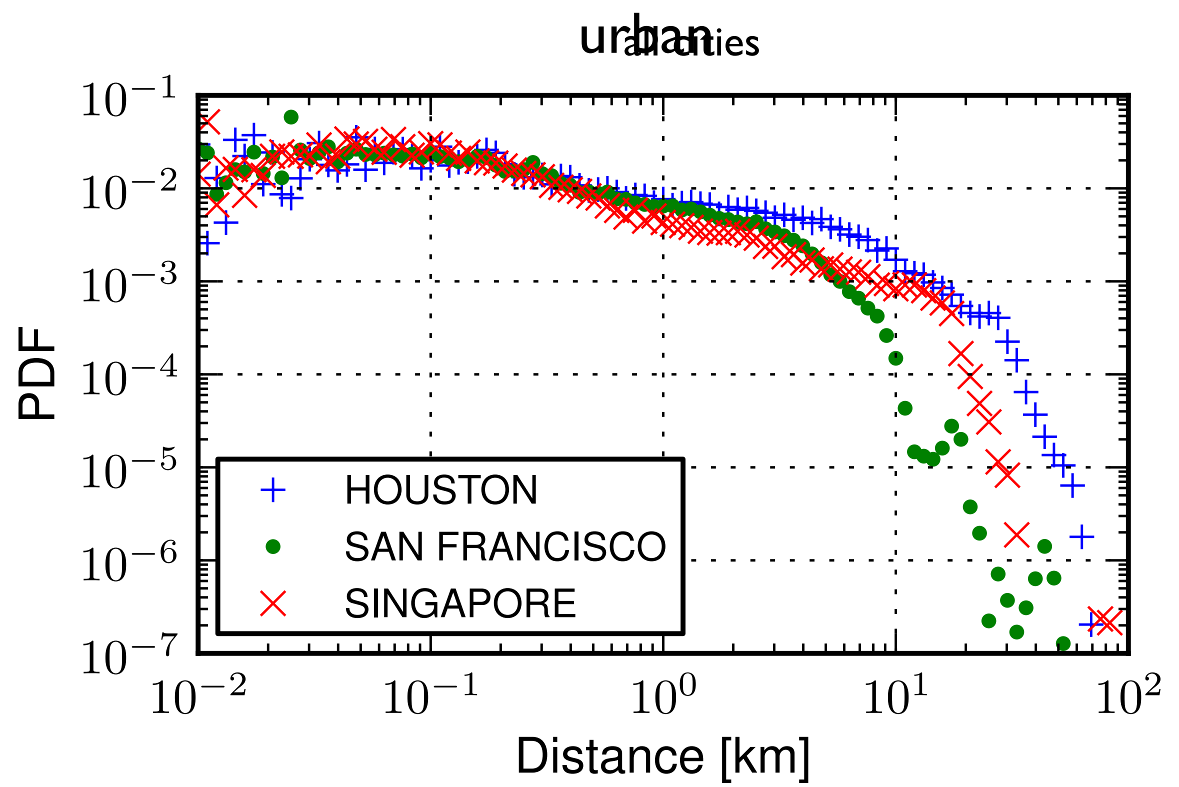
Defining Rank-Distance



$$rank_u(v) = |\{w : d(u,w) < d(u,v)\}|$$

Rank universality



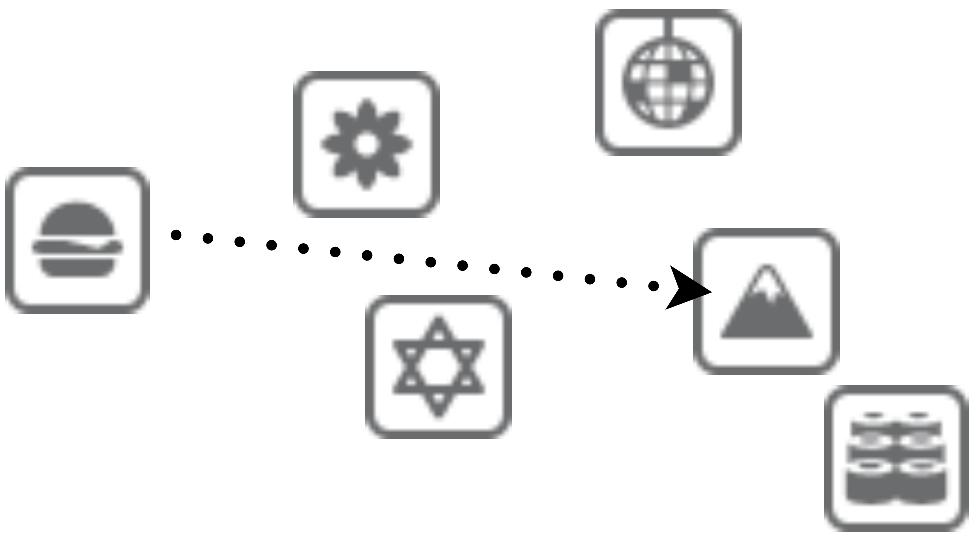


The rank of all cities collapse to a single line.

We have measured a power law exponent $\alpha = 0.84 \pm 0.07$

A new model for urban mobility

soil...



and mind!

$$Pr[u \to v] \propto \frac{1}{rank_u(v)^a}$$

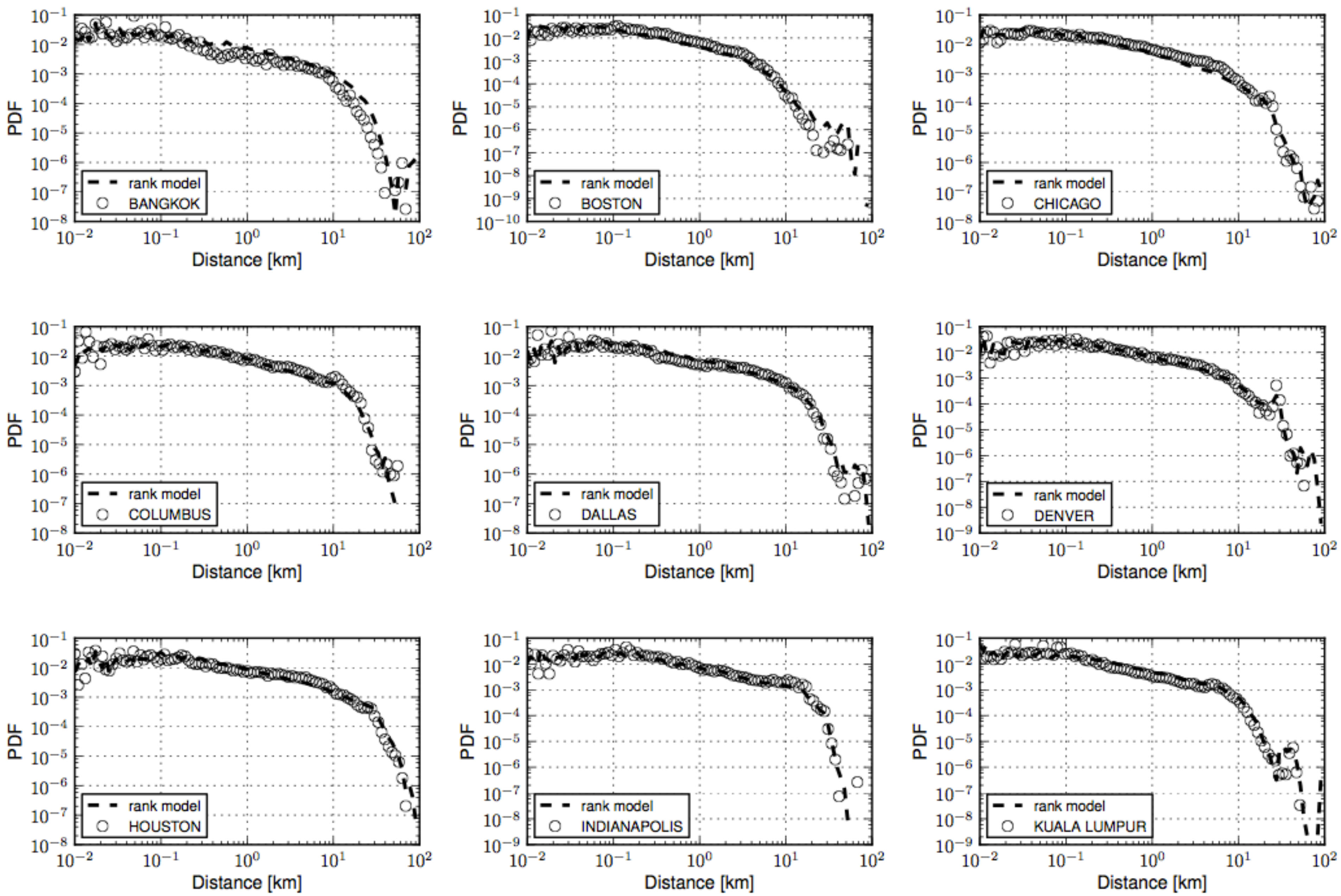
Set … and go!

for all cities we have used the average value $\alpha = 0.84$ for the rank exponent.

all places in the city used as potential starting points for our agents.

the rank element is universal, only the set of places differs from city to city.

Simulation Results …



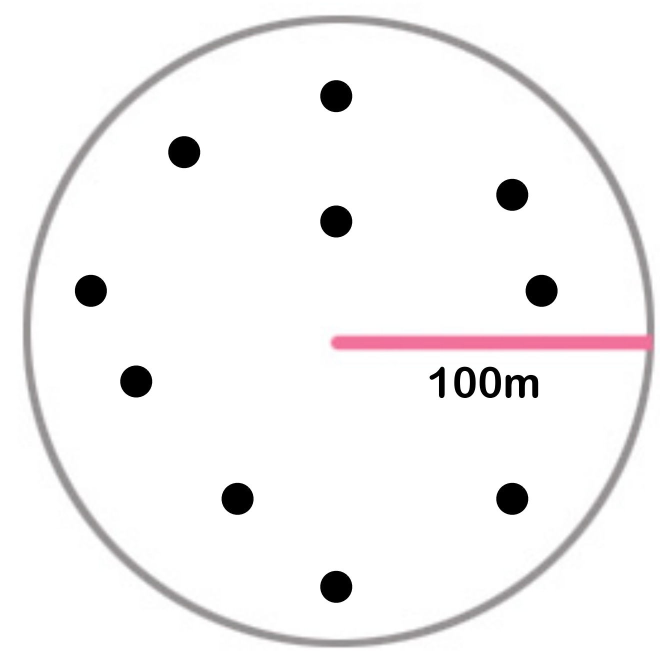
“Zero” Gravity

We have also built a gravity model in the urban context!

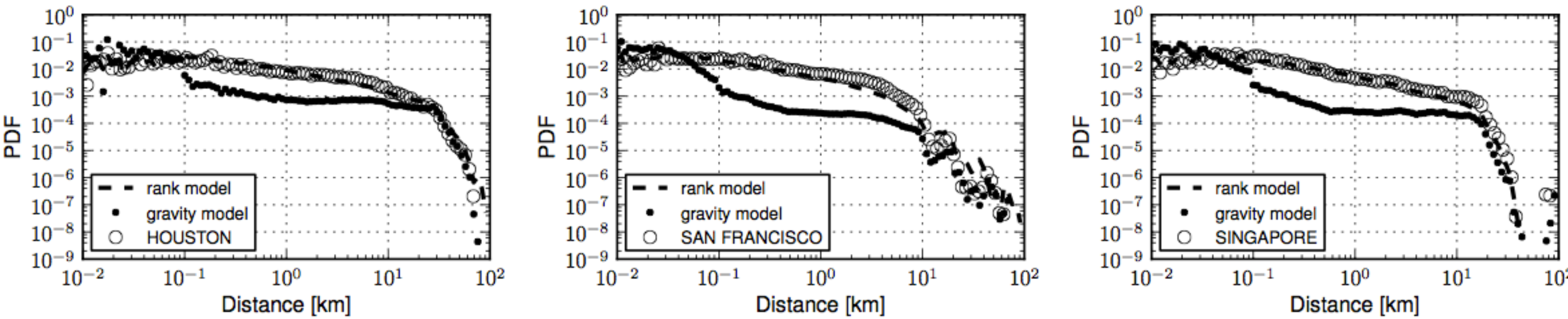
$$P_g[u \to v] \propto \frac{m_u . m_v}{d(u,v)^b}$$

Issue #1: how do we define “mass” in the urban context.

Issue #2: how do we set its parameters?



Rank vs Gravity

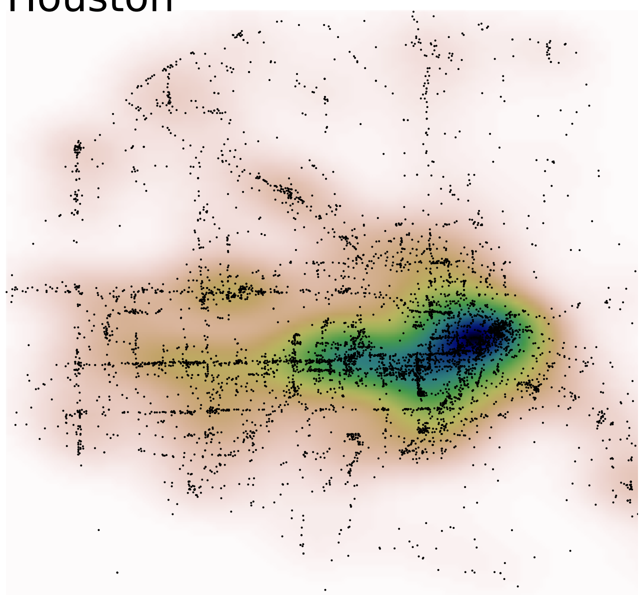


Rank is simpler and achieves better quality fits for all cities.

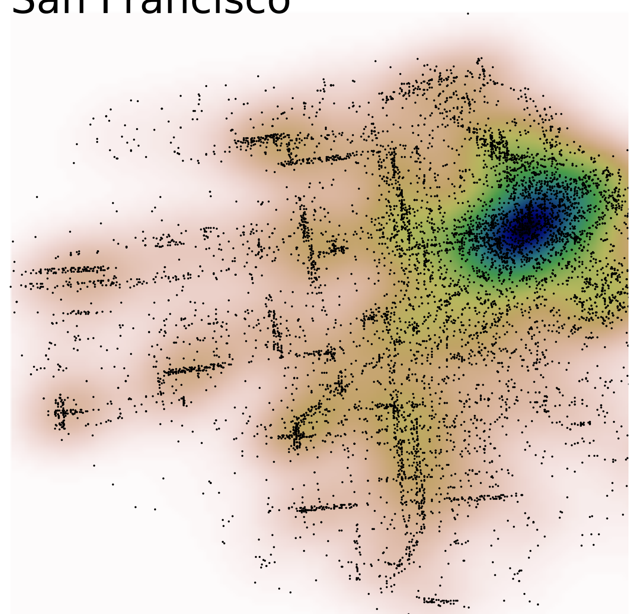
Gravity overestimates short transitions ...

The importance of Geography

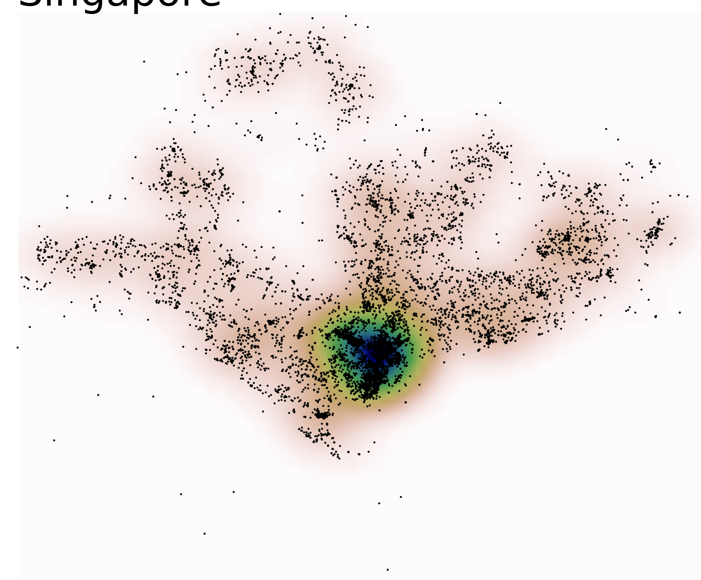
Houston



San Francisco



Singapore



Heterogeneities observed in human mobility is due to geographic variations. Cultural, organisational or other factors do not appear to play a role in urban movements.

The rank model, although simple, can cope with the complex spatial variations in densities observed in urban environments.

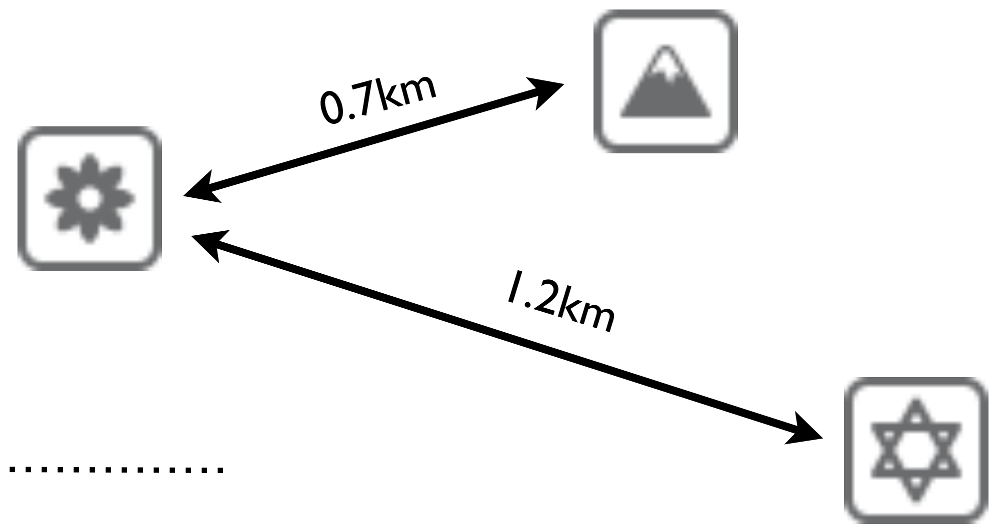
Computer Science at your Service

User Specific features

- historic visits
- friend check-ins
- preferred venue types

Geographic

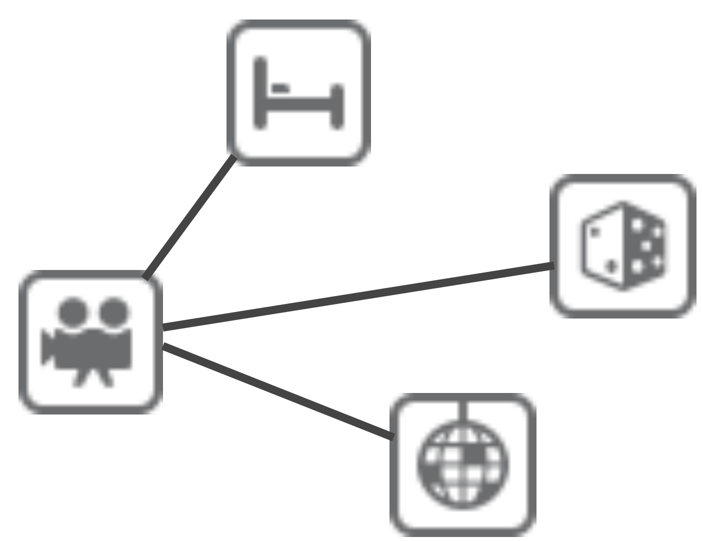
- distance and rank-distance



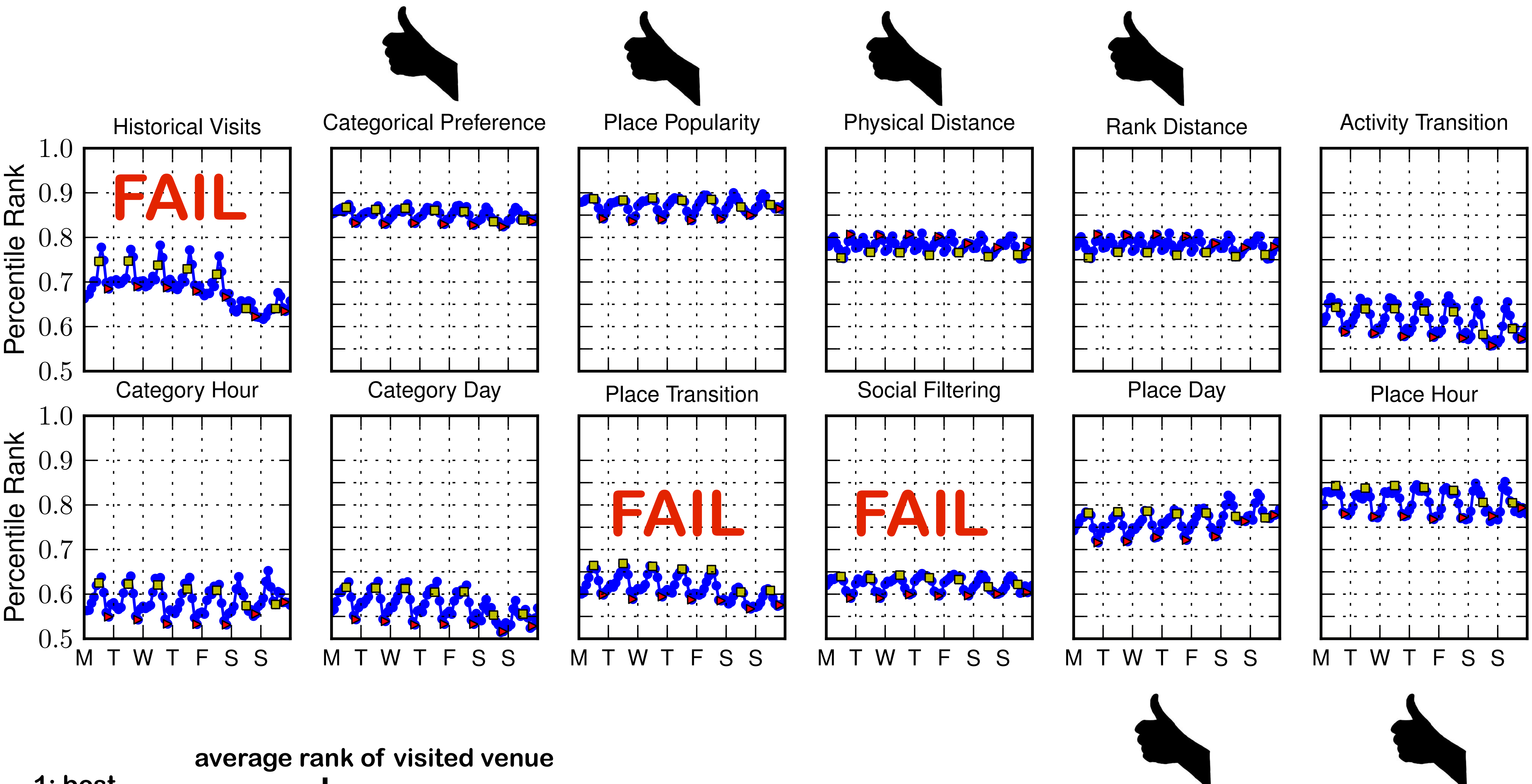
Temporal

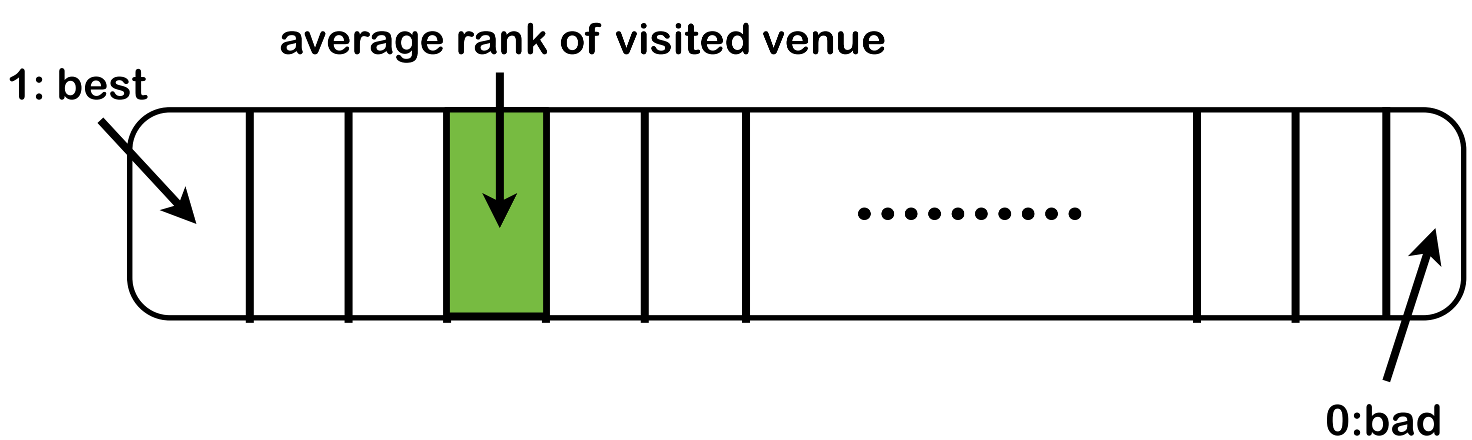
- trending places (hour/day)
- trending place types (eg. cinema at nights)

Place Network



Feature Performance #2





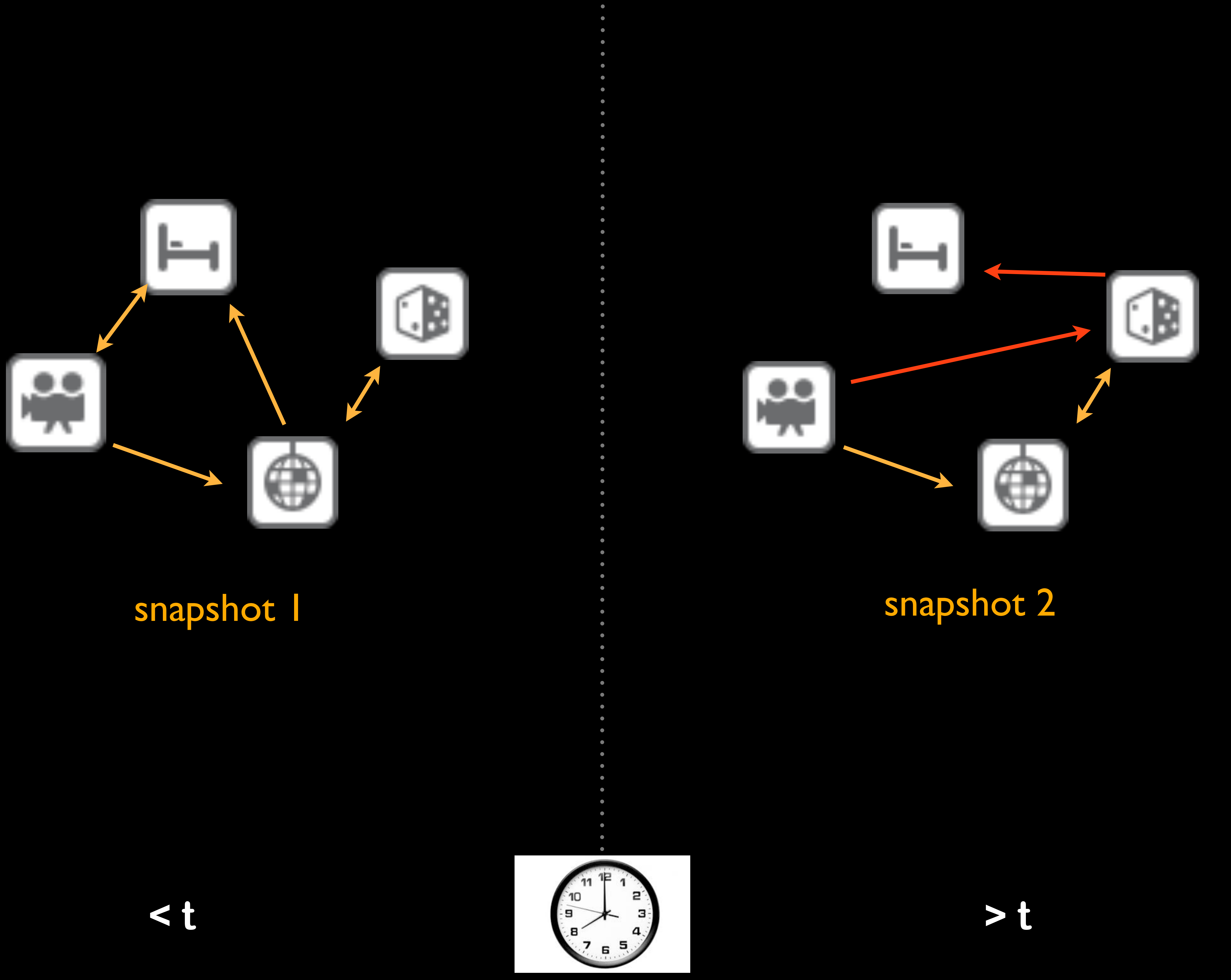
Supervised Training: teaching the good and the bad!

[visited place] +1

[randomly selected NOT-visited place] -1

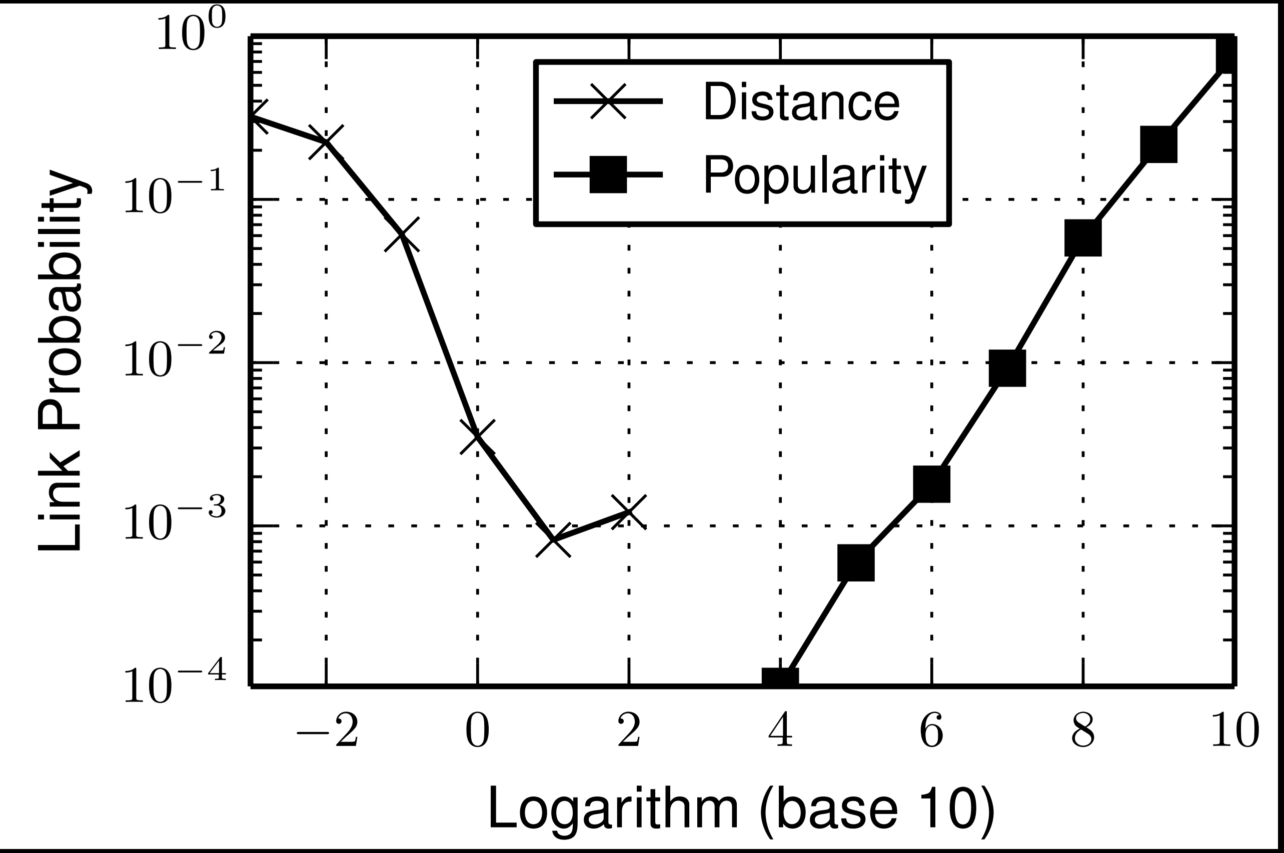
Key Insight: Every time little Amy checks-in she expresses a direct preference at a place and implicitly ignores all the rest!

learning: supervised classifier trains on millions of check-ins generated by populations of users.

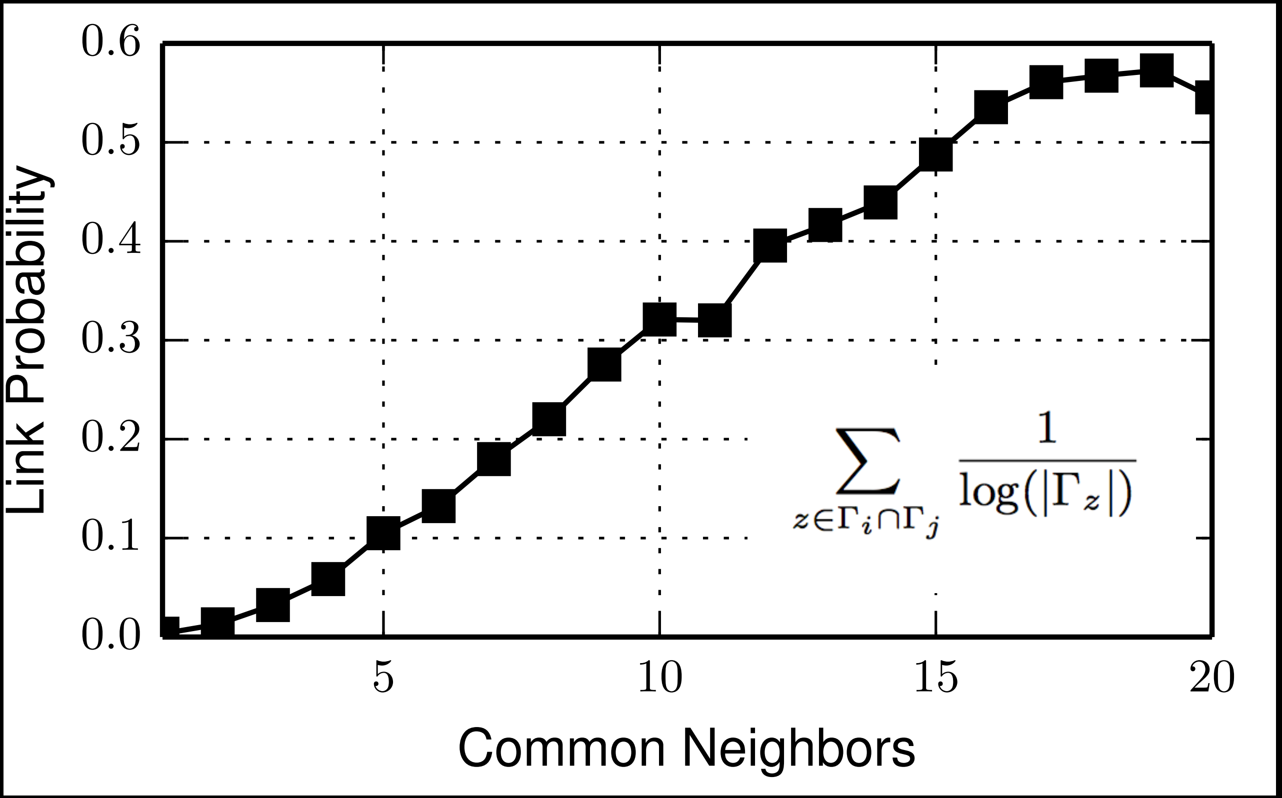


snapshot 1
snapshot 2
< t
> t

human mobility

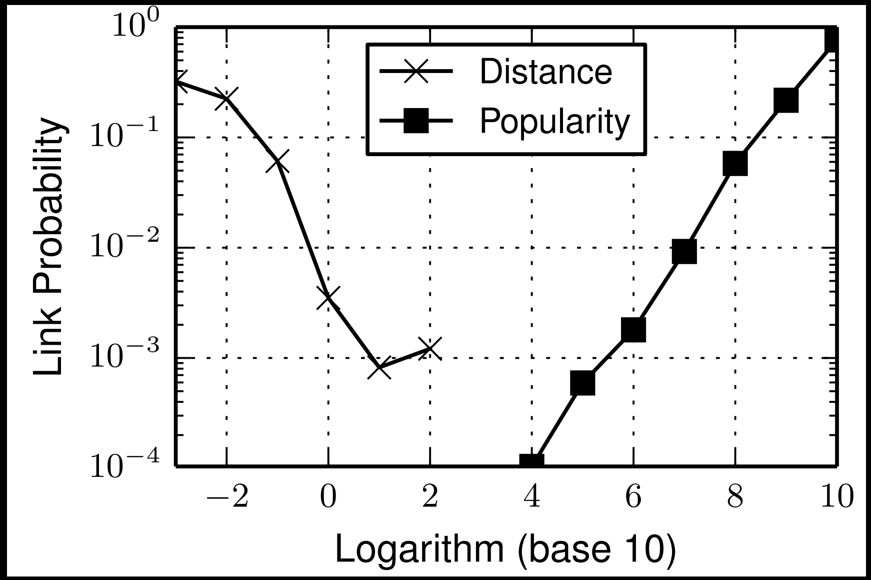


network form



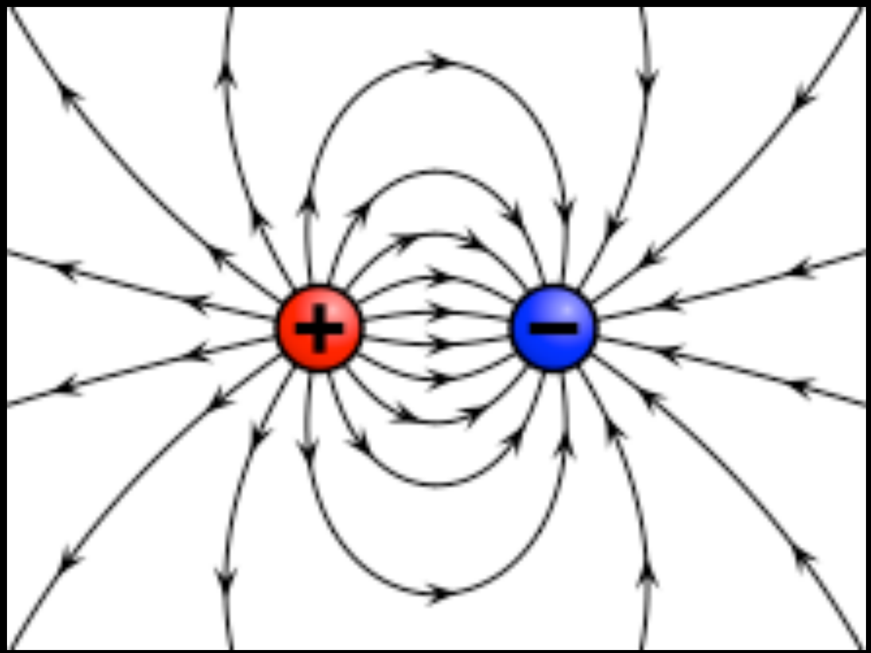
static gravity

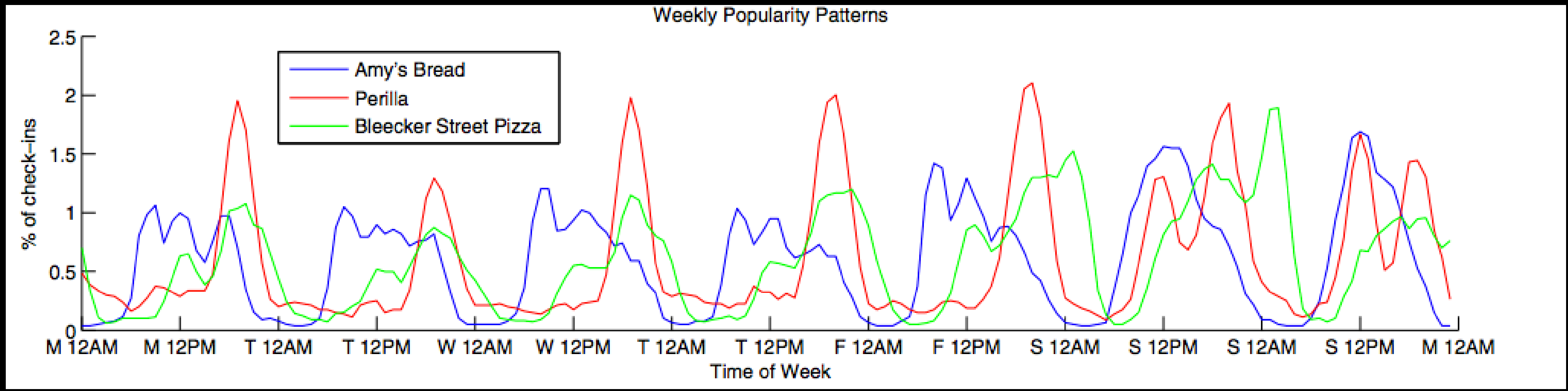
$$\frac{c_i c_j}{d(i,j)^{\beta}}$$



dynamic gravity

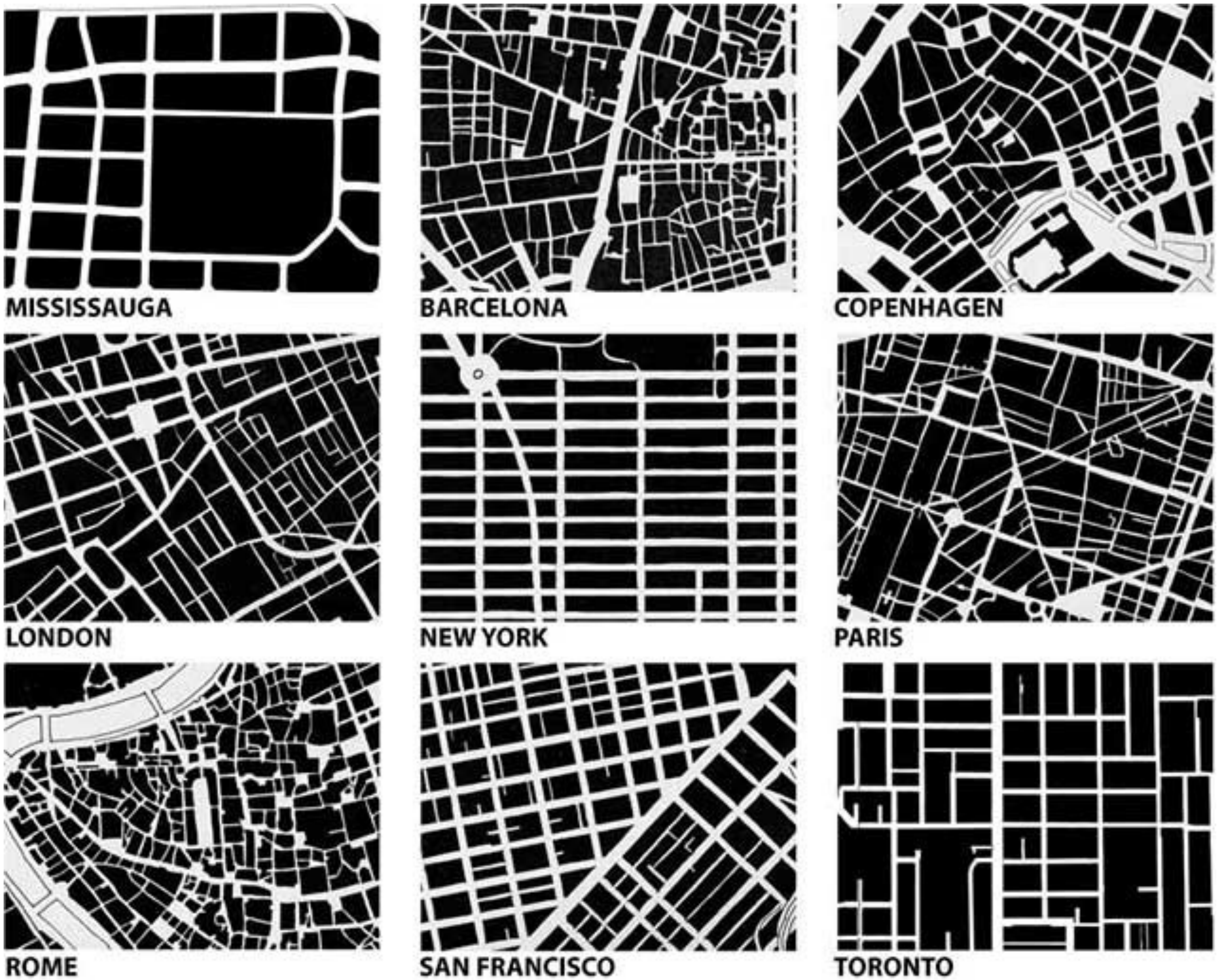
$$\frac{a_{ij} \sum_{\tau=1}^{T} c_i(\tau)^{+} c_j(\tau)^{-}}{d(i,j)^{\beta}}$$





awesome fact :
when T=1 and a_ij = 1
we fall back to the
static gravity model

Urban Morphology



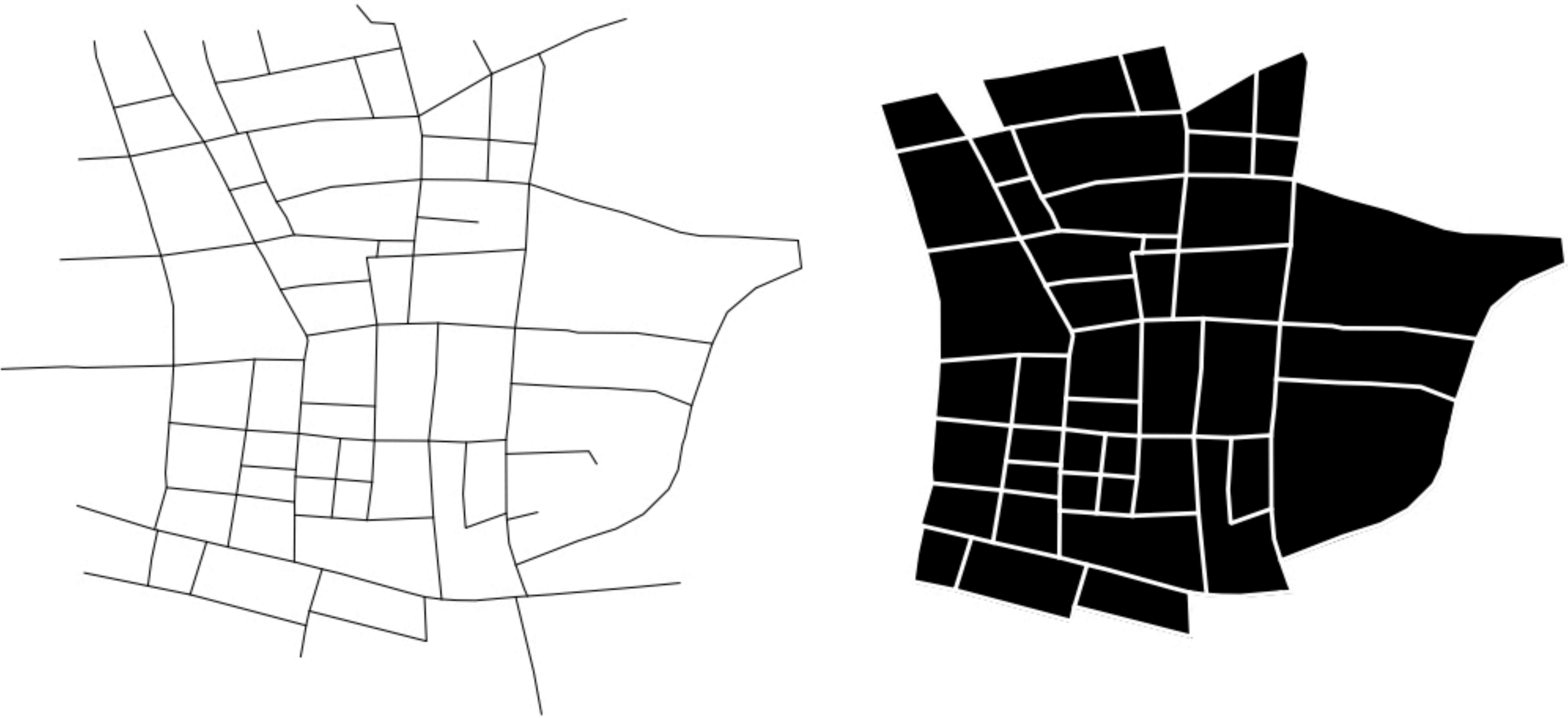
http://urbagram.stdio-london.com/v1/show/Network

https://nextcity.org/daily/entry/city-street-grid-maps-visualize-density

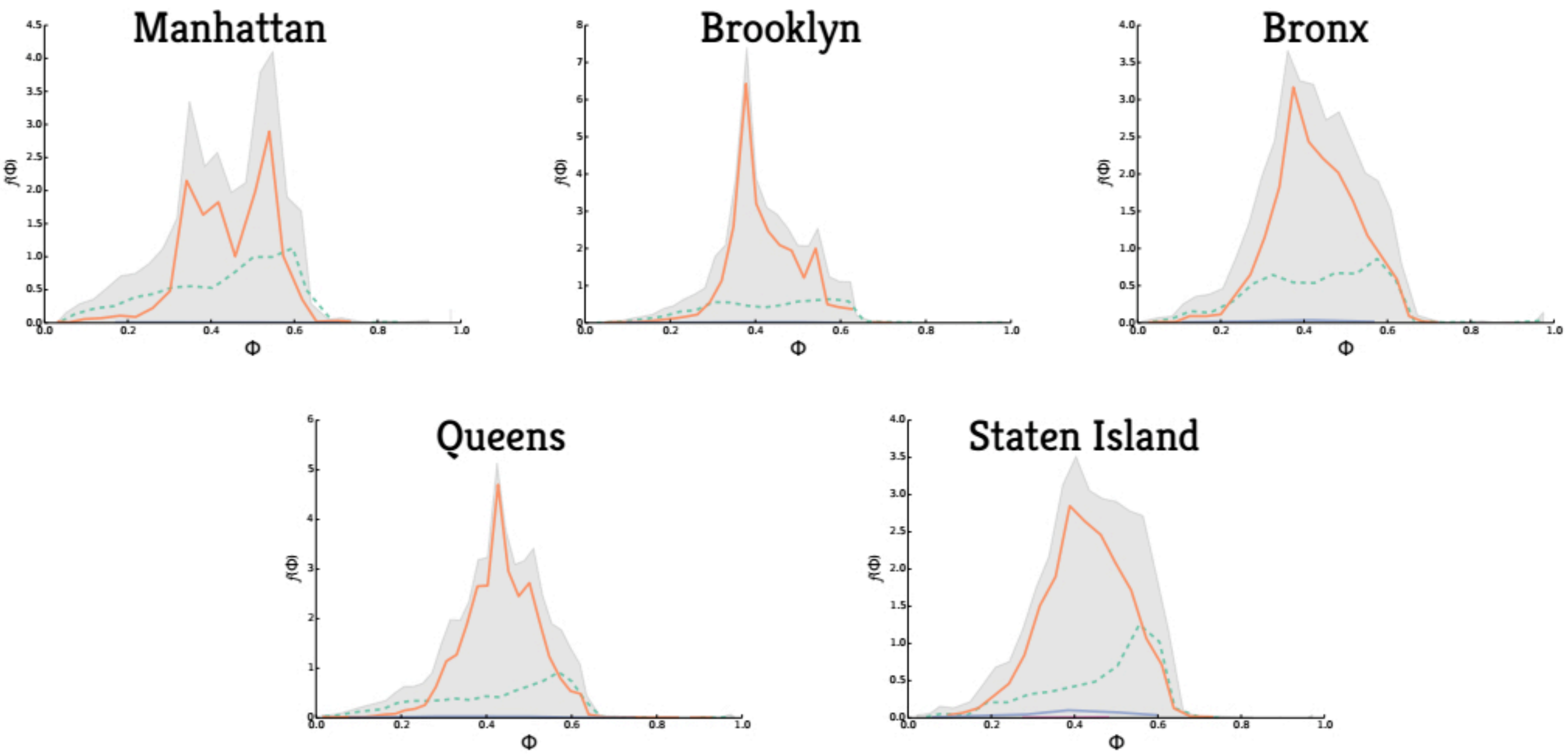
Connecting the Fractal City.

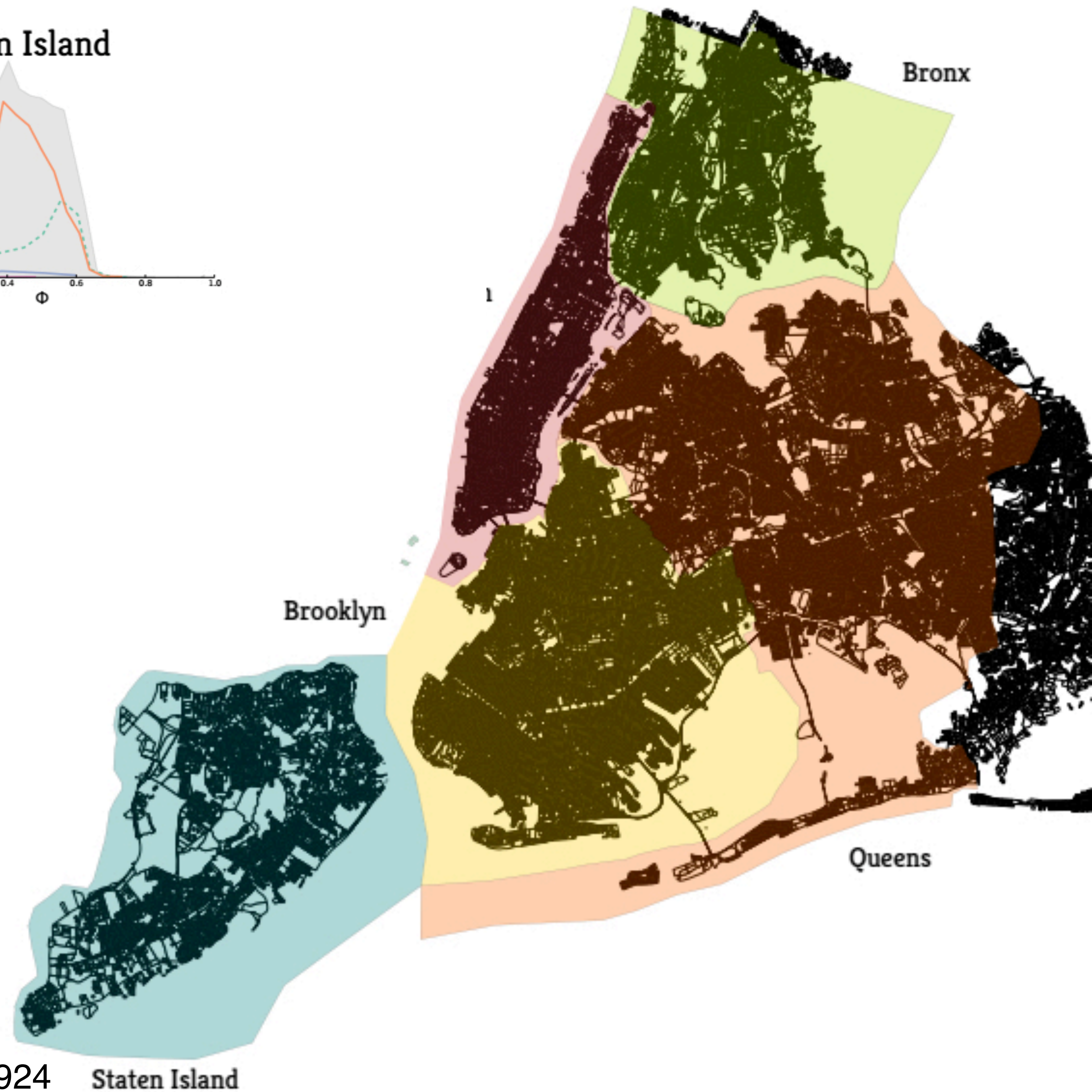
http://zeta.math.utsa.edu/~yxk833/connecting.html

Extracting land patches

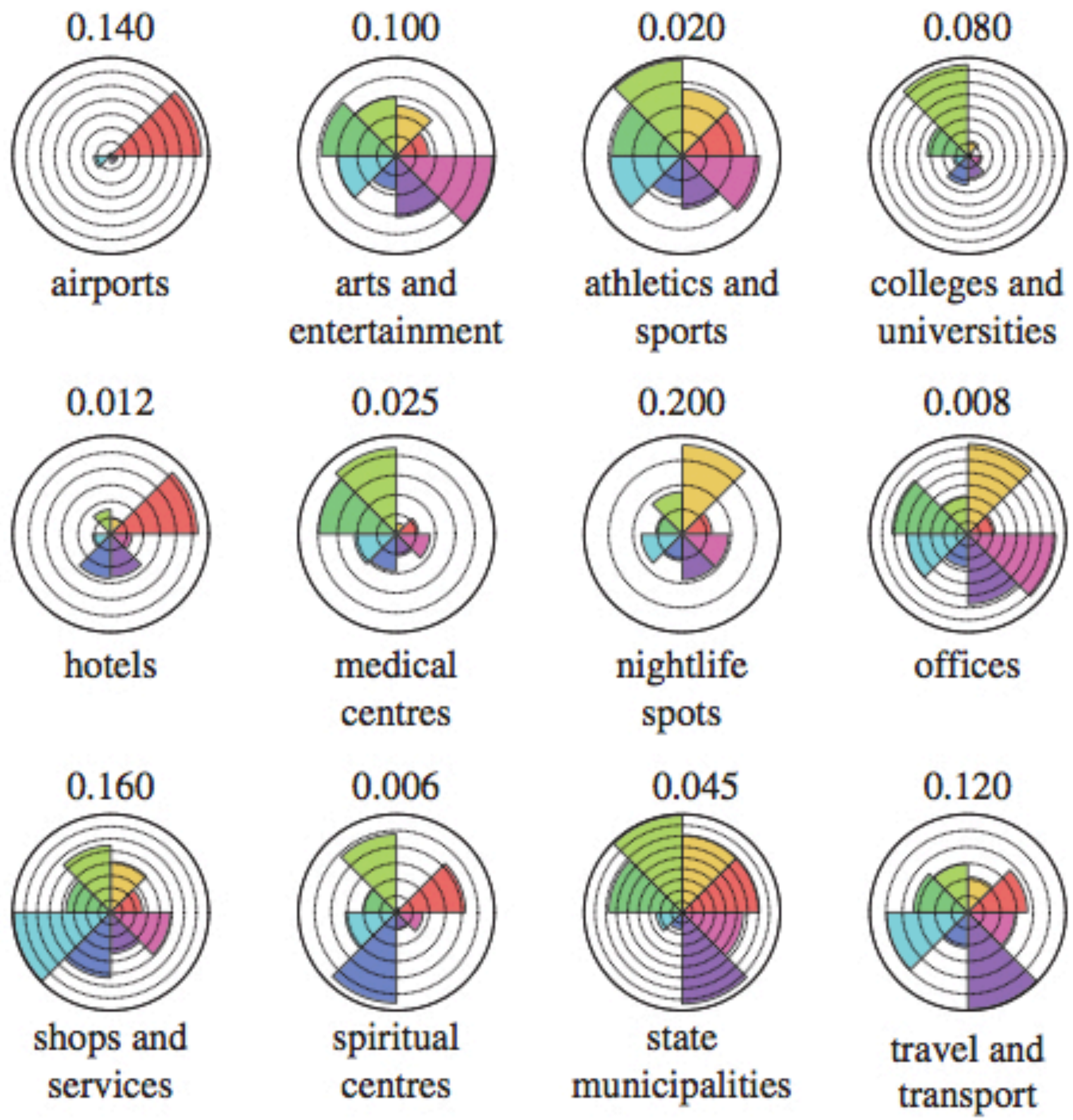


A typology of street patterns
R Louf, M Barthelemy
Journal of The Royal Society Interface 11 (101), 20140924





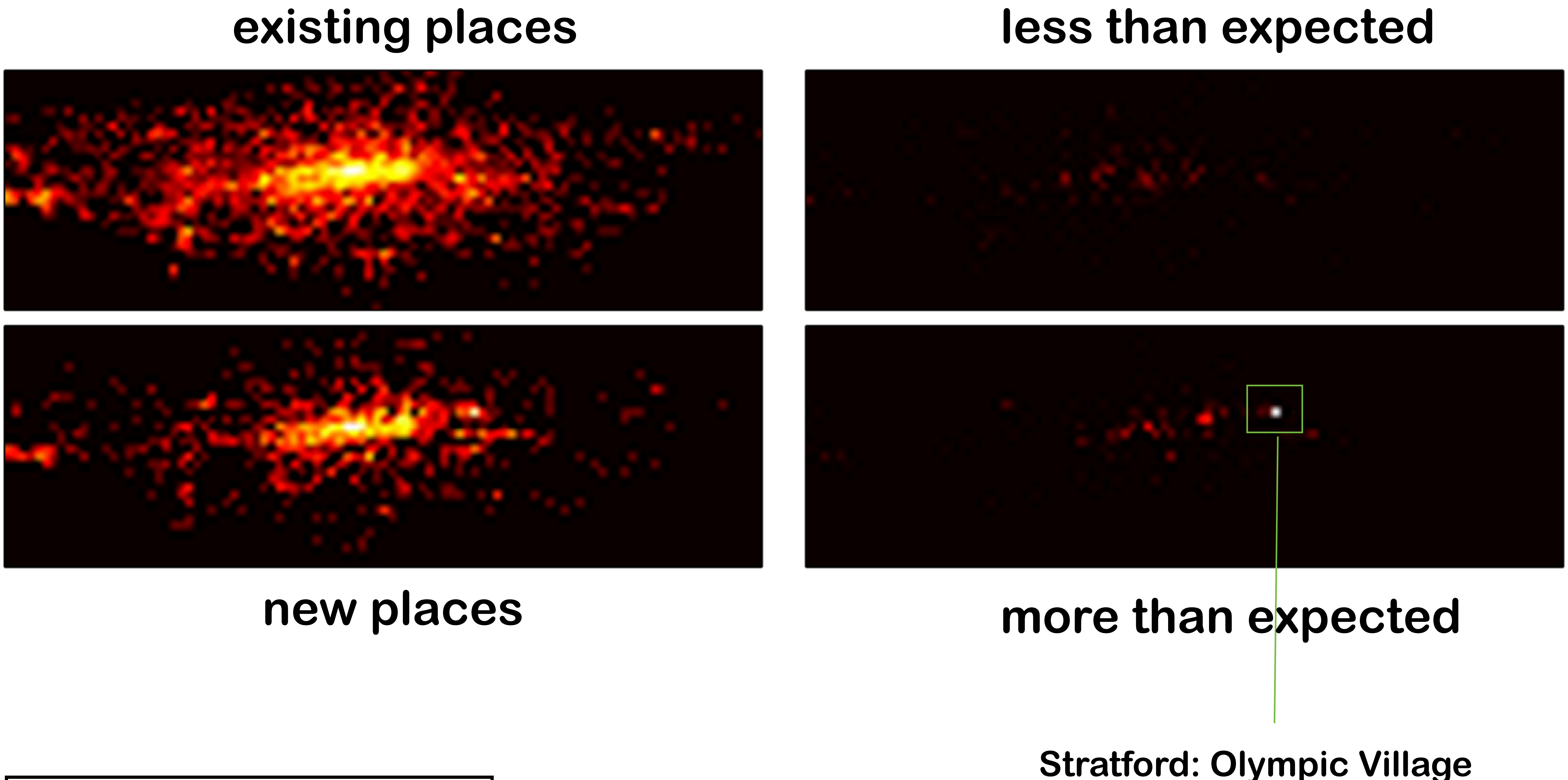
A typology of street patterns
R Louf, M Barthelemy
Journal of The Royal Society Interface 11 (101), 20140924



$$\mathbf{v}_i = \frac{N_{\text{new}}(i)}{\sum_i N_{\text{new}}(i)}.$$

Tracking Urban Activity Growth Globally with Big Location Data
M Daggitt, A Noulas, B Shaw, C Mascolo
Royal Society Open Science

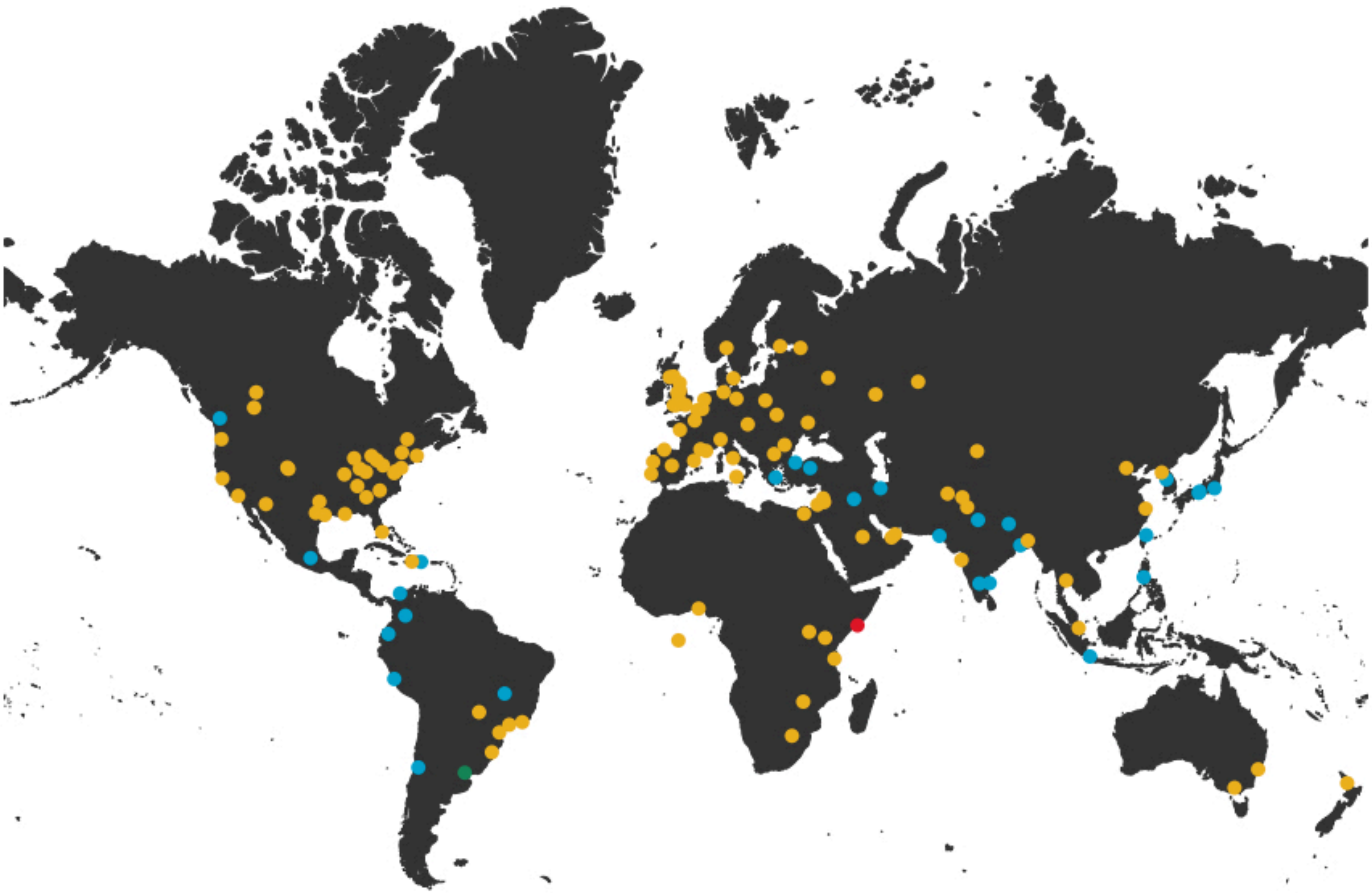
Tracking the birth of places [London 2010-2014]



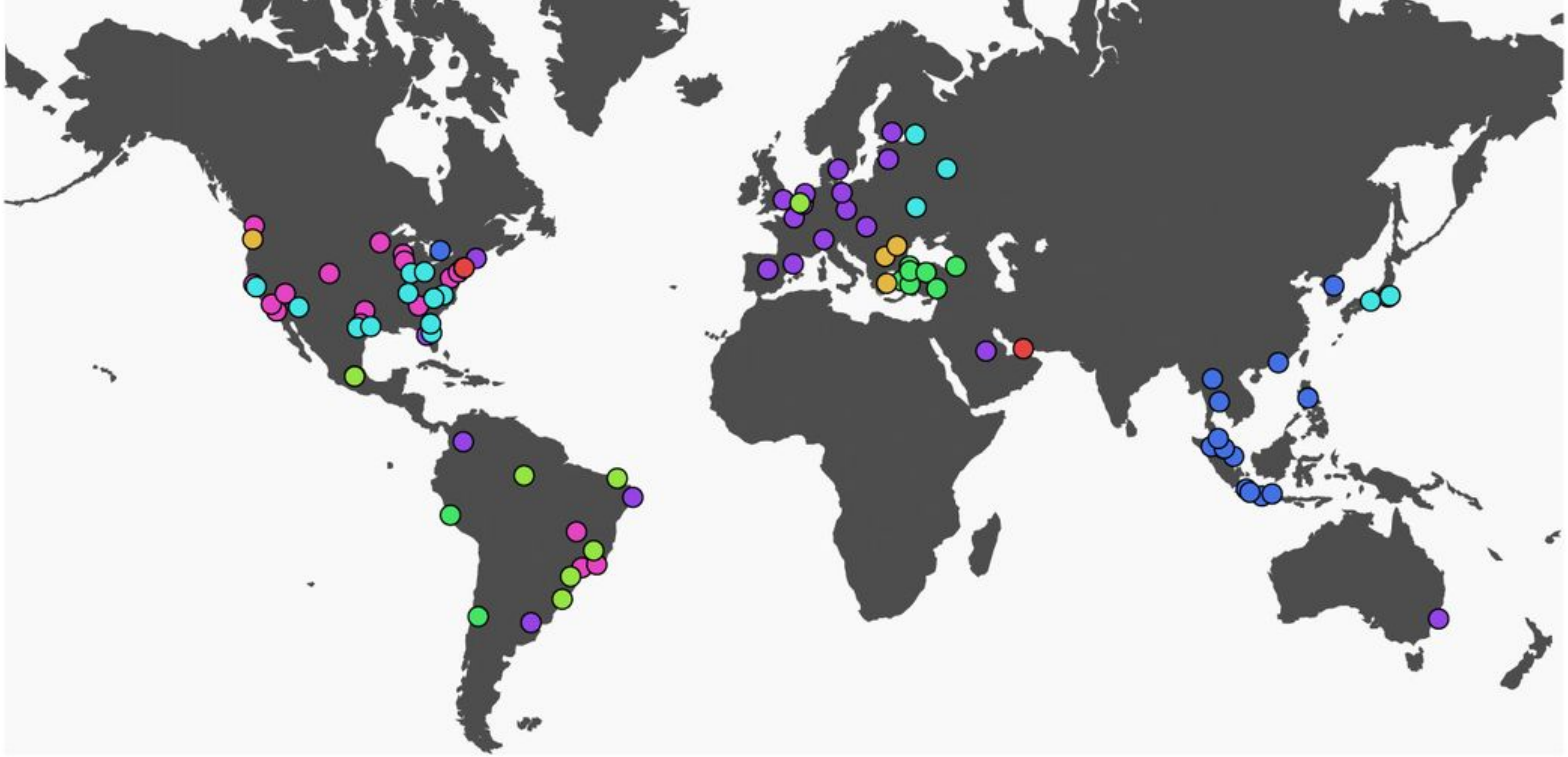
$$n_{i,j}^{\text{null}} = \frac{n_{i,j}^{\text{existing}} n^{\text{new}}}{n^{\text{existing}}}$$

$$v_{i,j} = n_{i,j}^{\text{null}} - n_{i,j}^{\text{new}}$$

Tracking Urban Activity Growth Globally with Big Location Data
M Daggitt, A Noulas, B Shaw, C Mascolo
Royal Society Open Science



A typology of street patterns
R Louf, M Barthelemy
Journal of The Royal Society Interface 11 (101), 20140924



no.	MPD (km)	cities
1	10 996	Dubai, Borough of Queens
2	5800	Athens, Brooklyn, Bucharest, Portland, Sofia
3	5250	Belo Horizonte, Coyoacán, Curitiba, Fortaleza, Gent, Manaus, Porto Alegre
4	4924	Adana, Ankara, Bursa, Denizli, Eskişehir, İstanbul, İzmir, Lima, Santiago, Trabzon
5	5887	Charlotte, Chiba, Columbus, Houston, Indianapolis, Jacksonville, Kiev, Moscow, Nashville, Orlando, Osaka, Phoenix, Raleigh, Saint Petersburg, San Antonio, San Jose, Yokohama
6	3537	Bandung, Bangkok, Chiang Mai, George Town, Hong Kong, Jakarta, Kuala Lumpur, Makati City, Medan, Petaling Jaya, Pineda, Quezon City, Seoul, Shah Alam, Singapore, Surabaya, Tokyo, Toronto, Yogyakarta
7	5790	Amsterdam, Barcelona, Berlin, Bogotà, Boston, Brussels, Budapest, Buenos Aires, Copenhagen, Helsinki, London, Madrid, Milano, Paris, Prague, Recife, Riga, Riyadh, Sydney, Tampa
8	4276	Antwerpen, Atlanta, Austin, Brasília, Chicago, Dallas, Denver, Las Vegas, Los Angeles, Mexico City, Milwaukee, Minneapolis, New York, Philadelphia, Rio de Janeiro, San Diego, San Francisco, São Paulo, Seattle, Washington DC

The “gig” economy

Mobile web and digital mapping technologies have brought a revolution on private resource utilization....

Rent out your extra space to travelers or use your car to drive them around... this is just the beginning of a big revolution that brings together the physical and digital space.

Home

About TLC

» TLC Mission Statement
» Commission Room
» TLC Facilities
» TLC Staff
» Research and Statistics
» Annual Reports
» Employment Opportunities
» Interagency MOUs

TLC Rules and Local Laws

Licensing/Industry Information

Passenger Information

Frequently Asked Questions

TLC News

TLC Site Map

Contact/Visit TLC

Online Transactions (LARS)

Apply for a License

Pay Renewal Fee

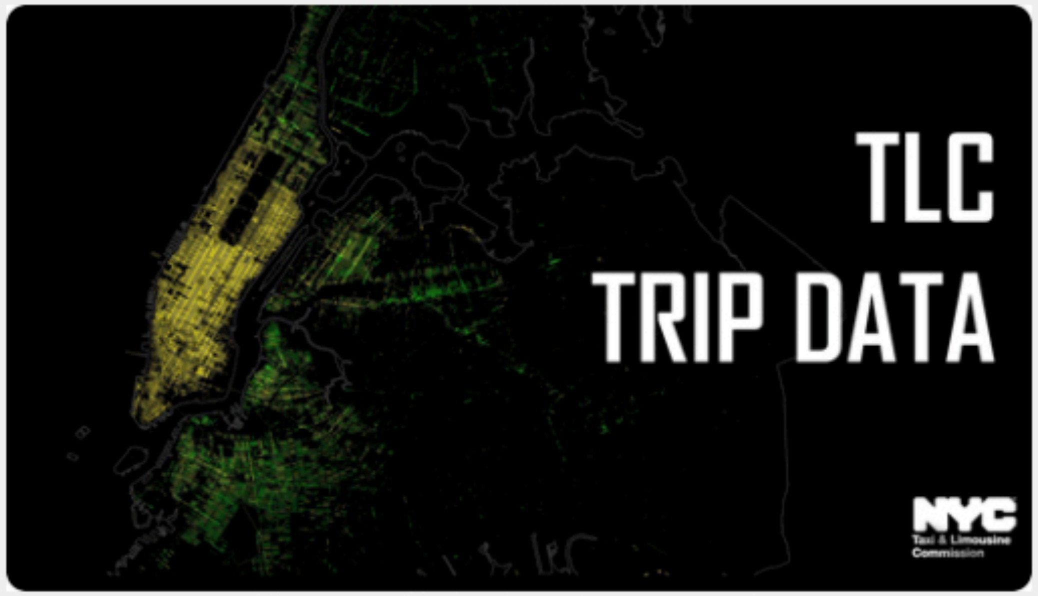
Pay Summons

Pay Other Fees

Update License Information

Additional Information

TLC Trip Record Data



The yellow and green taxi trip records include fields capturing pick-up and drop-off dates/times, pick-up and drop-off locations, trip distances, itemized fares, rate types, payment types, and driver-reported passenger counts. The data used in the attached datasets were collected and provided to the NYC Taxi and Limousine Commission (TLC) by technology providers authorized under the Taxicab & Livery Passenger Enhancement Programs (TPEP/LPEP). The trip data was not created by the TLC, and TLC makes no representations as to the accuracy of these data.

The For-Hire Vehicle ("FHV") trip records include fields capturing the dispatching base license number and the pick-up date, time, and taxi zone location ID (shape file below). These records are generated from the FHV Trip Record submissions made by bases. Note: The TLC publishes base trip record data as submitted by the bases, and we cannot guarantee or confirm their accuracy or completeness. Therefore, this may not represent the total amount of trips dispatched by all TLC-licensed bases. The TLC performs routine reviews of the records and takes enforcement actions when necessary to ensure, to the extent possible, complete and accurate information.

For trip record data including TLC taxi zone location IDs, location names and corresponding boroughs for each ID can be found here. A shapefile containing the boundaries for the taxi zones can be found here.

Trip Sheet Data (CSV Format)

2016

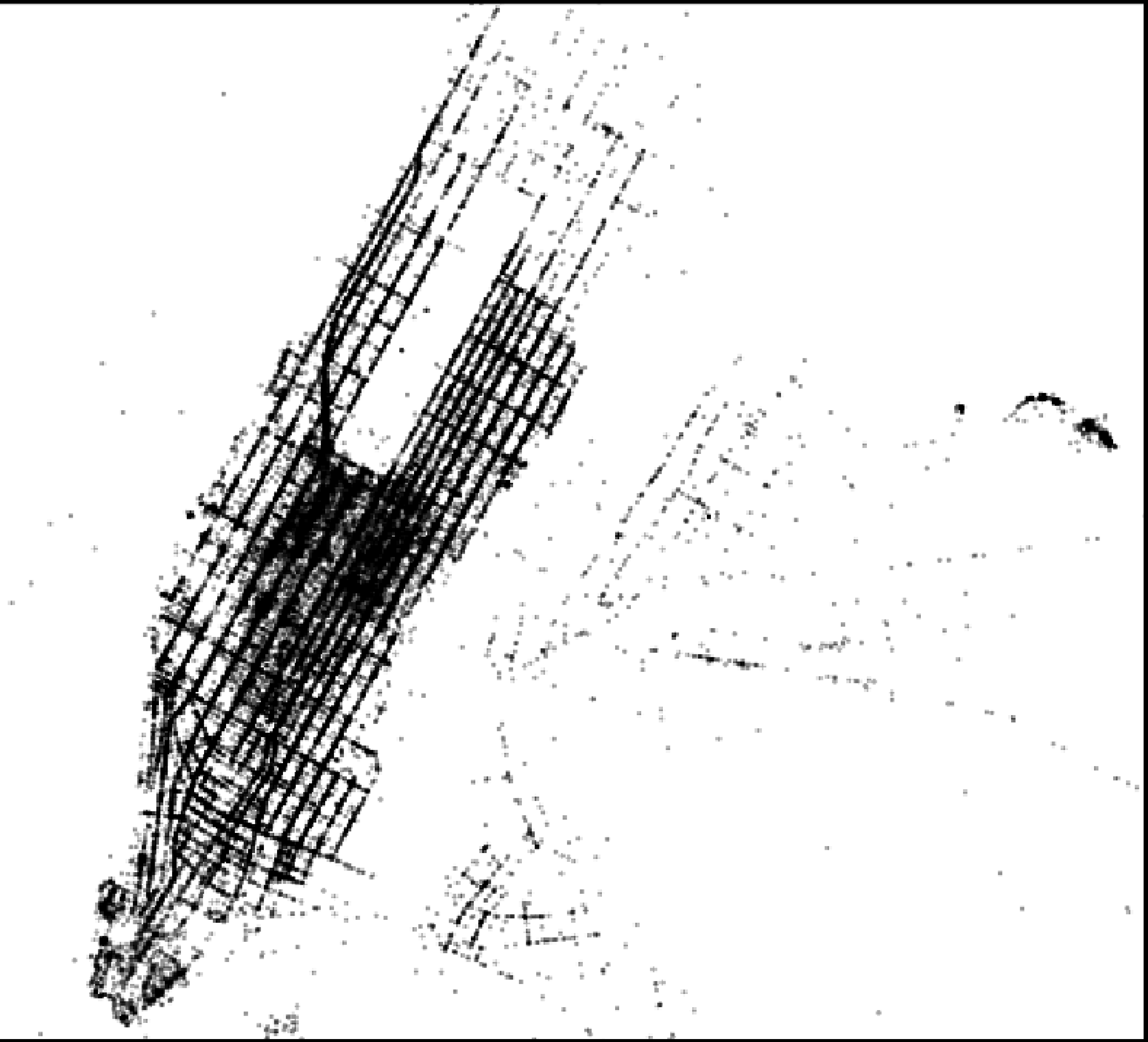
January	Yellow	Green	FHV
February	Yellow	Green	FHV
March	Yellow	Green	FHV
April	Yellow	Green	FHV

Taxi News

The new TLC Driver License will authorize holders to operate all taxi and for-hire vehicles. Find out more about this new license!

< 3/6 >

THE NEW YORK CITY TAXI DATASET



FOILing NYC's Taxi Trip Data

Freedom of Information Law

2013 Trip Data, 11GB, zipped!

2013 Fare Data, 7.7GB

Idea: Uber Vs Yellow Taxi Price Comparison.

July 7, 2014

uberX

NOW CHEAPER THAN A NEW YORK CITY TAXI

HOW THESE PRICES COMPARE

Williamsburg to East Village

KEEP IN MIND

These prices are only in effect for a limited time. The more you ride, the more likely we can keep them this low!

We know you may be asking yourself how this affects our partner drivers. What we've seen in cities across the country is that lower fares mean greater demand, lower pickup times and more trips per hour – increasing earning potential and creating better economics for drivers. What does what mean in the long run? They'll be making more than ever!

THE EXPERIMENT

x1,y1

x2,y2

1. For every trip in NYC taxi dataset.

2. Record origin & destination coordinates.

3. Retrieve total fare paid.

4. Query Uber API price for the same trip.

5. Compare yellow taxi VS uber prices.

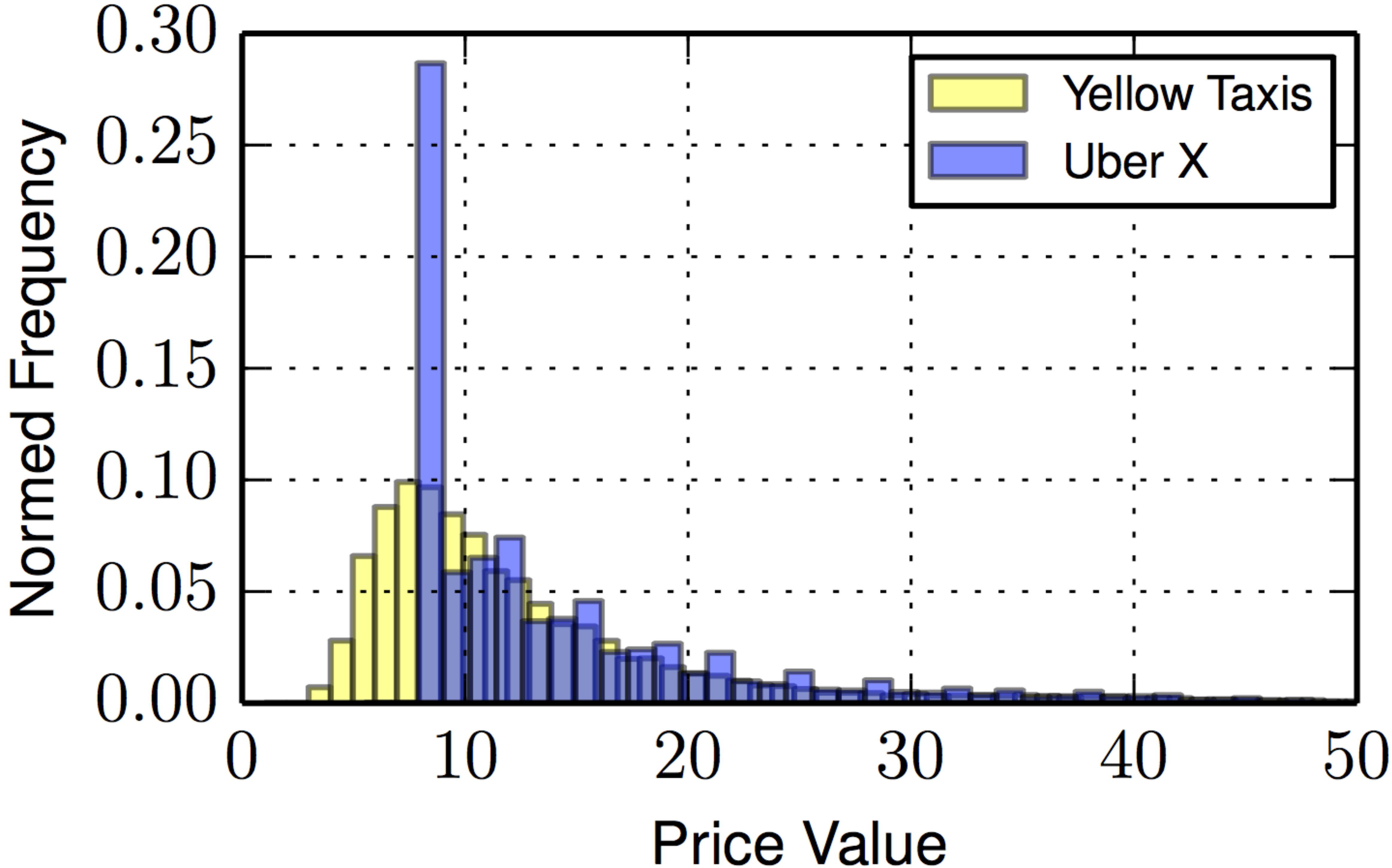
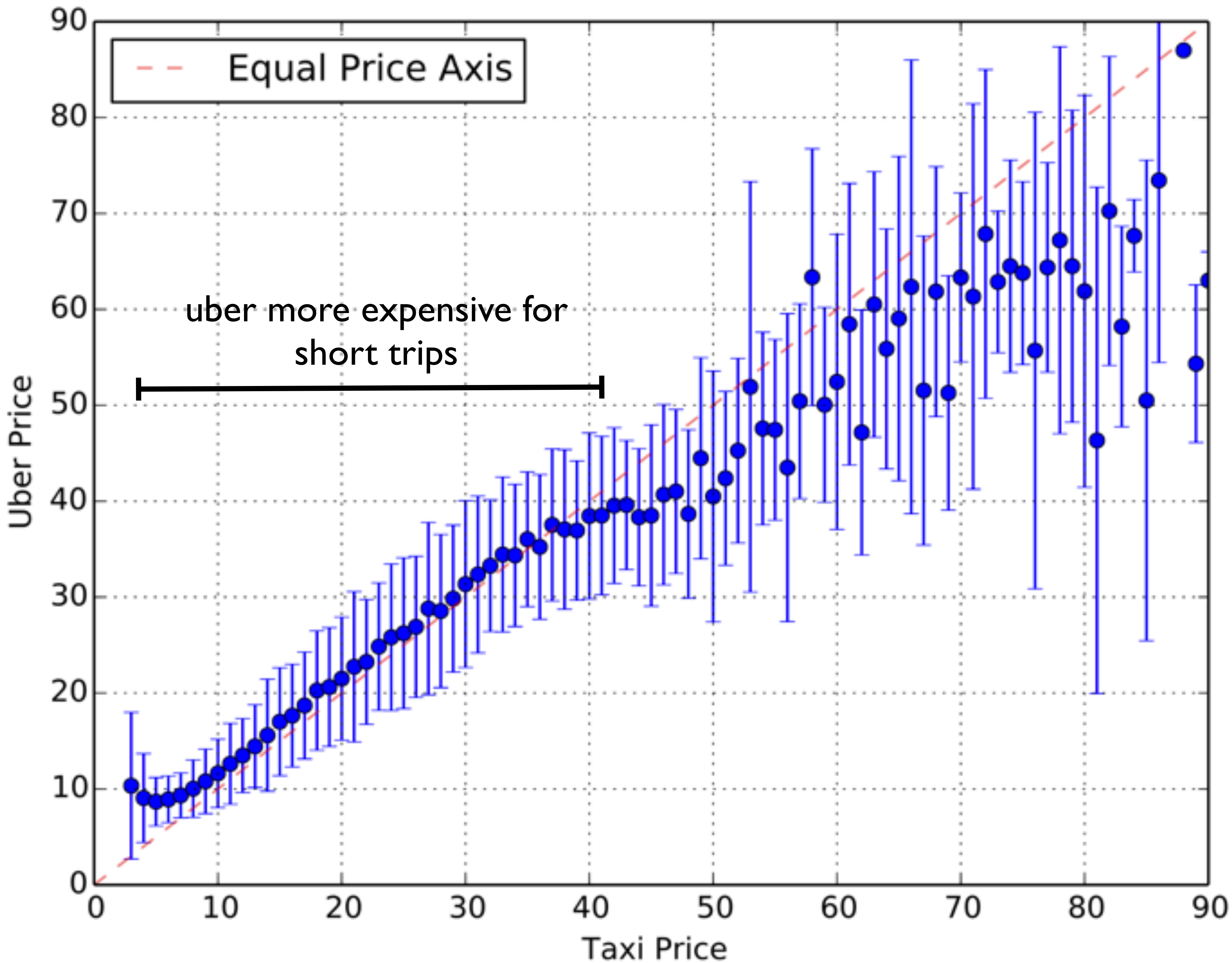
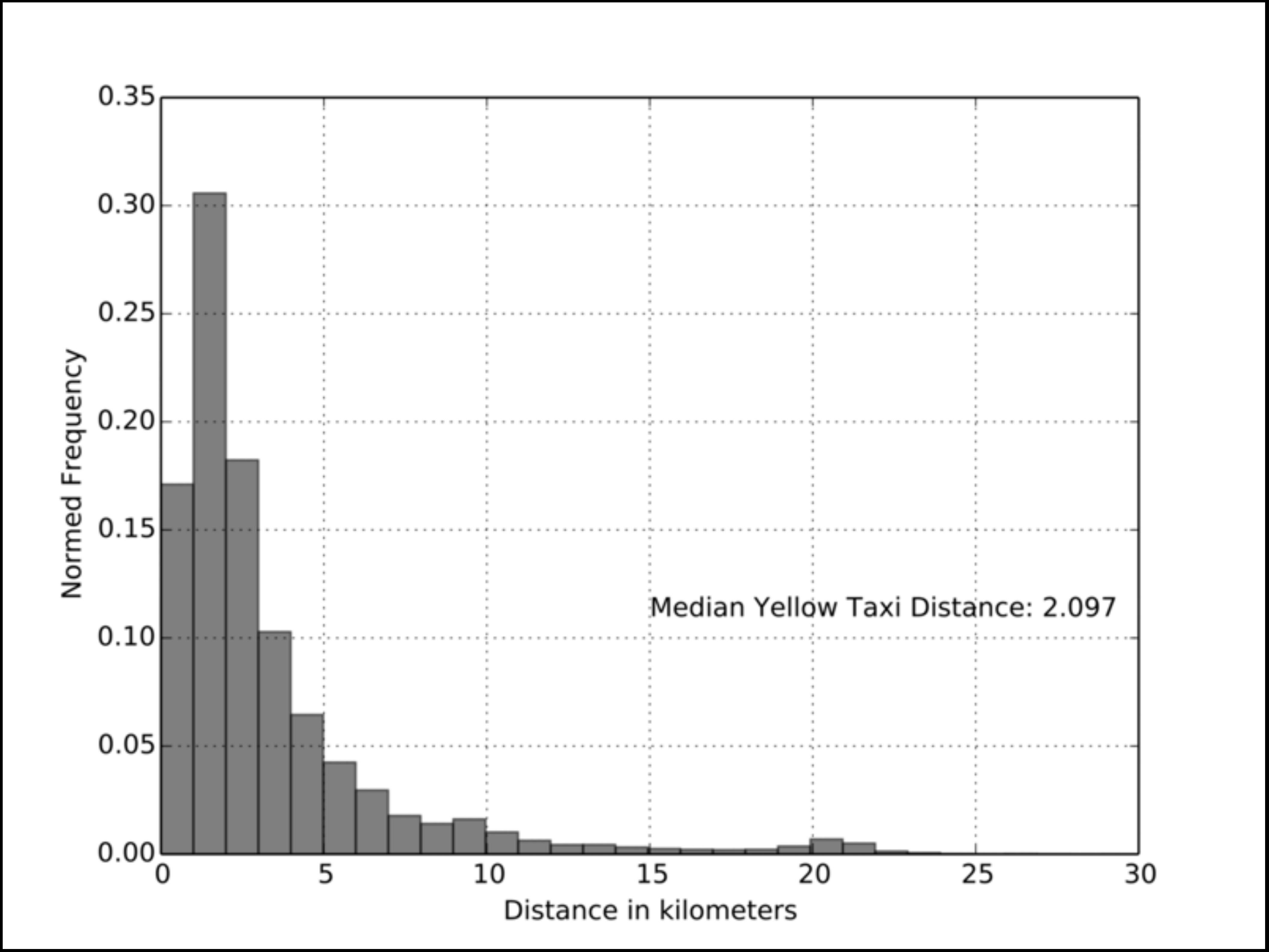


Figure 2: Distribution of prices per journey for Uber X and Yellow Taxis in New York City.

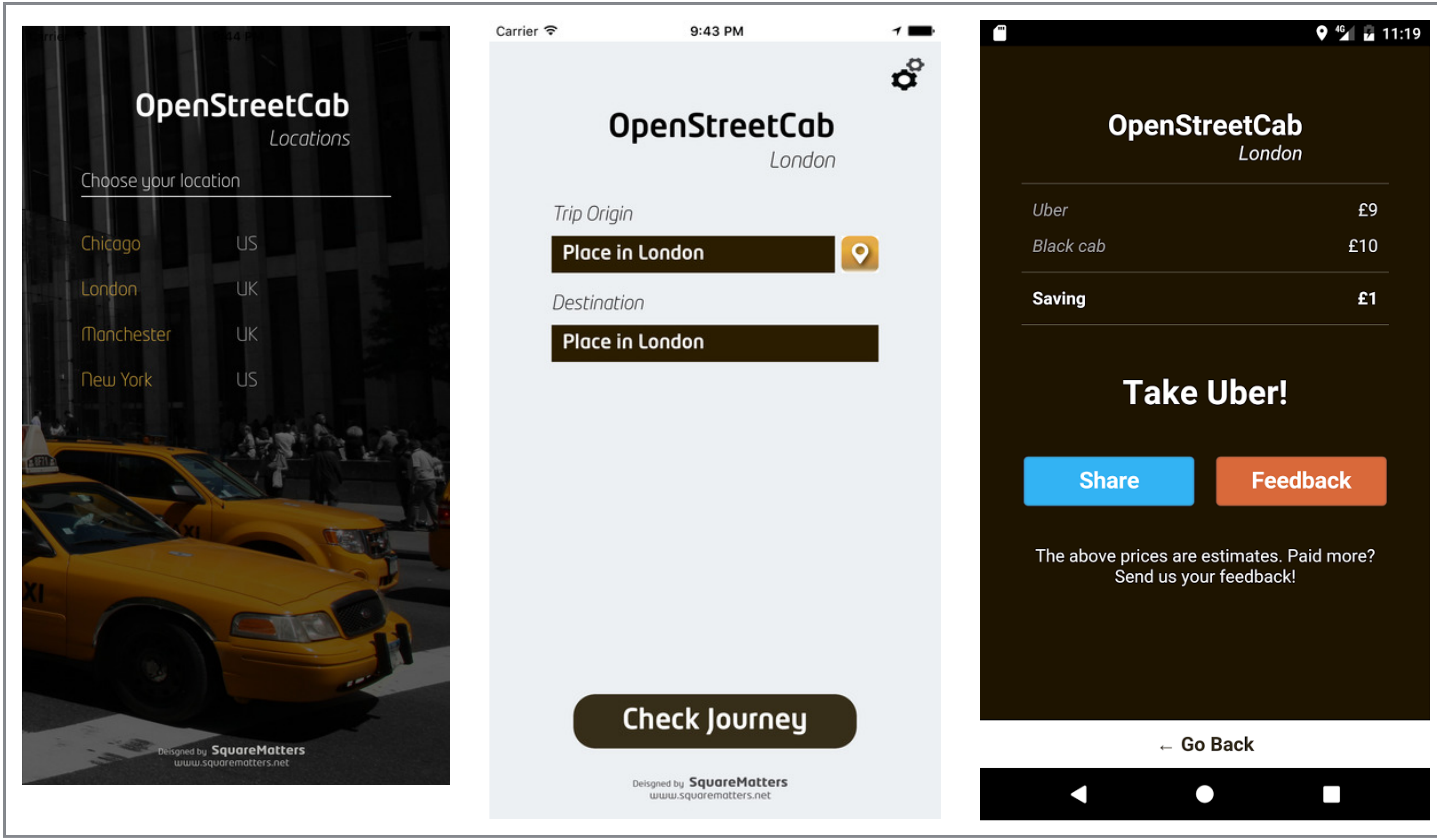

Equal Price Axis
uber more expensive for
short trips
Uber Price
Taxi Price
0
10
20
30
40
50
60
70
80
90

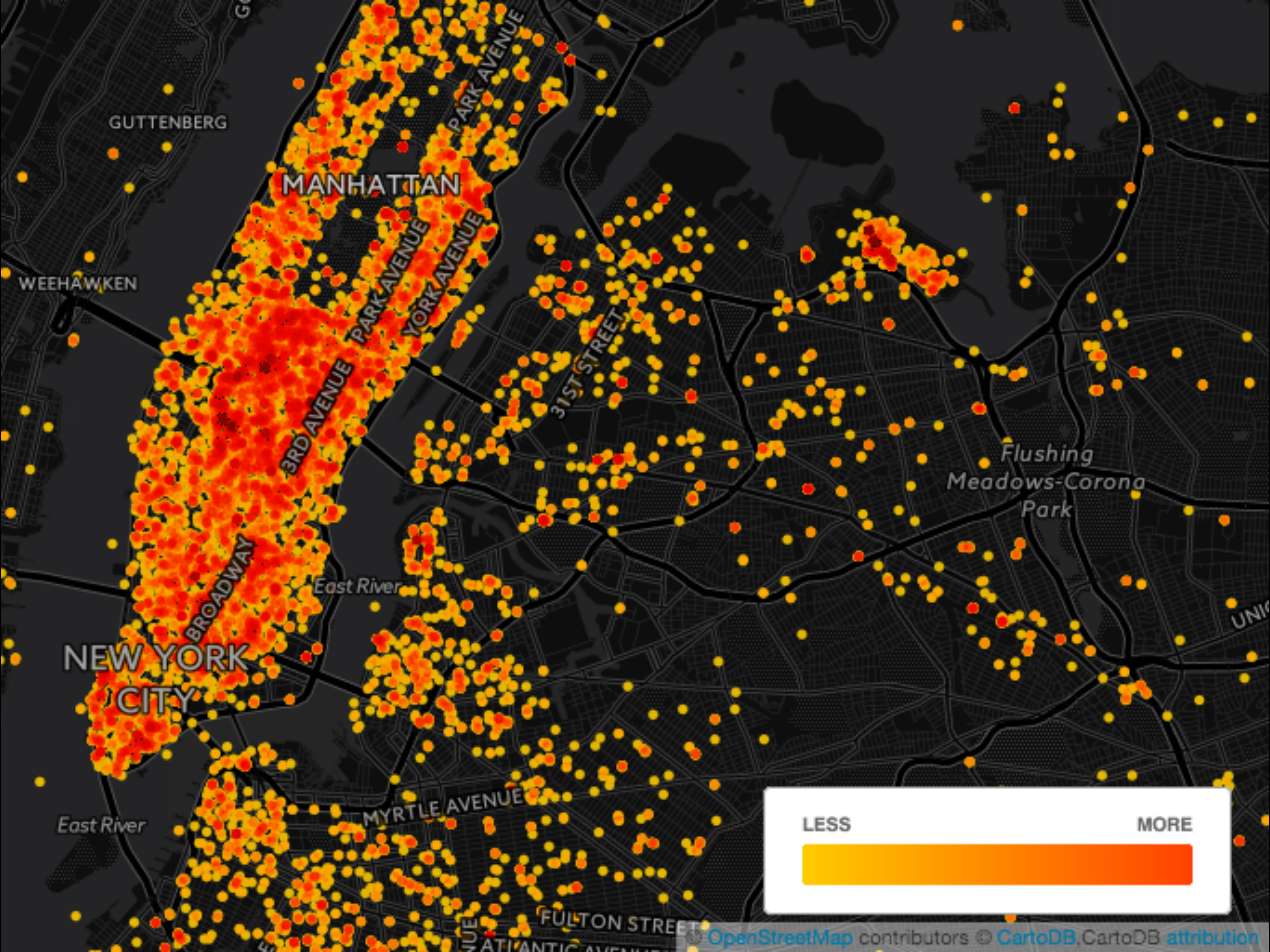


Most taxi movements are within a short distance range with longer movements occurring less frequently in the data

OpenStreetCab

OpenStreetCab





GUTTENBERG
MANHATTAN
PARK AVENUE
WEEHAWKEN
YORK AVENUE
3RD AVENUE
31ST STREET
Flushing Meadows-Corona Park
East River
BROADWAY
NEW YORK CITY
MYRTLE AVENUE
FULTON STREET
LESS
MORE
© OpenStreetMap contributors © CartoDB, CartoDB attribution

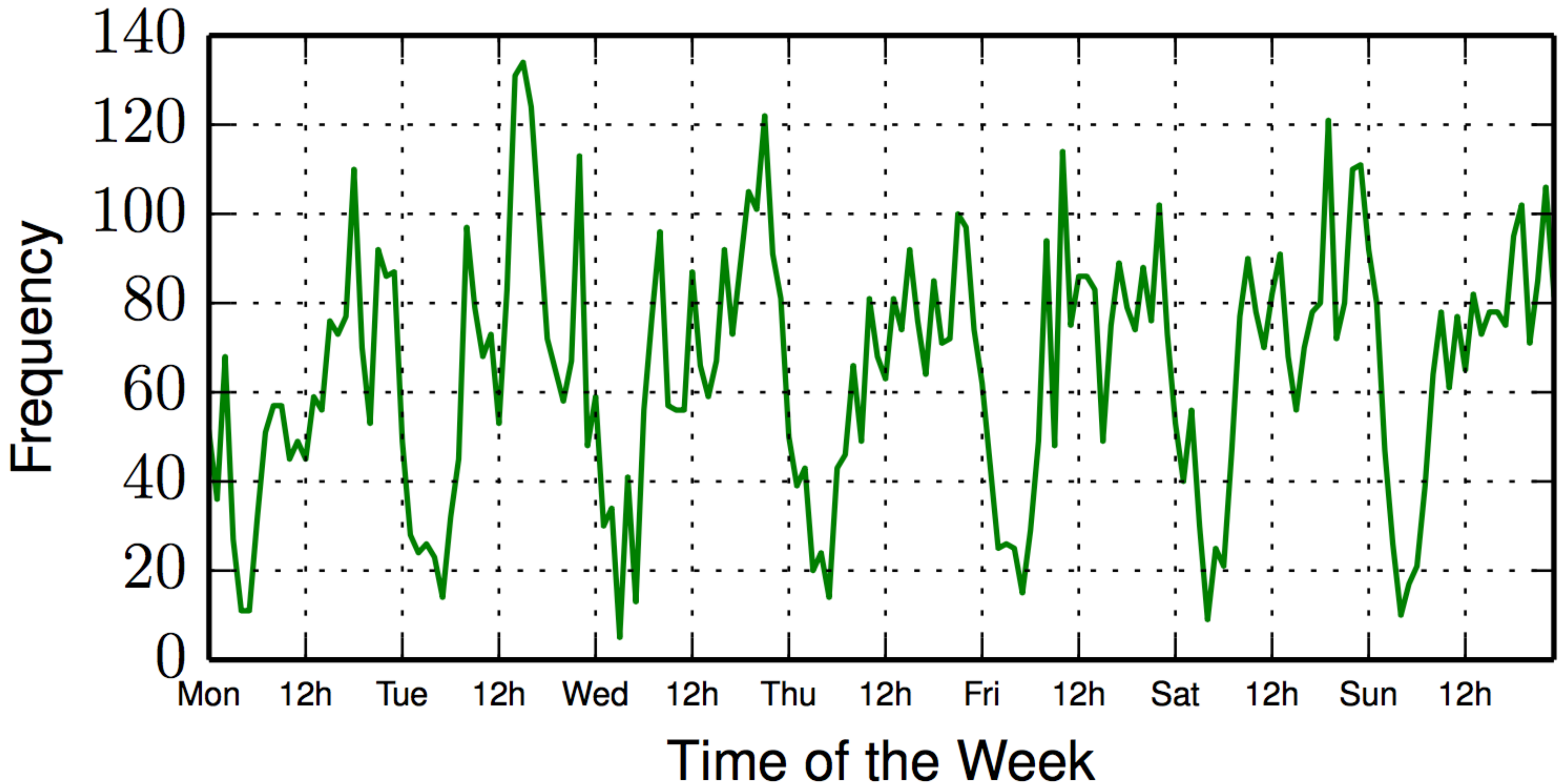
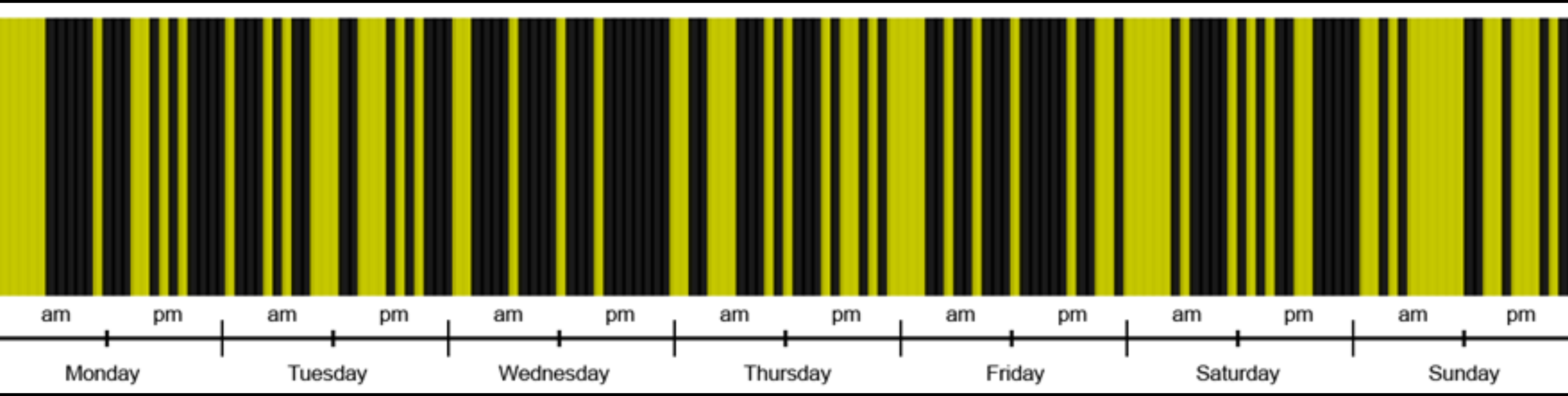
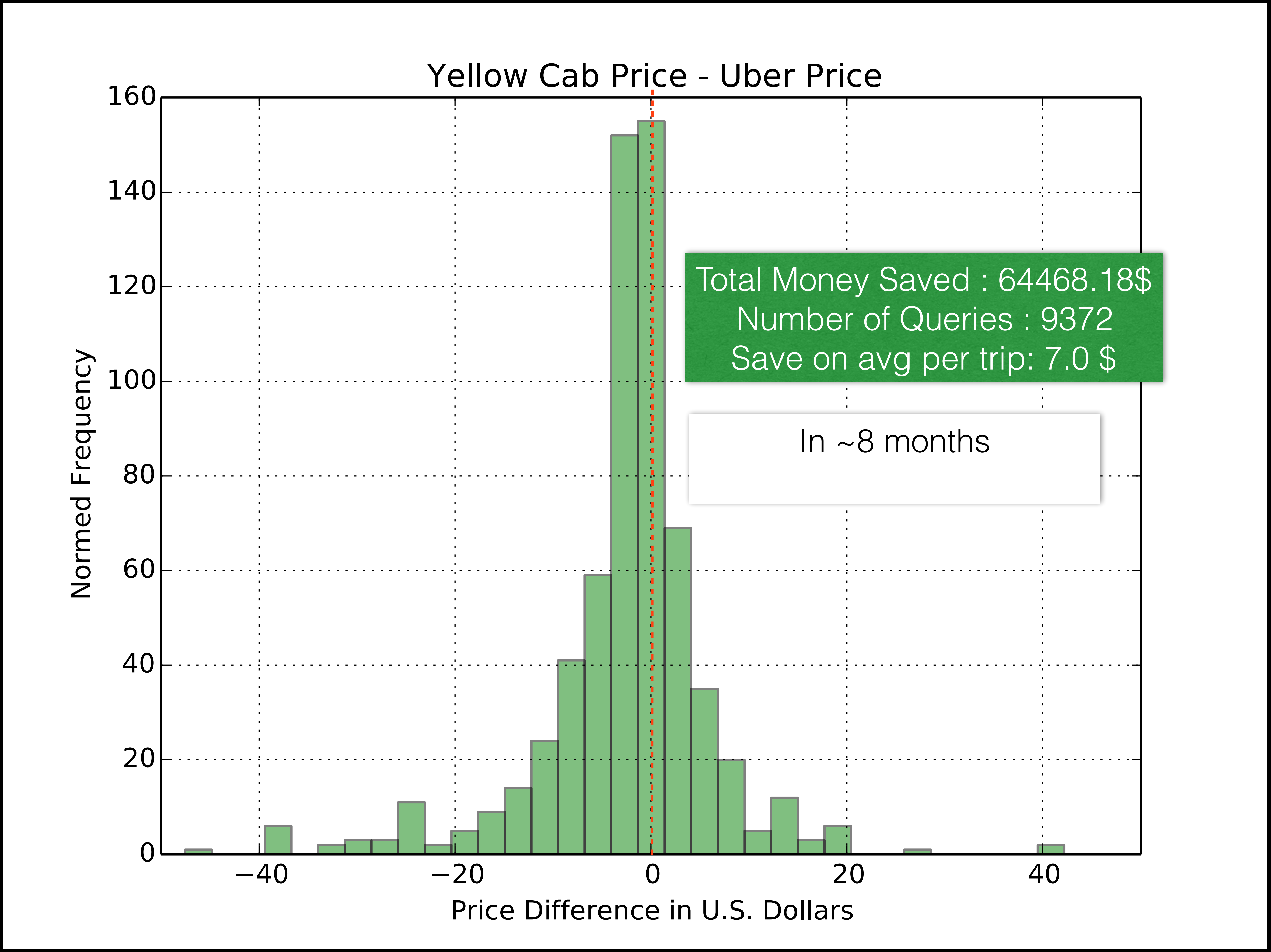


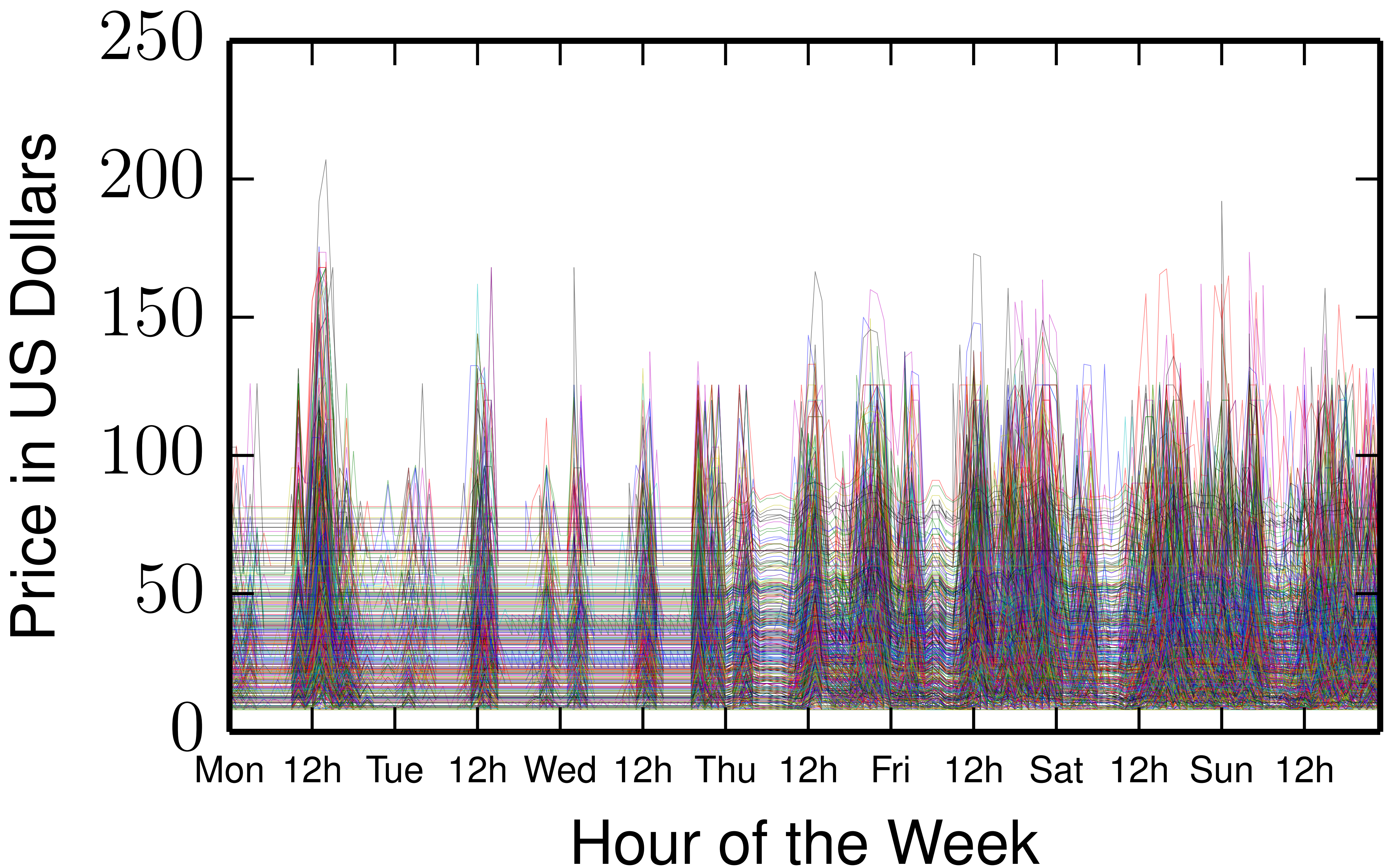
Figure 9: User Query Frequency in terms of weekly temporal evolution patterns.

> 5K only in New York (~8K downloaded the app)

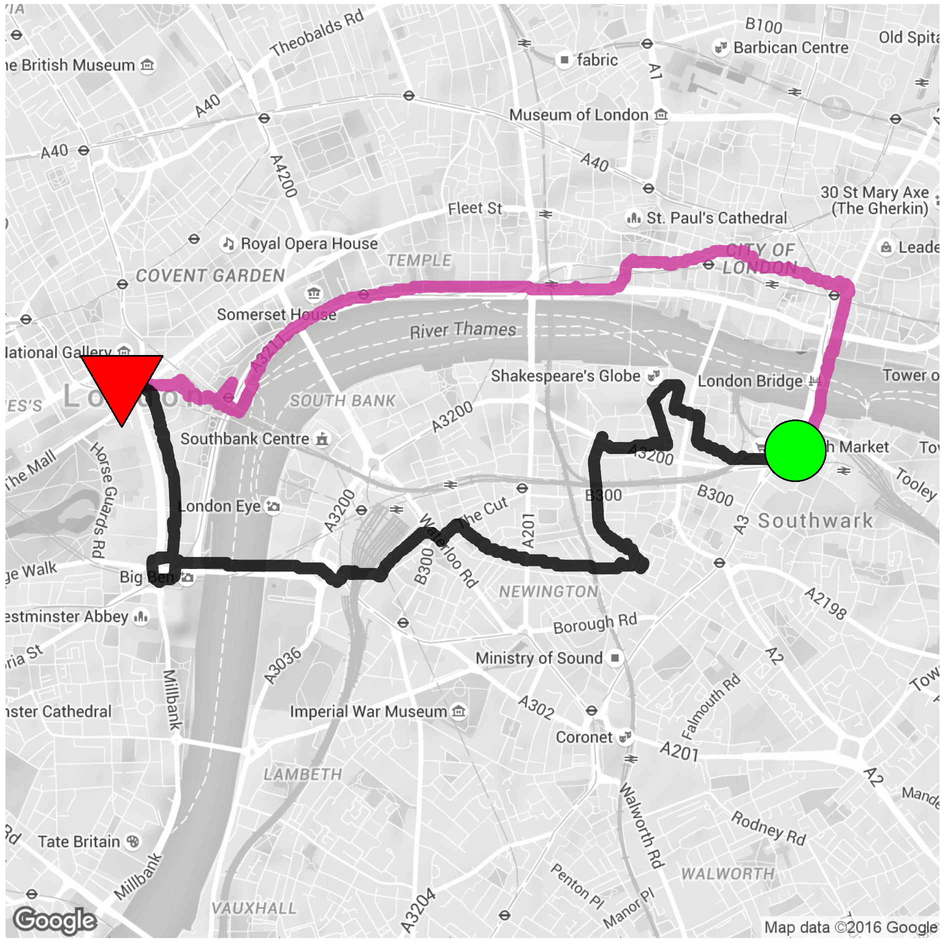
> 14000 search queries generated



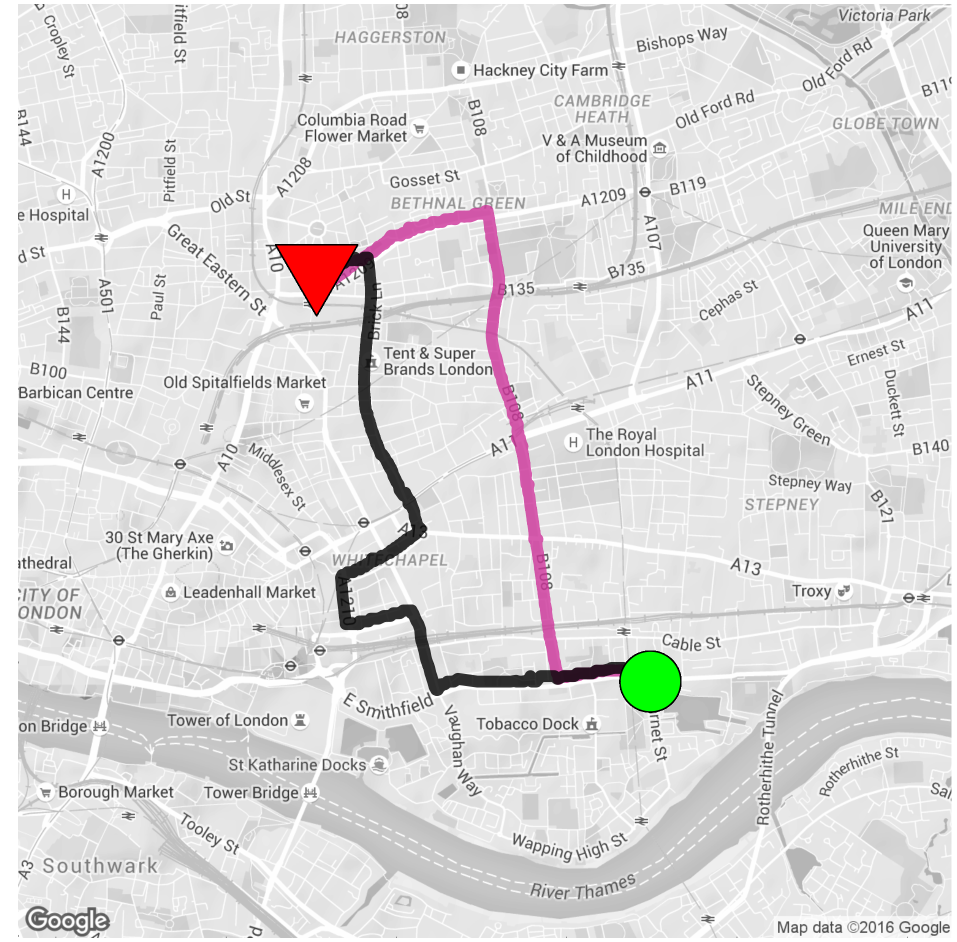

Yellow Cab Price - Uber Price
Normed Frequency
Price Difference in U.S. Dollars
160
140
120
100
80
60
40
20
0
−40
−20
0
20
40
Total Money Saved : 64468.18$
Number of Queries : 9372
Save on avg per trip: 7.0 $
In ~8 months



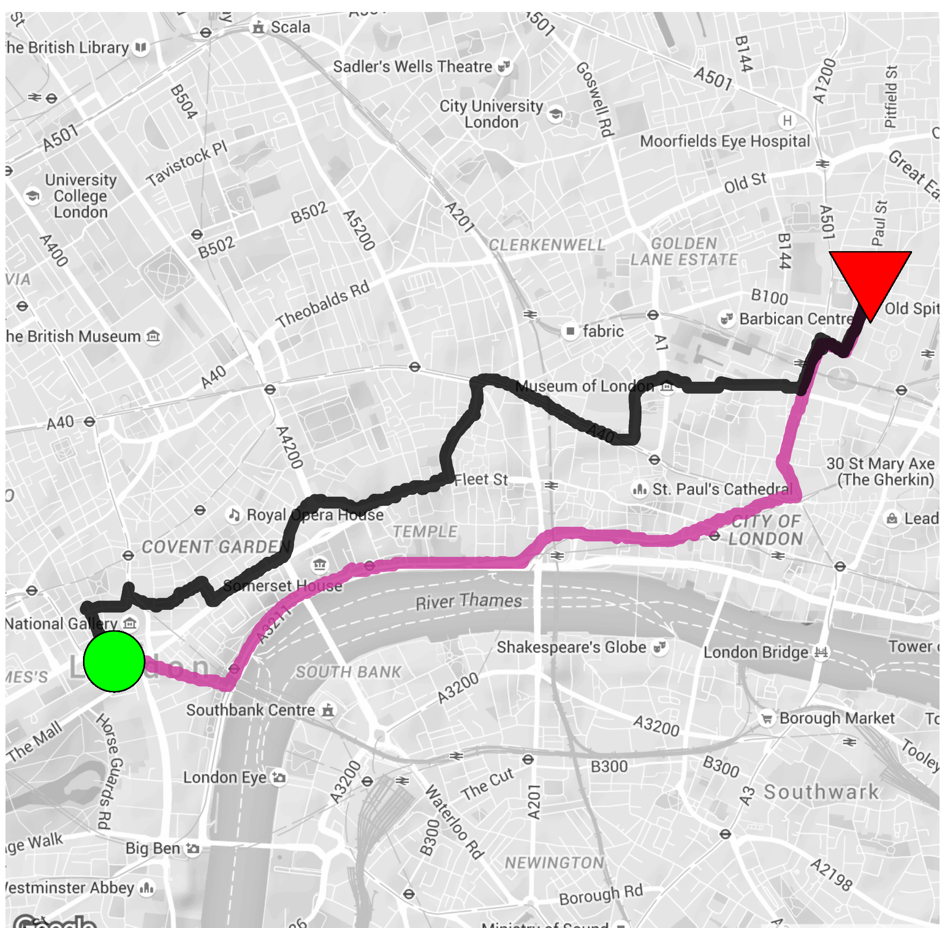
250
200
150
100
50
0
Price in US Dollars
Mon
12h
Tue
12h
Wed
12h
Thu
12h
Fri
12h
Sat
12h
Sun
12h
Hour of the Week



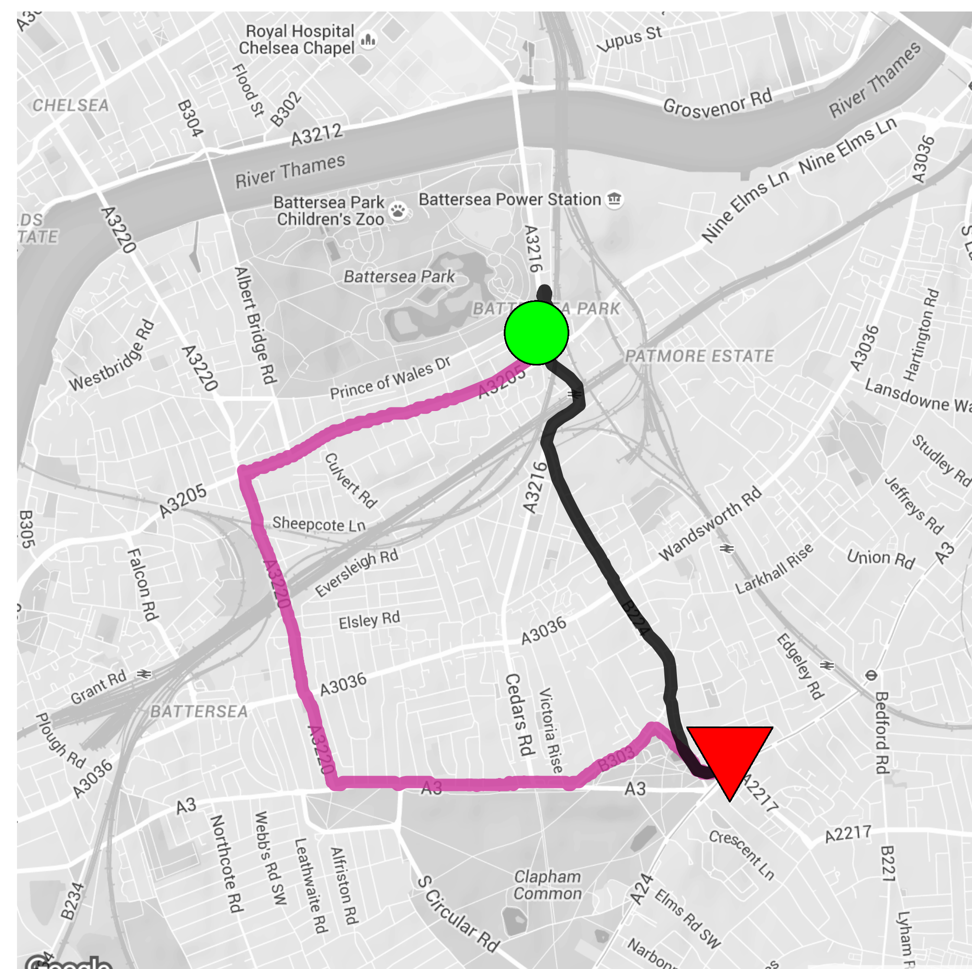
Theobalds Rd
Barbican Centre
fabric
Museum of London
Fleet St
St. Paul's Cathedral
30 St Mary Axe (The Gherkin)
Royal Opera House
COVENT GARDEN
TEMPLE
CITY OF LONDON
Somerset House
River Thames
Shakespeare's Globe
London Bridge
SOUTH BANK
Southbank Centre
London Eye
Big Ben
Westminster Abbey
Southwark
NEWINGTON
Borough Rd
Ministry of Sound
Imperial War Museum
Coronet
LAMBETH
Tate Britain
VAUXHALL
WALWORTH
Rodney Rd
Map data ©2016 Google



HAGGERSTON
Hackney City Farm
Bishops Way
Victoria Park
Columbia Road Flower Market
CAMBRIDGE HEATH
V & A Museum of Childhood
GLOBE TOWN
Gosset St
BETHNAL GREEN
Queen Mary University of London
Great Eastern St
Tent & Super Brands London
Old Spitalfields Market
Barbican Centre
The Royal London Hospital
STEPNEY
30 St Mary Axe (The Gherkin)
WHITECHAPEL
Leadenhall Market
CITY OF LONDON
Troxy
Cable St
Tower of London
E Smithfield
Tobacco Dock
St Katharine Docks
Borough Market
Tower Bridge
Wapping High St
River Thames
Southwark
Map data ©2016 Google



Scala
The British Library
Sadler's Wells Theatre
City University London
Moorfields Eye Hospital
University College London
CLERKENWELL
GOLDEN LANE ESTATE
Theobalds Rd
Barbican Centre
The British Museum
fabric
Museum of London
Fleet St
St. Paul's Cathedral
30 St Mary Axe (The Gherkin)
COVENT GARDEN
TEMPLE
CITY OF LONDON
Somerset House
River Thames
Shakespeare's Globe
London Bridge
National Gallery
SOUTH BANK
Southbank Centre
Borough Market
London Eye
Big Ben
Southwark
NEWINGTON
Borough Rd
Ministry of Sound
Map data ©2016 Google

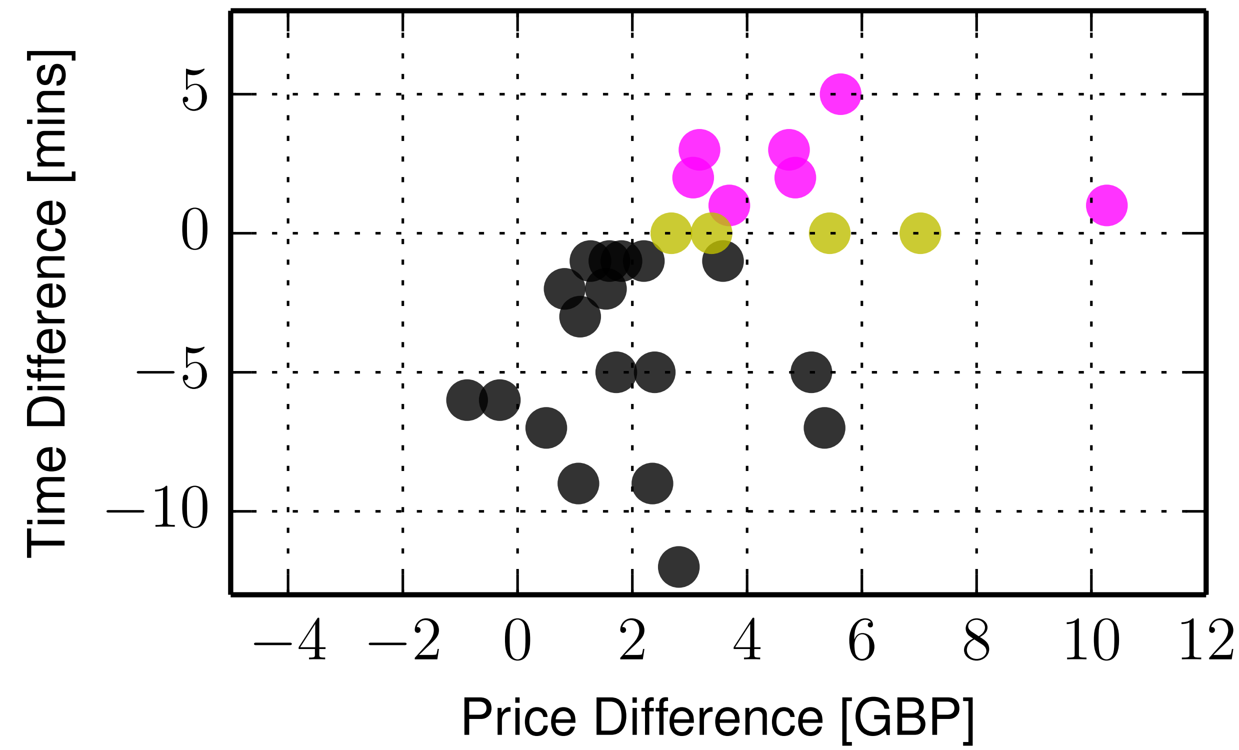


Royal Hospital Chelsea Chapel
CHELSEA
Grosvenor Rd
River Thames
Battersea Park Children's Zoo
Battersea Power Station
Nine Elms Ln
Battersea Park
PATMORE ESTATE
Westbridge Rd
Prince of Wales Dr
Lansdowne Way
Culvert Rd
Sheepcote Ln
Wandsworth Rd
Eversleigh Rd
Elsley Rd
Union Rd
BATTERSEA
Cedars Rd
Victoria Rise
Clapham Common
S Circular Rd
Map data ©2016 Google

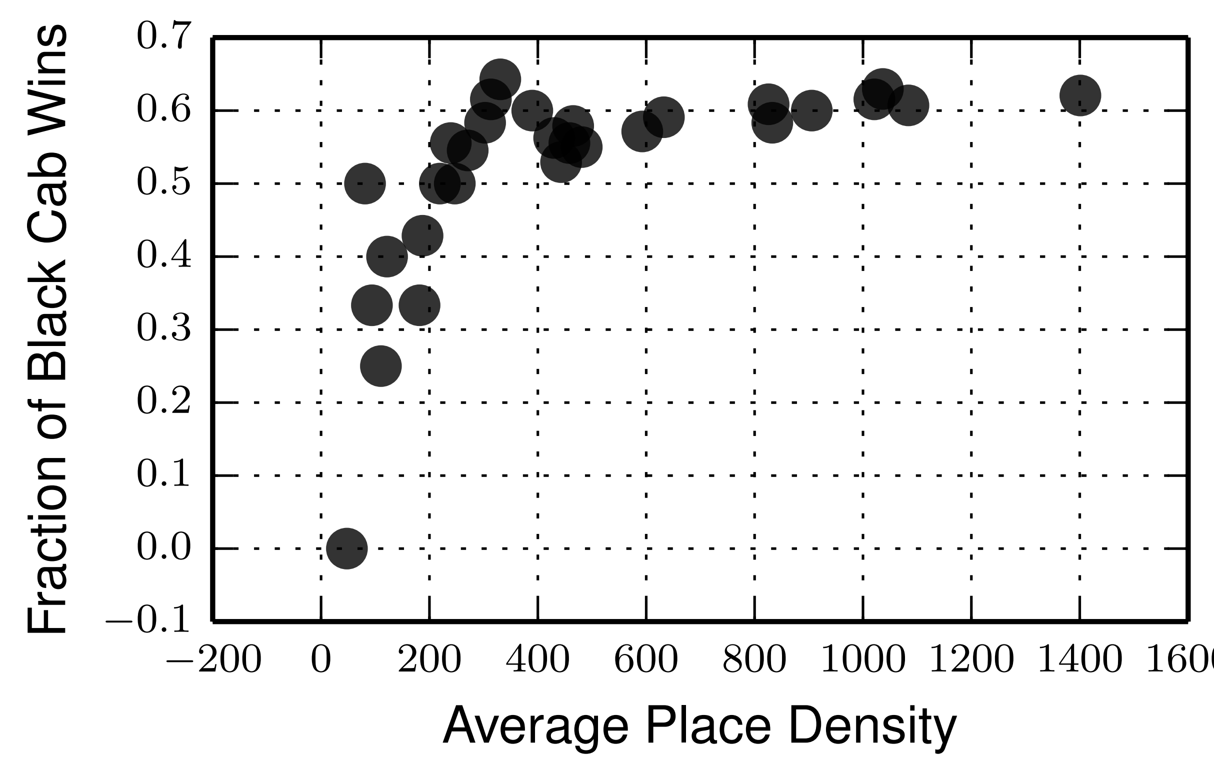

12:00
2016/02/22
Monday

Urban Complexity & Performance

- Uber faster
- Black Cab faster
- Journey duration tie



$$Trip\_Density = \frac{1}{|T|} \frac{\sum_{i=1}^{|T|} P(x = lng_i, y = lat_i, r = 200m)}{\pi r^2}$$

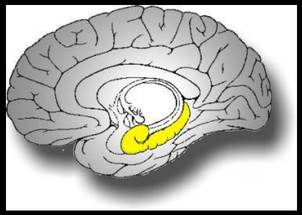


Drivers:Black Cab vs Yellow vs Uber

uses his (big) brain

Does not know where
is Brooklyn!

blindly follows the GPS



Maguire, Eleanor A., et al. "Navigation-related structural change in the hippocampi of taxi drivers." Proceedings of the National Academy of Sciences 97.8 (2000): 4398-4403.

Thanks!
Questions?

email: a.noulas@lancaster.ac.uk