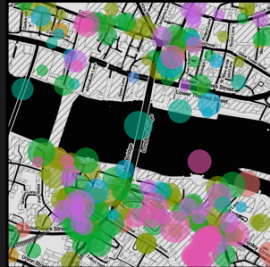
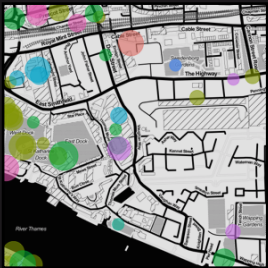


SOCIAL NETWORK ANALYSIS FOR LOCAL AND GLOBAL DEVELOPMENT

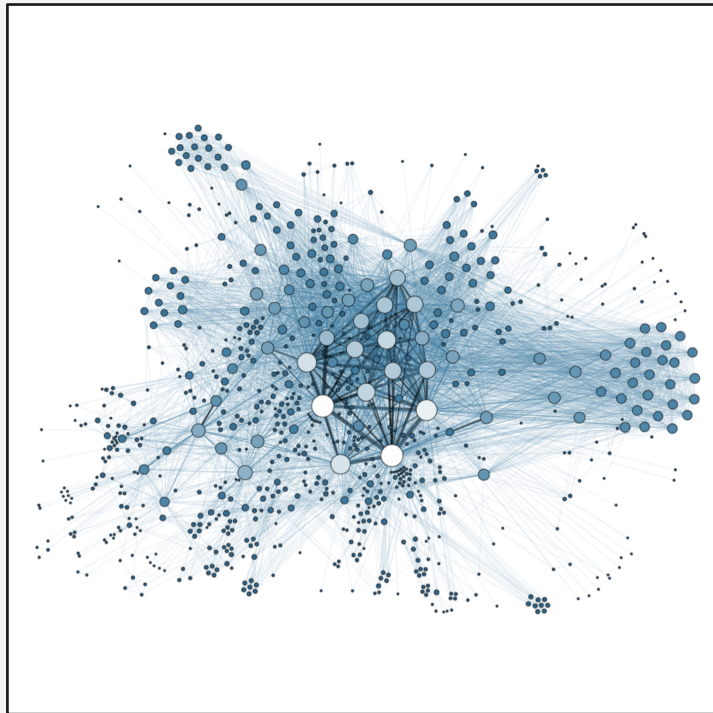


DESI HRISTOVA
DATA SCIENTIST
FUTURE CITIES CATAPULT
PHD COMPUTER LAB



CATAPULT
Future Cities

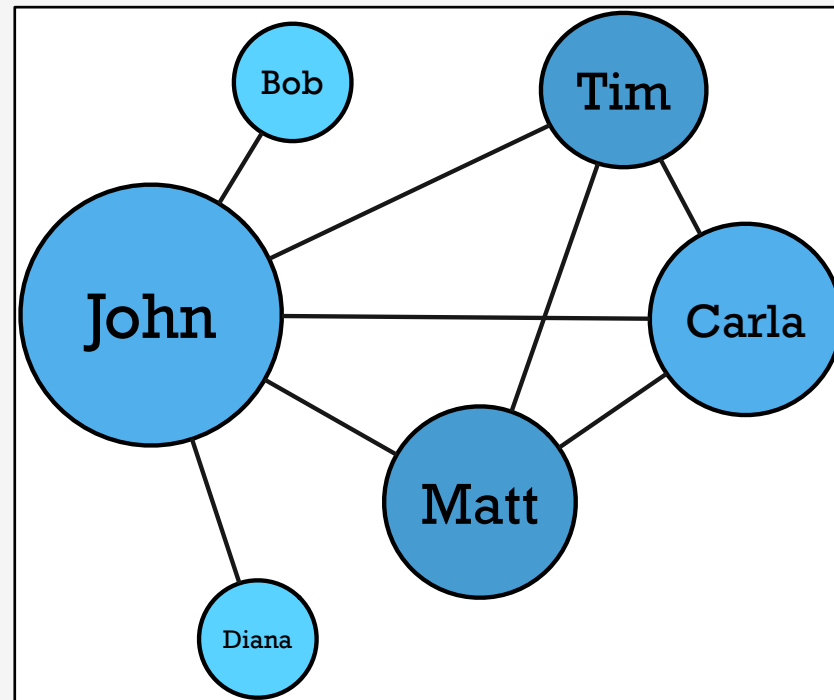
Outline



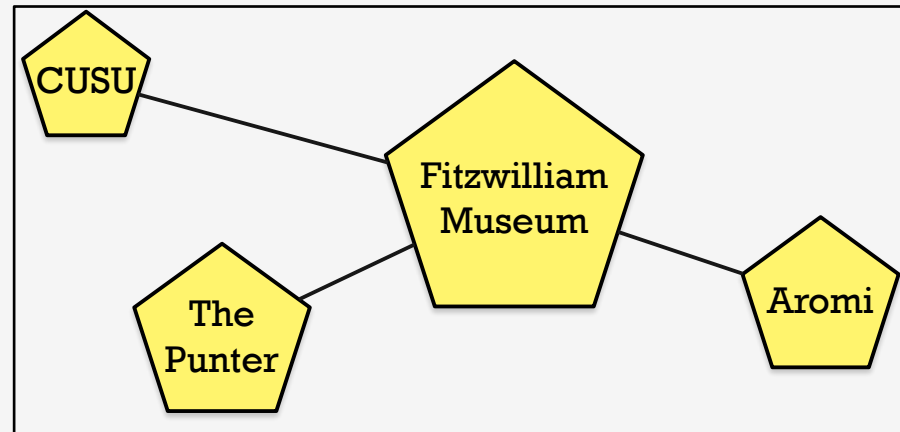
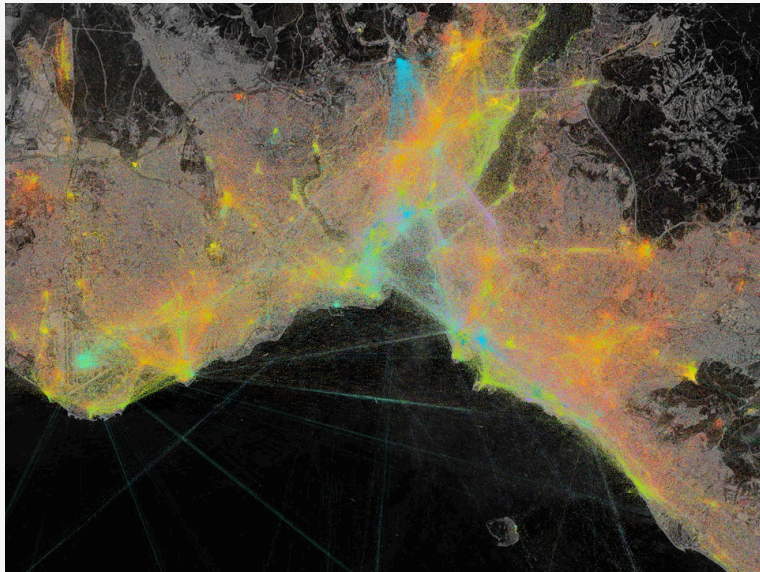
- Multilayer networks
- International development & network metrics
- Urban development & social networks
- Social, cultural & economic capital in cities

Social Networks

Social networks are usually
reduced to a single
interaction type



Place Networks



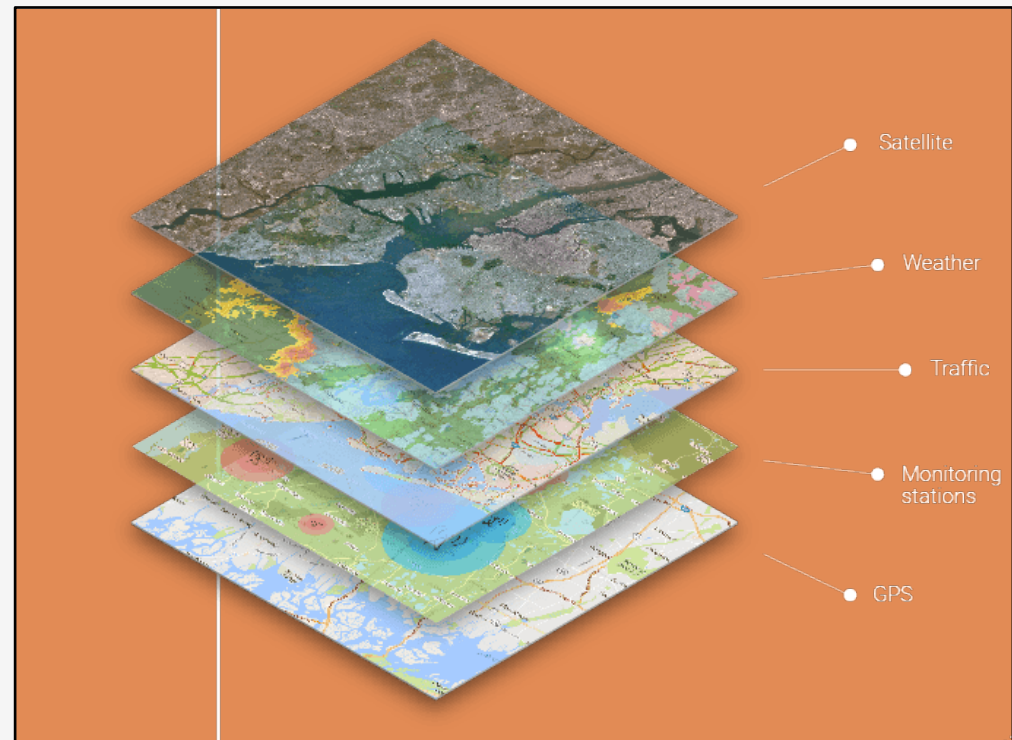
Place networks can be constructed from human mobility traces between locations.

Noulas, Anastasios, et al. "Topological properties and temporal dynamics of place networks in urban environments." *Proceedings of the 24th International Conference on World Wide Web*. ACM, 2015.

Real World

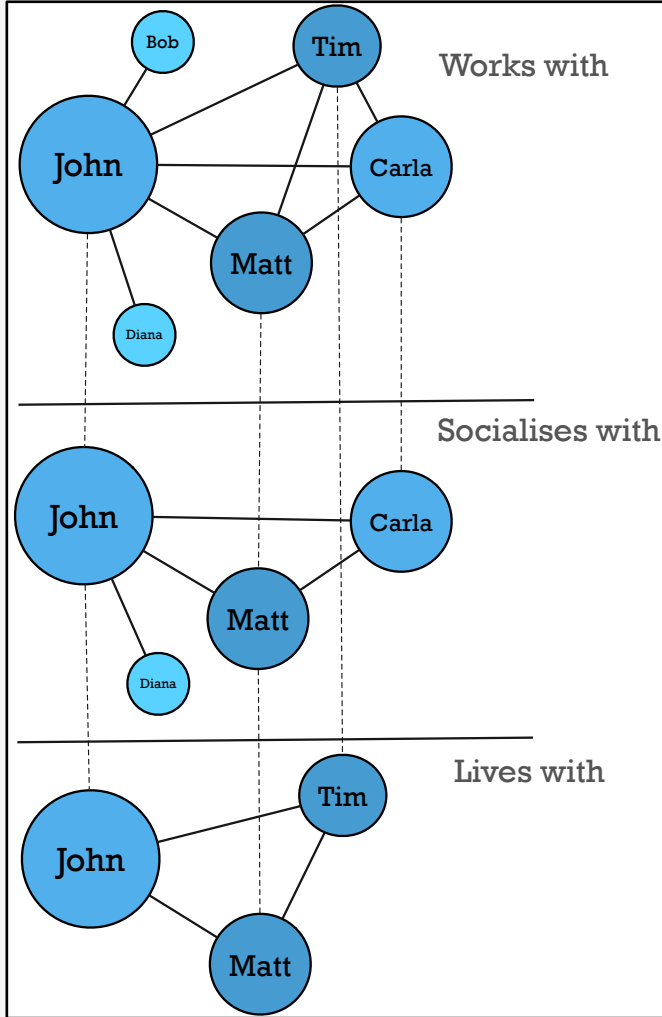
In reality, human relationships are multifaceted and places can be linked in more than one way.

We need augmented network models to capture this...



Multilayer Networks

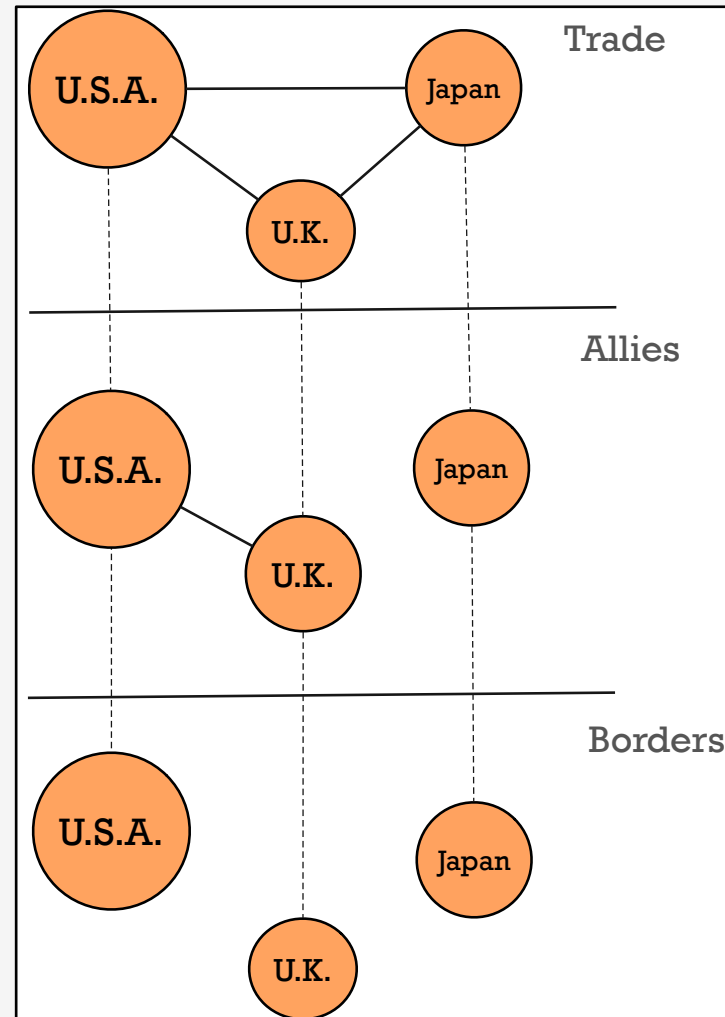
Capturing multiple heterogeneous interactions.



Kivelä, Mikko, et al. "Multilayer networks." *Journal of complex networks* 2.3 (2014): 203-271.

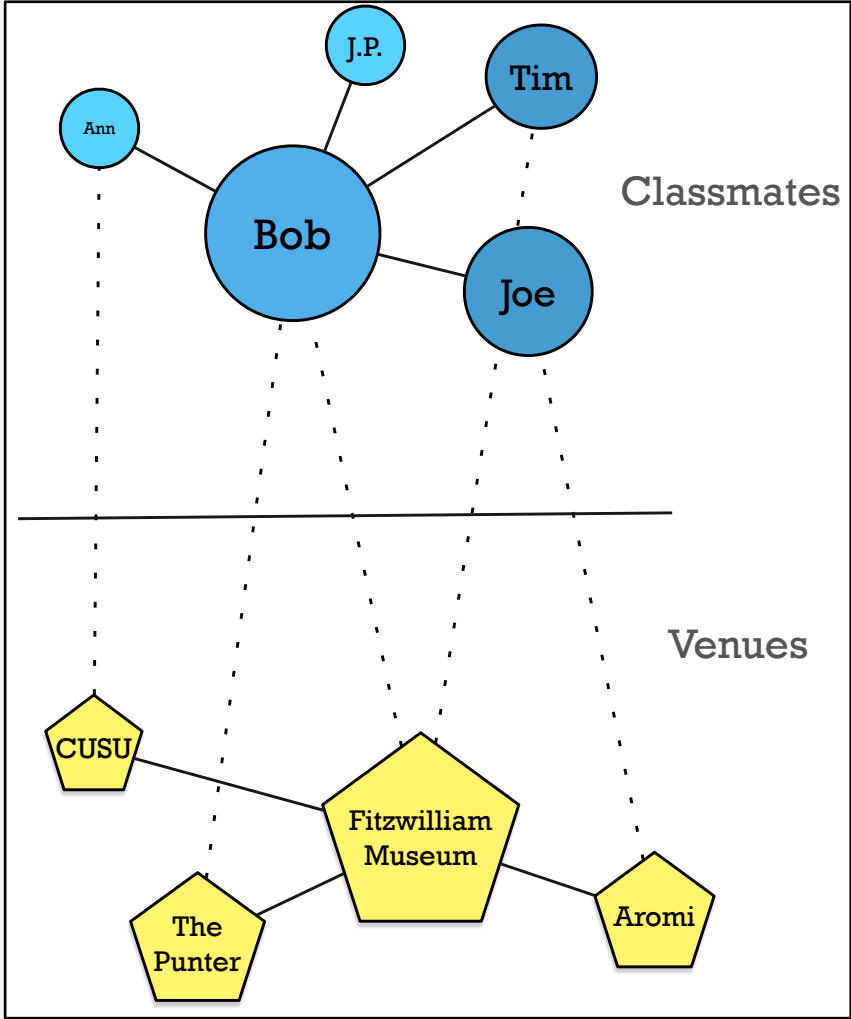
Multiplex Networks

"A multiplex network is a special type of multilayer network in which the only possible types of interlayer connections are ones in which a given node is connected to its counterpart nodes in the other layers."



Interconnected Networks

Interconnected network models can capture interactions between heterogeneous entities across layers and therefore be used to study cascading effects.



Buldyrev, Sergey V., et al. "Catastrophic cascade of failures in interdependent networks." *Nature* 464.7291 (2010): 1025-1028.

*Can we use
network data as
a proxy for
international
development?*



Challenges: measurement is costly, coverage is incomplete, time window is 5-10 years for some indicators.



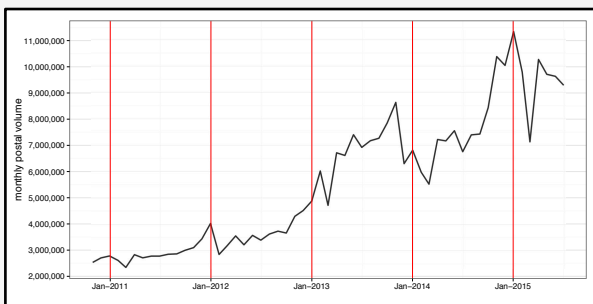
Socioeconomic Indicators



- Gross Domestic Product (GDP)
 - Gini Index
 - Economic Complexity Index
 - CO₂ Emissions
 - Life Expectancy
 - Poverty Rate
 - Literacy Rate
 - Human Development Index
 - Mobile phone penetration rate
 - Internet penetration rate
 - Corruption Perception Index
 - World Happiness Index
- Economic development
- Environmental development
- Human development
- Technological development
- Subjective wellbeing



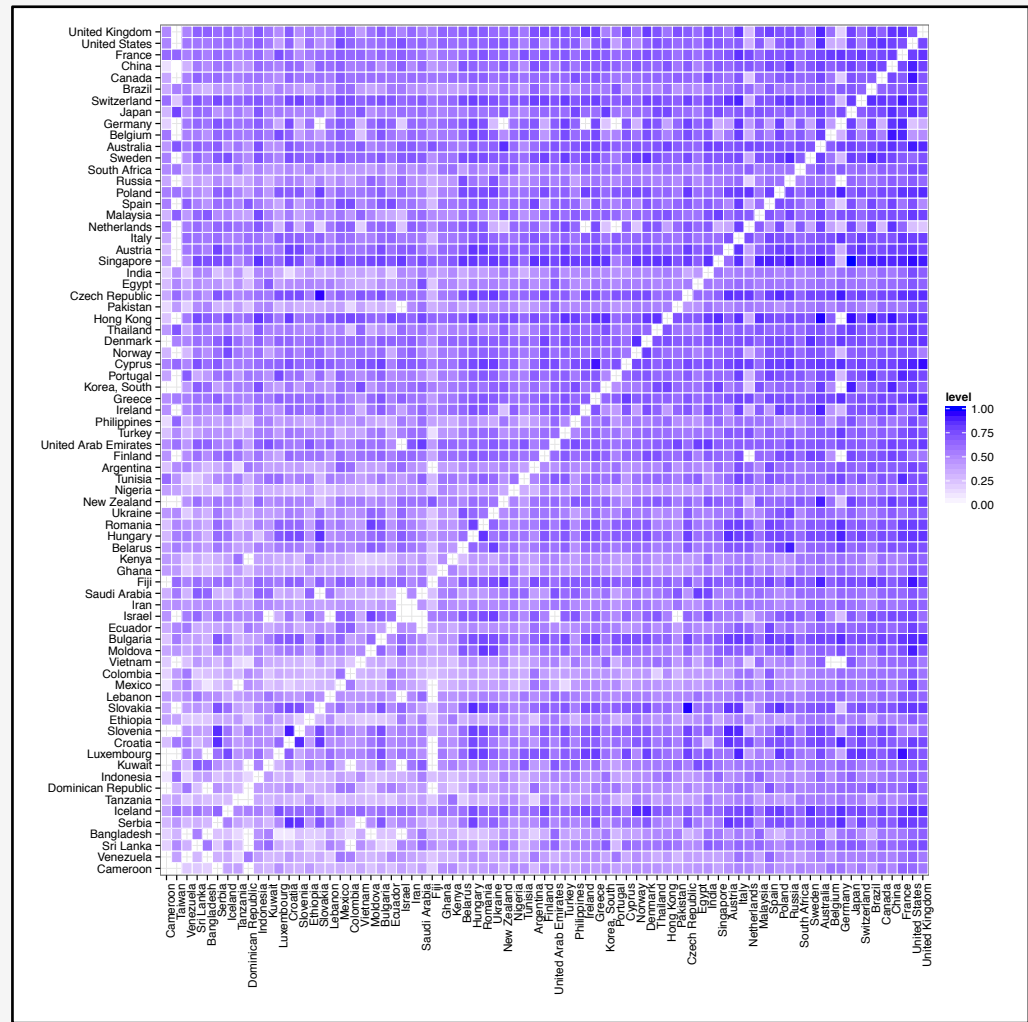
International Postal Network



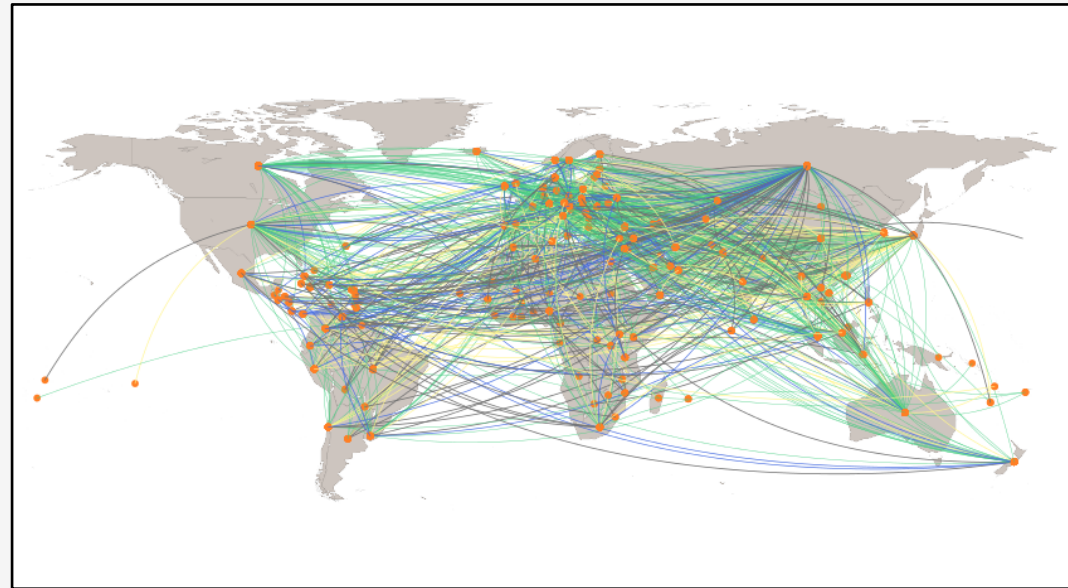
The oldest social media network is a good candidate for a proxy Measurements because of its **sustainability** (growing), **cost** (already being collected anyway), **coverage** (global) and **granularity** (daily)

Postal Network Degree

Countries can be ranked by the number of countries they exchange post with and its volume.



Global Interaction Network

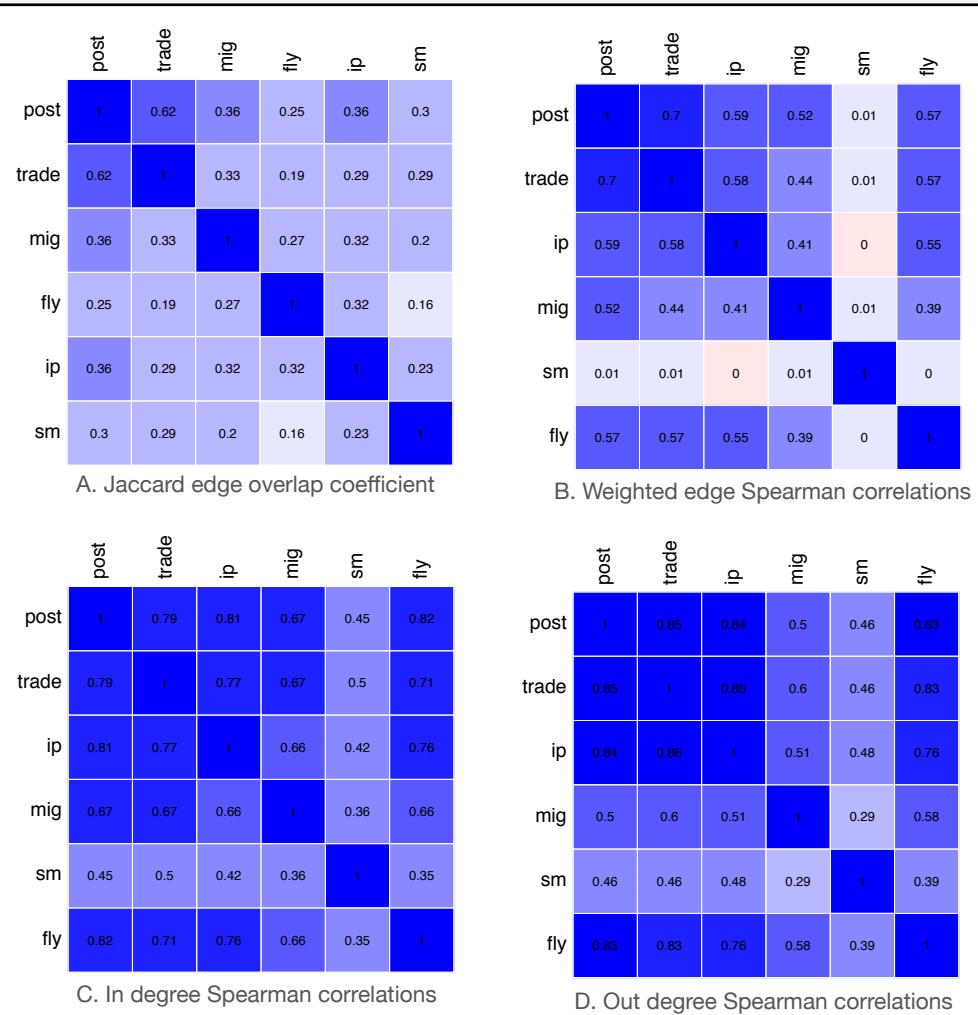


In reality, countries have many other types of interactions for which we have data...

network	weight	years	V	E	$\langle k \rangle$	assort	d	cc
Post	postal items	2010 – 15	201	22,280	110.85	-0.26	0.55	0.79
Trade	export value	2007 – 12	228	30,235	132.6	-0.39	0.58	0.84
Migration	migrants	2005 – 10	193	11,431	59.22	-0.33	0.31	0.68
Flights	flights	2010 – 15	223	6,425	28.81	-0.1	0.13	0.49
IP	IPs	2007 – 11	225	9,717	43.19	-0.42	0.19	0.6
SM	density	2009	147	10,667	145.13	-0.02	0.98	0.99

Comparing networks

- ★ A. Low edge overlap means that we have non-redundant network structure across layers.
- ★ B. High edge correlation indicates that countries exhibit multiplex interaction.
- ★ C+D. High in/out degree correlation means that countries hold a similarly central or peripheral rank across layers.



Lets build a multiplex



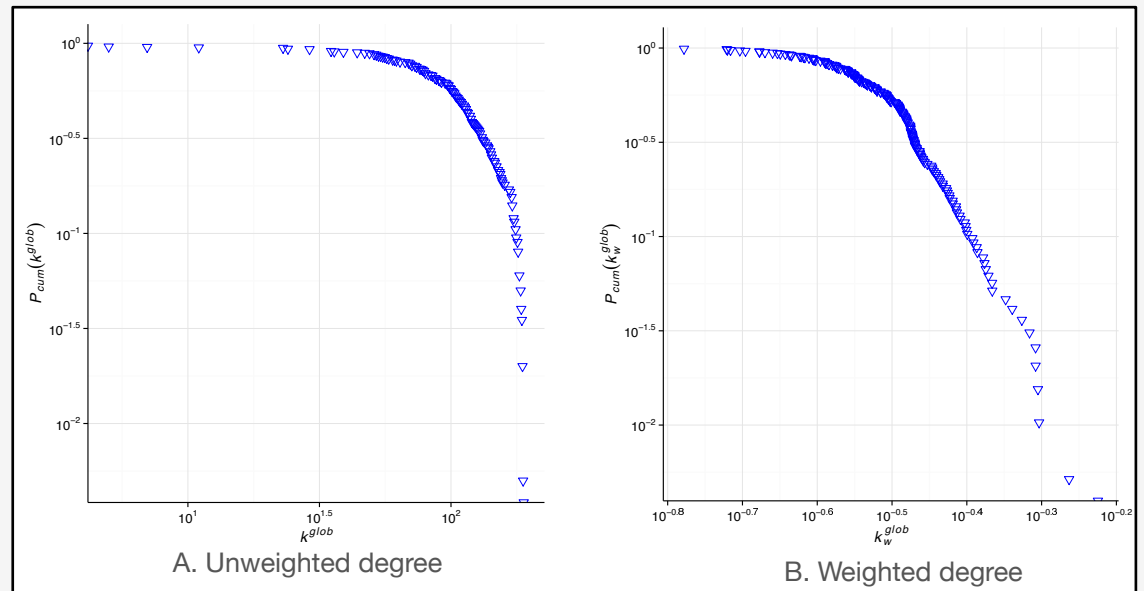
A multiplex network can be defined as a simple collection of graphs.

$$\mathcal{M} = \{G^1(V^1, E^1), \dots, G^\alpha(V^\alpha, E^\alpha), \dots, G^M(V^m, E^m)\}$$

Global Network Degree

The multilayer neighbourhood can be derived from the union of single-layer neighbourhoods.

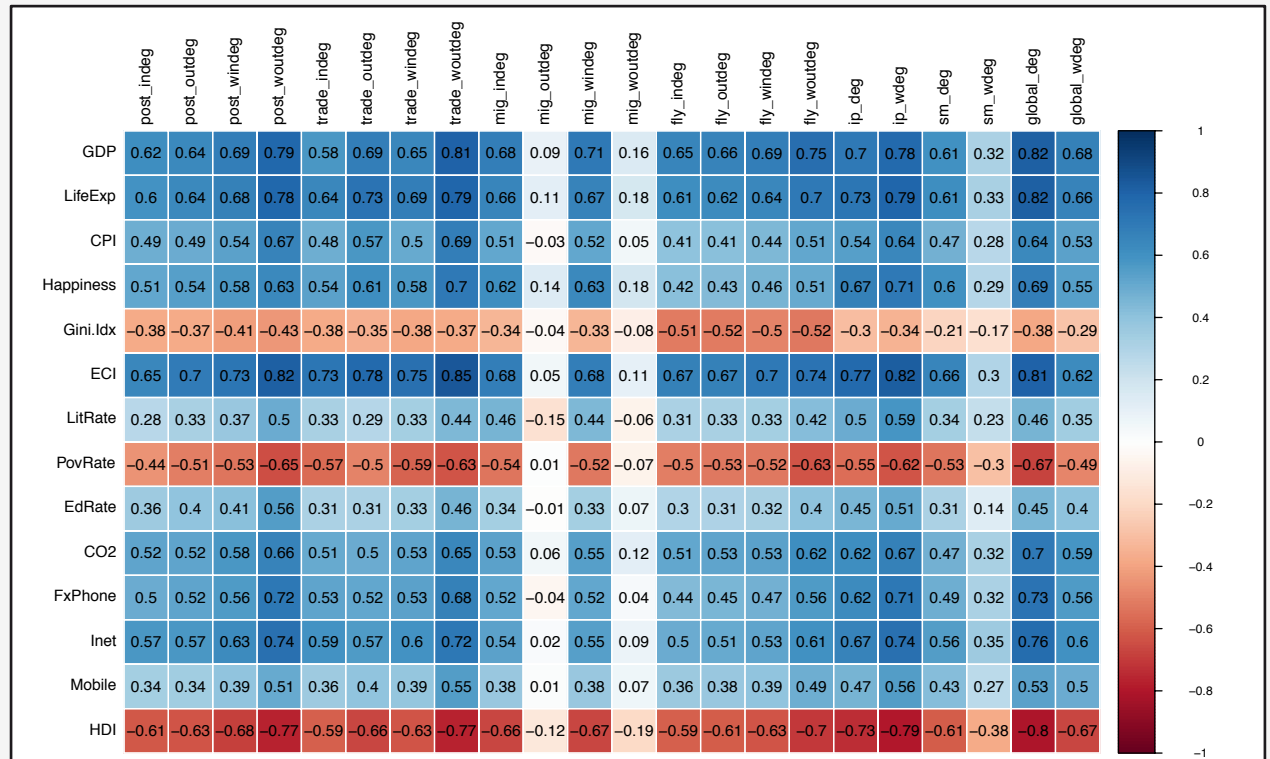
$$N_{\mathcal{M}}(i) = \{N_{\alpha}(i) \cup N_{\beta}(i) \dots \cup N_m(i)\}$$



$$k^{glob}(i) = |N_{\mathcal{M}}(i)|$$

$$k_w^{glob}(i) = \frac{\sum_{j \in N_{\mathcal{M}}(i)} \sum_{G \in \mathcal{M}} e_{ji}}{n * m}$$

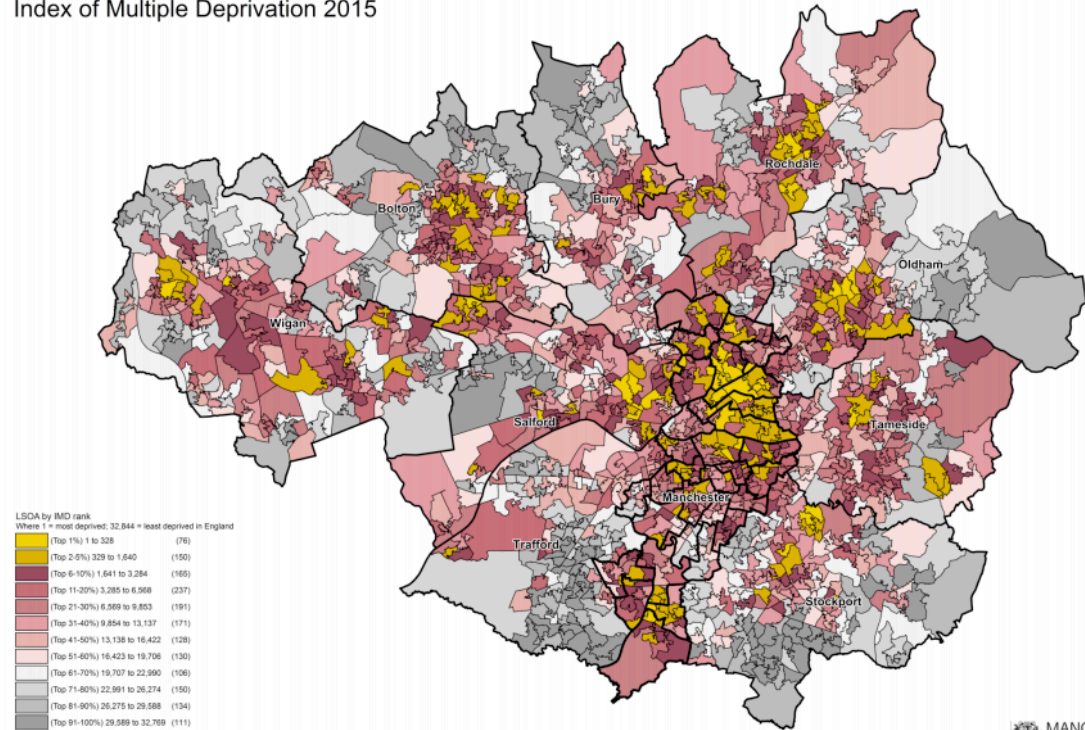
Can we use network metrics to approximate socioeconomic indicators?



Hristova, Desislava, et al. "The international postal network and other global flows as proxies for national wellbeing." *PloS one* 11.6 (2016).

Can we model urban development with multilayer networks?

Greater Manchester
Index of Multiple Deprivation 2015

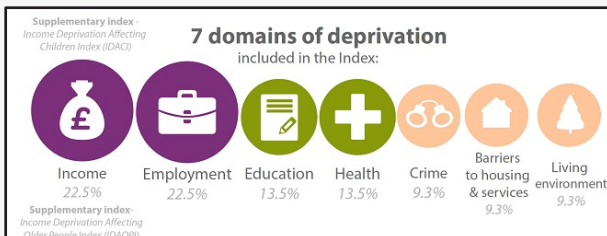


Source: DCLG, Crown copyright

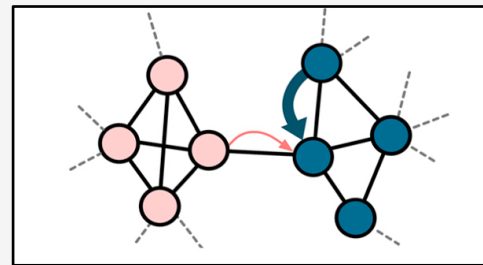
MANCHESTER CITY COUNCIL

Crown copyright. All rights reserved. Manchester City Council 100019568 (2015). Analysis by Public Intelligence, PRI

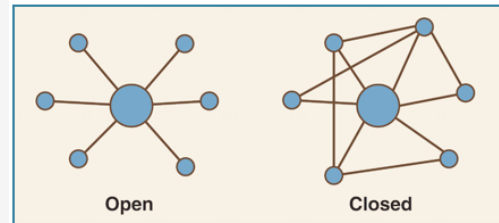
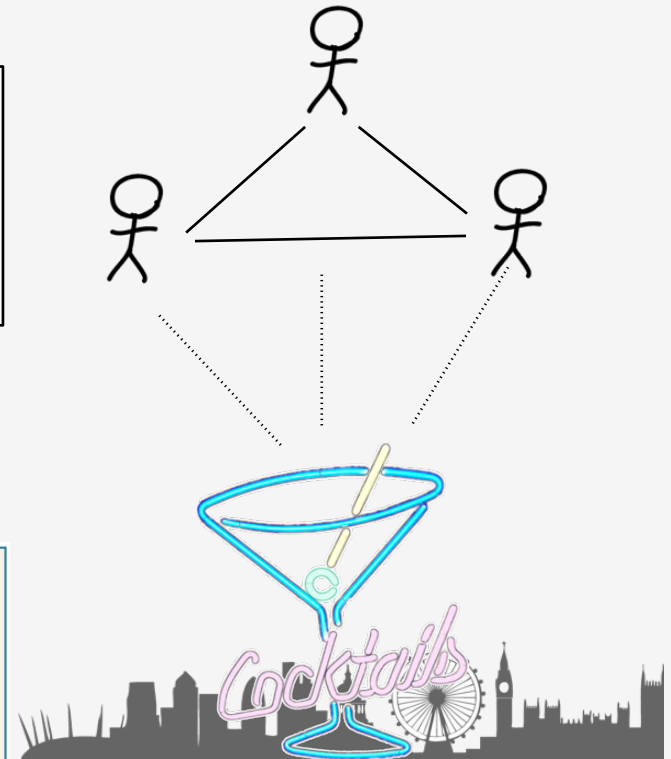
Urban development is the process of **social, cultural, economic** and **physical** development of cities, along with the underlying causes of this process...



Social networks and social capital



Bridging vs Bonding



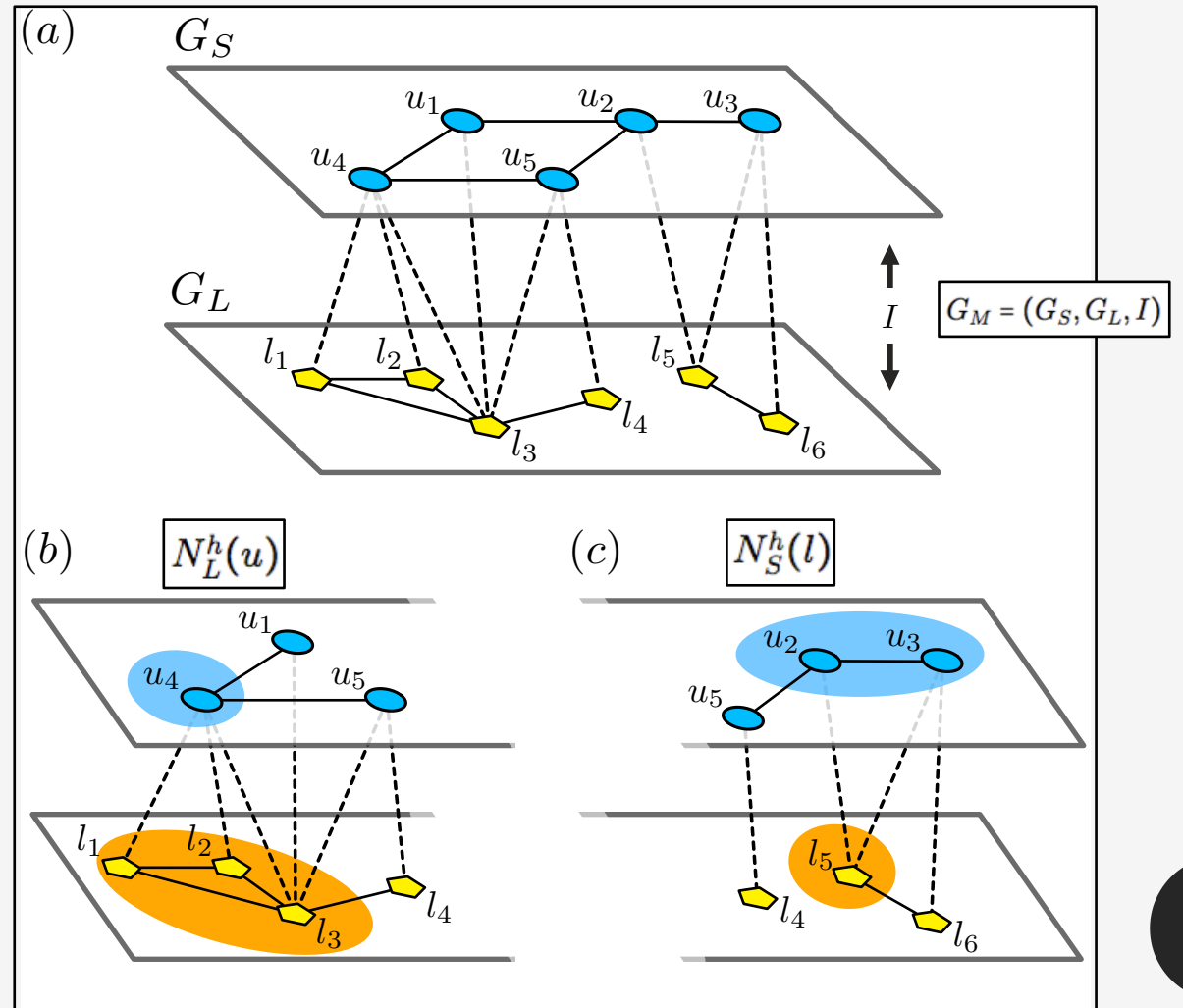
Social Brokerage

Geo-Social Brokerage

Open network structures (aka "structural holes") are associated with having an advantageous position in the social network but what about the place network?

Urban Geo-Social Network Model

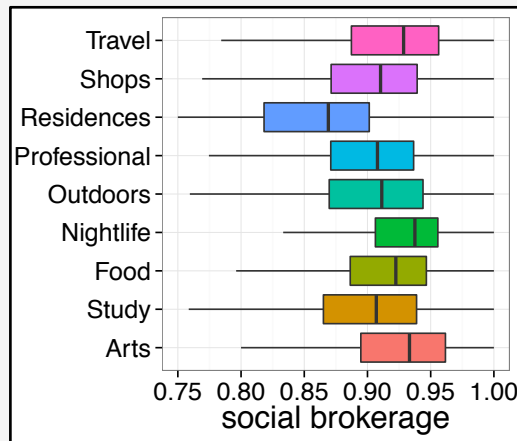
We can model urban areas as a multilayer interconnected network of people and the places they visit.



Deriving the social brokerage of places from Foursquare



$$B(l) = |N_S^h(l)| - \frac{\sum_{u,v \in N_S^h(l)} e_{u,v}}{|N_S^h(l)|}$$

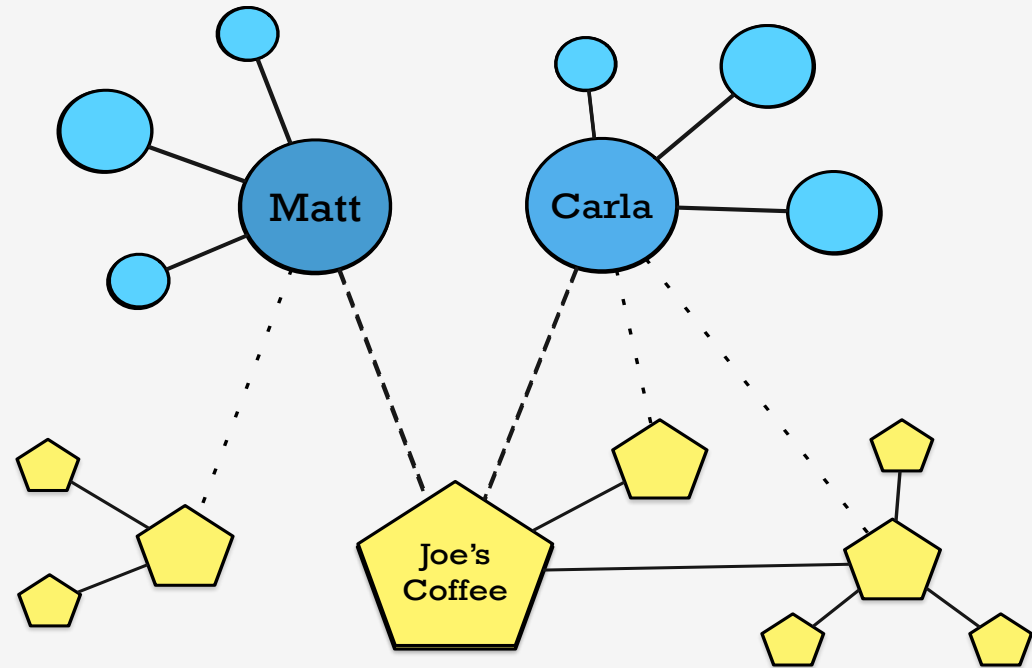
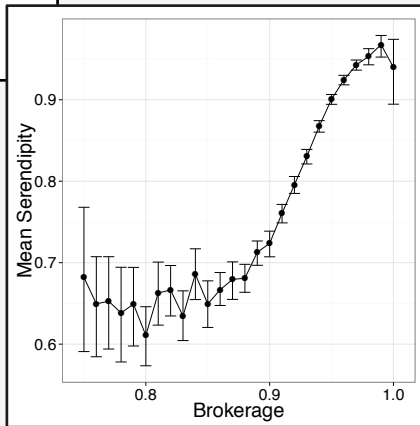


Category	Bridging role	Bonding role
Travel	Motel	B&B
Shops	Mall	Laundry
Residences	Apartment Building	Home
Professional	Courthouse	Emergency Room
Outdoors	Bridge	Vineyard
Nightlife	Gay Bar	Strip Club
Food	Dumplings	Fried Chicken
Study	Bookstore	Classroom
Arts	Art Museum	Football

Serendipity

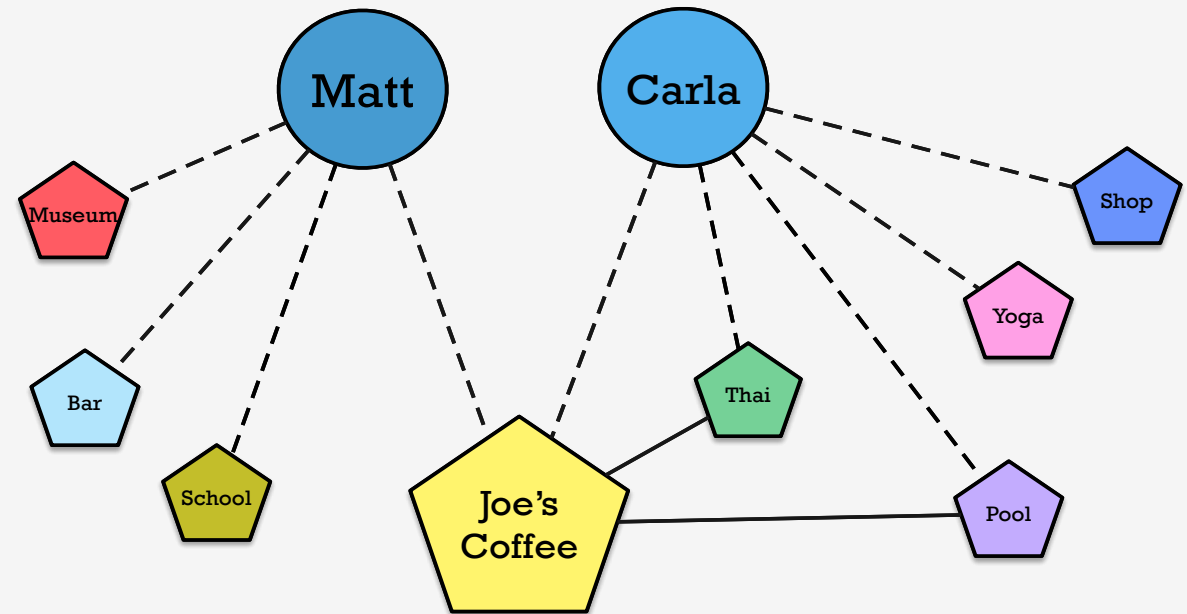
$$D(l) = 1 - \frac{\sum_{u \in N_S^h(l)} p_i^t(u)}{|N_S^h(l)|}$$

$$p_i^t(u) = \frac{\sum_{v \in N_L^h(u) < t} w_{v,l}}{\sum_{v \in N_L^h(l) < t} w_{v,l}}$$

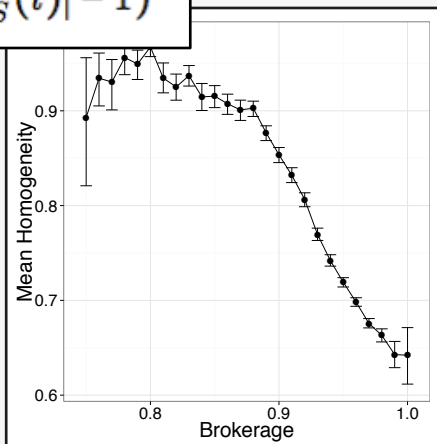


Places with high serendipity values are those which have the potential to induce chance encounters among its visitors.

Visitor Diversity



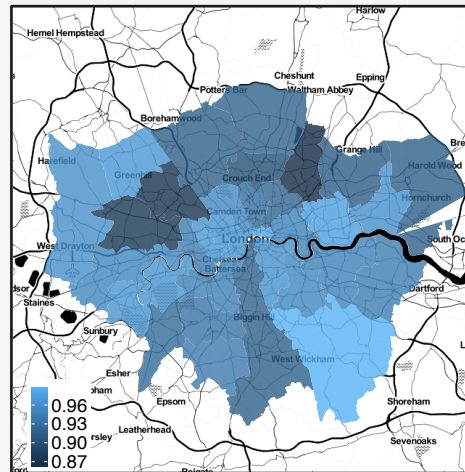
$$S(l) = \frac{\sum_{u,v \in N_S^h(l)} \text{sim}(u,v)}{|N_S^h(l)|(|N_S^h(l)| - 1)}$$



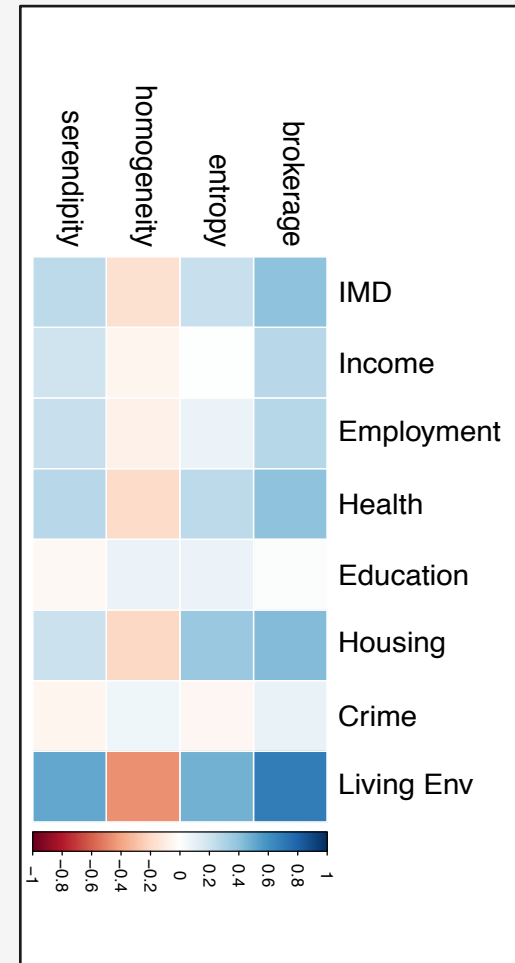
How diverse are visitors based on their category preferences?

Urban development & social place diversity metrics

Geographical distribution of brokerage values.

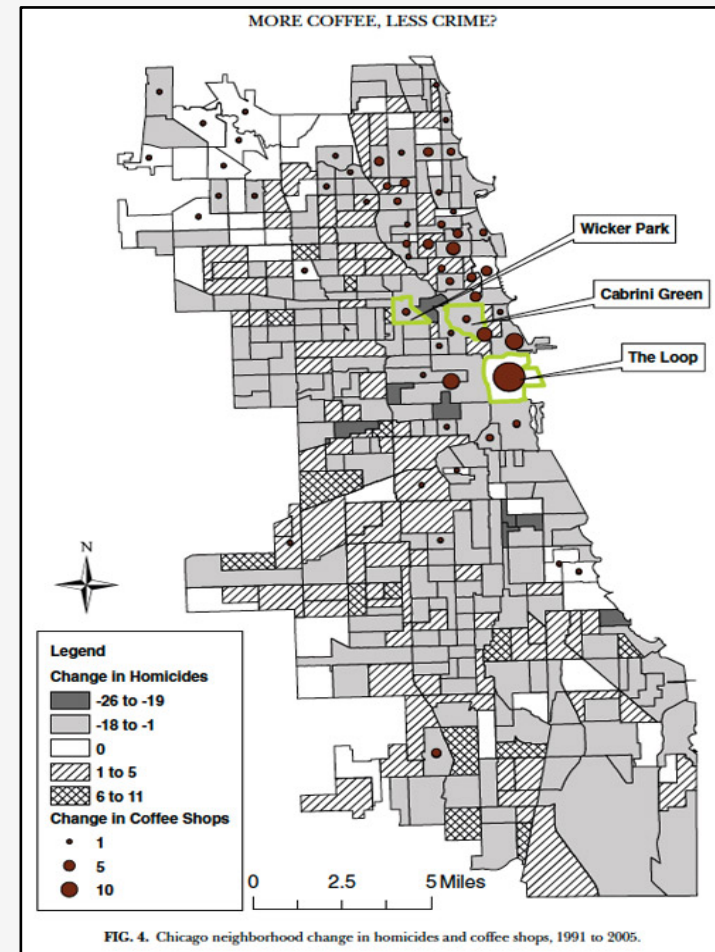


Neighbourhood brokerage correlates with deprivation. Higher brokerage means higher deprivation? Hmm..



Symptoms of Gentrification

- "Organic" change
- Falling crime rates
- Coffee shops (amenities)
- Lowering deprivation
- Employment industry of residents
- Increase in housing prices
- Population demographics
- Economic growth
- Social media use?



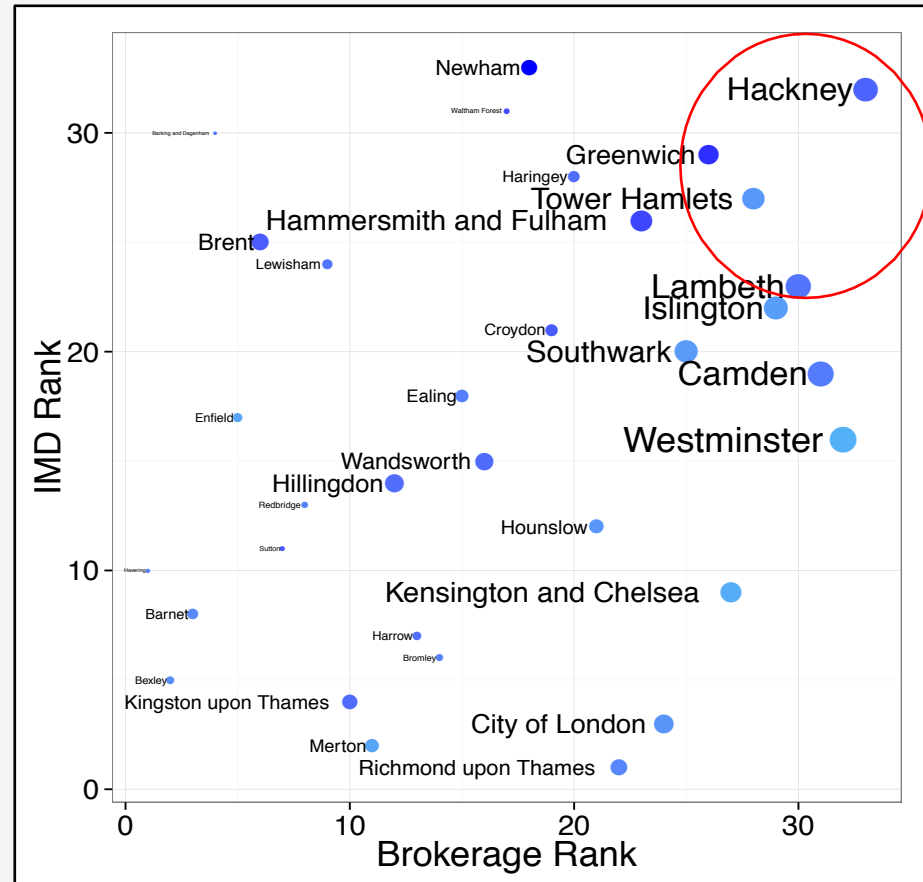
Papachristos, Andrew V., et al. "More coffee, less crime? The relationship between gentrification and neighborhood crime rates in Chicago, 1991 to 2005." *City & Community* 10.3 (2011): 215-240.

Gentrification: Diversity & Deprivation?

High deprivation and high diversity
in 2010 signal gentrification in 2015.

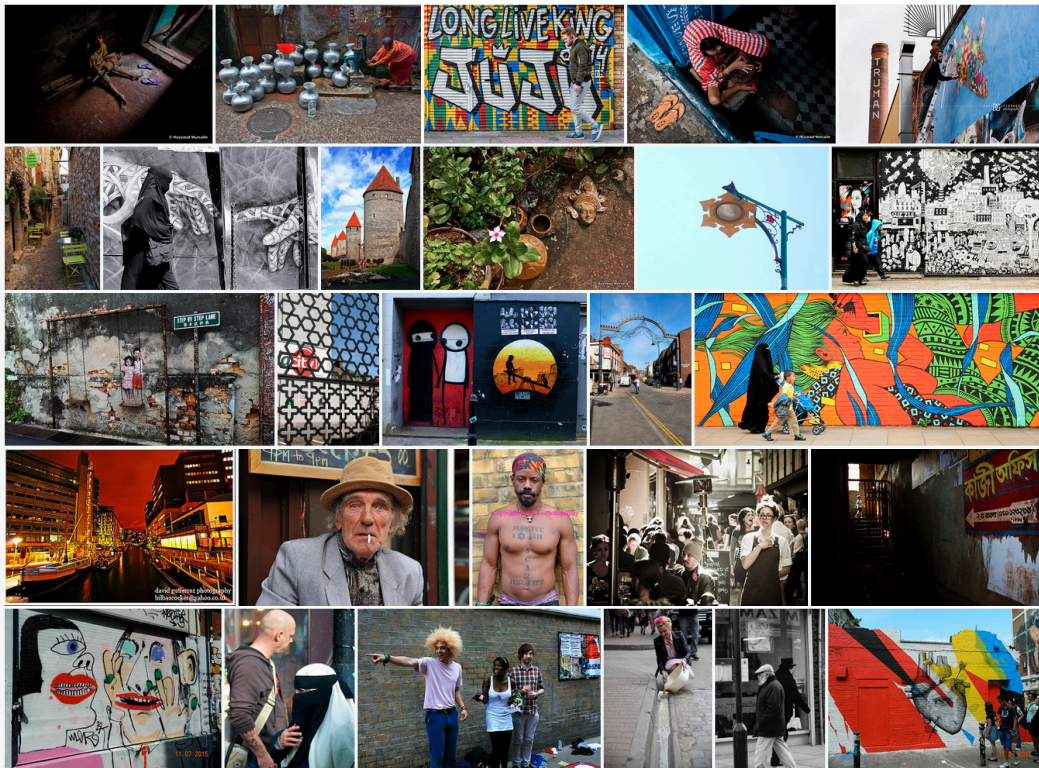


The rank of IMD vs Brokerage for 2010.
Node size indicates the percent change in IMD over the next 5 years.

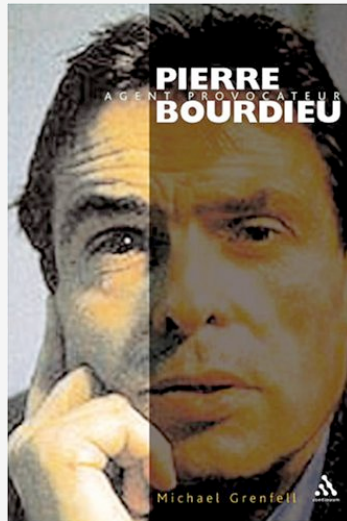


Hristova, Desislava, et al. "Measuring urban social diversity using interconnected geo-social networks."
Proceedings of the 25th International Conference on World Wide Web, 2016.

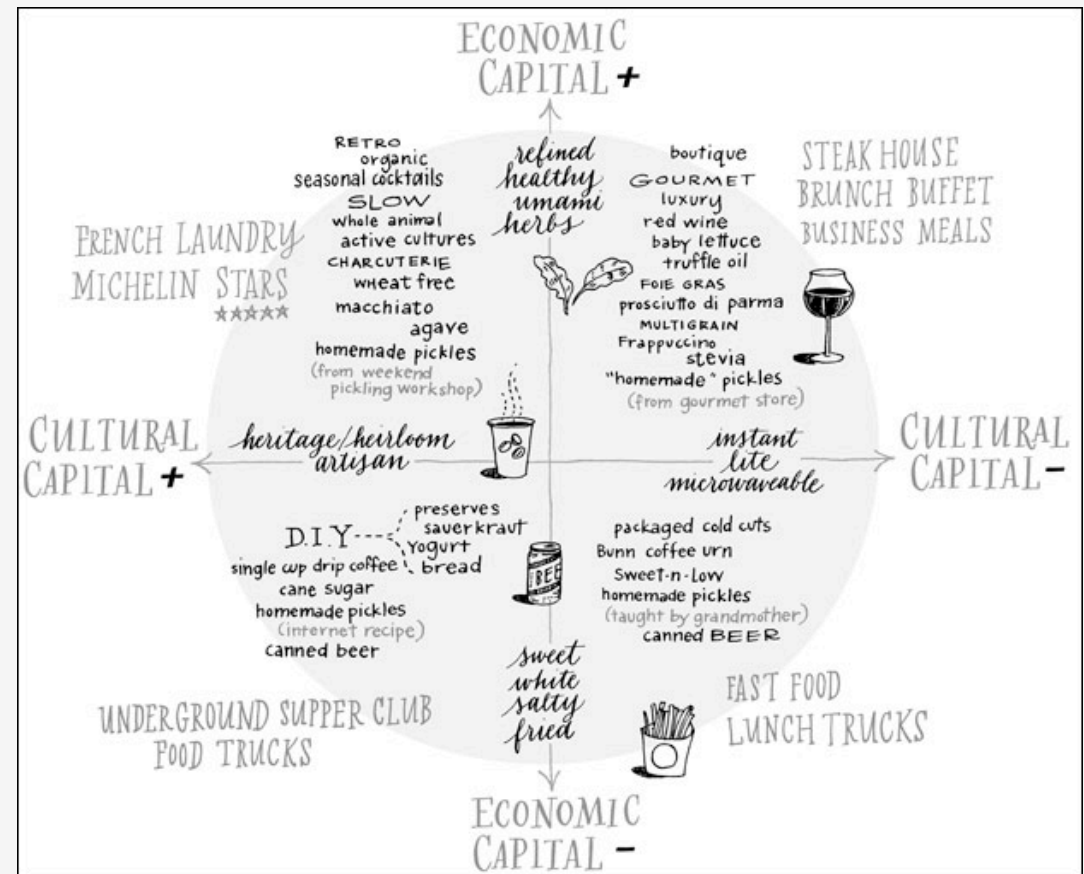
*Culture & Regeneration:
Can we use
Flickr photos to
predict
gentrification?*



Cultural Capital



Cultural and economic capital can both be acquired and one can be converted into the other.

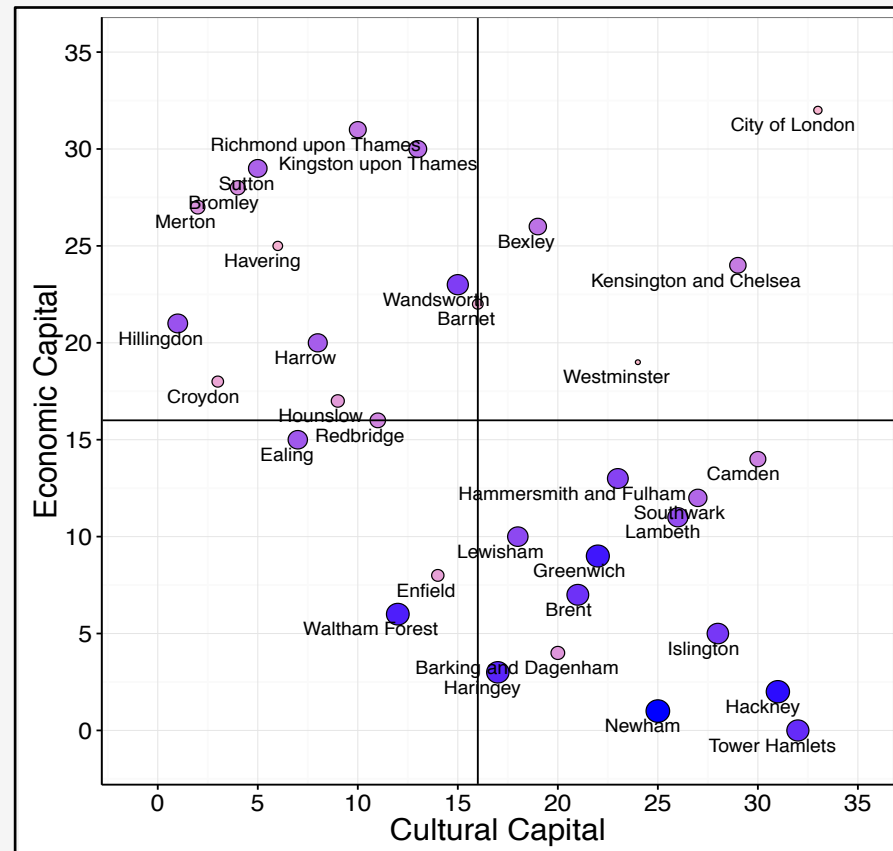


Cultural vs Economic Capital

Recreating Bordieu's cultural vs economic capital plots for neighbourhoods shows that high cultural capital but low economic capital in 2010 is indicative of gentrifying neighbourhoods in 2015.

$$capital(l) = \frac{capital_l - \mu(capital)}{\sigma(capital)}$$

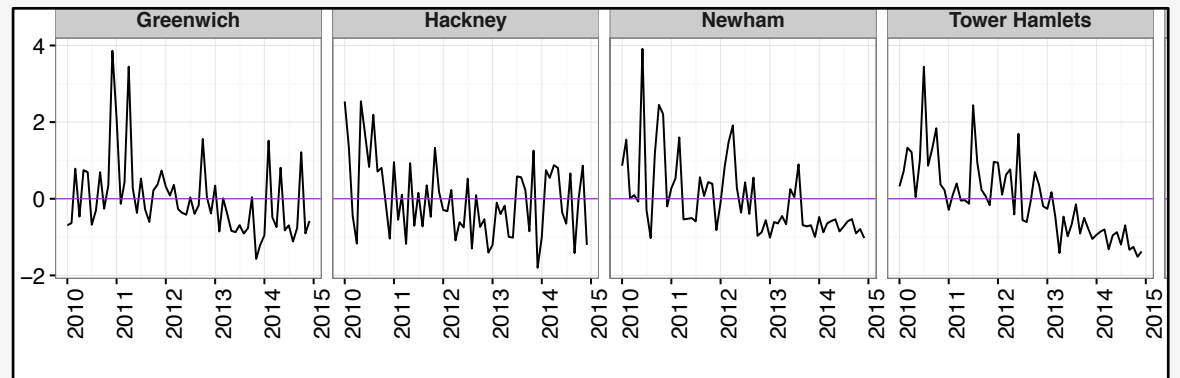
Economic capital (income) rank vs Cultural capital rank (Flickr) for neighbourhoods in 2010. Color and size indicate the percent change in IMD.



Predicting up-and-coming neighbourhoods

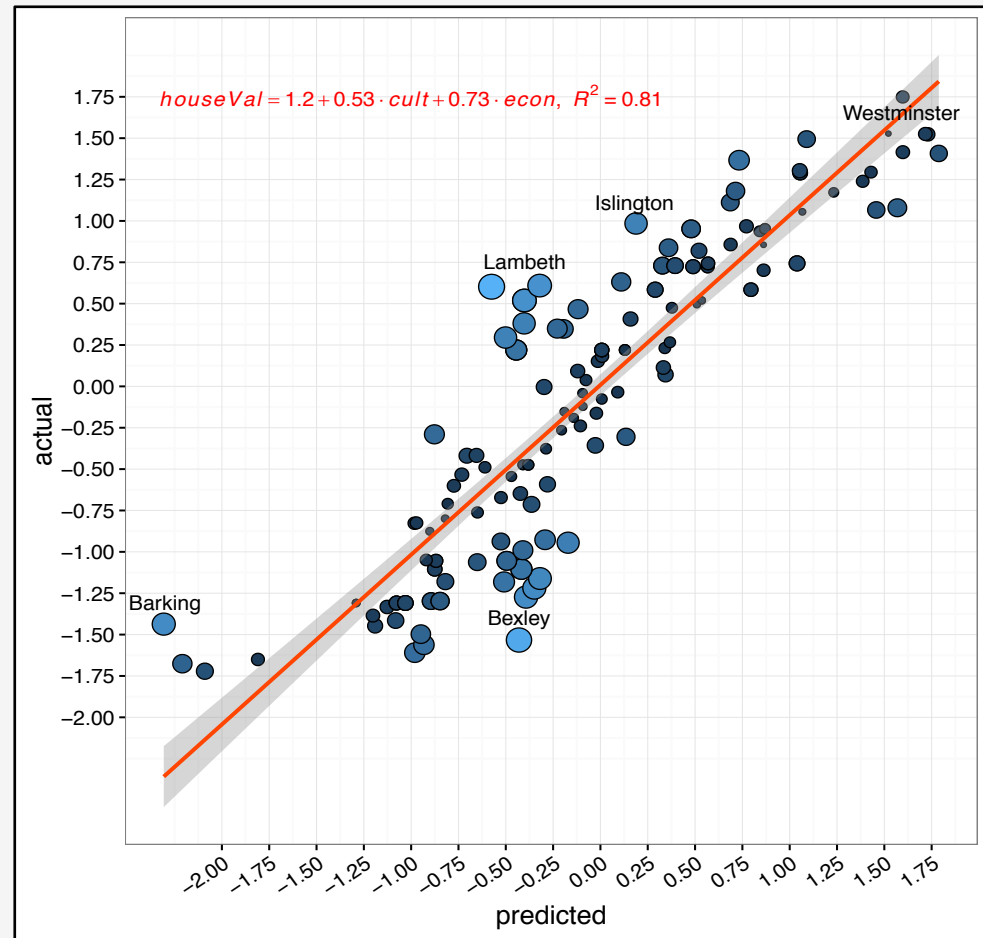


While IMD data is available only once every 5 years, social media data is real-time. This way, we can track changes in urban social networks and anticipate both positive and negative aspects of gentrification.



Housing price prediction

One of the most severe effects of gentrification is sudden housing price rises. Using only public income data and Flickr photos related to cultural activities, we can predict a great deal of these price fluctuations.



Summary & Questions

- Multilayer network models
- Multiple heterogeneous data sources
- Inspiration from urbanism and sociology
- Approximating critical socioeconomic indicators
- Quantifying urban processes of development



desislava.hristova@cl.cam.ac.uk



@desi_hristova