The Age of the Smart City
Big Data, Mobility, and Transport

Michael Batty

m.batty@ucl.ac.uk
@jmichaelbatty

9 March 2019

http://www.spatialcomplexity.info/

Outline

- A Long Preamble about Digital Technologies
- Space and Time and Cities
- *The Smart City: Using Computers to Explore Cities Which Are Increasingly Composed of Computers*
- A Short History of Big Data: How Big is Big?
- Mobility, Transit, & Real-Time Streaming: The Oyster Card Data Set
- Learning about Mobility from the Data *Variabilities – Disruptions – Dynamics of Demand*
- Related Real-Time Data: Bikes, Social Media
- What Has All This to Do with LED?

A Long Preamble about Digital Technologies

- How did we move and how far did we travel in the past
- We have always been able to walk about 3 miles a day which takes about an hour
- In nomadic times, man spent 3-4 hours a day hunting and foraging
- But around about 10,000 BCE, man moved to settled agriculture and the amount of time reduced to about one hours travel per day – at 3 miles per hour
- These limits lasted until the industrial revolution
- They did not change with the appearance of cities in 5000 BCE or with the Roman and Chinese Empires

- In fact, the one hour limit still survives but what changed about 200 years ago was the speed at which we could travel
- In other words technology – the invention of steam power – railways – so in the early 19th century we began to travel much greater distances in one hour – 30-40 mph. New technologies were all about speed
- And as new technologies were invented – the car, the plane, the mag lev train, we were able to travel ever further in one hour. In fact the one hour limit in the biggest cities has been gradually creeping up
- All this means that space is annihilated by time, as Marx amongst others spelt out so clearly in the middle of the 19th century. And cities began to grow, for the first time, they grew broke the 1 million limit

Space and Time and Cities

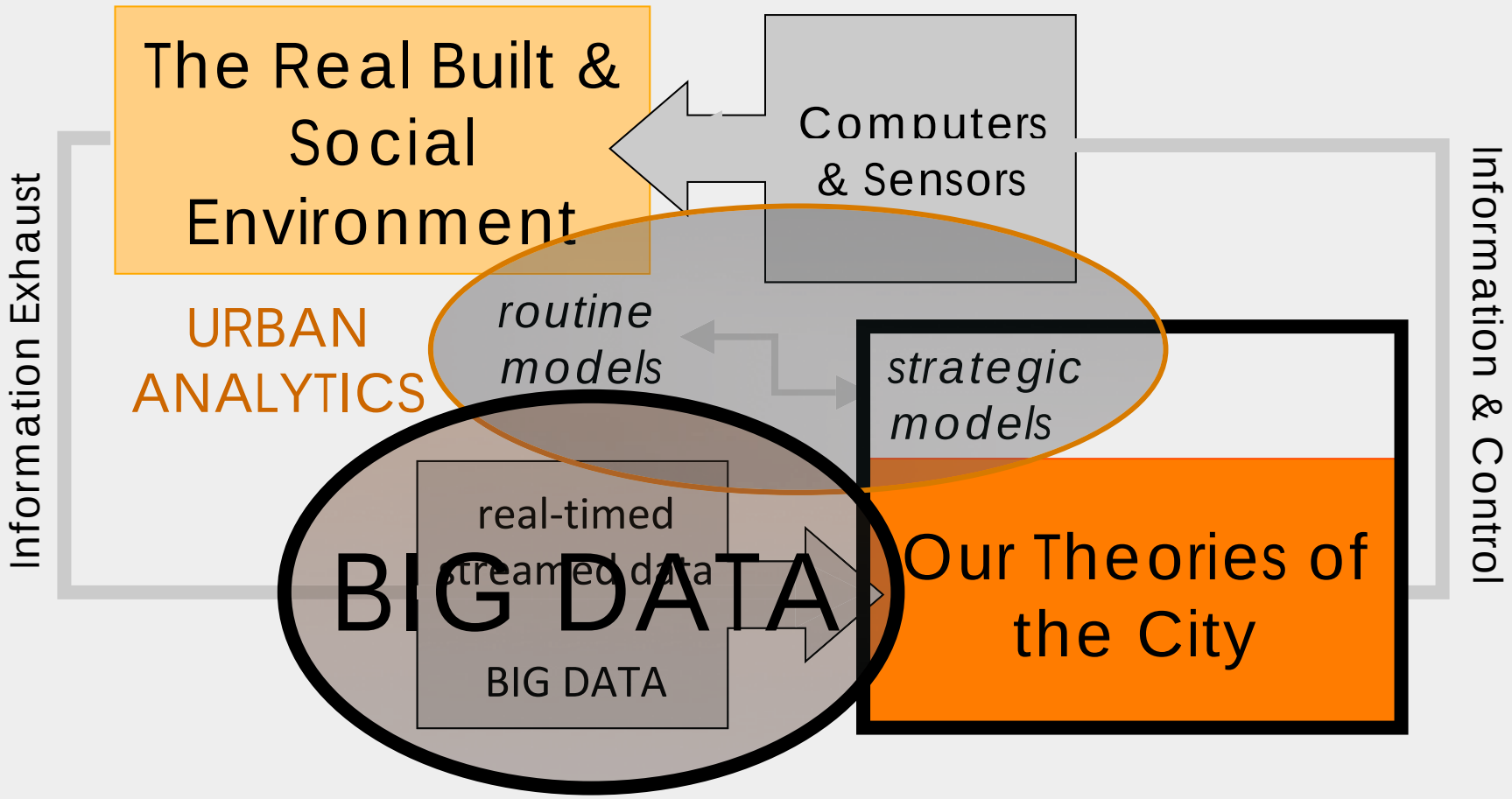
- Now about 75 years ago, the digital computer was invented – the universal machine as it was called
- No one ever believed it was universal, even Turing who said so, never believed it
- But as the 20th century moved on, computers got ever smaller
- And communications began to converge with them
- To all intents and purposes we can now shift information at the speed of light and as we travel usually to move information, not just for the hell of it, all the limits on how far we can move information are up for grabs
- We can communicate more cheaply & effortlessly

- Essentially as the economy becomes more and more about information and as industries based on energy get ever more automated and underpinned by information, increasingly we can act on things at a distance without worrying about travelling
- In terms of our hour a day, if we can move information at zero cost and immediately the amount of ‘travel’ become infinite
- In other words, ever since mankind emerged as a primate, there has always been a cost to movement and this has always dictated space. This constraints is not longer there
- The are a million implications – to city size, to the very definition of what a city is, but the implications for development and for work are profound.

The Smart City: Using Computers to Explore Cities Which Are Increasingly Composed of Computers

- OK, so enter the smart city.
- Now there is another issue: as information comes to underpin our economy and social life, computers are being embedded into the very fabric of our cities
- We originally used computers to study cities composed of materials but increasingly we are using computers to study cities composed of information which means they are composed of computers
- This is a strange recursion and it provides a basis for me to begin to look at some examples of how IT is changing everything

A framework to think about the smart city – computers change the city and change the way we study it

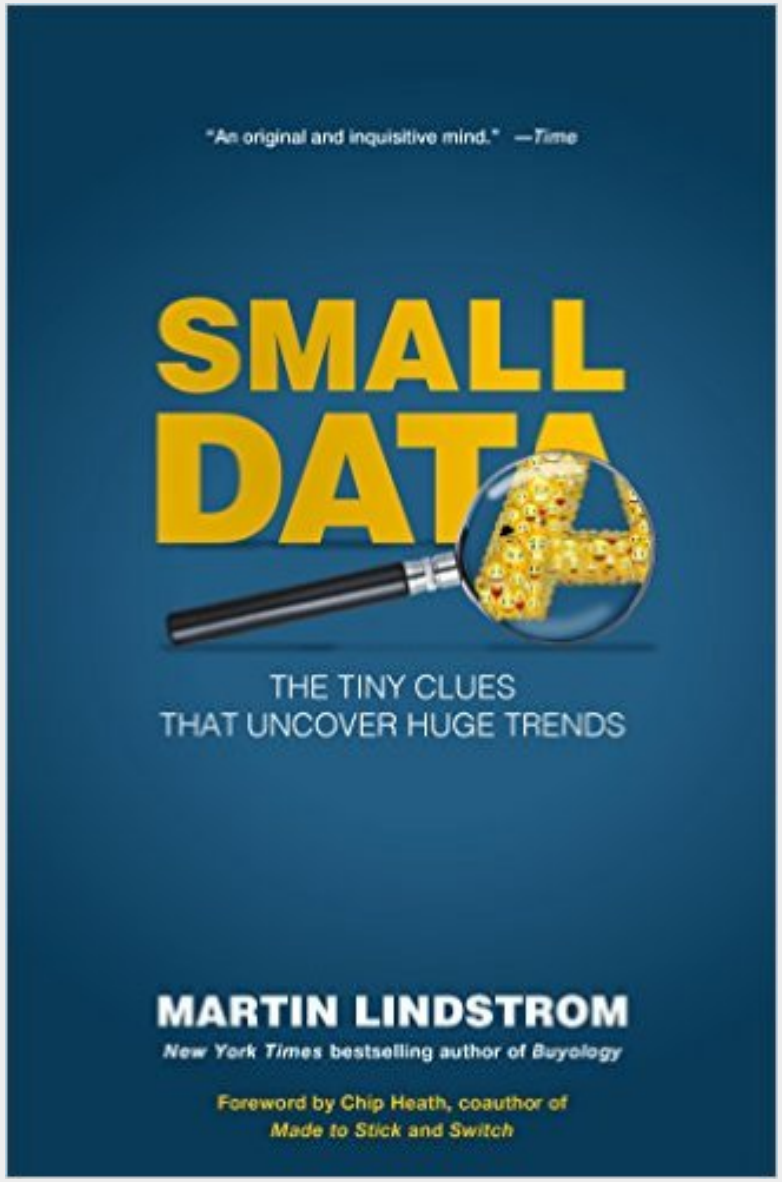


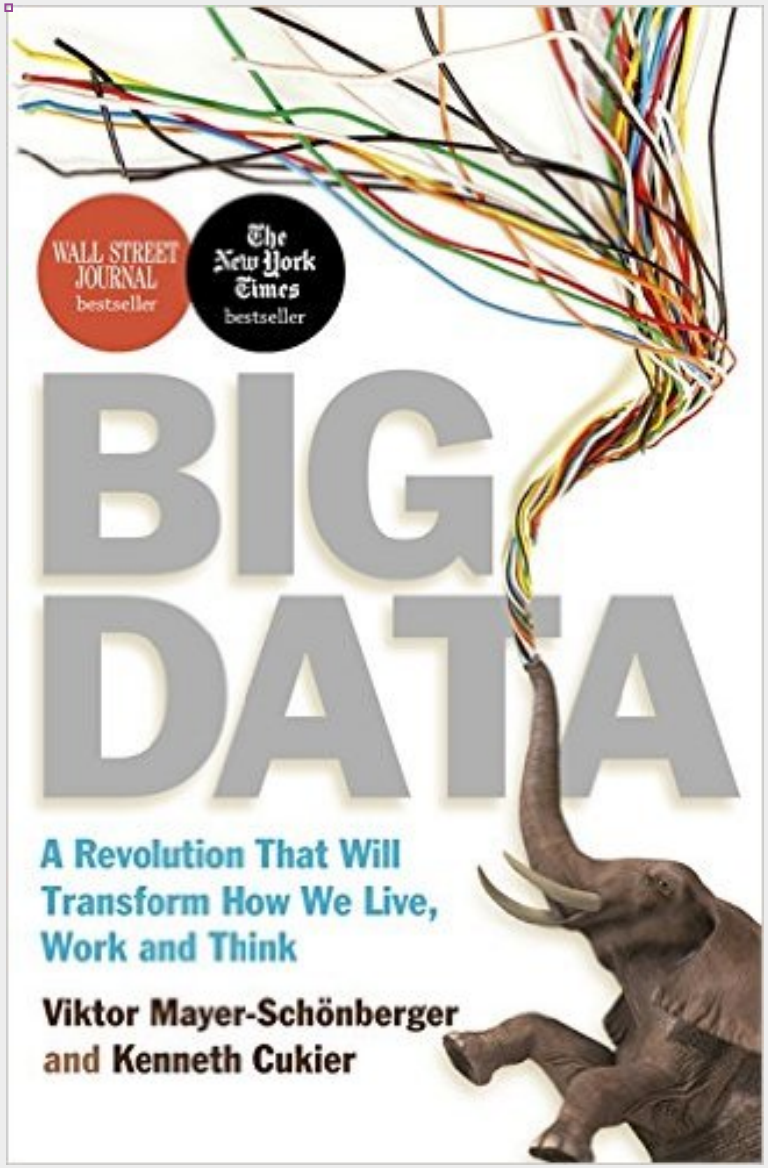
A couple more points before we continue

- The way we access the smart city is through technologies that let us generate and use data and its useful equivalent – information (data) is key
- Access through mobile and fixed devices like phones, smart cards, through fixed sensors which record transactions and so on
- These usually complement rather than substitute for data which we collected and used in the past
- This has introduced time into our thinking
- This is all part and parcel of increasing complexity; more time scales, more opportunities, more diversity ……

A Short History of Big Data: How Big is Big?

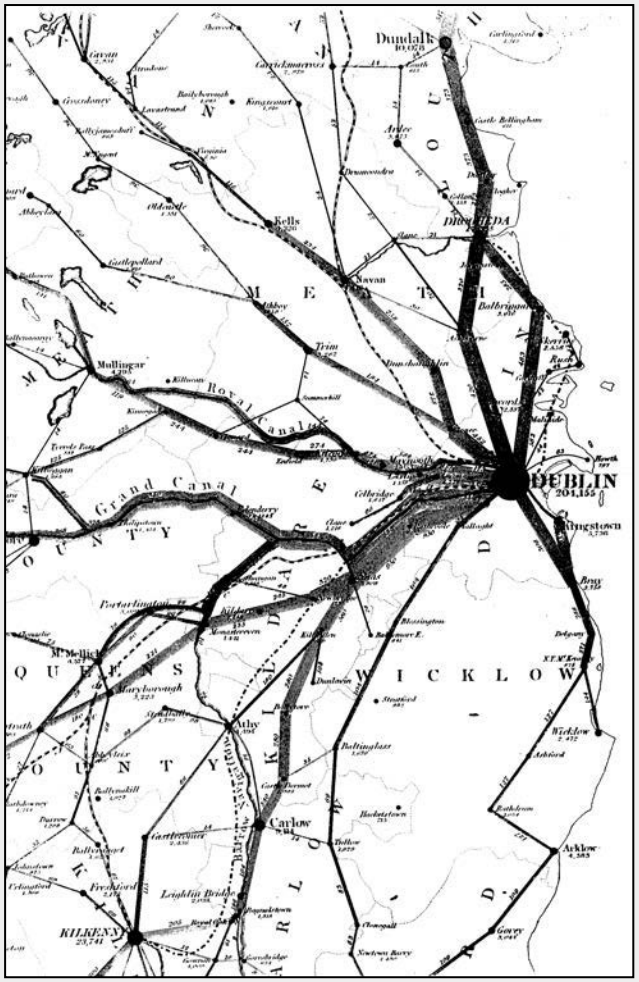
- Data is big with respect to its volume. I know there are other definitions – velocity, variety etc. but to me, data is big if it requires large use of computer memory implying volume.
- The conventional definition in business is the Five V's – volume, velocity, variety, veracity, value
- In cities, data usually implies numbers of locations and their attributes but locations imply interactions.
- Thus data are relations between locations and in essence if we have *n locations*, we have n^2 *interactions*. Thus small data can become big, and there is a caveat, small data can be good data too



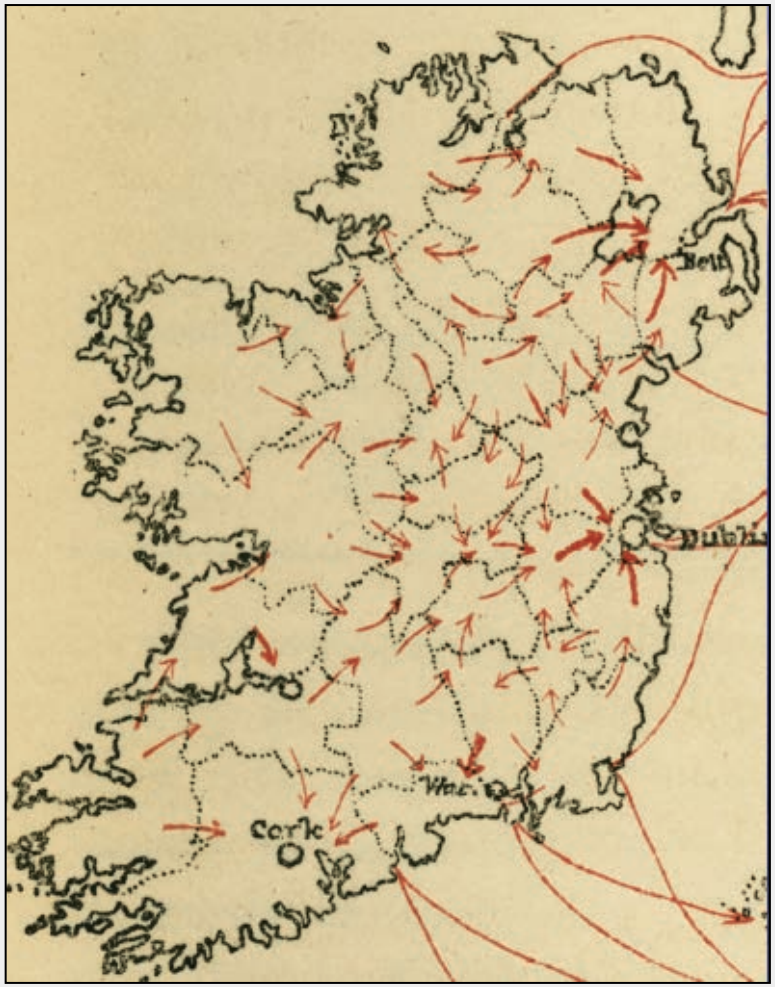


The Age of the Smart City

Examples: Dublin 1837, Ireland 1888, London 1955



Harness, 1837



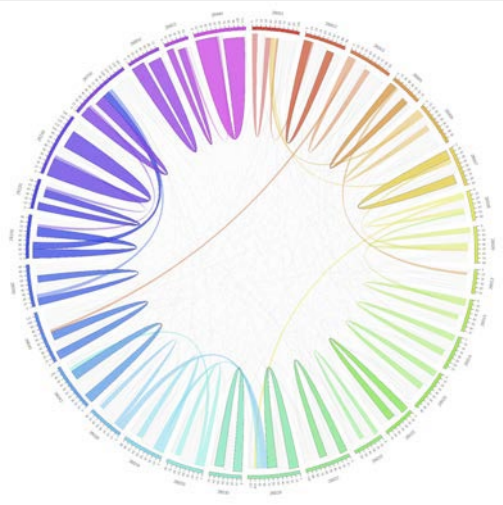
Ravenstein 1888



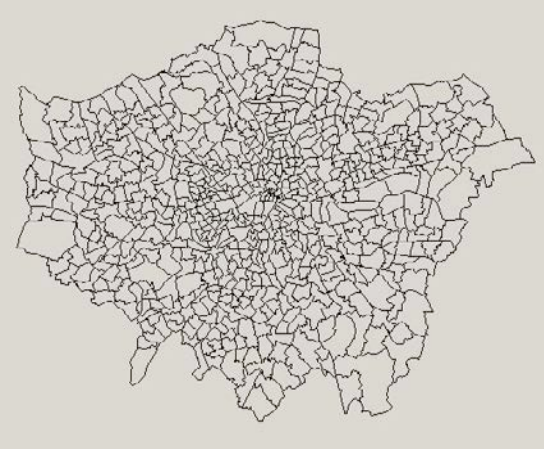


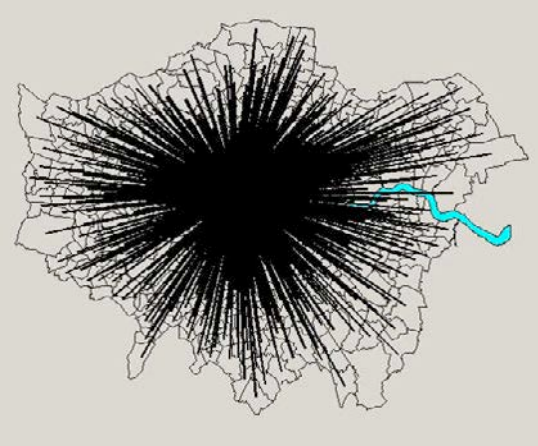


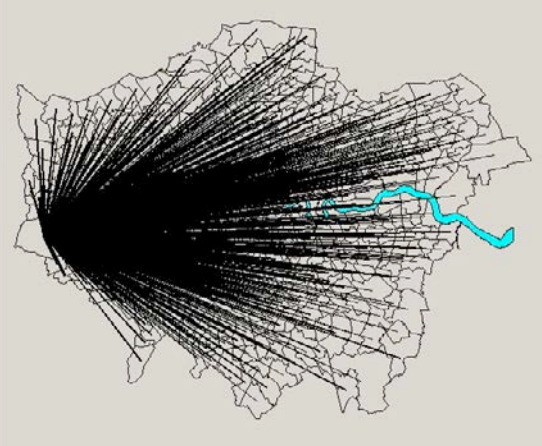
https://www.youtube.com/watch?v=pcBJfkE5UwU



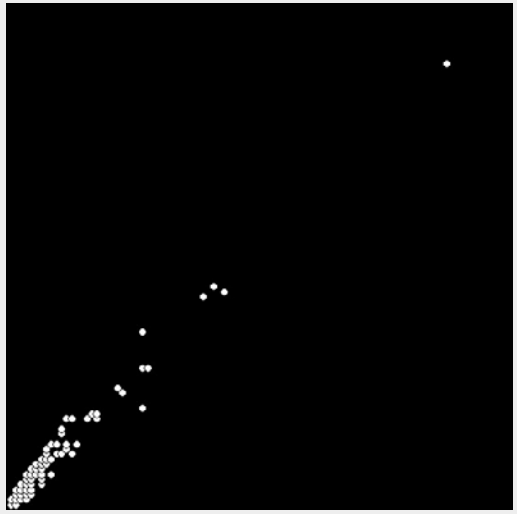


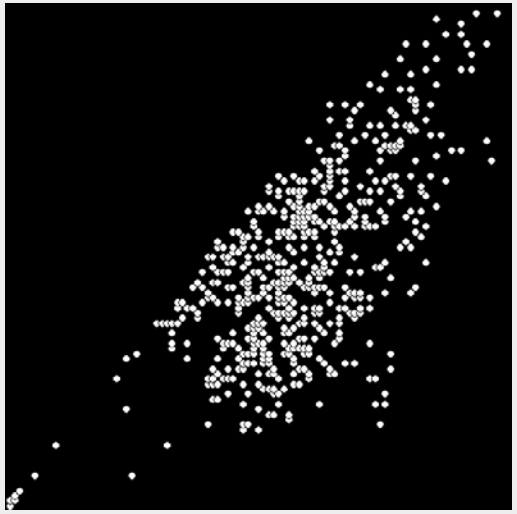


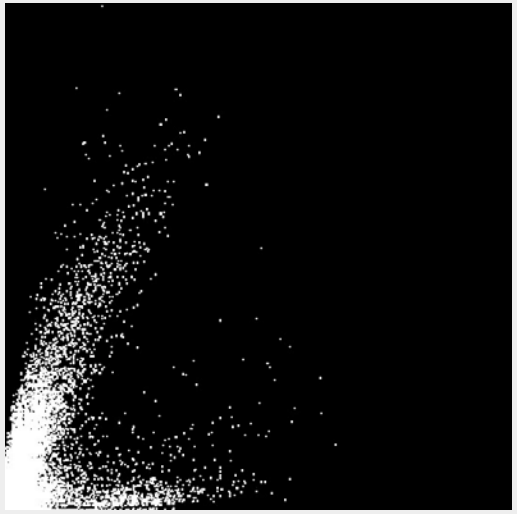




$n^2=633^2=400,689$, bigger but impossible to visualise





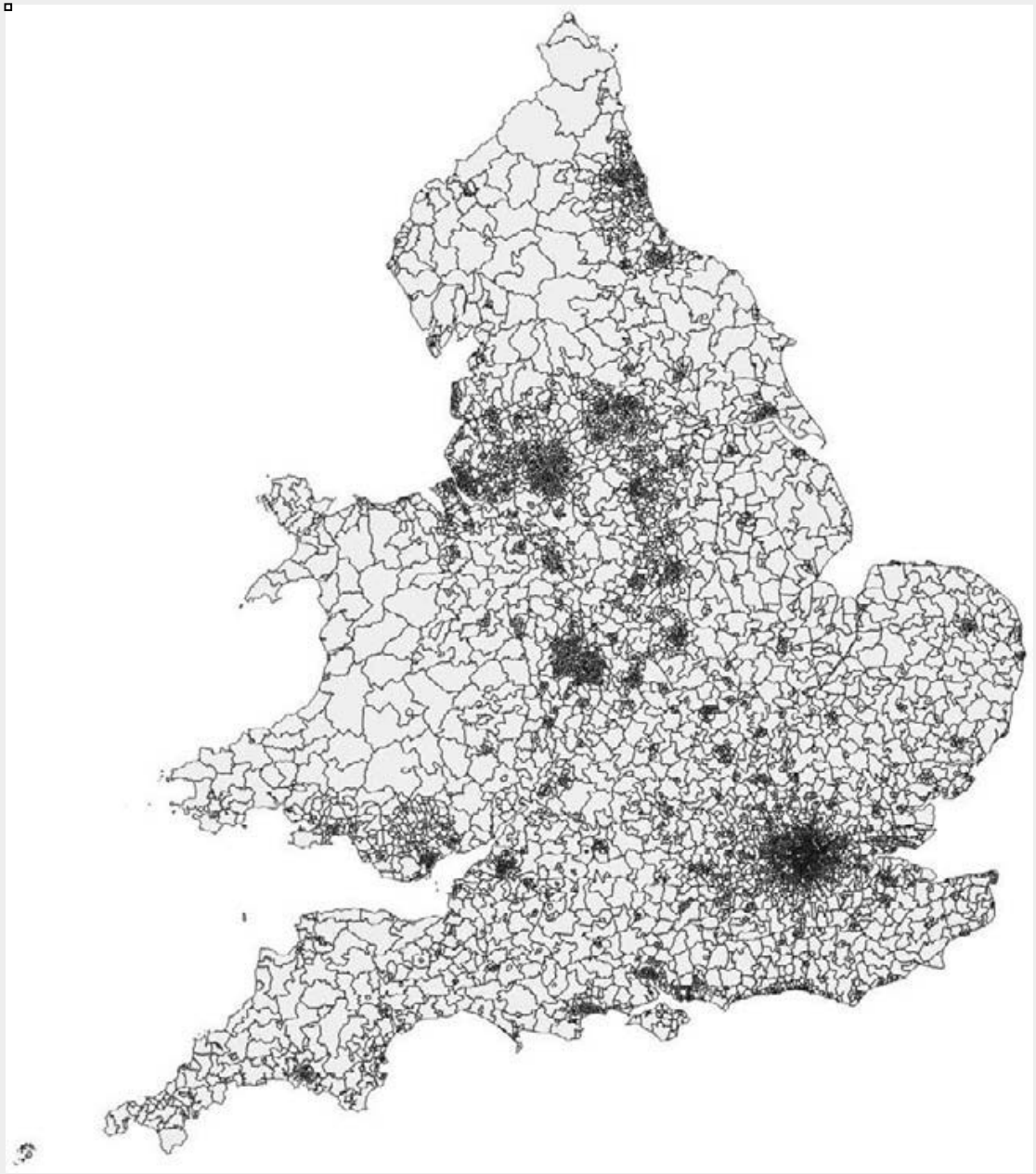




Even our statistics breaks down when we get large numbers like over several thousand as you can see on the left and above right for 400K data points where the pattern is highly convoluted. This is from a gravity model.

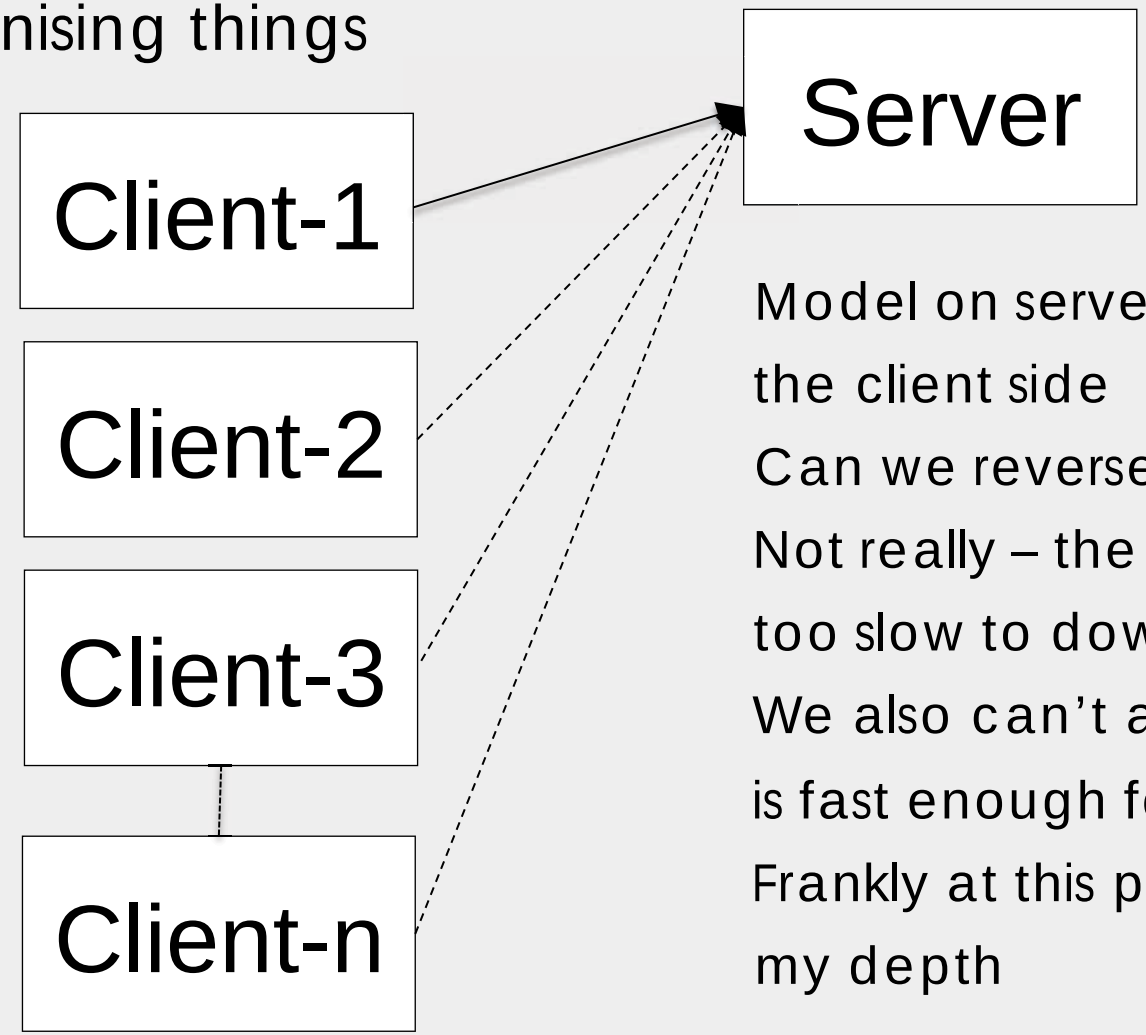
Now what happens when we really do scale up to the level of MSOAs of which there are 7201 in the UK – do we partition and argue we don't need to scale up to $n^2=7201^2=51,854,401$.

Circa 52 million points is an issue but our models run in a matter of seconds but that is a lot of data to store – ok it is sparse but sparsity isn't structured so we can't easily partition and in any case we want to compute any possible flows between central London say and Newcastle. Here is the problems scaled up and this is what we are grappling with at present.



The Age of the Smart City

Here is a block diagram of how we are currently organising things



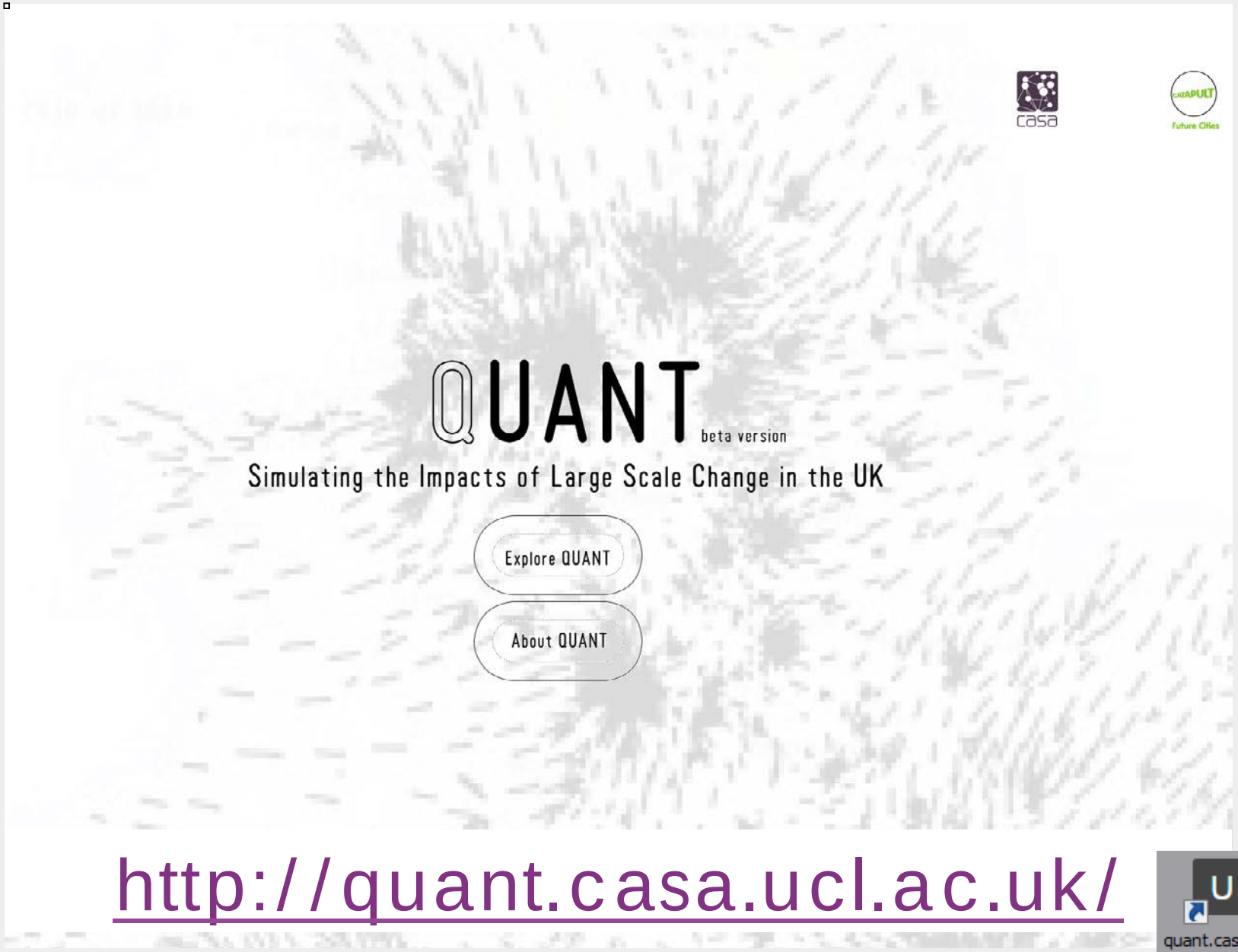
Model on server side; Maps on the client side

Can we reverse this?

Not really – the matrices are too slow to download to client?

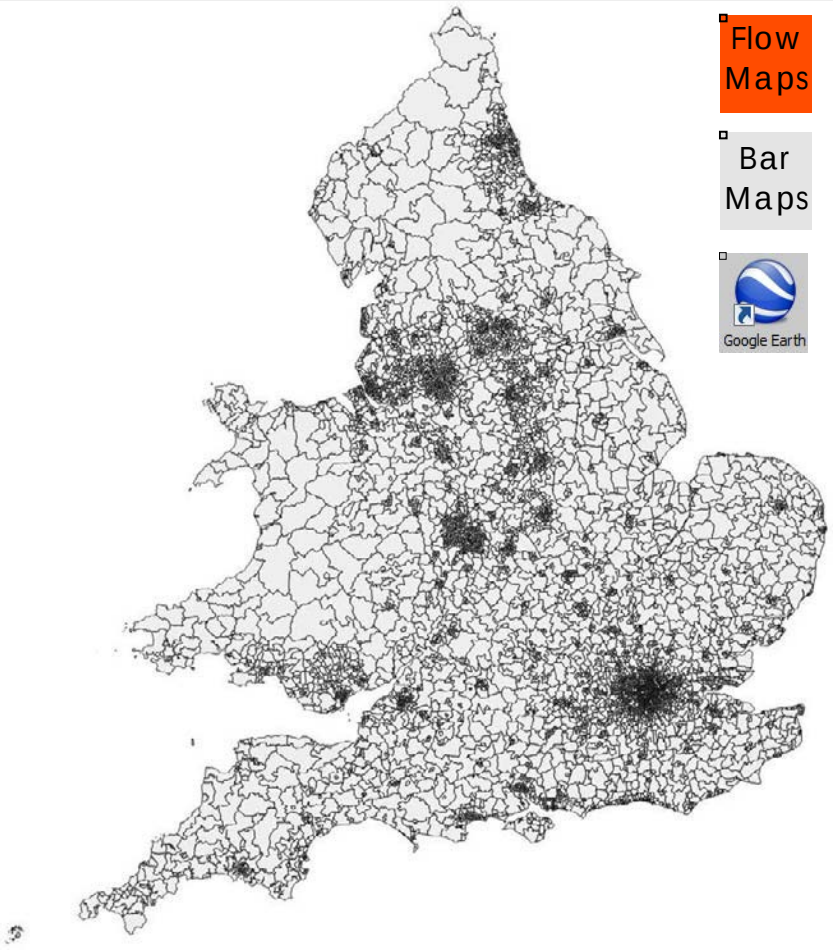
We also can't assume the client is fast enough for computation.

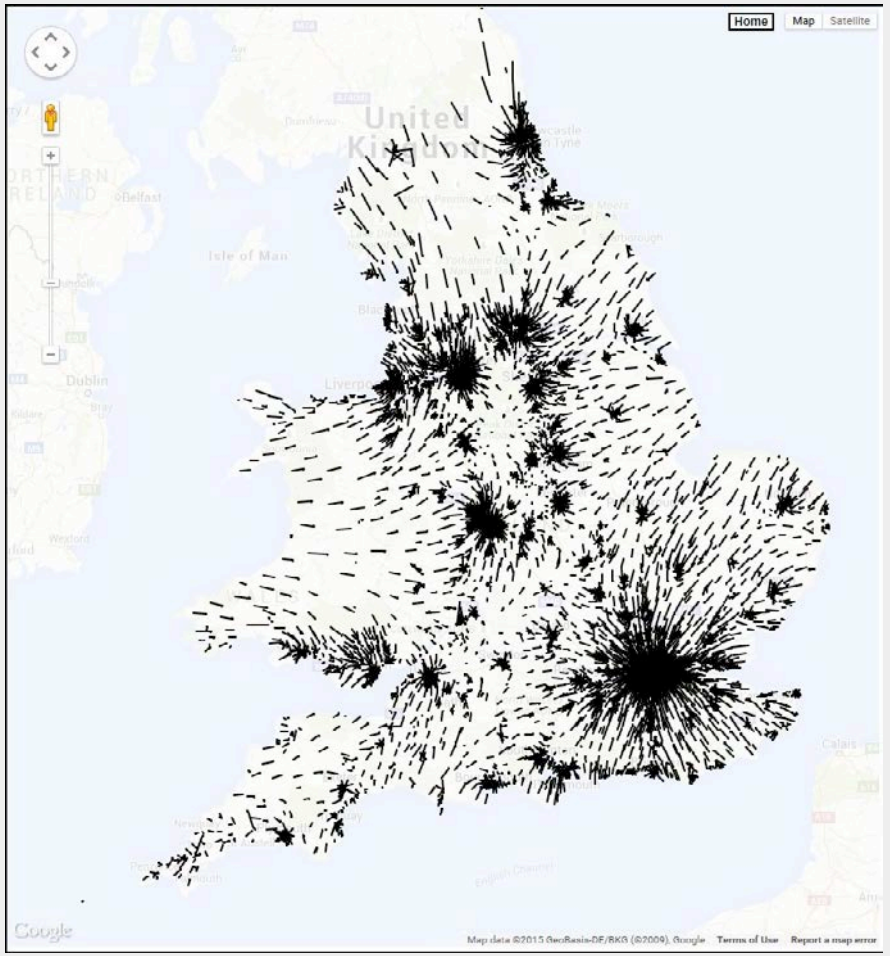
Frankly at this point, I am out of my depth



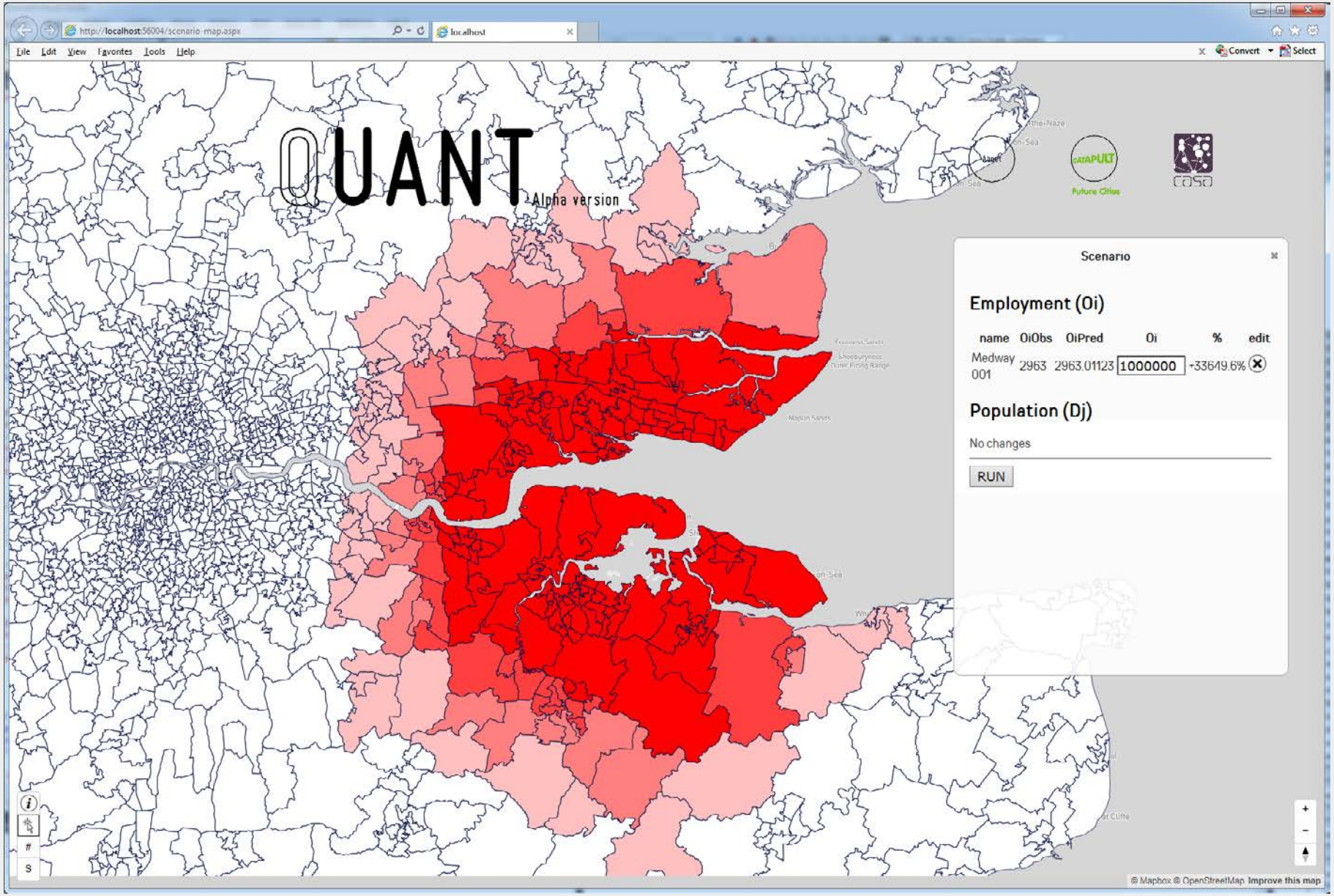
http://quant.casa.ucl.ac.uk/







$$[x_i, y_i] = [[x_i, y_i], \left[\left[x_i + \frac{\sum_j T_{ij}\left[x_i - x_j\right]}{n}\right], \left[y_i + \frac{\sum_j T_{ij}\left[y_i - yy_j\right]}{n}\right]\right]]$$



The Age of the Smart City

Mobility, Transit, & Real-Time Streaming: The Oyster Card Data Set

First though something about real time streamed data



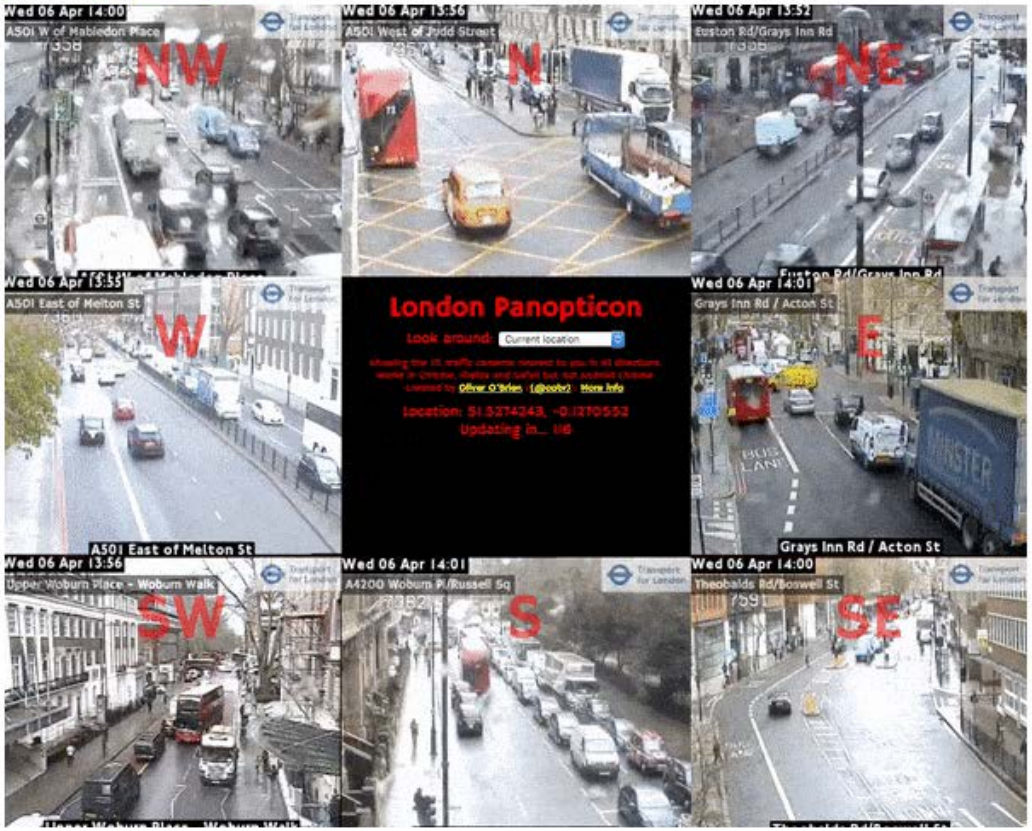


http://www.citydashboard.org/

London Panopticon

6 April 2016 London

http://vis.oobrien.com/panopticon/





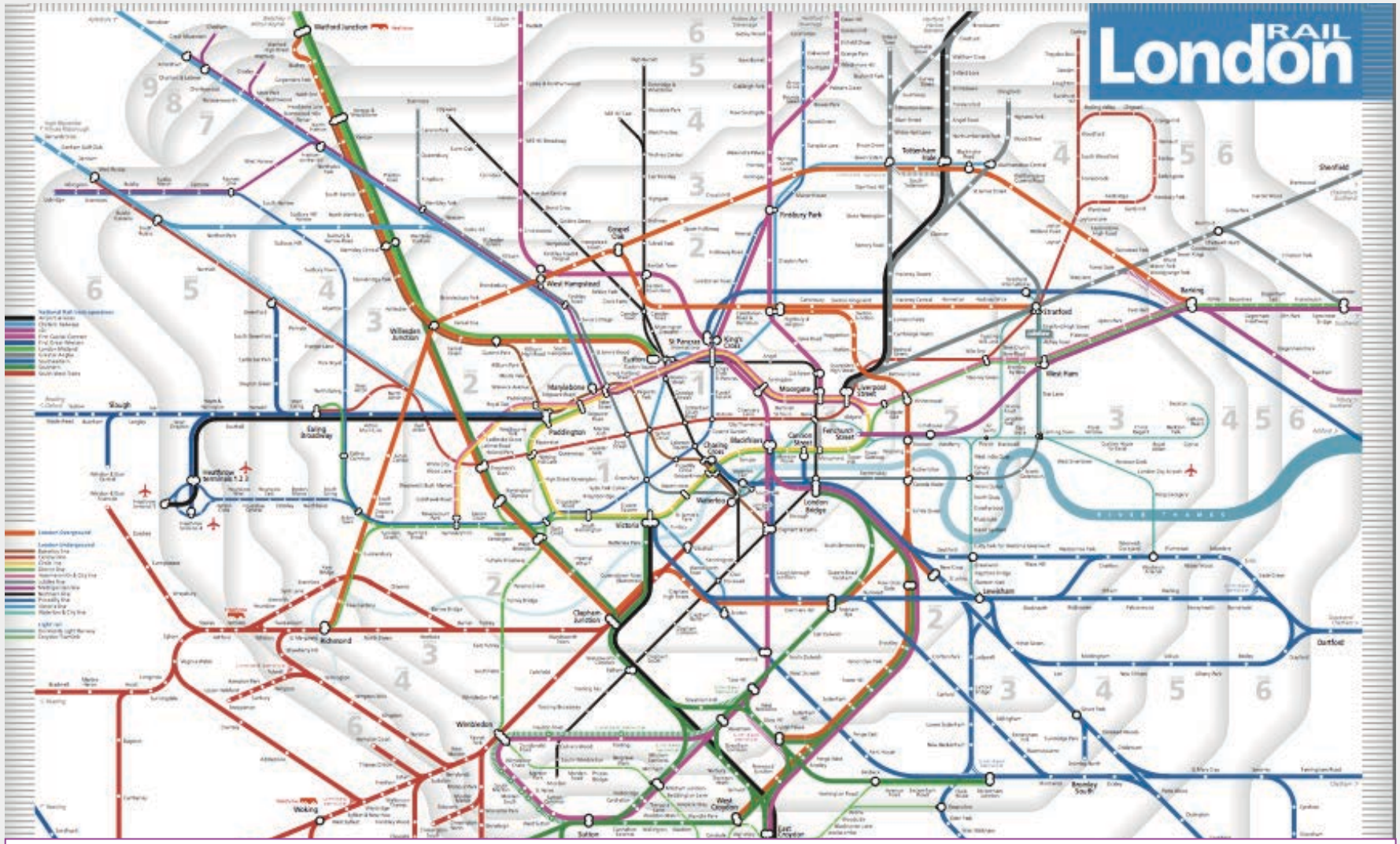
Real-Time Streaming: The Oyster Card Data Set

- Tap at start and end of train journeys
- Tap at start only on buses
- Accepted at 695 Underground and rail stations, and on thousands of buses
- Many Variants of the Data Sets
- 991 million Oyster Card taps over Summer 2012 – this is big data
- Quality of Data
- What Can We Use It For
- Missing Data and Noise



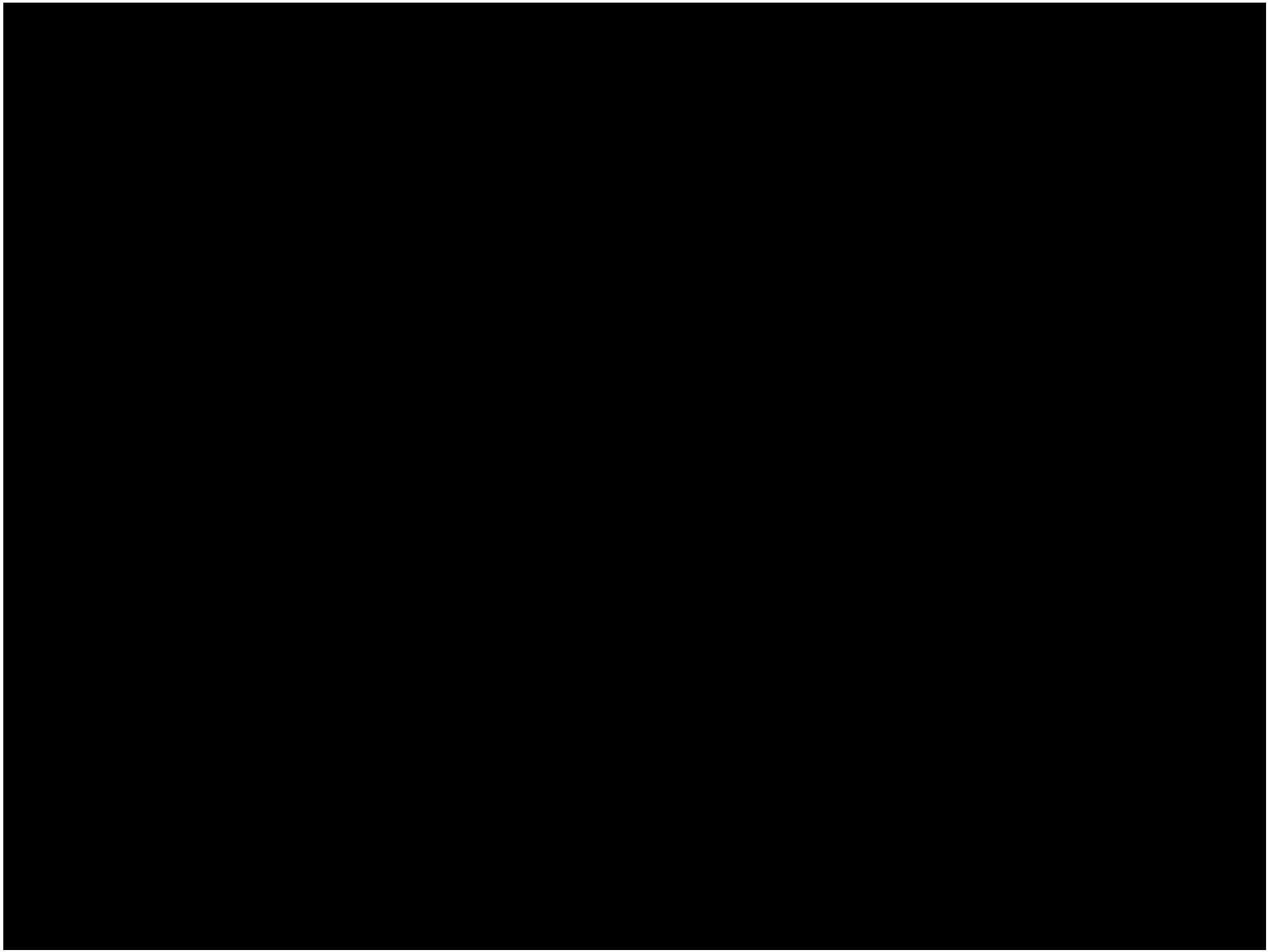






Tube, Overground and National Rail Networks in London where Oyster cards can be used


UCL ENGINEERING
Change the world
UCL
Menu +
Home – Projects – Oyster gives up pearls
OYSTER GIVES UP PEARLS
How studying millions of Oyster Card journeys reveals London's 'polycentres'
Researchers from UCL have analysed millions of Oyster Card journeys in a bid to understand how, why and where we travel in London.
Professor Michael Batty (UCL Centre for Advanced Spatial Analysis) and Dr Soong Kang (UCL Management Science and Innovation) applied the techniques of statistical physics to their mountain of raw data.
The pair joined forces with a computational social scientist and a physicist, both based in Paris, to explore patterns of commuting by tube into central London.
UCL Engineering - Oyster Gives up Pearls
0:00 / 3:32
They used Transport for London's database of 11 million records taken over one week from the Oyster Card electronic ticketing system.
Latest news from UCL Engineering
New web privacy system could revolutionise the safety of surfing
UCL host Google Girls Coding Programme with Generating Genius and University of West indies
Professor Polina Bayvel to Give Royal Society Lecture
Twitter feed
RT @markmiodownik: Am giving a ENGins seminar today for @UCLEngineering @UCLENGins all UCL engineers welcome - Roberts G06, 6:30pm. http:/...
8:58am Thu 9th October 2014
RT @Centre4EngEdu: We're hiring! Multi-talented Centre Administrator required to help us launch and expand! bit.ly/ZsERSM
10:54am Wed 8th October 2014
Join our mailing list

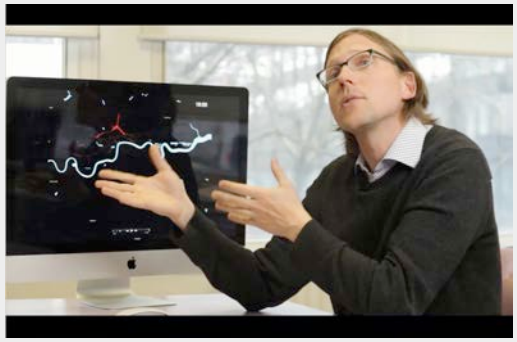


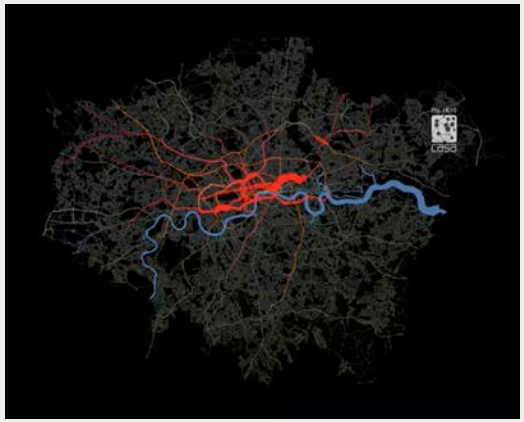
And how can we make sense of this



http://www.simulacra.info/



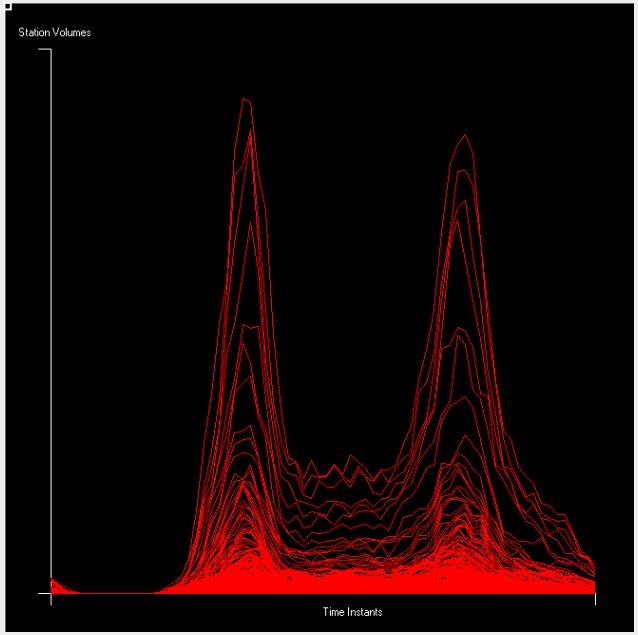



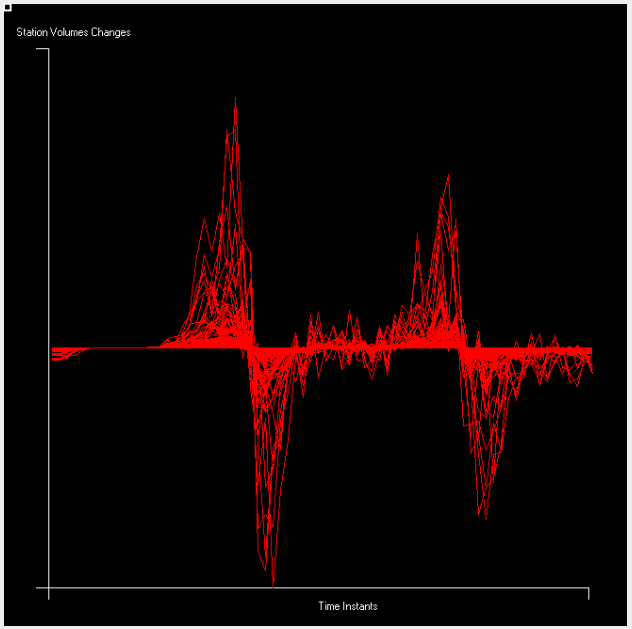



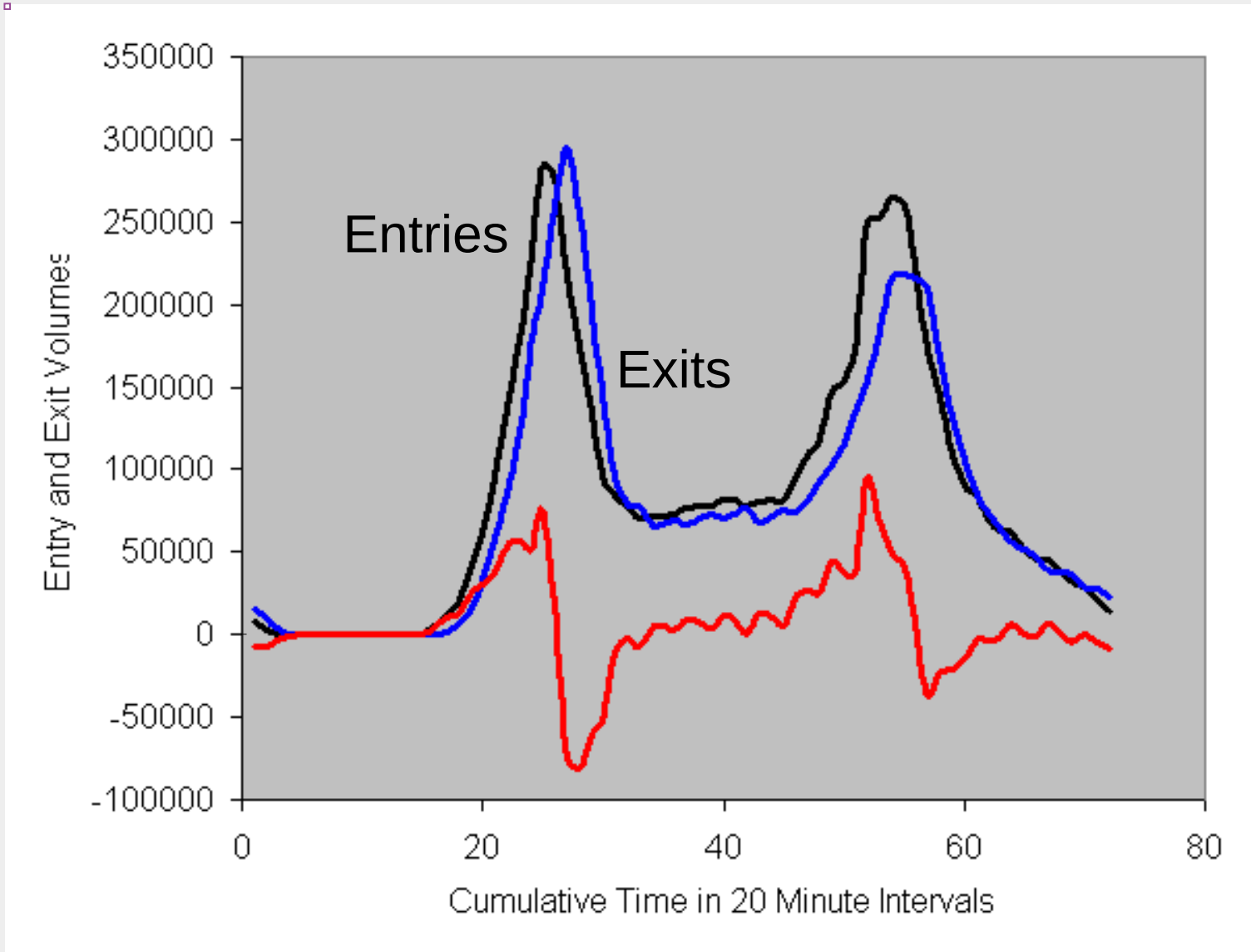
This of course was the thing that Lt Henry Harness did in Dublin in 1837 and what Minard et al. did a little later. In our LUTI models, this is an enormous problem as the scale of this assignment to networks is different

Variabilities – Heterogeneity and Travel Profiles

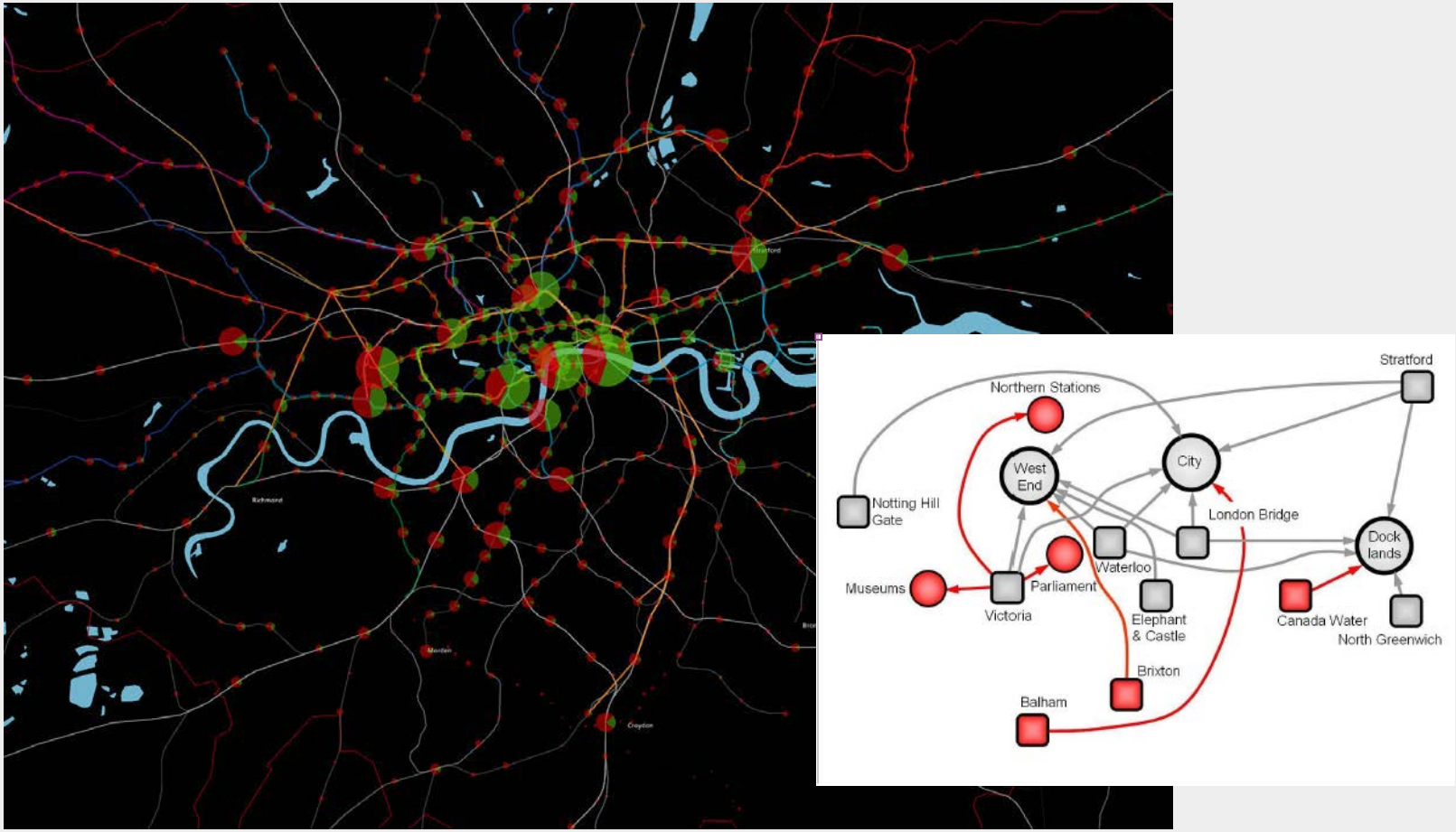
First we will look at some of the data and how it varies in terms of the diurnal flows usually morning and evening peaks, with a small blip (peak) around 10pm at night





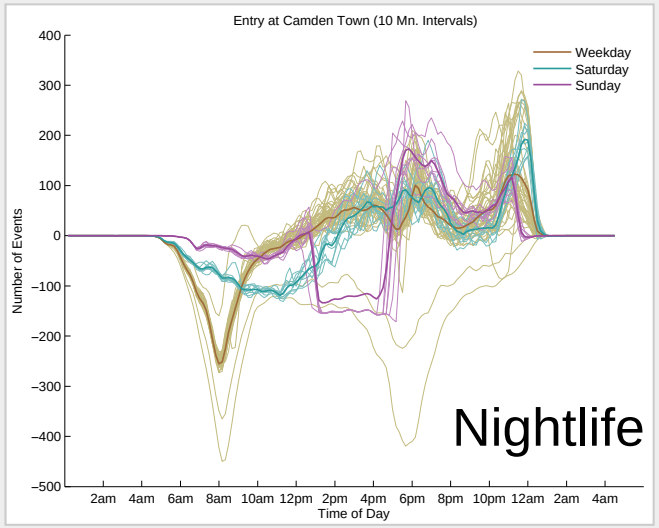

350000
300000
250000
200000
150000
100000
50000
0
-50000
-100000
Entry and Exit Volumes
Entries
Exits
0
20
40
60
80
Cumulative Time in 20 Minute Intervals

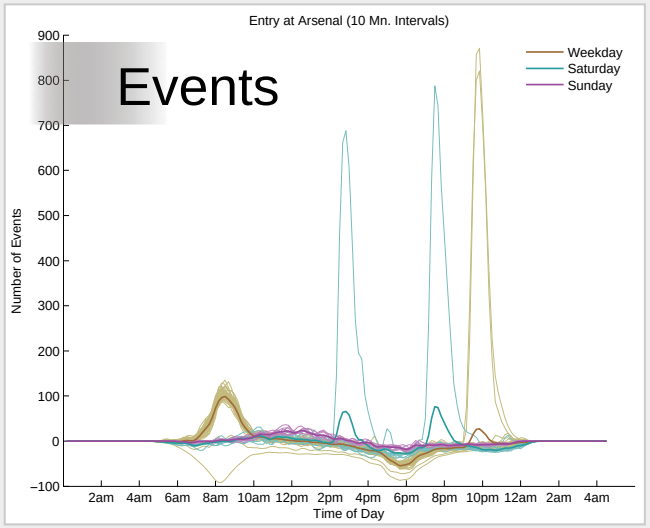
Oyster Card Data – interpreting urban structure, multitrips, etc.

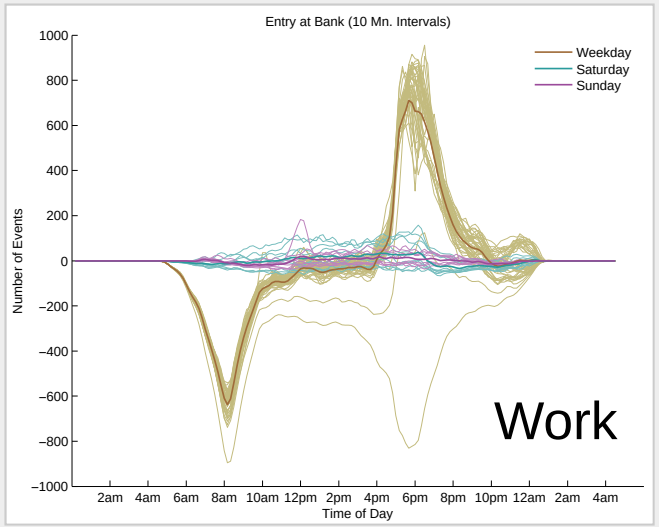


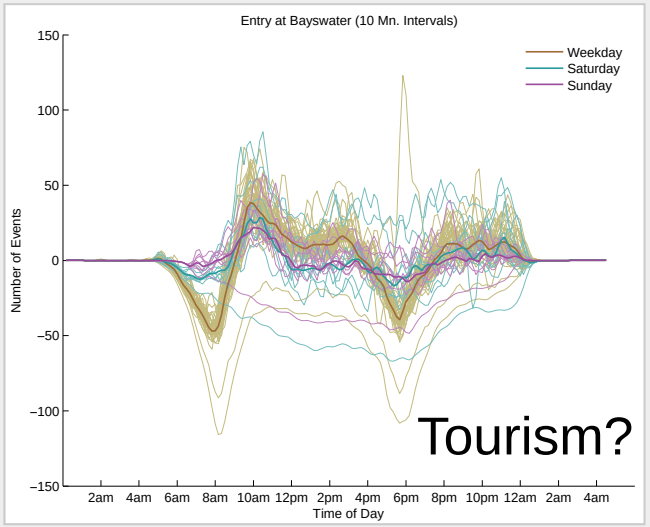
Roth C., Kang S. M., Batty, M., and Barthelemy, M. (2011) Structure of Urban Movements: Polycentric Activity and Entangled Hierarchical Flows. PLoS ONE 6(1): e15923. doi:10.1371/journal.pone.0015923

Particular Events: Weekdays, Saturdays and Sundays









Learning about Mobility from the Data Variabilities – Disruptions – Dynamics

Comparing Variability for different time intervals for Three World Cities: London, Beijing and Singapore

Table 1. Summary statistics of one-week of smart-card data (metro trips only)

	London	Singapore	Beijing
Monday	3,457,234	2,208,173	4,577,500
Tuesday	3,621,983	2,250,597	4,421,737
Wednesday	3,677,807	2,277,850	4,564,335
Thursday	3,667,126	2,276,408	4,582,144
Friday	3,762,336	2,409,600	4,880,267
Number of stations (1)	400	130	233
Number of tube line	13	4	17
Area (2)	1,572 km^2	718.3 km^2	2267 km^2
Total population (3)	8.63 million	5.3 million	21.15 million
Ridership of Metro	20%	35%	21%
Length of metro lines	402km	182km (MRT+LRT)	465 km

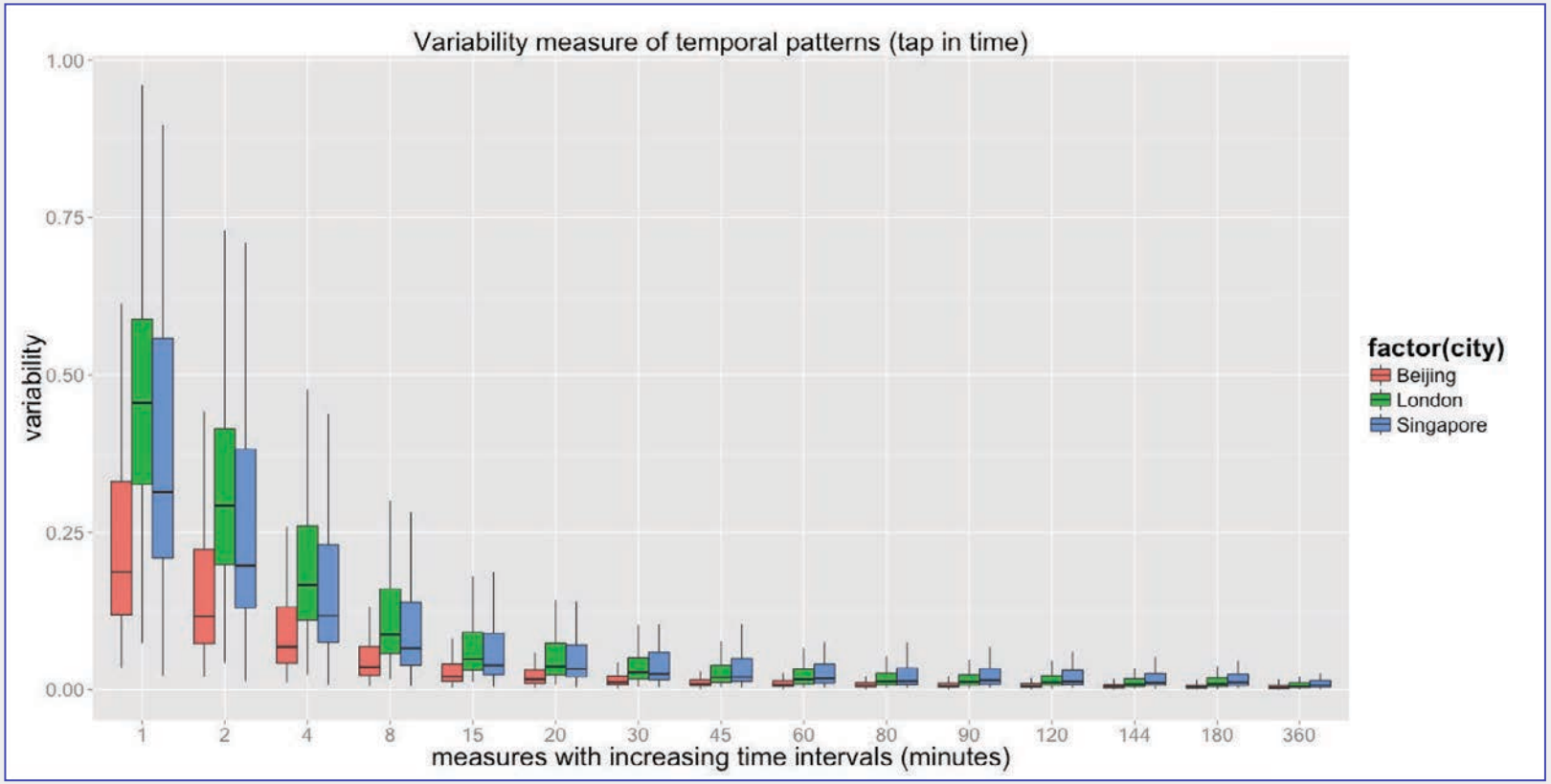
(1) Number of stations is the number of stations with smart-card records generated.

(2) The area of Beijing only counts the area enclosed by the 6th ring road for a fair comparison.

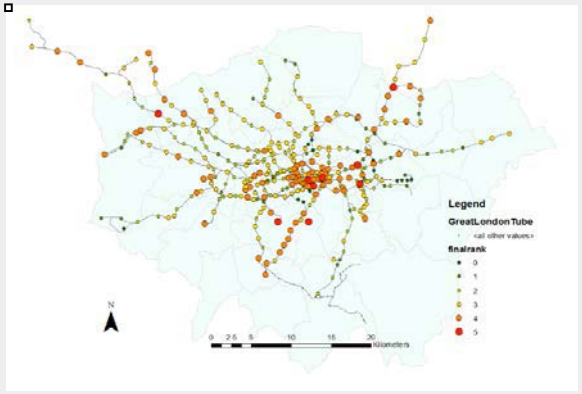
(3) From the World Population Review, http://worldpopulationreview.com/world-cities/ accessed 17 January 2016

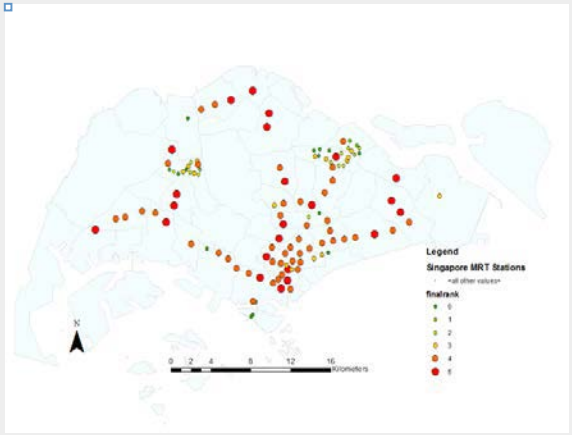
Zhong, C., Batty, M., Manley, E., Wan, J., Wang, Z., Che, F., and Schmitt, G. (2016) Variability in Regularity: Mining Temporal Mobility Patterns in London, Singapore and Beijing using Smart-Card Data., **PLOS One**, http://dx.doi.org/10.1371/journal.pone.0149222

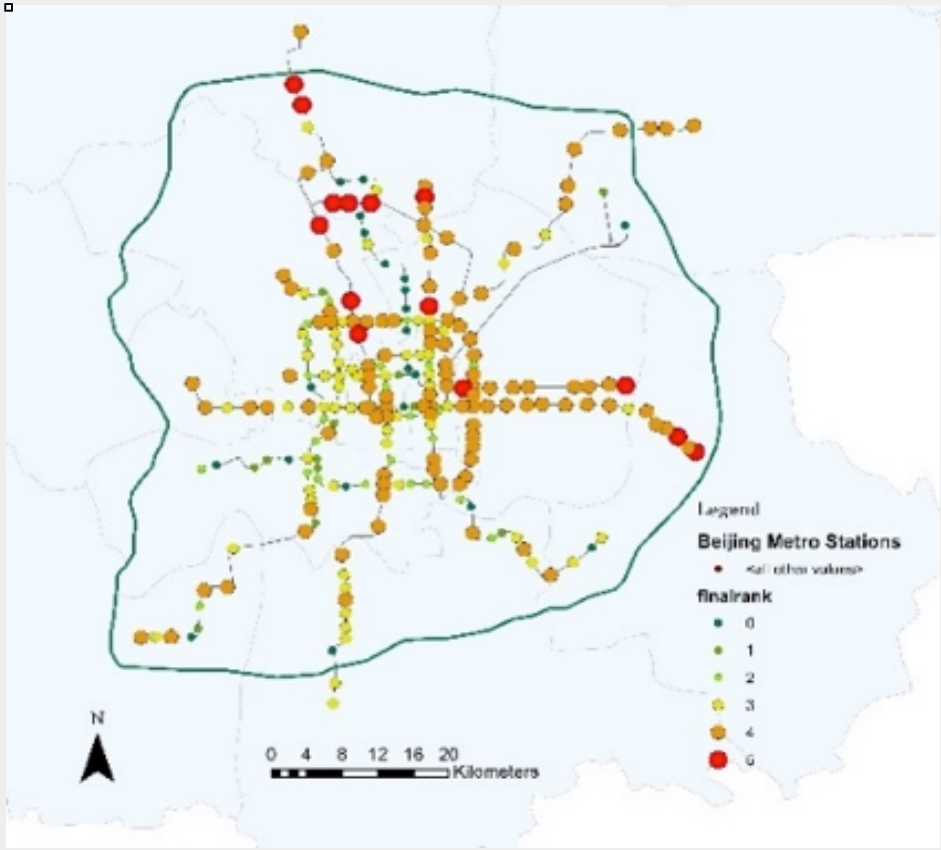
From 1 minute intervals to the whole day



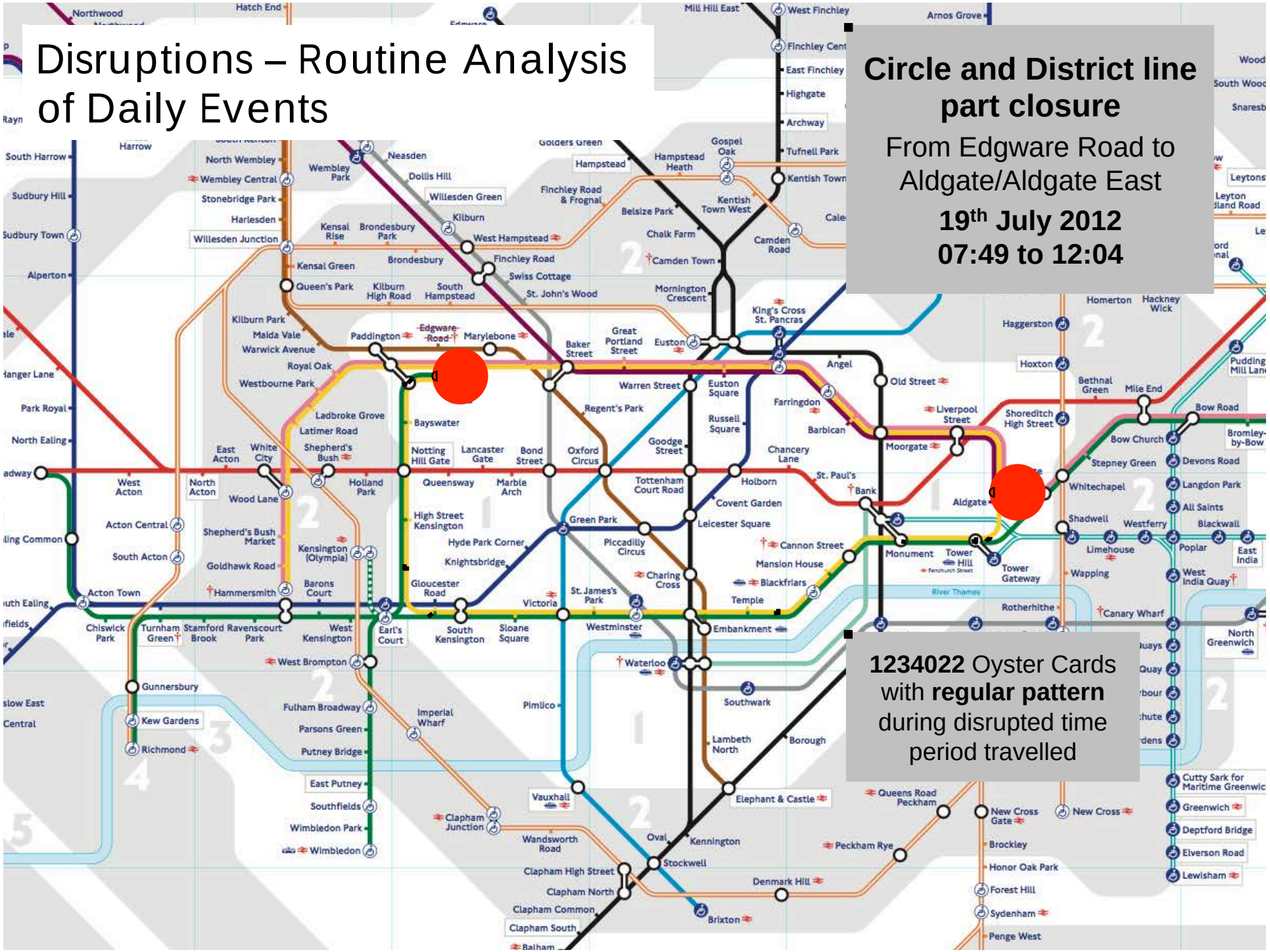
Comparing Variability for different time intervals for Three World Cities: London, Beijing and Singapore

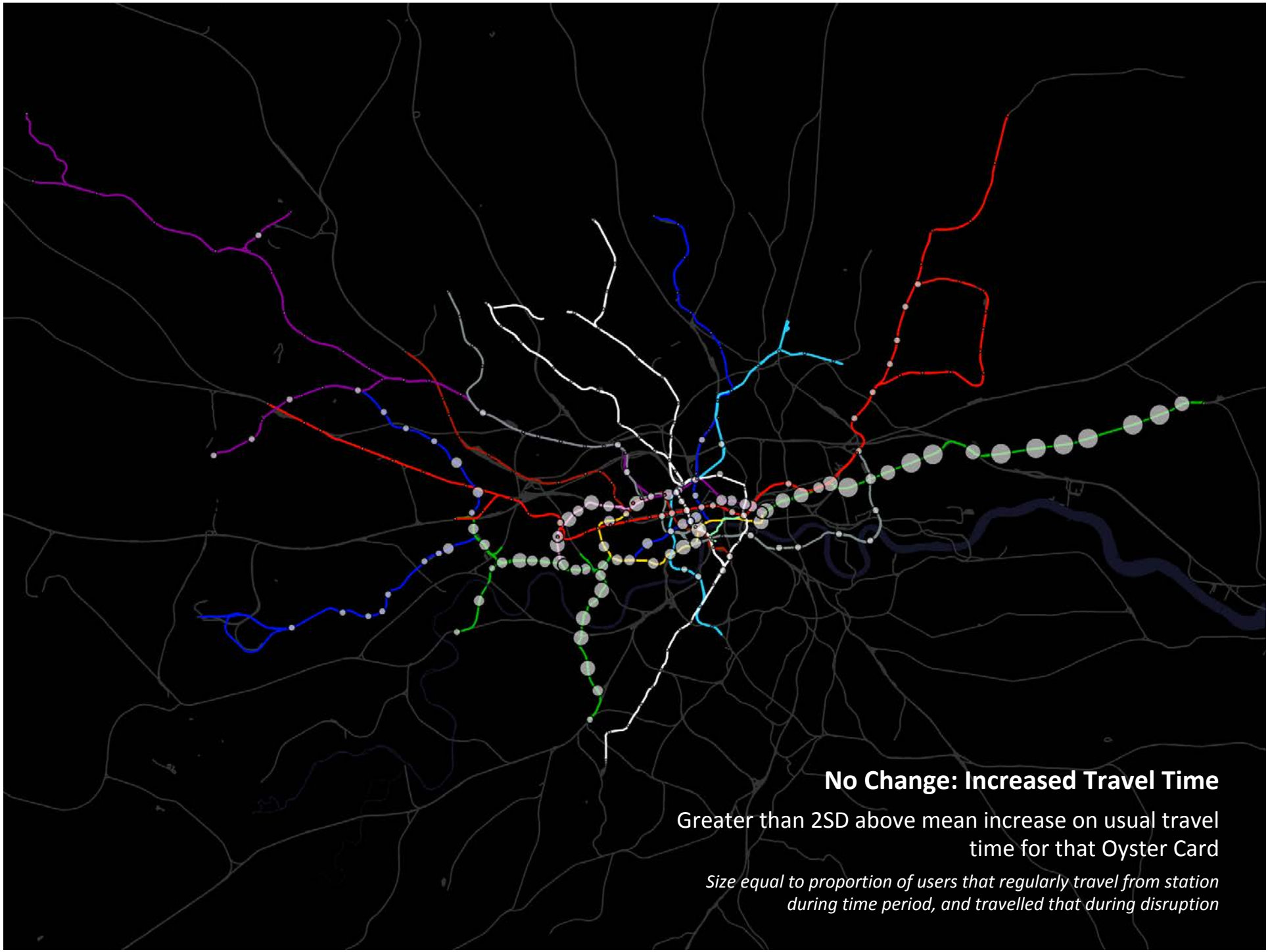


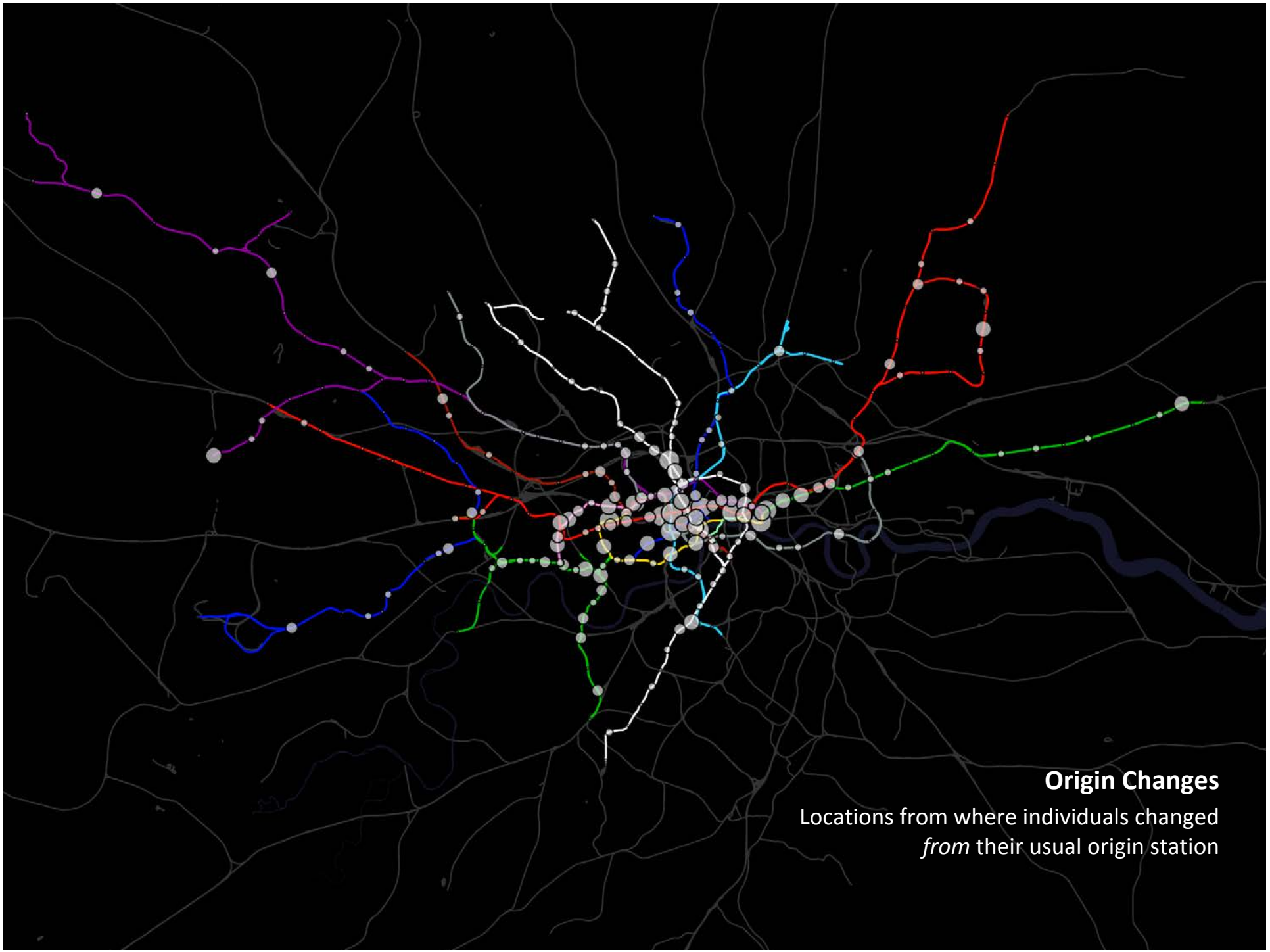




Disruptions – Routine Analysis of Daily Events

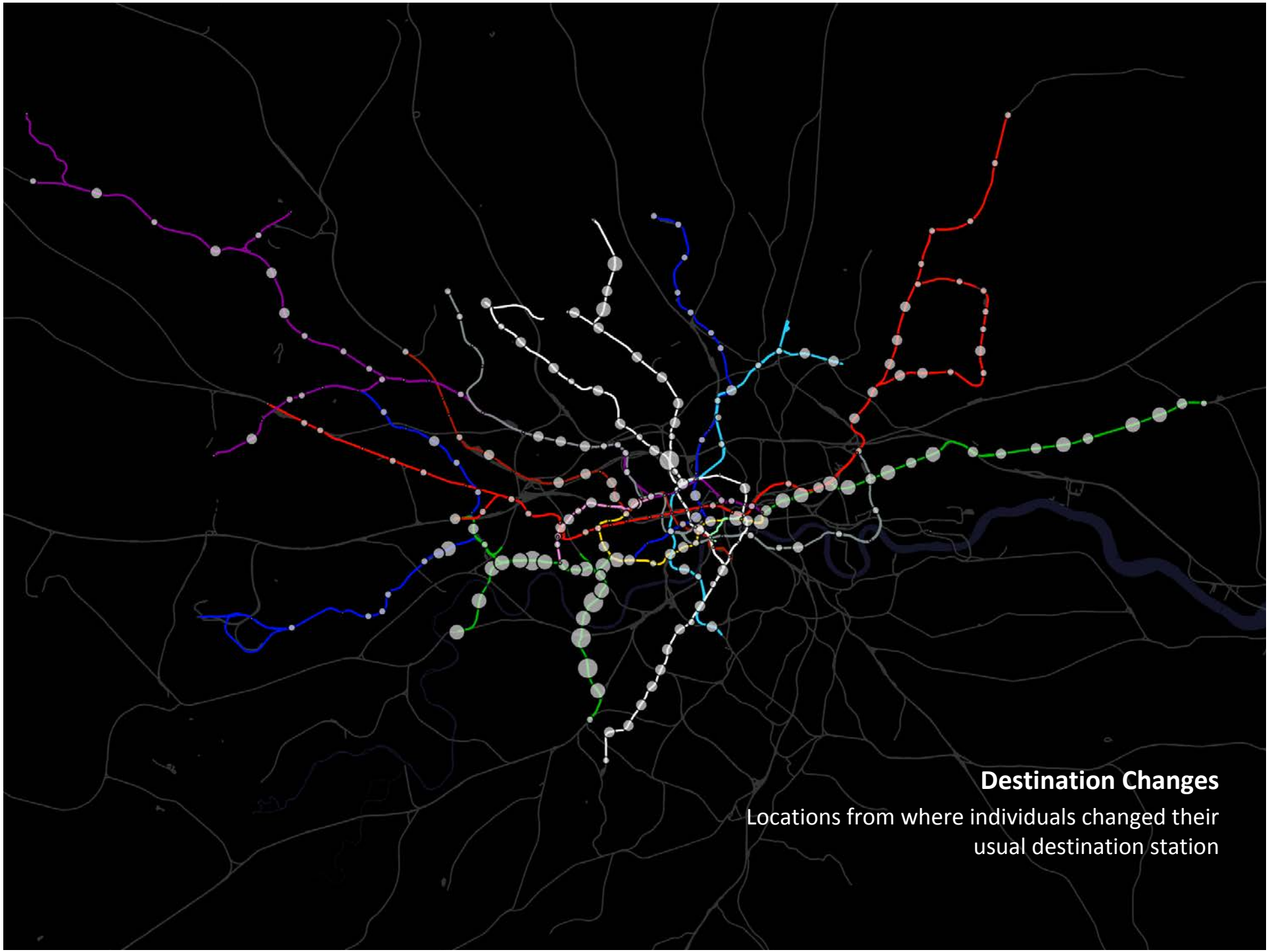



No Change: Increased Travel Time
Greater than 2SD above mean increase on usual travel time for that Oyster Card
Size equal to proportion of users that regularly travel from station during time period, and travelled that during disruption



Origin Changes

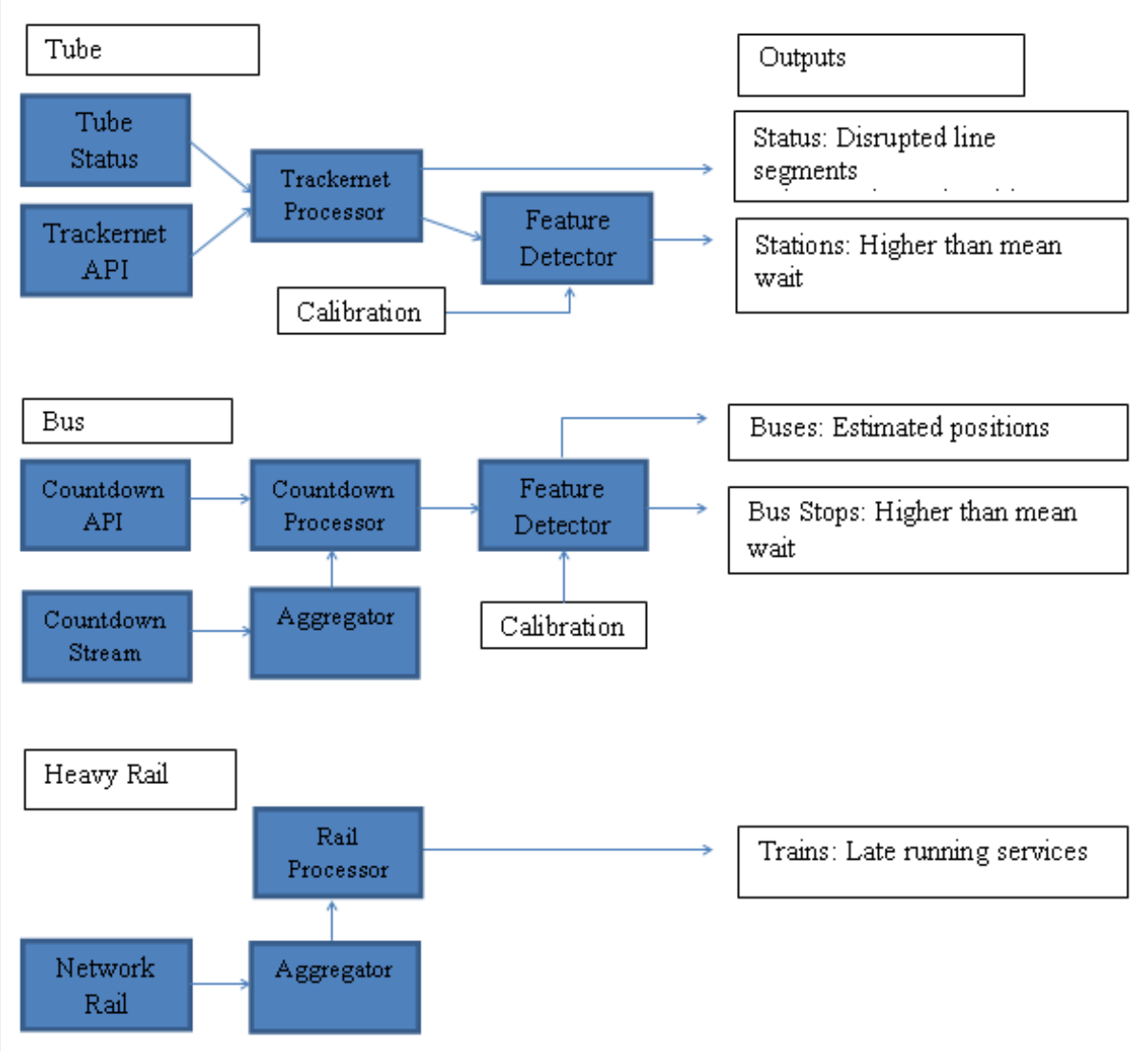
Locations from where individuals changed *from* their usual origin station

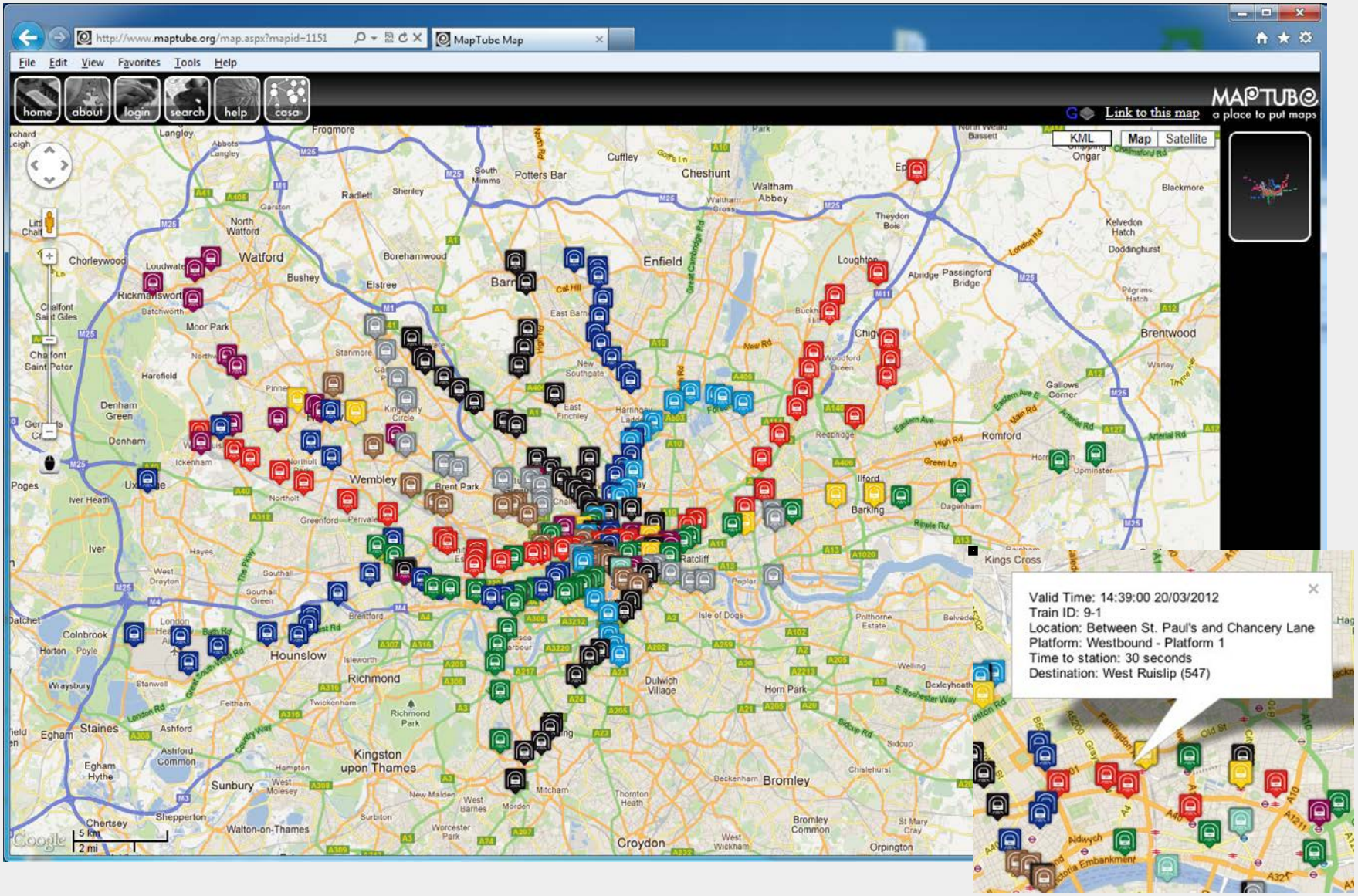


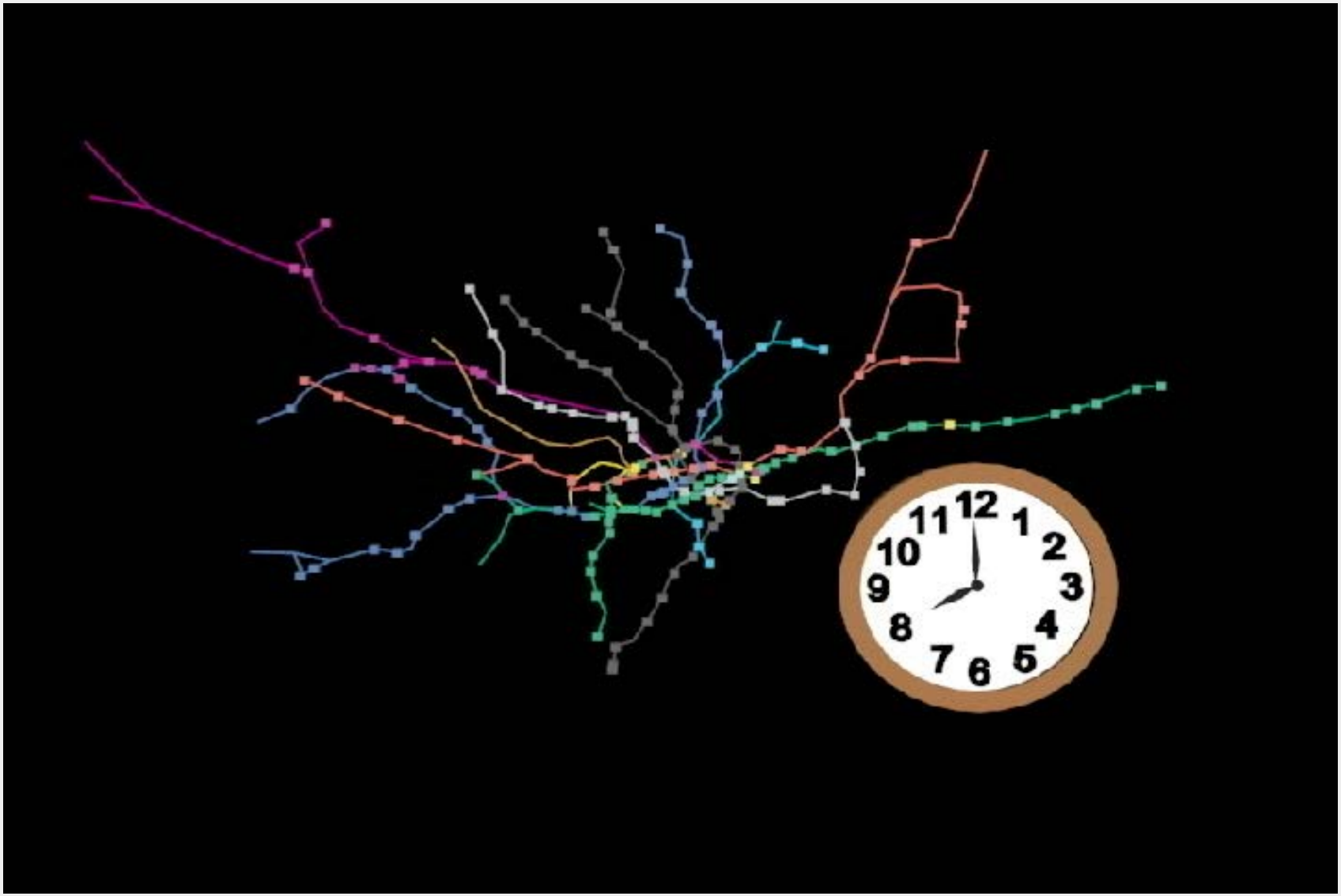
Destination Changes

Locations from where individuals changed their usual destination station

The Public Transport System in Terms of Vehicle Flows

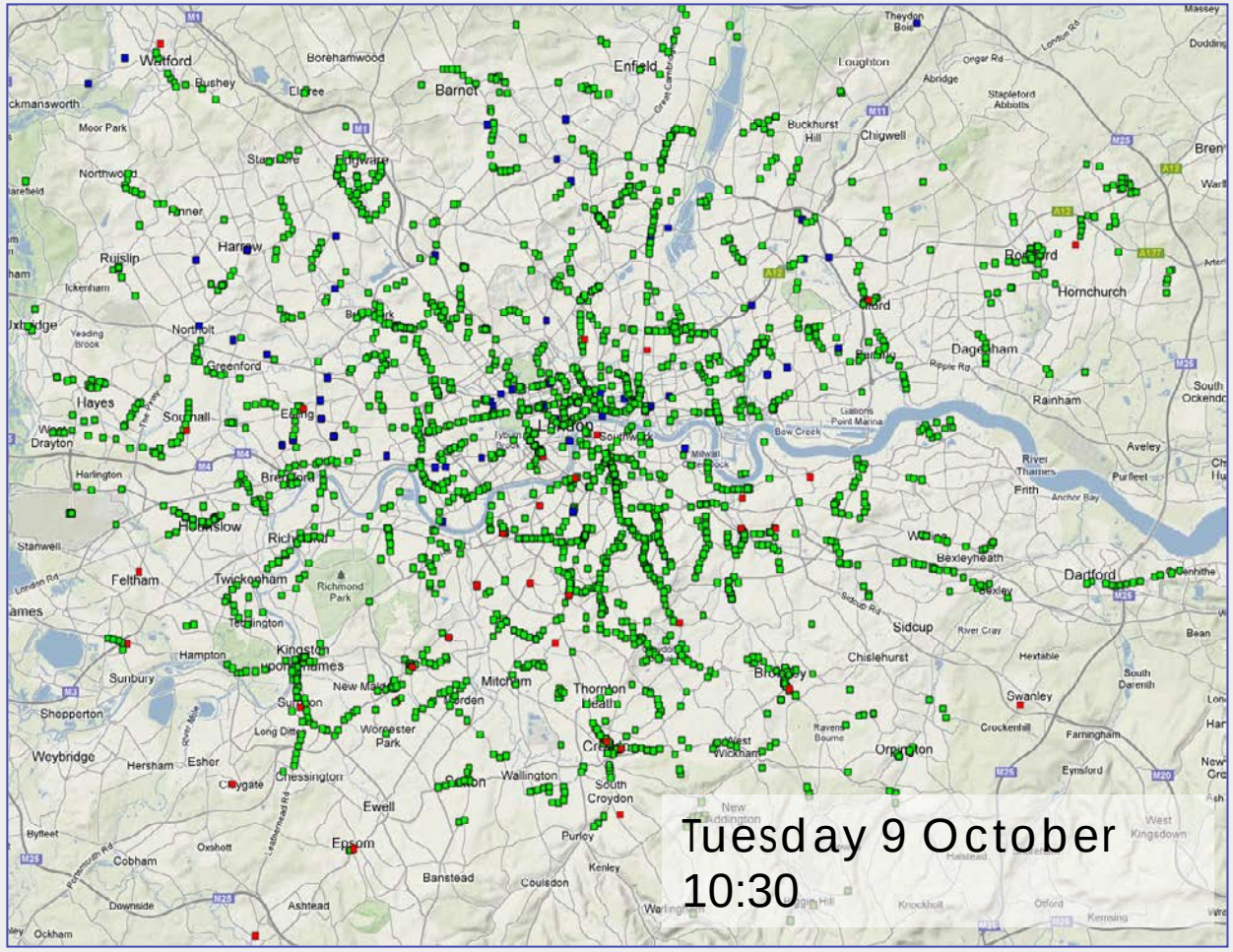



http://www.maptube.org/map.aspx?mapid=1151
MapTube Map
File Edit View Favorites Tools Help
home about login search help casa
MAPTUBE
a place to put maps
Link to this map
KML Map Satellite
Valid Time: 14:39:00 20/03/2012
Train ID: 9-1
Location: Between St. Paul's and Chancery Lane
Platform: Westbound - Platform 1
Time to station: 30 seconds
Destination: West Ruislip (547)



The Age of the Smart City

Delays from Tube, National Rail and Bus Fused



Key

- National Rail more than 5 minutes late
- Tube stations showing a wait time 15% above expected
- Bus stops showing a wait time 20% above expected

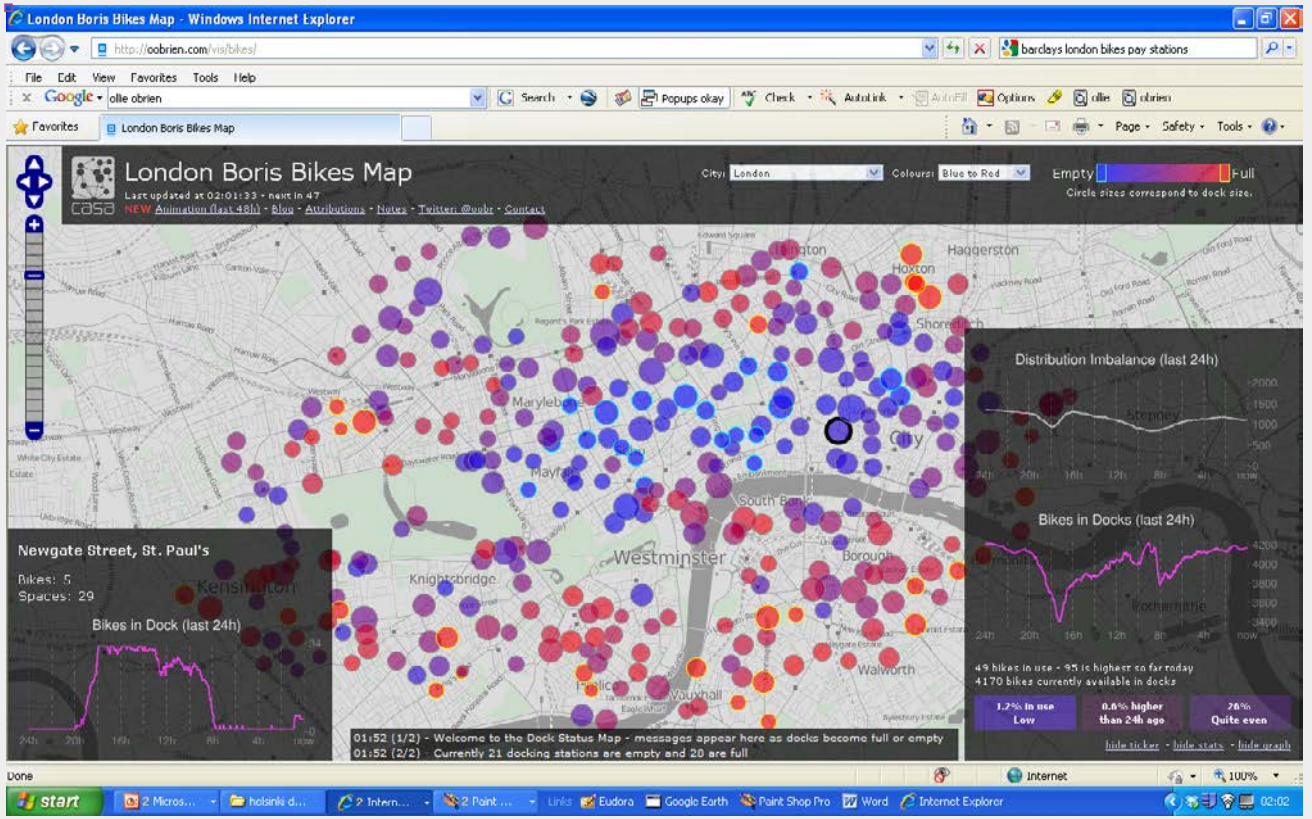
Tube delays from the TfL status feed are also plotted as lines

Related Real -Time Data: Bikes, Social Media

A lot of data is now coming online for travel and one of our group Oliver O'Brien has some 97 bike schemes world wide for which he has online data in real time - Bikes Data – 4200 bikes, started Nov 2010, all the data– everything – all trips, all times, all stations/docks

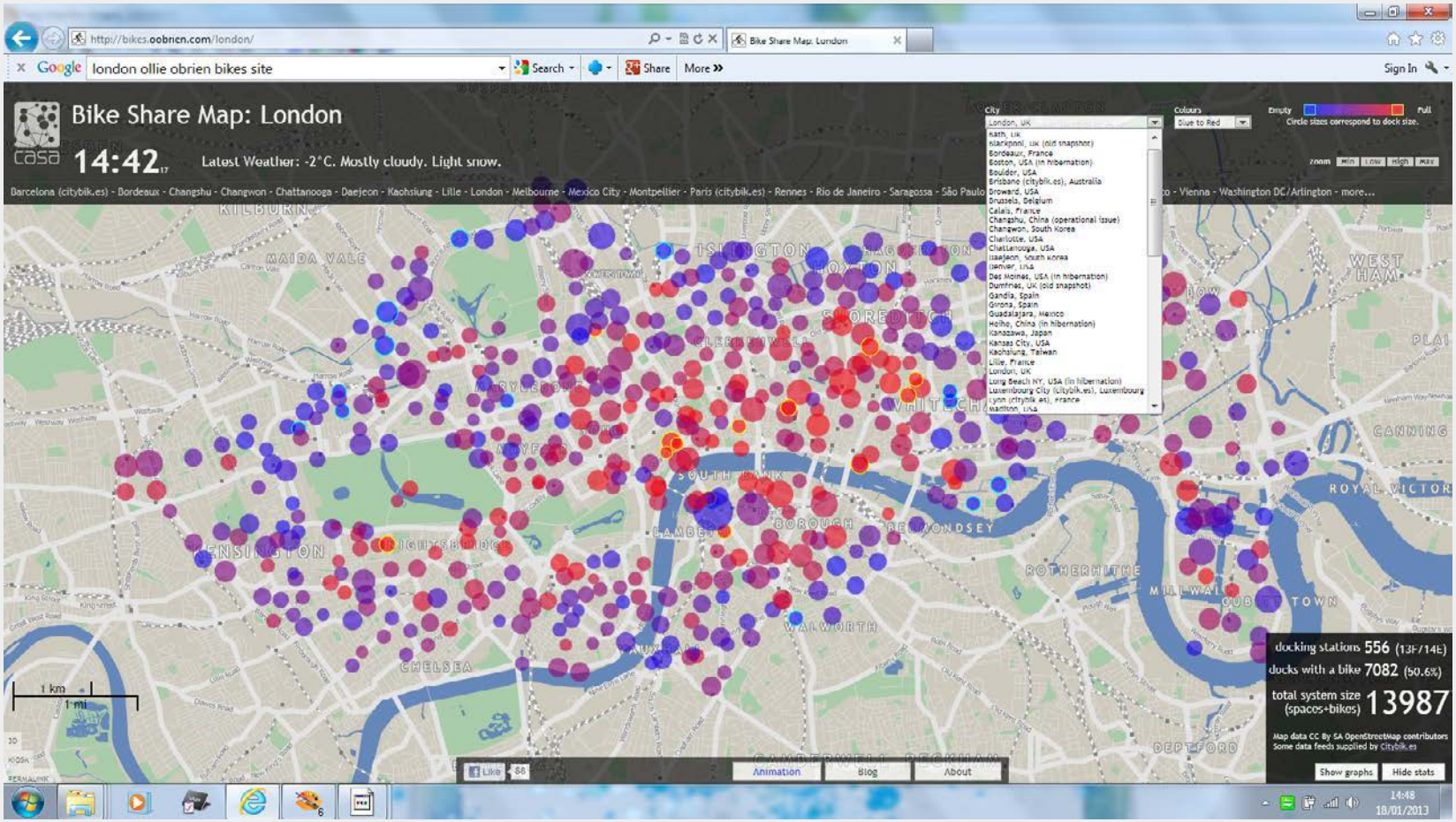


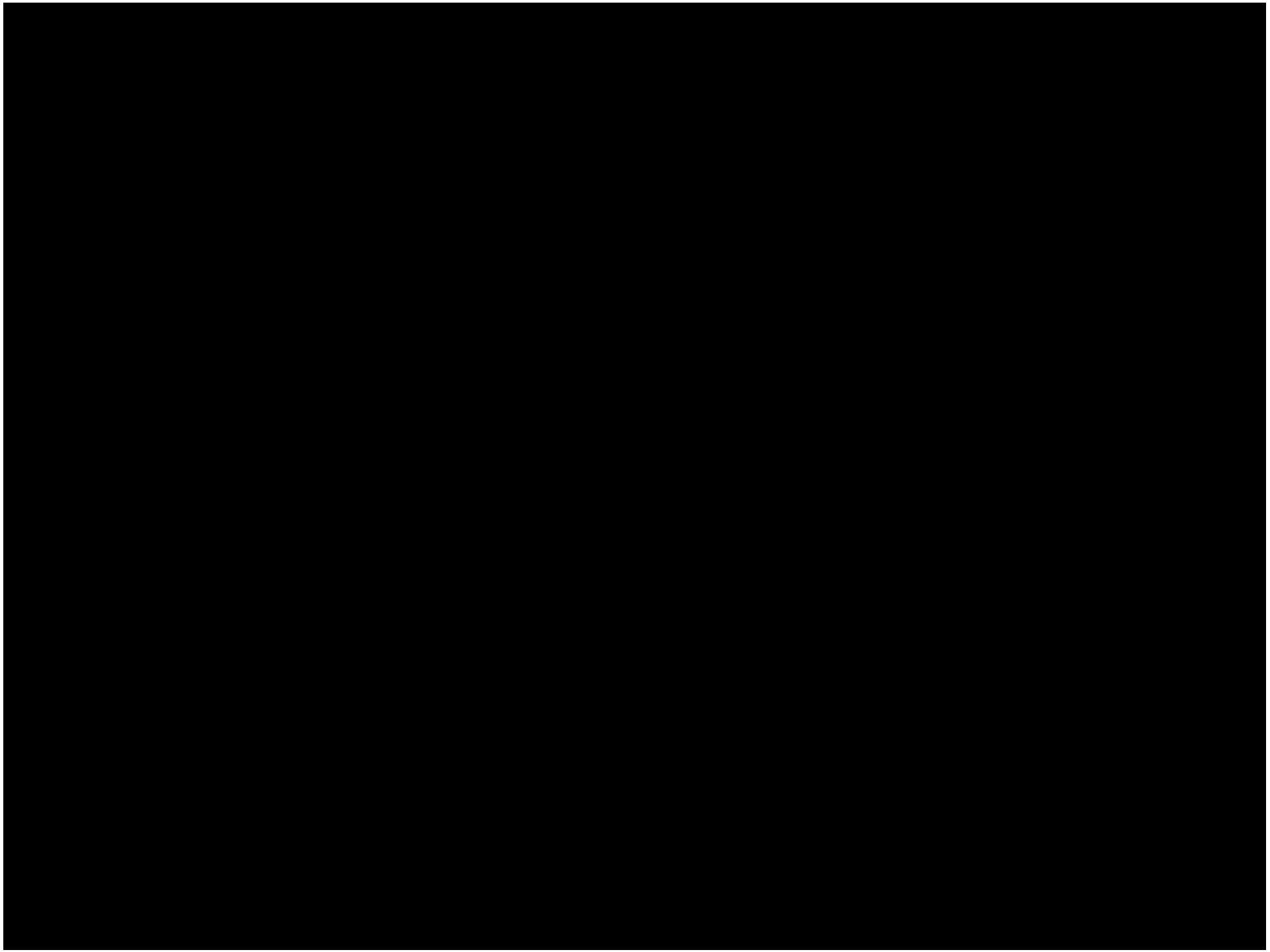




The Website: Real Time Visualisation of Origins and Destinations Activity

http://bikes.oobrien.com/london/





Creating Data: CrowdSourcing

Crowdsourcing is using the power of digital media to create data – this can be done in highly structured ways all the way through to routine ways. They range from online censuses to fully interactive ways of discovering new things online which might be regarded as a form of data

We will look at some online examples. In this we need a broadcast media to make the audience or community aware that we need them to interact and a digital online media to capture their responses – their data

home about login search help casa

MAPTUBE
a place to put maps

BBC RADIO 4

Radio 4: Mapping the Credit Crunch

Welcome to Radio 4 Listeners, below is the Credit Crunch question, simply select an option and then input the first part of your postcode - for example RG11

MapTube will then take your answer and every hour automatically create a map of the nation's mood.

What single factor is hurting you most about the credit crunch?

- Mortgage or Rent
- Petrol
- Food Prices
- Job Security
- Utility Bills
- Not Affected

Enter the first part of your postcode:

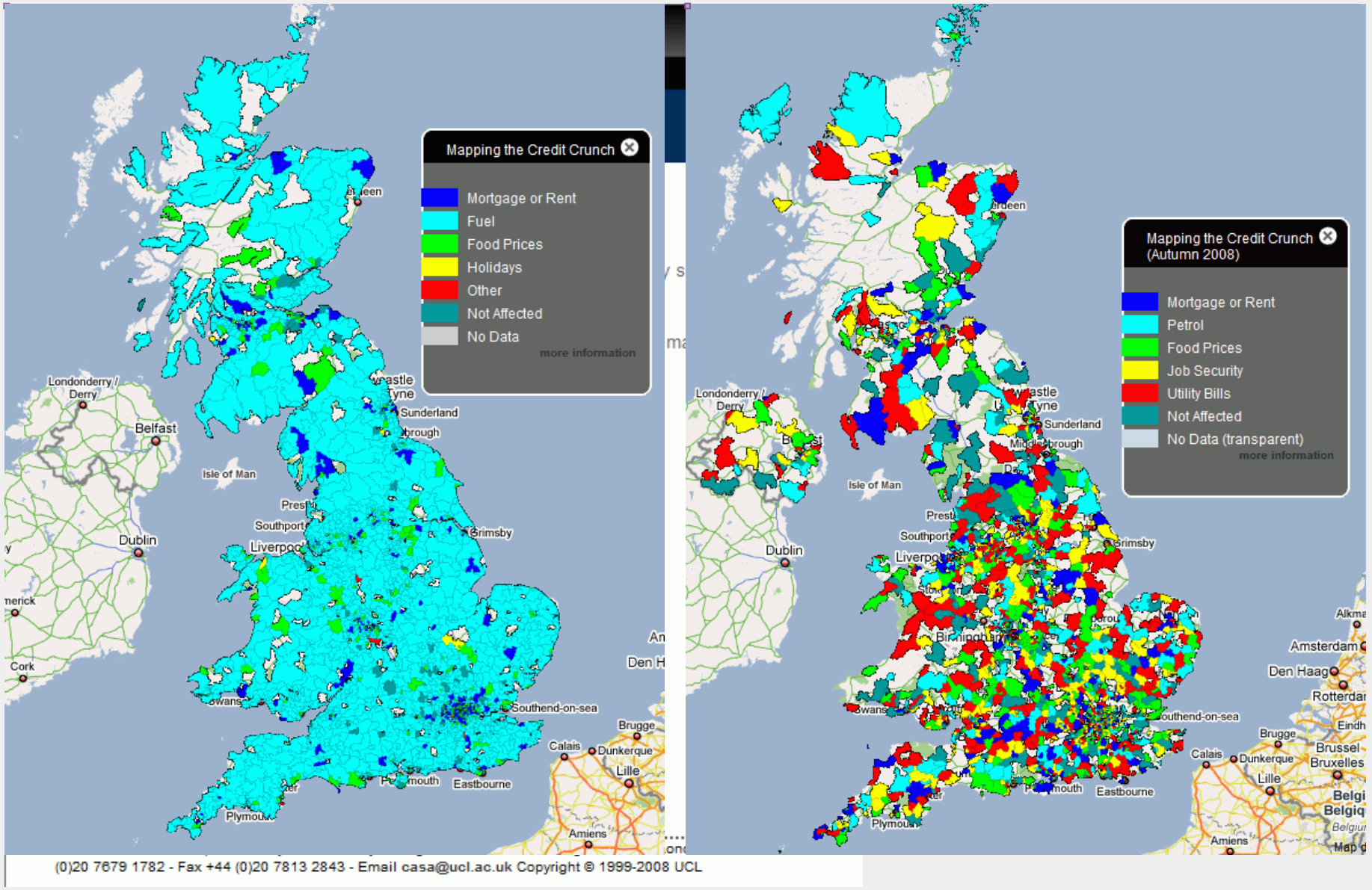
Submit

Centre for Advanced Spatial Analysis - University College London - 1-19 Torrington Place - London - WC1E 7HB - ☎ +44 (0)20 7679 1782 - Fax +44 (0)20 7813 2843 - Email casa@ucl.ac.uk Copyright © 1999-2008 UCL

23,475 responses
April, May, June 2008

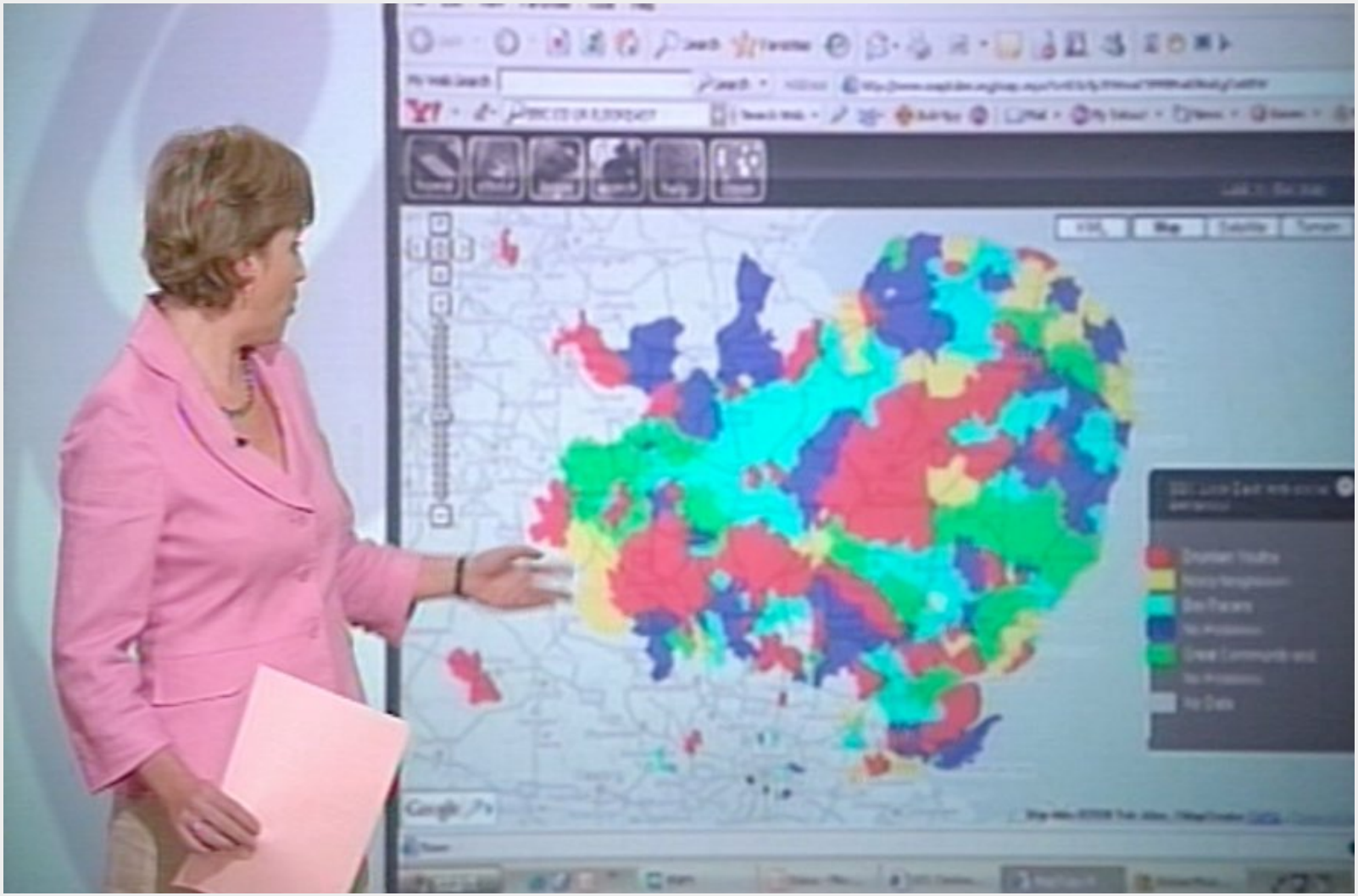
A new credit crunch survey started in October and currently has 3,802 responses.





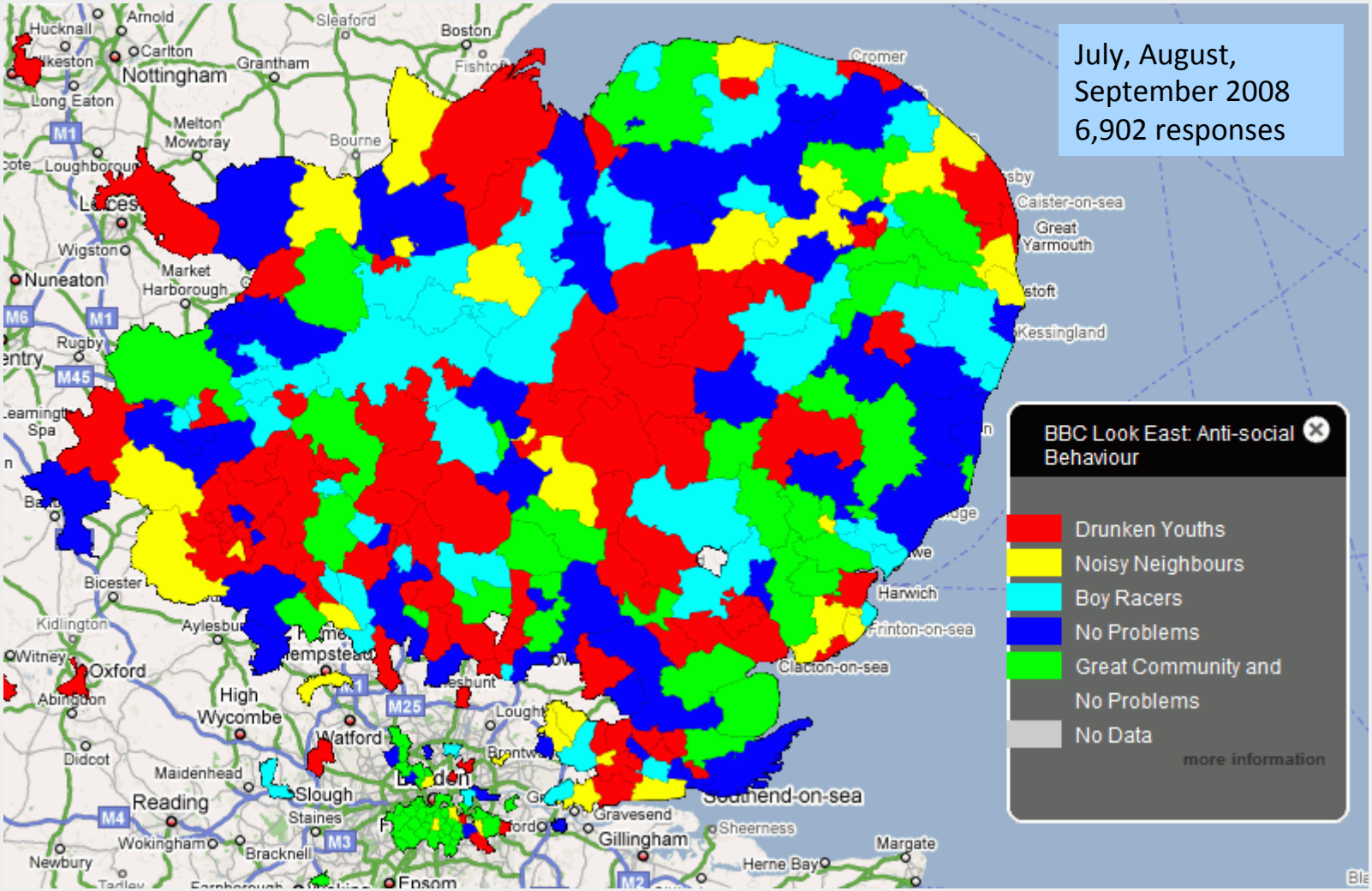
http://www.maptube.org/creditcrunch/

BBC Look East: Anti-Social Behaviour



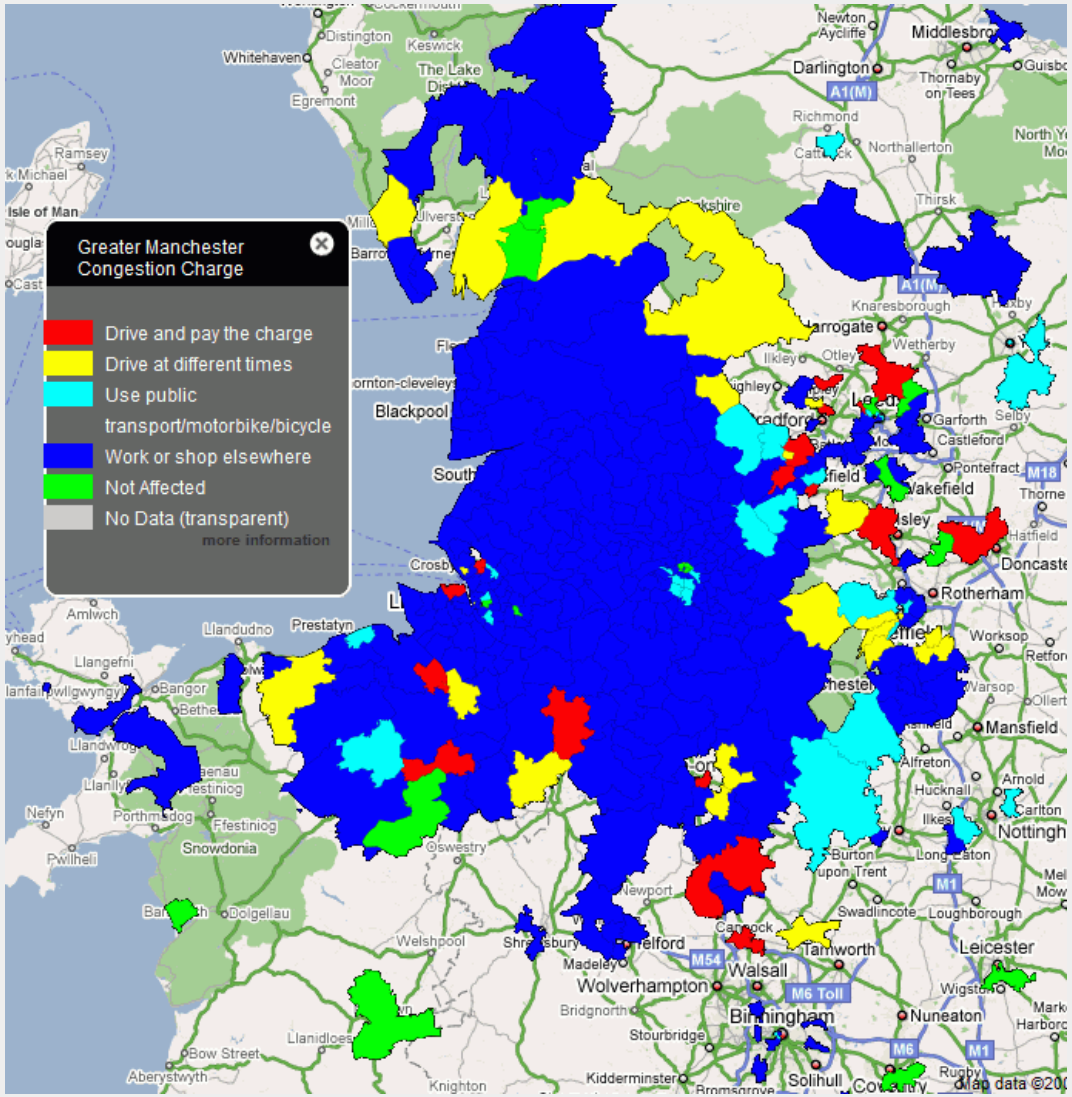
http://www.maptube.org/lookeast

BBC Look East: Anti-Social Behaviour

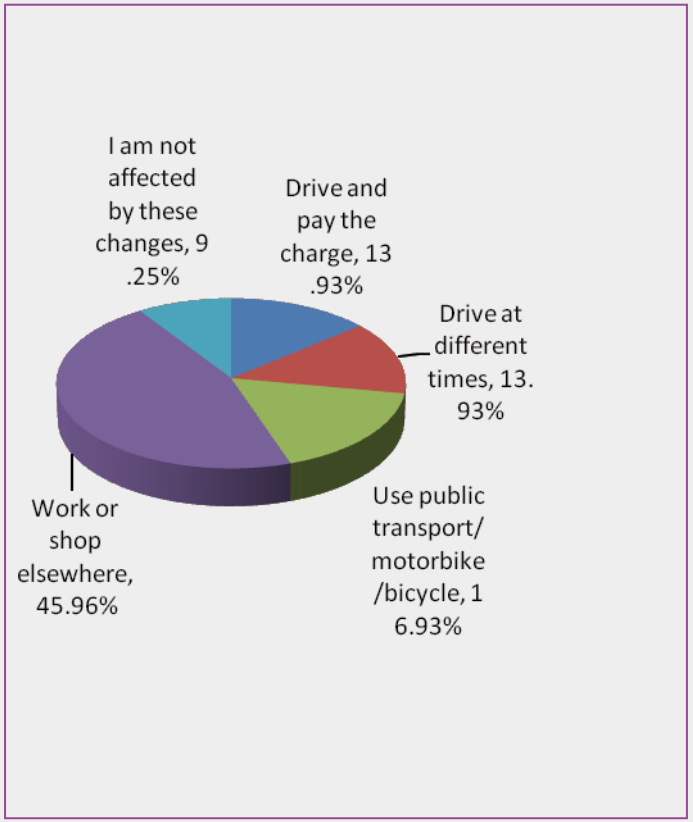


http://www.maptube.org/lookeast

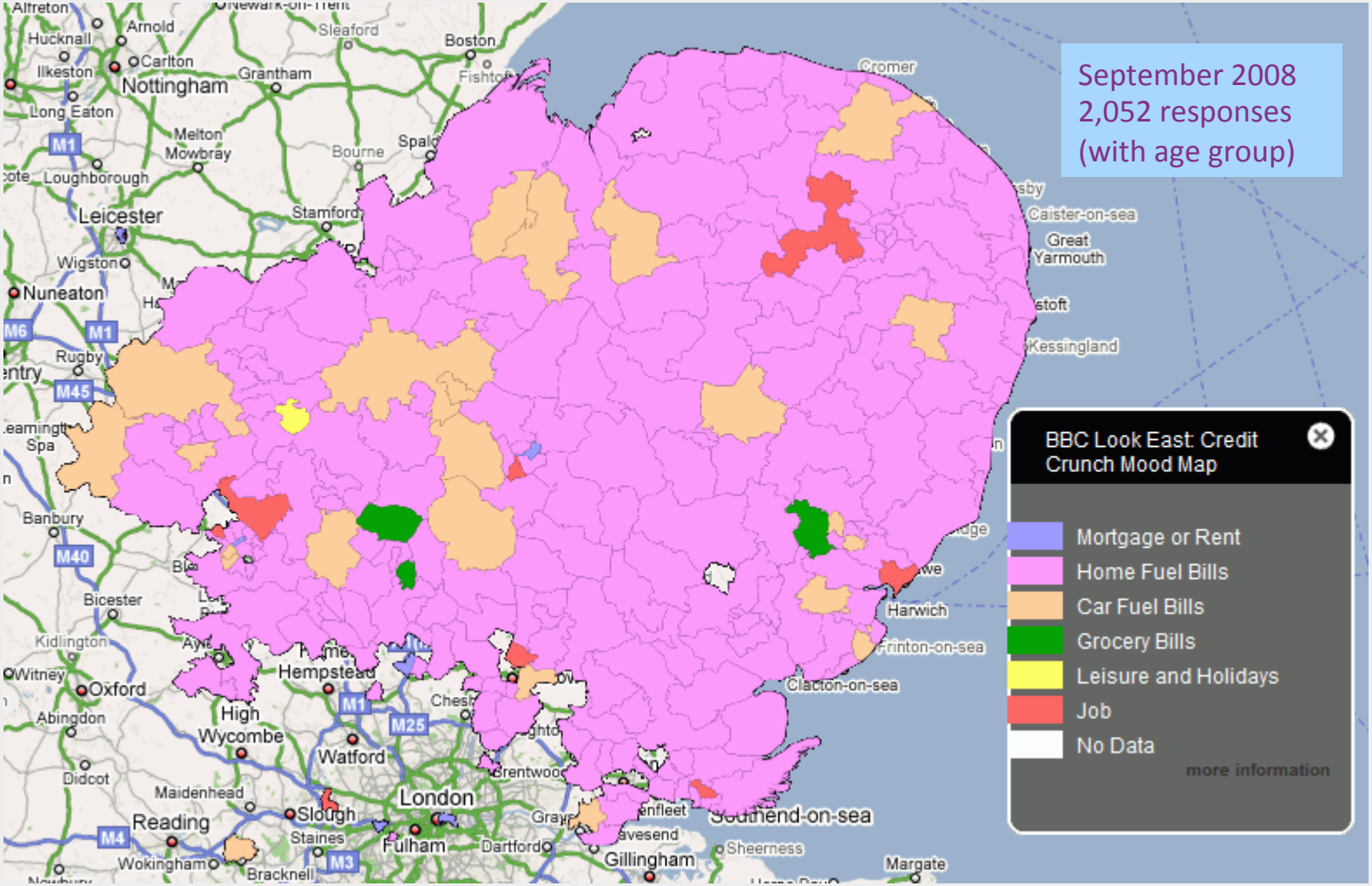
Manchester Congestion Charge



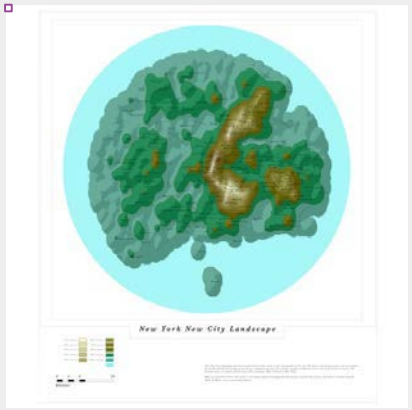
15,902 responses
October to December 2008



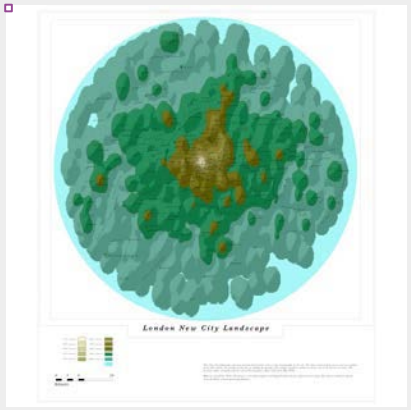
BBC Look East: Credit Crunch



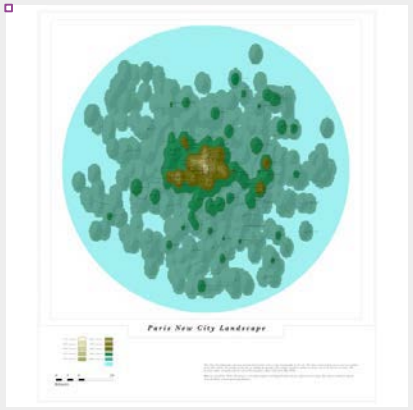
http://www.maptube.org/LookEastCreditCrunch/



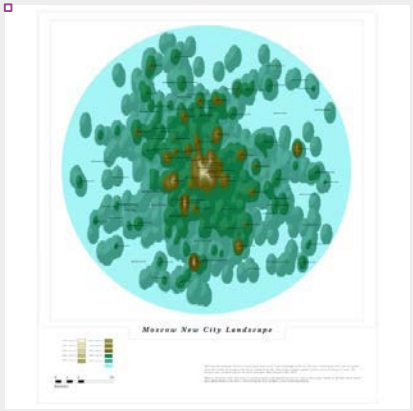
New York



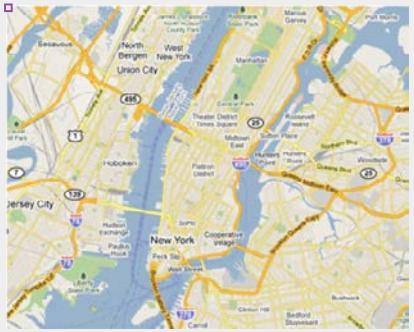
London



Paris

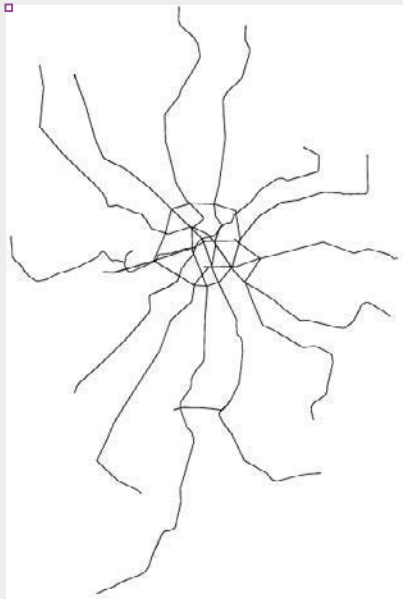


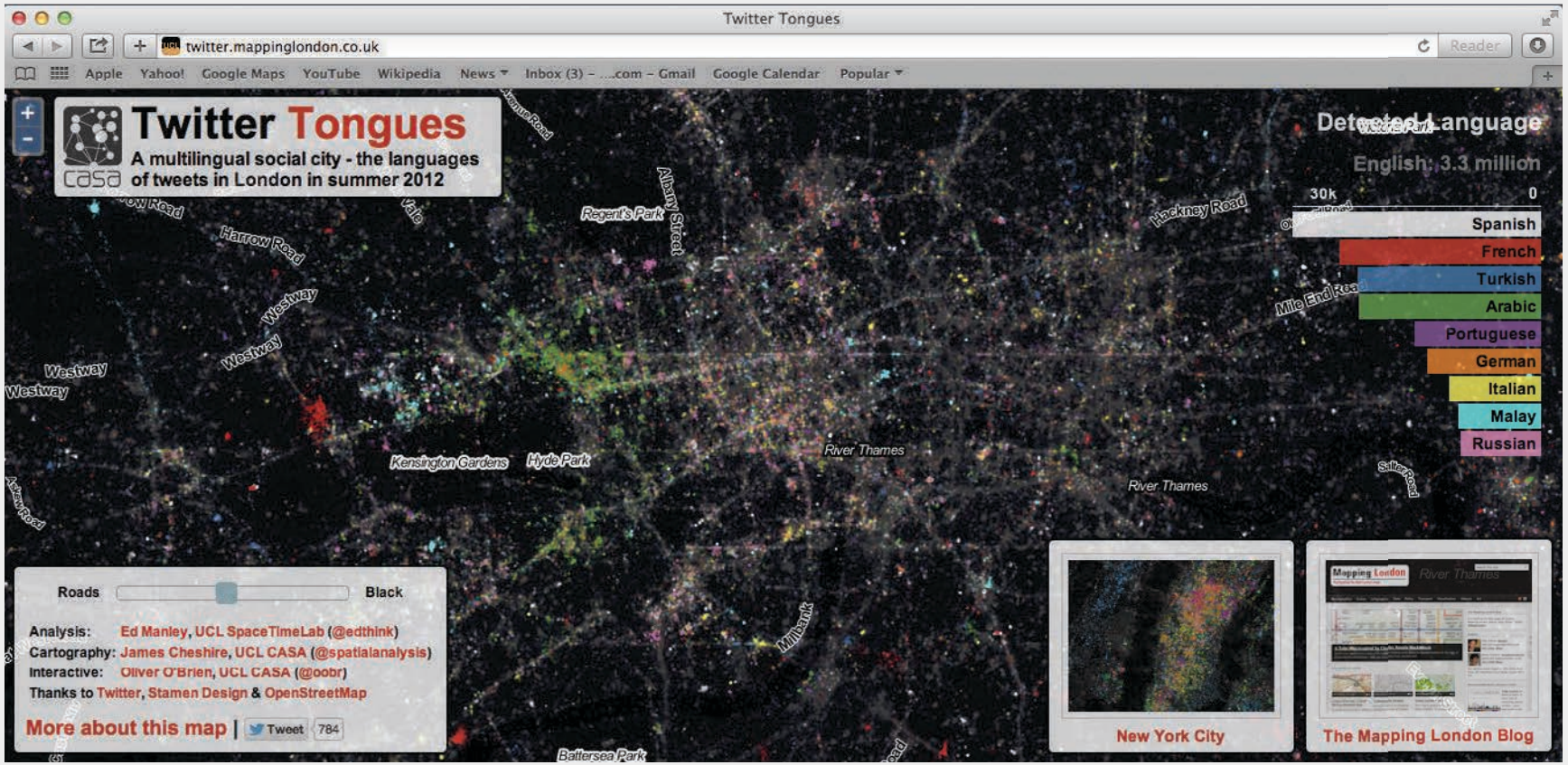
Moscow

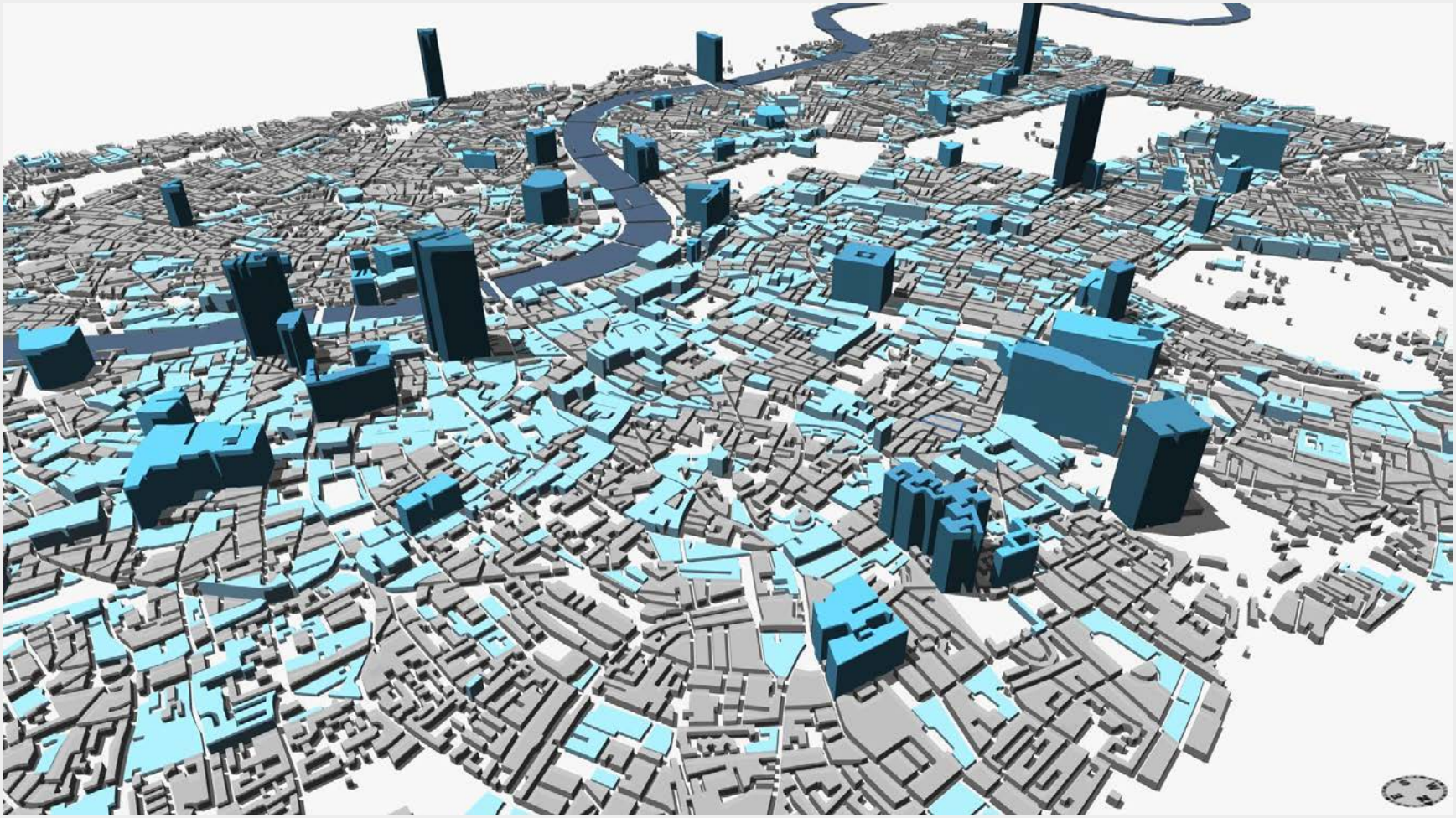




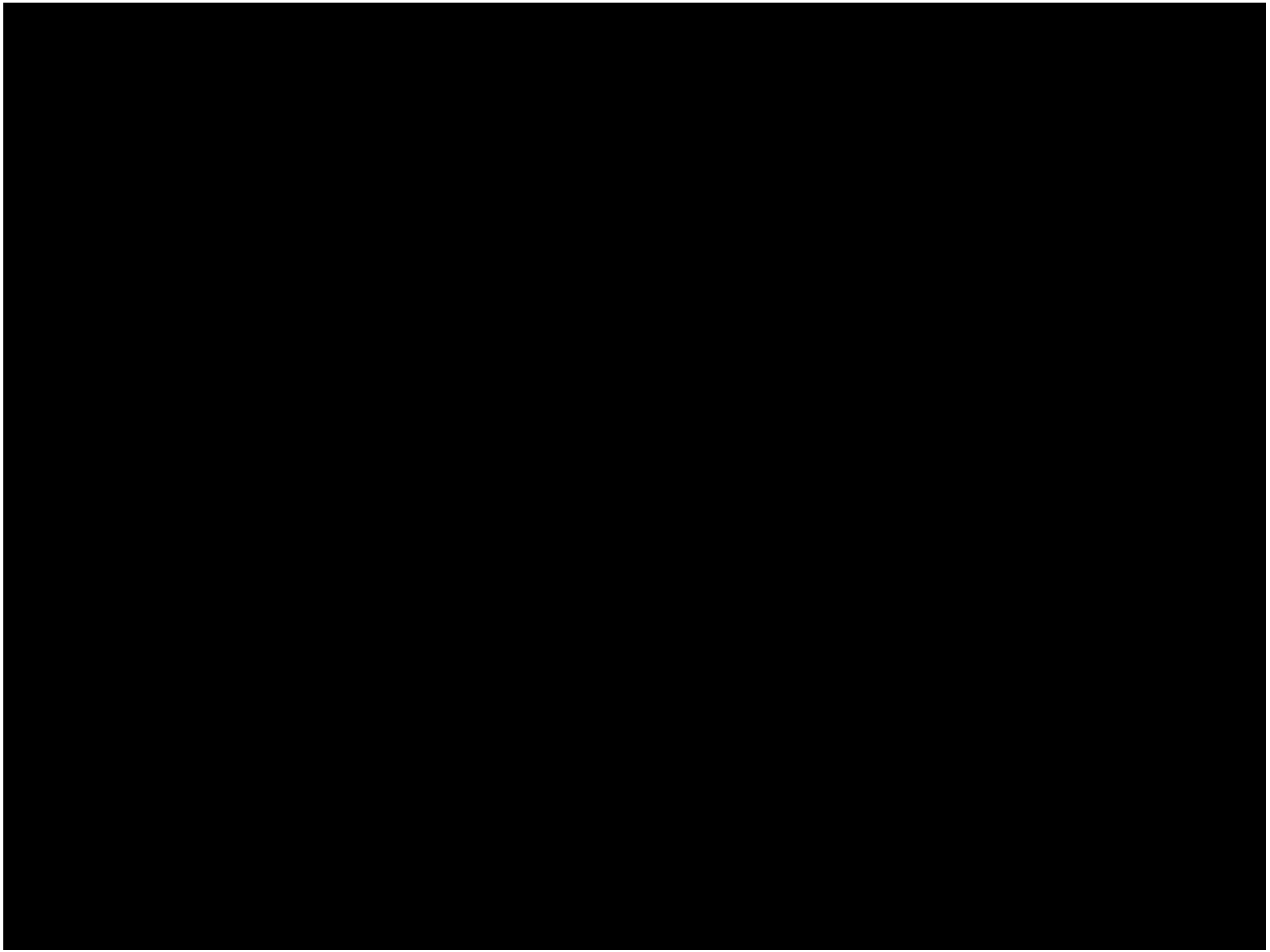





Twitter Tongues
twitter.mappinglondon.co.uk
Reader
Apple
Yahoo!
Google Maps
YouTube
Wikipedia
News
Inbox (3) - ....com - Gmail
Google Calendar
Popular
Twitter Tongues
A multilingual social city - the languages of tweets in London in summer 2012
casa
Detected Language
English: 3.3 million
30k
0
Spanish
French
Turkish
Arabic
Portuguese
German
Italian
Malay
Russian
Harrow Road
Regent's Park
Albany Street
Hackney Road
Mile End Road
Westway
Kensington Gardens
Hyde Park
River Thames
Millbank
Battersea Park
Roads
Black
Analysis: Ed Manley, UCL SpaceTimeLab (@edthink)
Cartography: James Cheshire, UCL CASA (@spatialanalysis)
Interactive: Oliver O'Brien, UCL CASA (@oobr)
Thanks to Twitter, Stamen Design & OpenStreetMap
More about this map | Tweet 784
New York City
The Mapping London Blog



The Age of the Smart City



What Has All This to Do with LED?

- I hear you asking – well I would like you to consider the death of distance that I began with
- But more important in my view – because distance will never die – it will change but place is still key to life and community – but more important is the nature of work that is part of all this
- And the question of what is the city? And what can we do in a global world to influence localities – this is key too
- And in terms of my own focus, then how should we study systems that are so complex and becoming ever more so when they are composed of the very things that we use in their dissection and study and their planning too.

Thanks

Acknowledgements

Melanie Bosredon, Gareth Simons, Roberto Murcio, Richard Milton, Oliver O'Brien, Stephen Gray, Fabian Neuhaus, Jon Reades, Ed Manley, Chen Zhong, Flora Roumpani & Stephan Hugel

http://www.complexcity.info/
http://www.spatialcomplexity.info/
http://blogs.casa.ucl.ac.uk/

m.batty@ucl.ac.uk
@jmichaelbatty

Big Data and the City

Editor: Michael Batty
Centre for Advanced Spatial Analysis, University College London

Built Environment

Volume 42, number 3, September 2016

Big data is everywhere, largely generated by automated systems operating in real time that potentially tell us how cities are performing and changing. A product of the smart city, it is providing us with novel data sets that suggest ways in which we might plan better,anddesignmore sustainable environments. The articles in this issue tell us how scientists and planners are using big data to better understand everything from new forms of mobility in transport systems to new uses of social media. Together, they reveal how visualization is fast becoming an integral part of developing a thorough understanding of our cities.



http://www.spatialcomplexity.info/archives/3026

CONTENTS

Editorial: Big Data, Cities and Herodotus
MICHAEL BATTY
Big Data and the City
MICHAEL BATTY
From Origins to Destinations: The Past, Present and Future of Visualizing Flow Maps
MATTHEW CLAUDEL, TILL NAGEL and CARLO RATTI
Towards a Better Understanding of Cities Using Mobility Data
MAXIME LENORMAND and JOSÉ J. RAMASCO
Finding Pearls in London's Oysters
JONATHAN READES, CHEN ZHONG, ED MANLEY, RICHARD MILTON and MICHAEL BATTY
A Classification of Multidimensional Open Data for Urban Morphology
ALEXANDROS ALEXIOU, ALEX SINGLETON and PAUL A. LONGLEY
User-Generated Big Data and Urban Morphology
A.T. CROOKS, A. CROITORU, A. JENKINS, R. MAHABIR, P. AGOURIS and A. STEFANIDIS
Sensing Spatiotemporal Patterns in Urban Areas: Analytics and Visualizations Using the Integrated Multimedia City Data Platform
PIYUSHIMITA (VONU) THAKURIAH, KATARZYNA SILA-NOWICKA, and JORGE GONZALEZ PAULE
Playful Cities: Crowdsourcing Urban Happiness with Web Games
DANIELE QUERCIA
Big Data for Healthy Cities: Using Location-Aware Technologies, Open Data and 3D Urban Models to Design Healthier Built Environments
HARVEY J. MILLER and KRISTIN TOLLE
Improving the Veracity of Open and Real-Time Urban Data
GAVIN MCARDLE and ROB KITCHIN
Wise Cities: 'Old' Big Data and 'Slow' Real Time
FABIO CARRERA
Collecting and Visualizing Real-Time Urban Data Through City Dashboards
STEVEN GRAY, OLIVER O'BRIEN AND STEPHAN HÜGEL

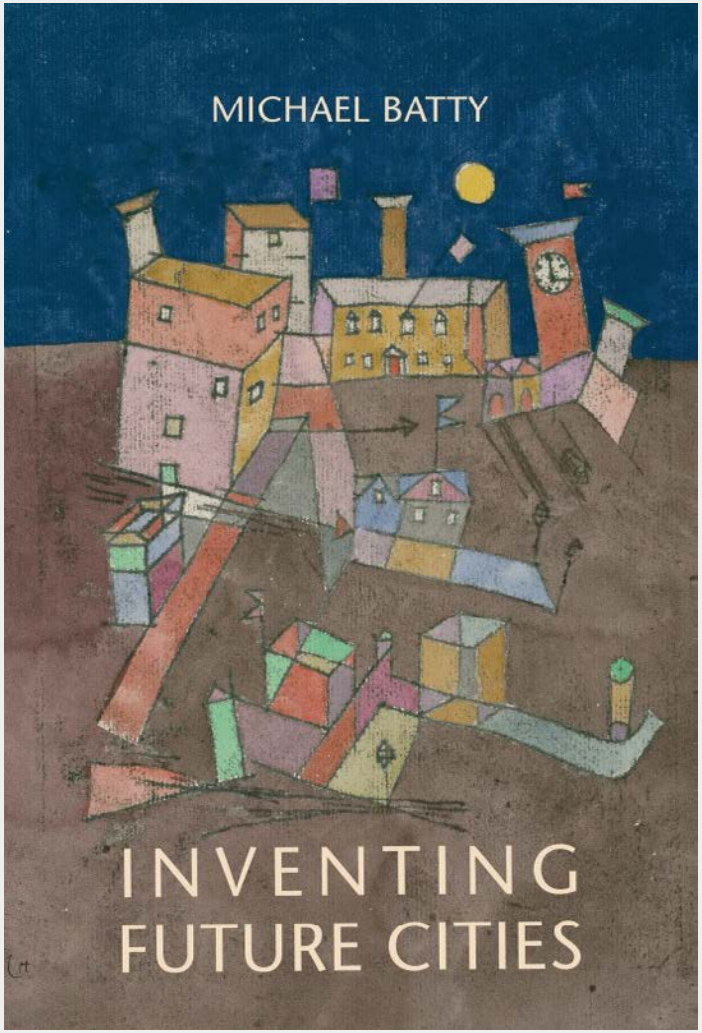
Edited by David Banister and Stephen Marshall with Reviews Editor Lucy Natarajan, *Built Environment* is published four times a year, both print and online. For more information or to order a copy of this issue contact:

Alexandrine Press, 1 The Farthings, Marcham, Oxfordshire OX13 6QD

phone : **01865 391518** fax : **01865 391687**

e-mail : **alexandrine@rudkinassociates.co.uk**

visit : **www.alexandrinepress.co.uk**



MIT Press, Cambridge, MA, 2018

The Age of the Smart City