中山大学地理科学与规划学院
School of Geography and Planning, Sun Yat-Sen University

University College London

Big Data and Urban Analytics:
New Tools for Planning the Smart City.

Michael Batty

m.batty@ucl.ac.uk
@jmichaelbatty

25 April 2017

http://www.spatialcomplexity.info/

I was first here in 1986.

I gave a lecture about urban modelling


海报
英国威尔斯理工大学Batty教授将于十一月二十四日下午二时半在管理楼202室作题为"城市模拟系统研究的进展"的报告，请经地八四、八三级学生及人文地理研究生参加，并欢迎其他感兴趣的师生光临。
中山大学地理系
一九八六年十一月二十一日
特别欢迎有独特学术

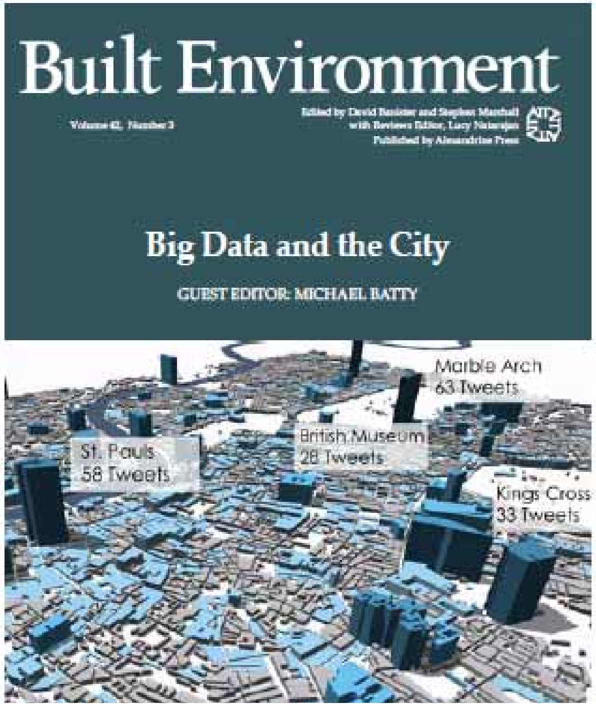
Big Data and the City

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

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**

Outline

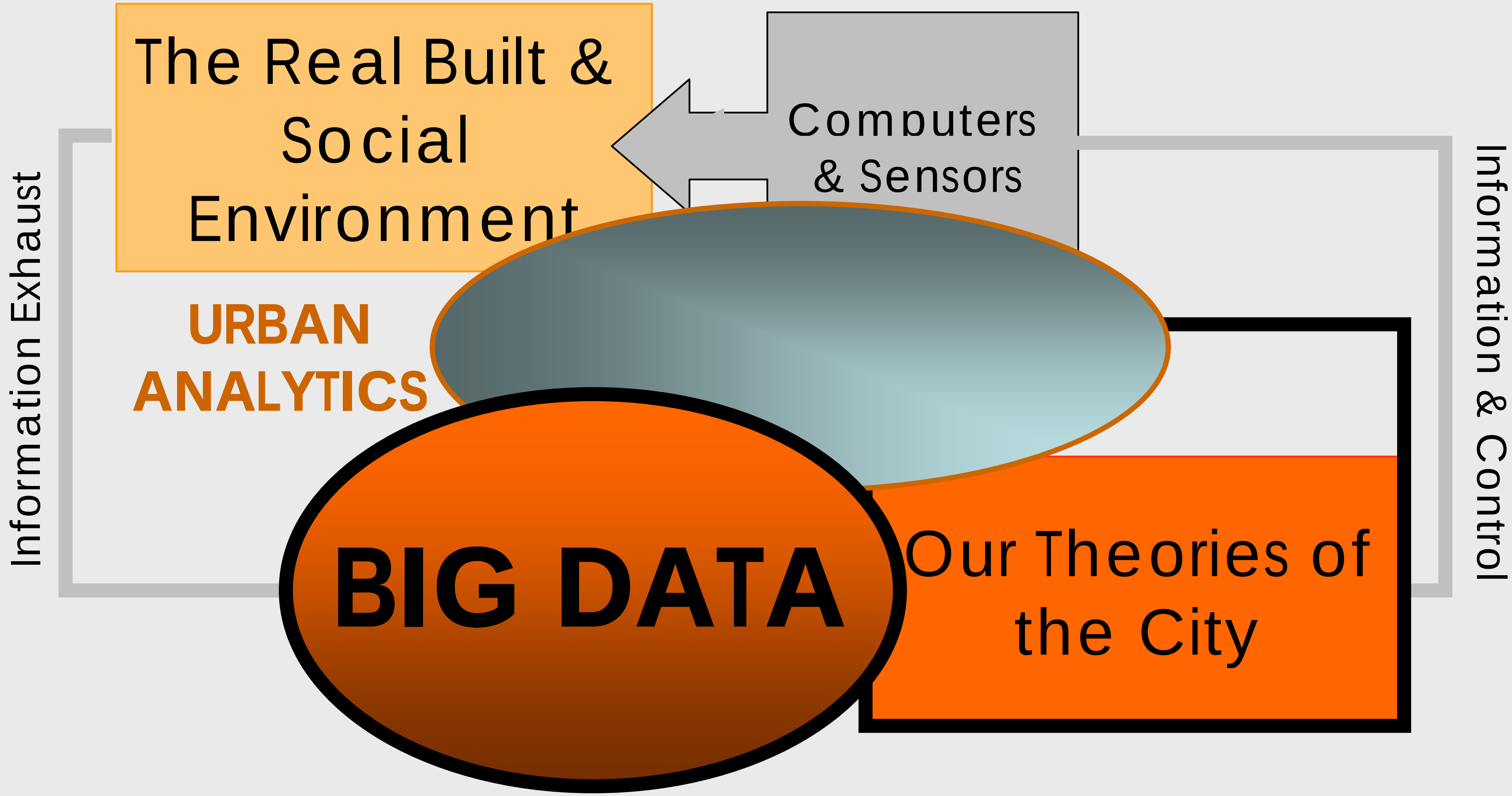
- The Smart City
- A Short History of Big Data: How Big is Big?
- Mobility, Transit, & Real-Time Streaming: The Oyster Card Data Set
- Learning about Mobility from the Data

  Variabilities – Heterogeneity and Travel Profiles

  Disruptions – Signal Failures, Stalled Trains

  Variable Locational Dynamics of Demand
- Related Real -Time Data: Bikes, Social Media
- What Can We Learn: The Limits to Big Data

Our Framework once again



Some Basic Points Again

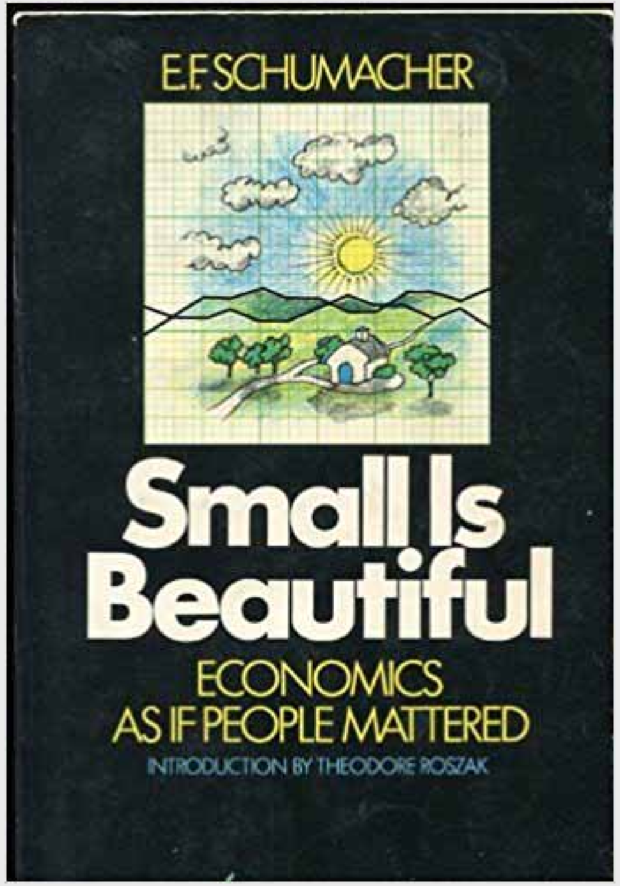
- The way we access the smart city is through technologies that let us generate and use data and its useful equivalent – **<u>information</u>** (data) is key
- Access through **<u>mobile</u>** and **<u>fixed devices</u>** like phones, smart cards, through fixed sensors which record transactions and so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

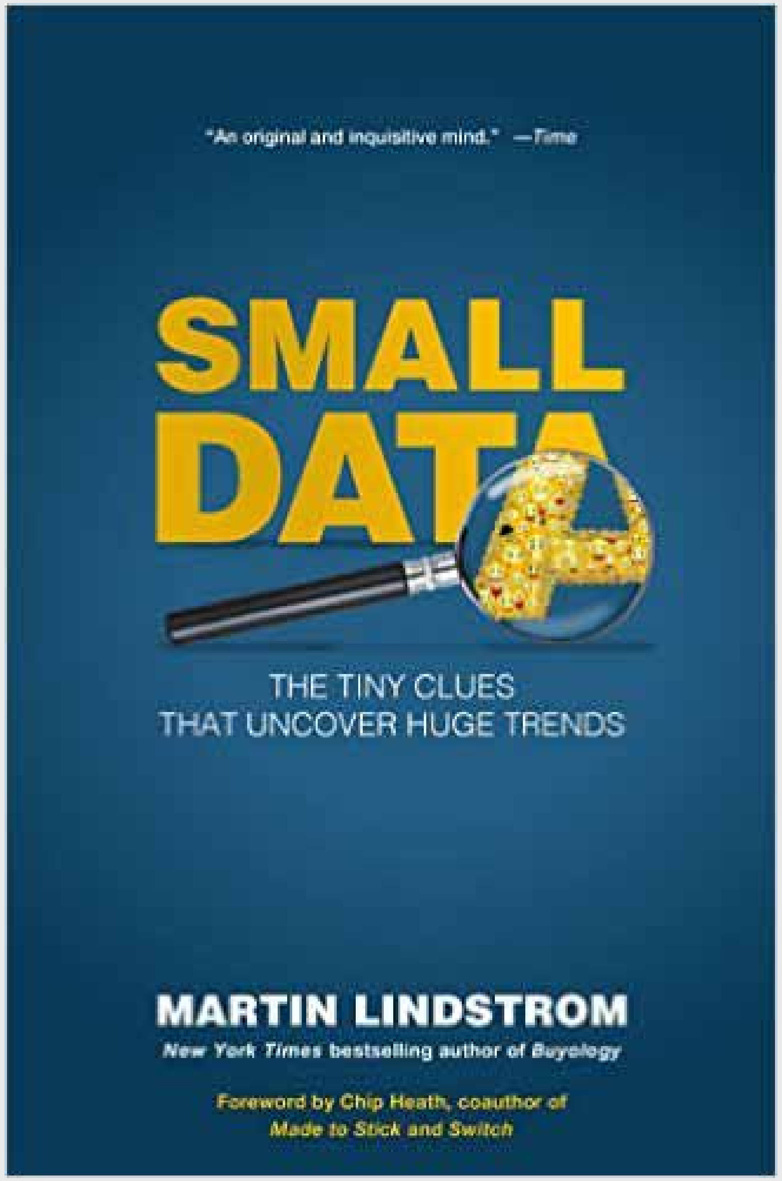
How Big is Data? Big Can Be Small & Small Big

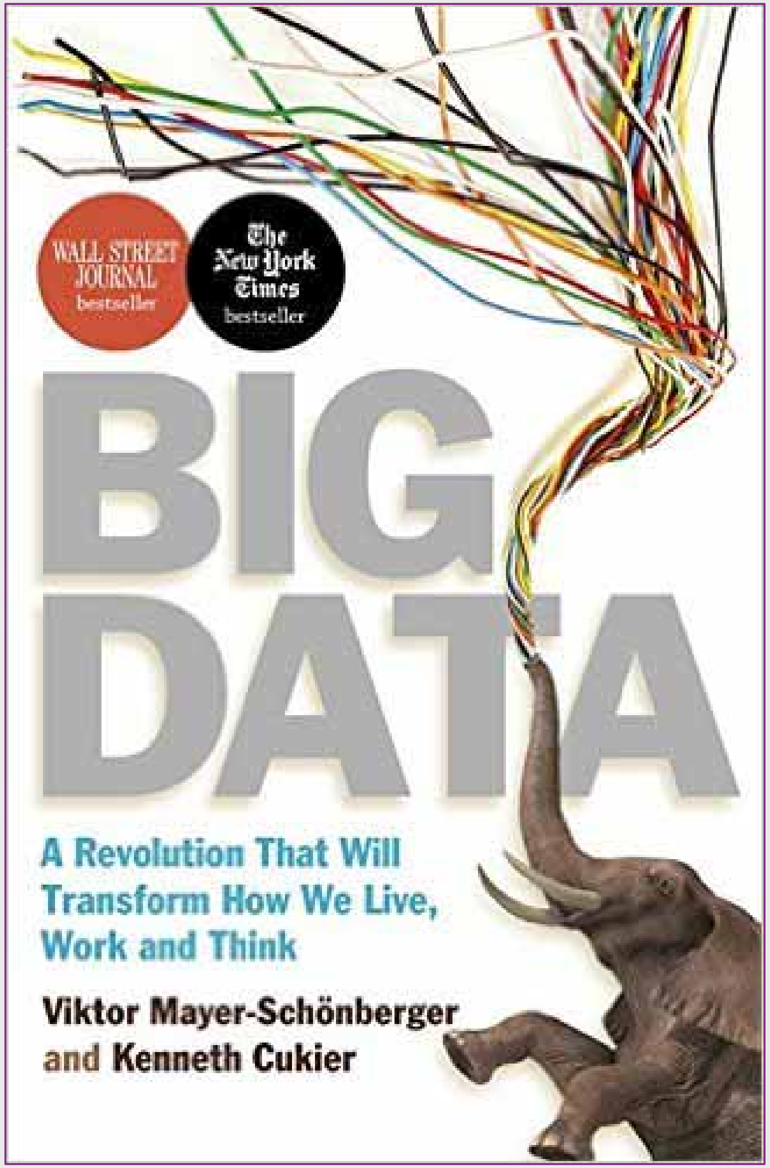
- Data is big with respect to its volume. I know there are other definitions – velocity, variety etc. but to me, data is big if it requires large use of computer memory implying volume.
- The conventional definition in business is the Five V's – volume, velocity, variety, veracity, value
- In cities, 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

- But there is controversy about how big.
- Big is not necessarily beautiful – small is beautiful – this was the watch word of the 1970s
- So we need to be careful

We can still develop good ideas and good theory with small data.
In fact the idea that the truth or even the path to progress lies in big data is problematic; ***there are limits to machine learning***




"An original and inquisitive mind." —Time
SMALL
DATA
THE TINY CLUES
THAT UNCOVER HUGE TRENDS
MARTIN LINDSTROM
New York Times bestselling author of Buyology
Foreword by Chip Heath, coauthor of
Made to Stick and Switch


WALL STREET JOURNAL bestseller
The New York Times bestseller
BIG
DATA
A Revolution That Will
Transform How We Live,
Work and Think
Viktor Mayer-Schönberger
and Kenneth Cukier

Examples: Dublin 1837, Ireland 1888, London 1953

Posts About Complexity Fractals Networks Simulations Media Books Articles Editorials

← Visualising Fast Flows

Movies Are Now Online →

The Oldest Flow Map

Posted on June 27, 2011 by Michael Batty

Tweet 3

.... according to the great cartographer Arthur Robinson, the two maps of traffic between Dublin and the rest of Ireland by Lt. Harness of the British Army in 1837, are the oldest. Mapped for the Irish Railway Commissioners prior to construction of the the railway. Are these actually the first?

Be Sociable, Share!

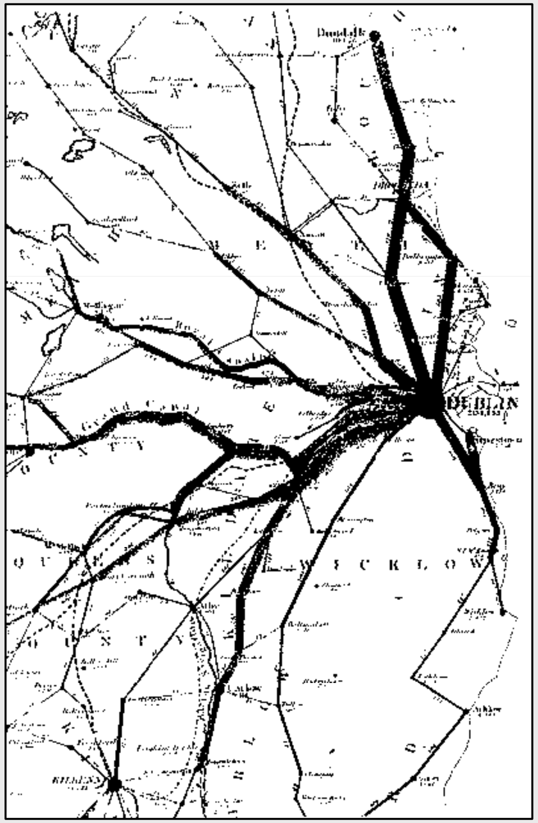
Tweet 3 Like 0 +1 0 Share Submit

About Michael Batty

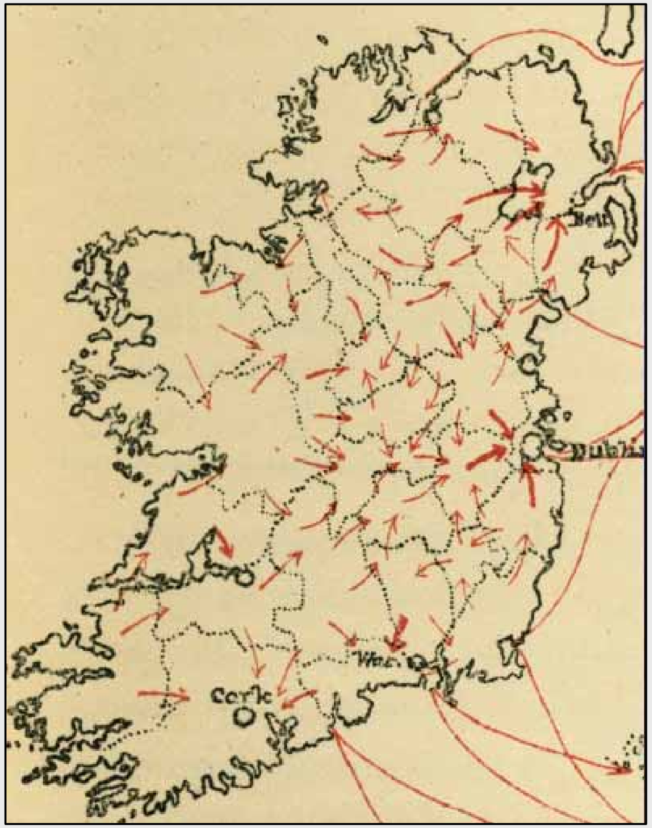
I chair CASA at UCL which I set up in 1995. I am Bartlett Professor In UCL.

View all posts by Michael Batty →

Examples: Dublin 1837, Ireland 1888, London 1955



Harness, 1837



Ravenstein 1888

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

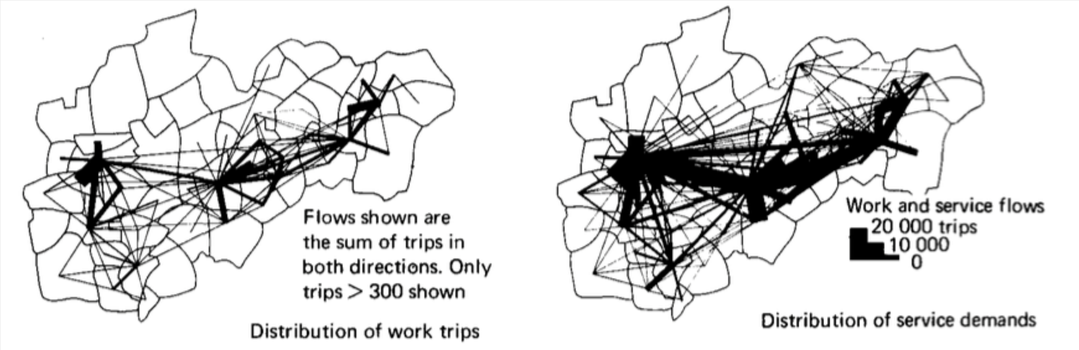
Locations and Interactions: Flow Systems in Cities

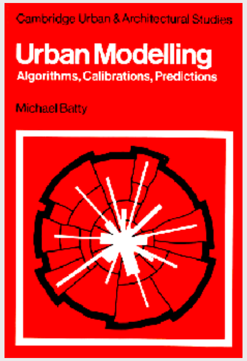
Elsewhere I have argued that we should treat cities as flow systems – as networks. This has been a focus for a long time in transport and land use and we have always been up against the problem of big data.

So let me begin my illustration of this dilemma and how we are thinking about it with some problems that have very small data. Problems of spatial interaction where our numbers of locations is small < 100, ~ 50

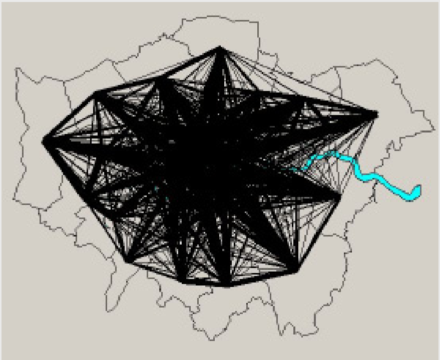
Understanding and Visualising Flows

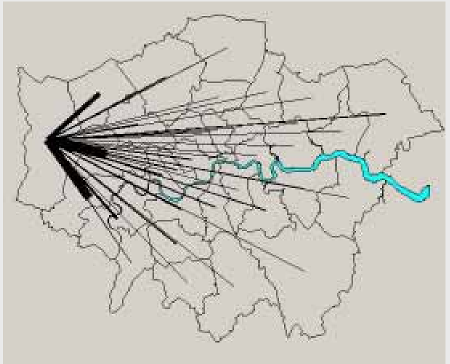
An early model circa 1967-8 Central and NE Lancs



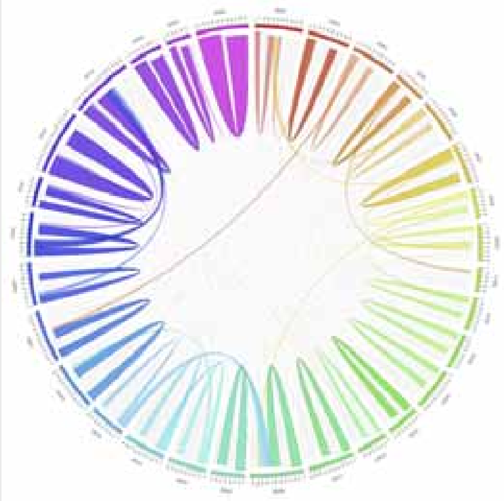


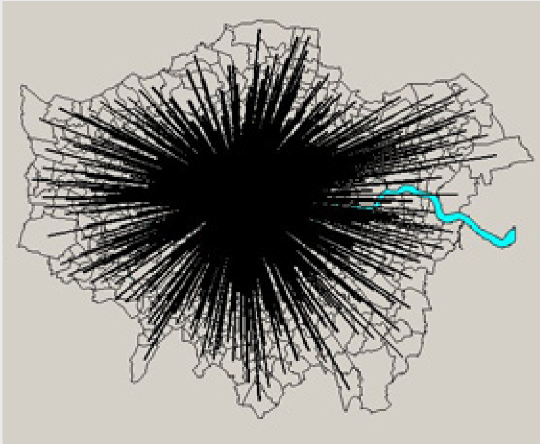
M. Batty (1976)
Urban Modelling
Cambridge UP

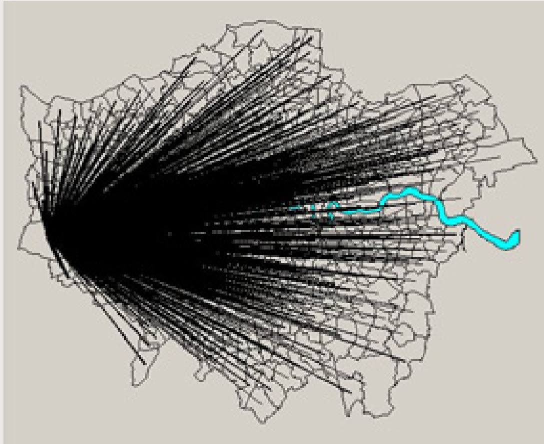




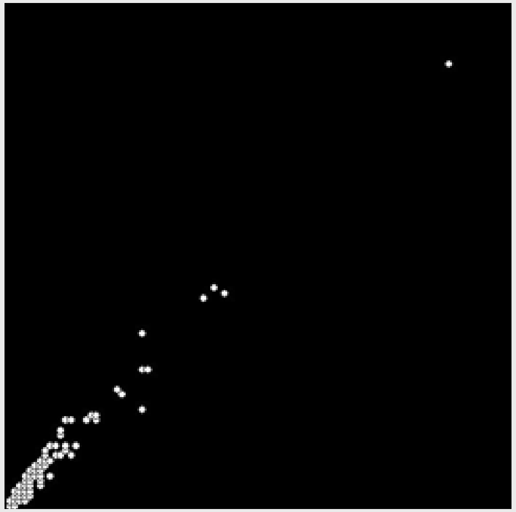
$n^2=33^2=1089$, not so big but hard to visualise

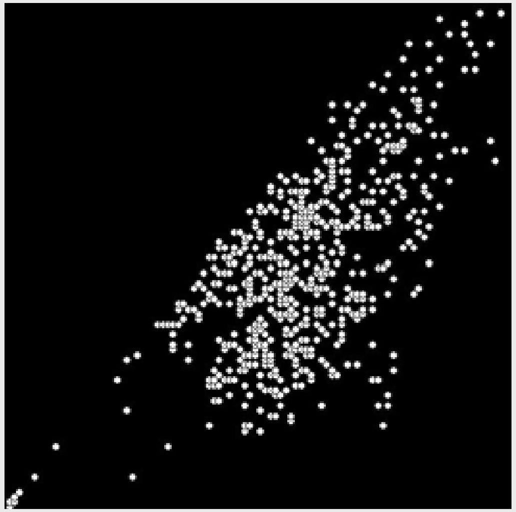


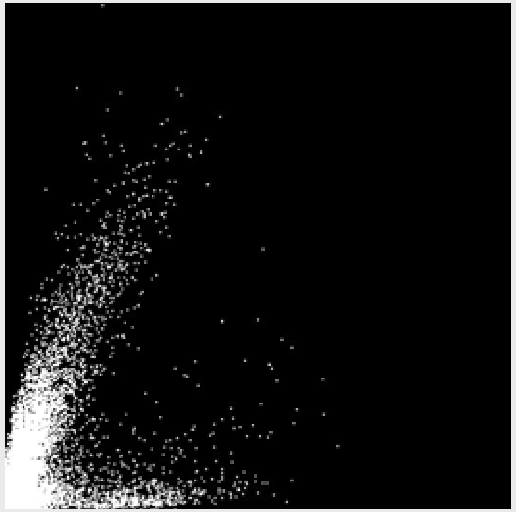


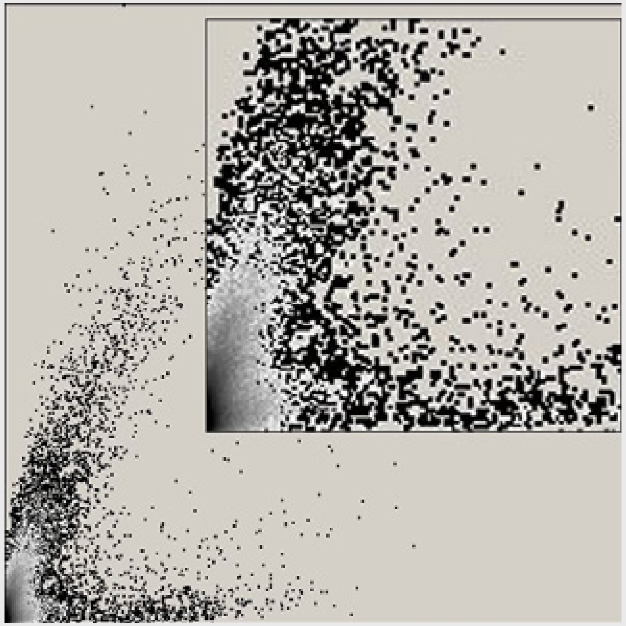


$n^2=633^2=400,689$, bigger but impossible to visualise





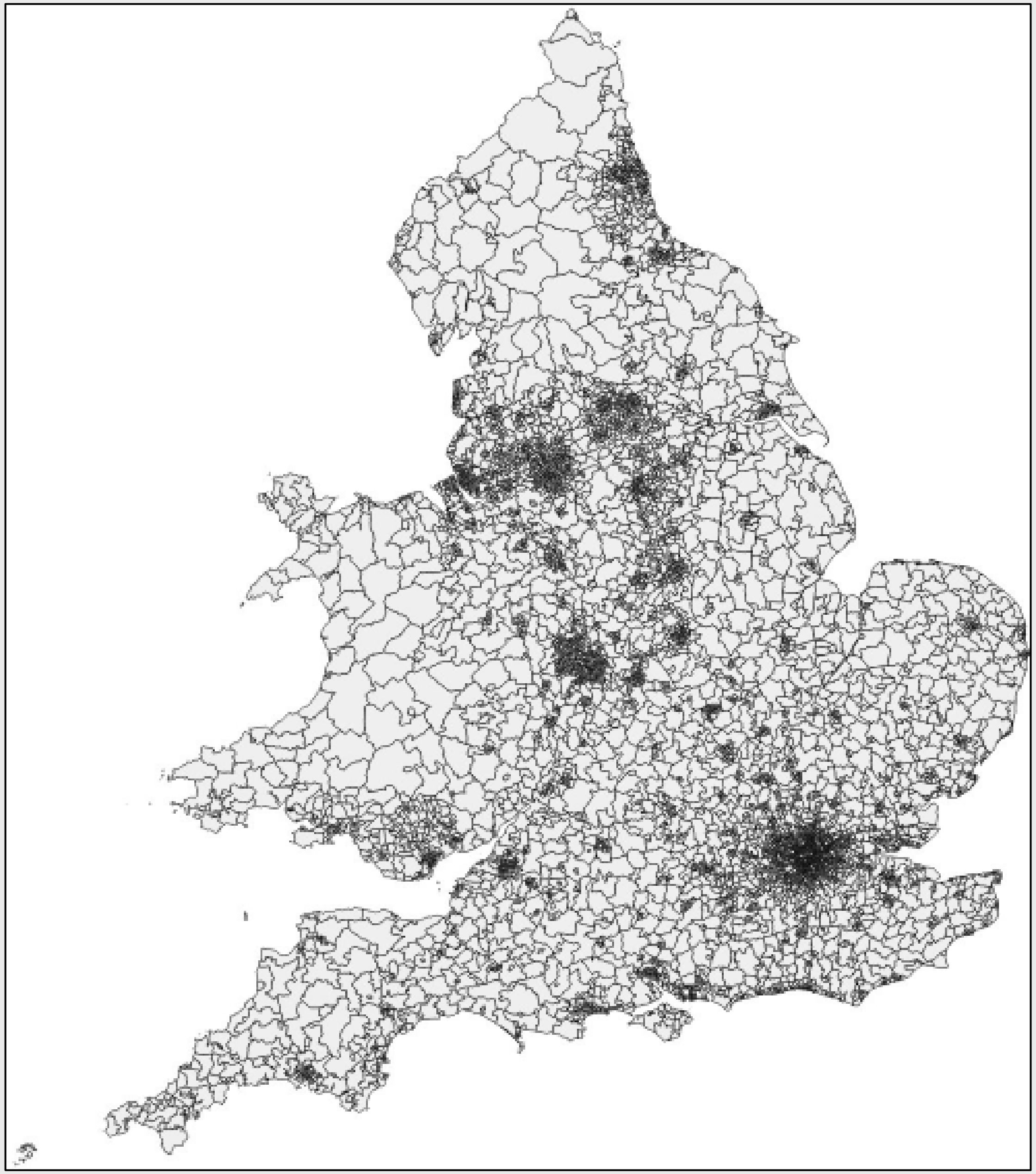




Even our statistics breaks down when we get large numbers like over several thousand as you can see on the left and above right for 400K data points where the pattern is highly convoluted. This is from a gravity model.

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.



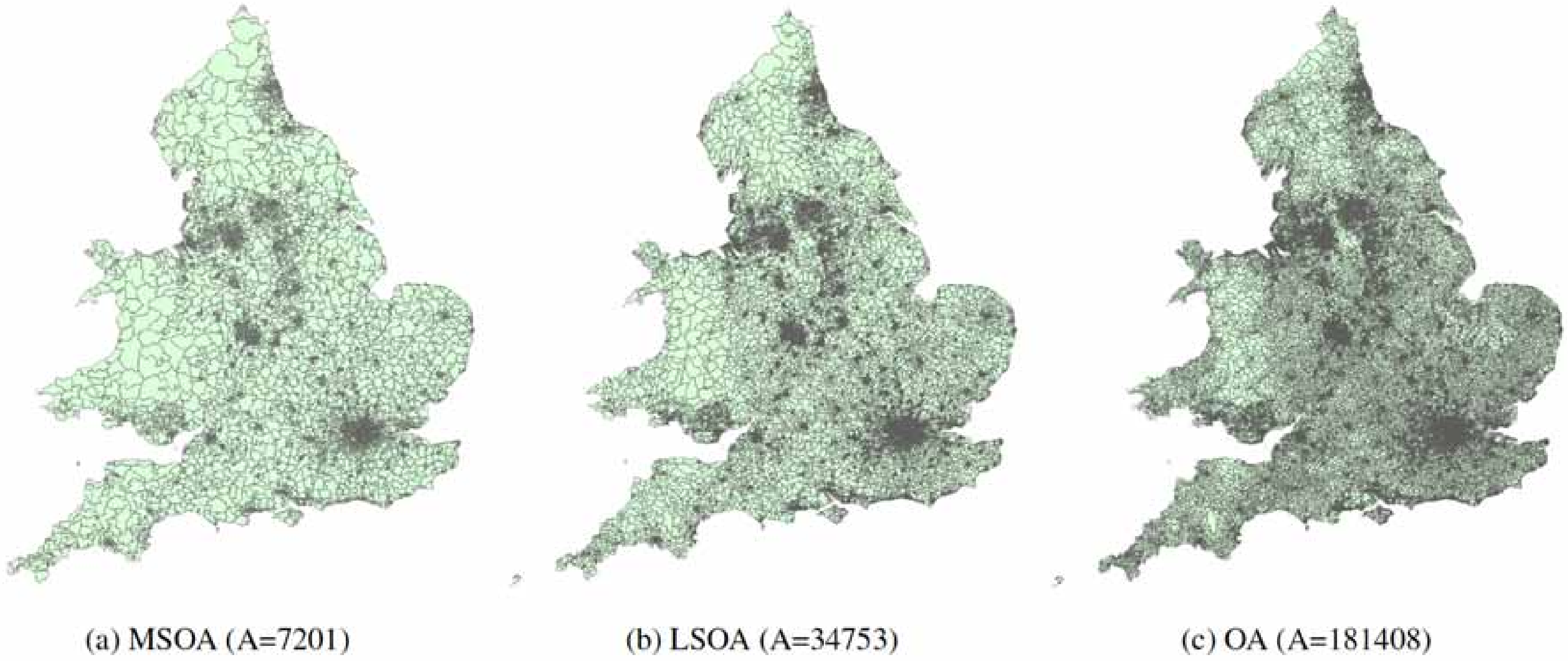
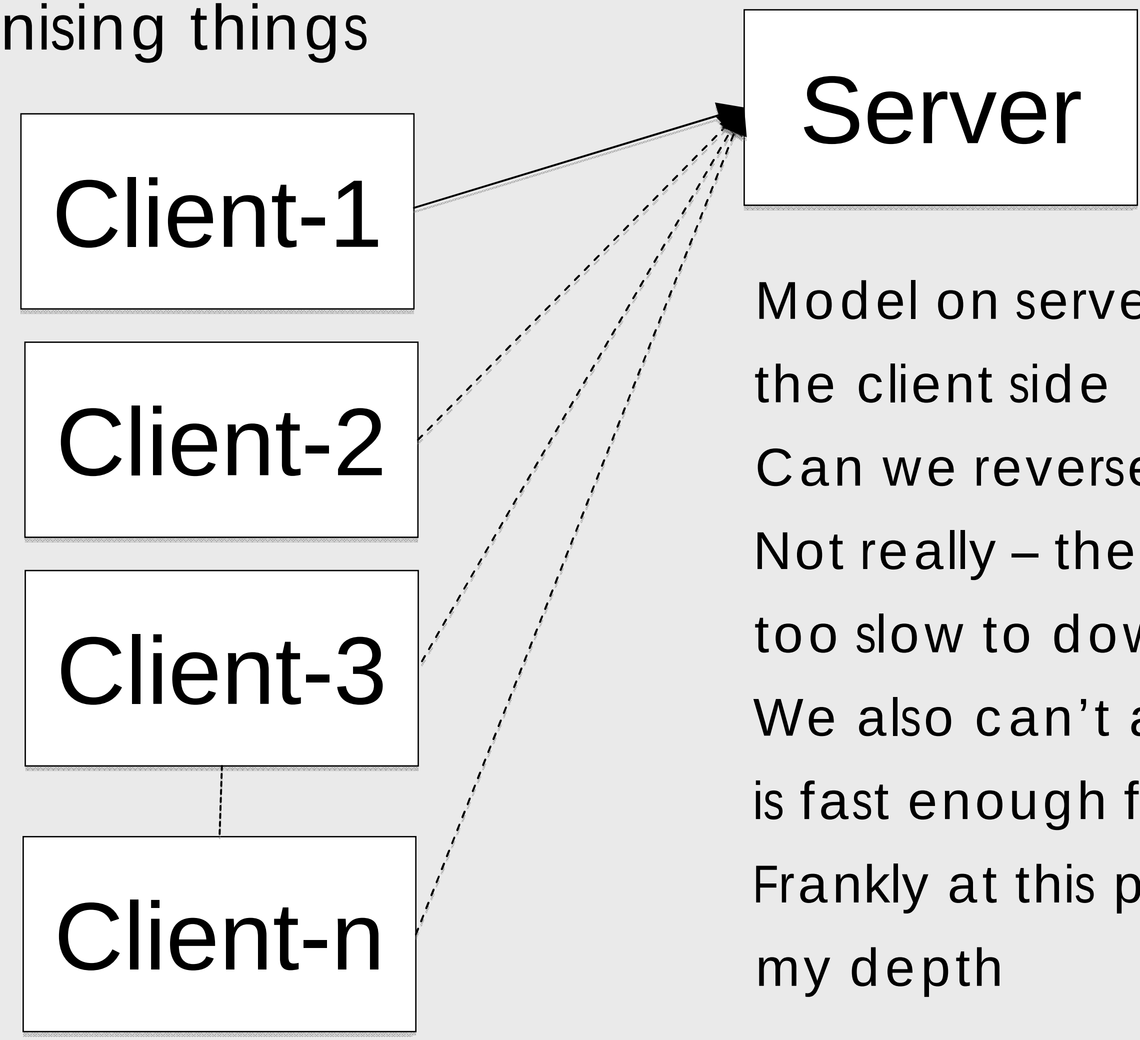


Figure 8.2: ONS Geographies for MSOA, LSOA and OA levels.

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



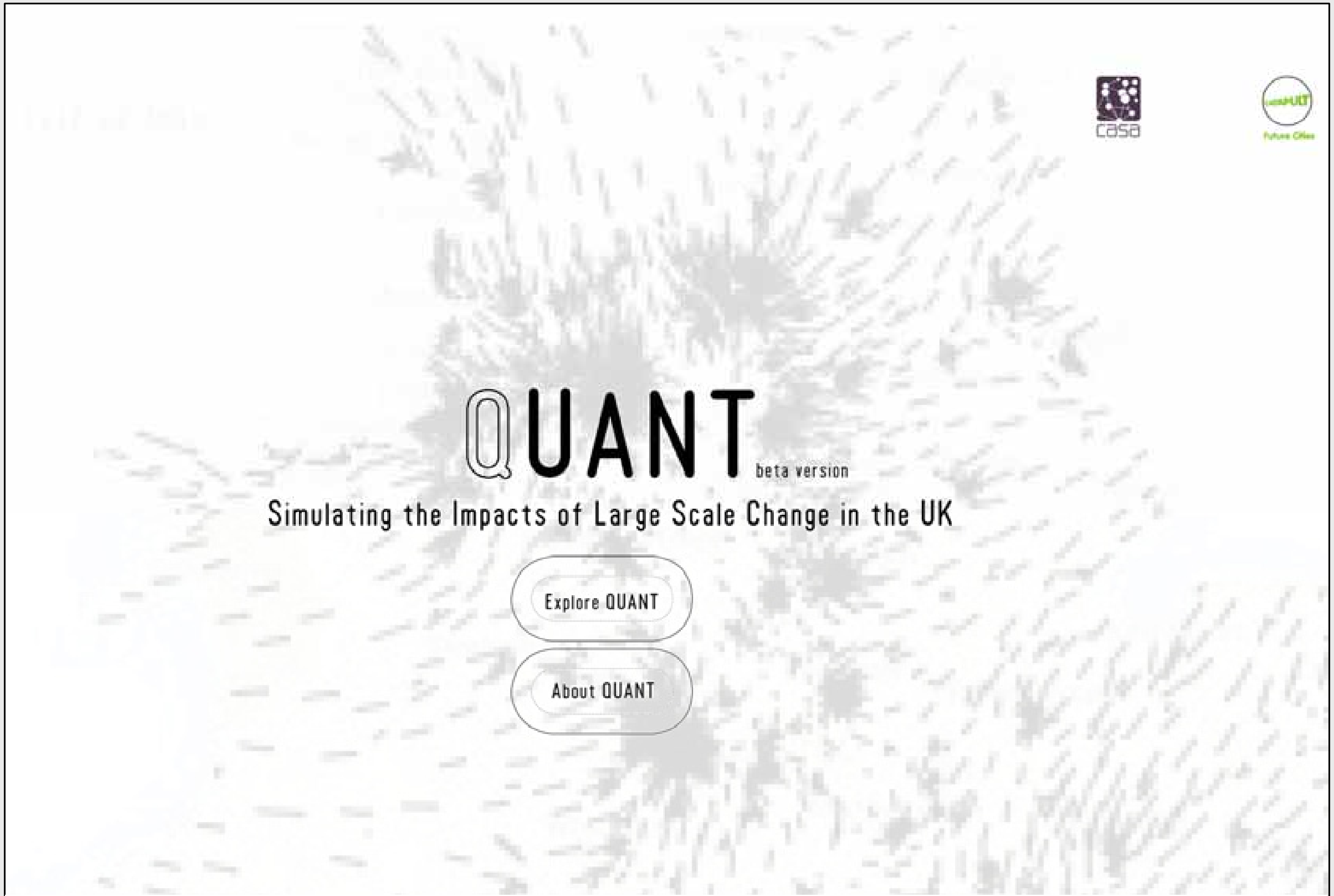
Model on server side; Maps on the client side

Can we reverse this?

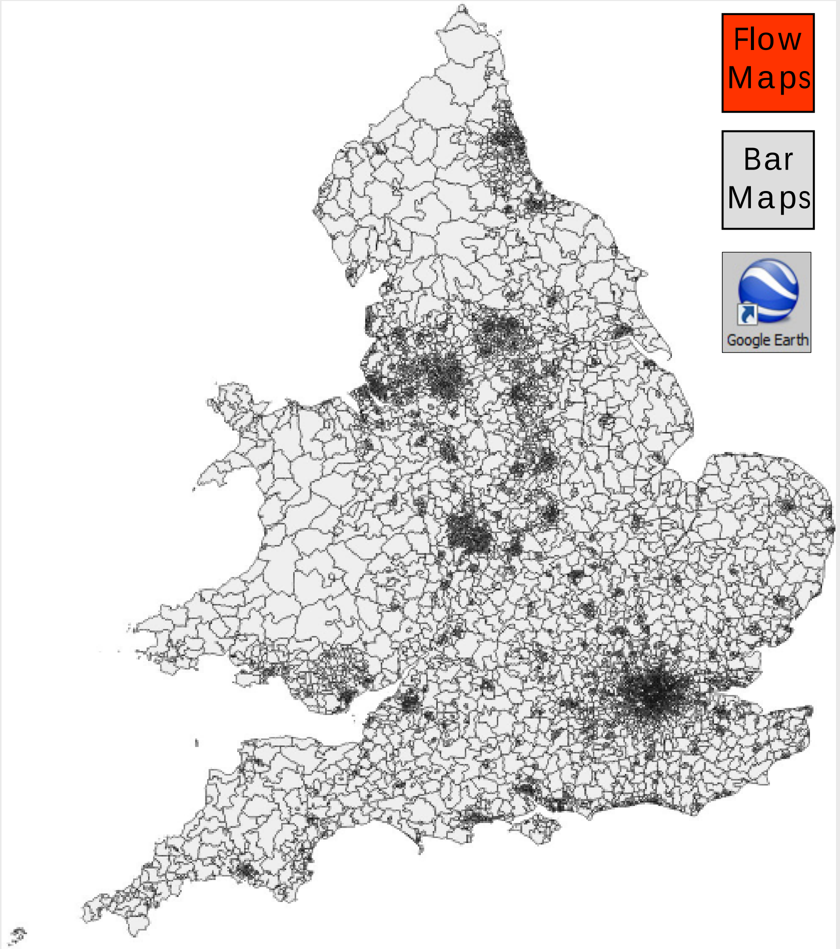
Not really – the matrices are too slow to download to client?

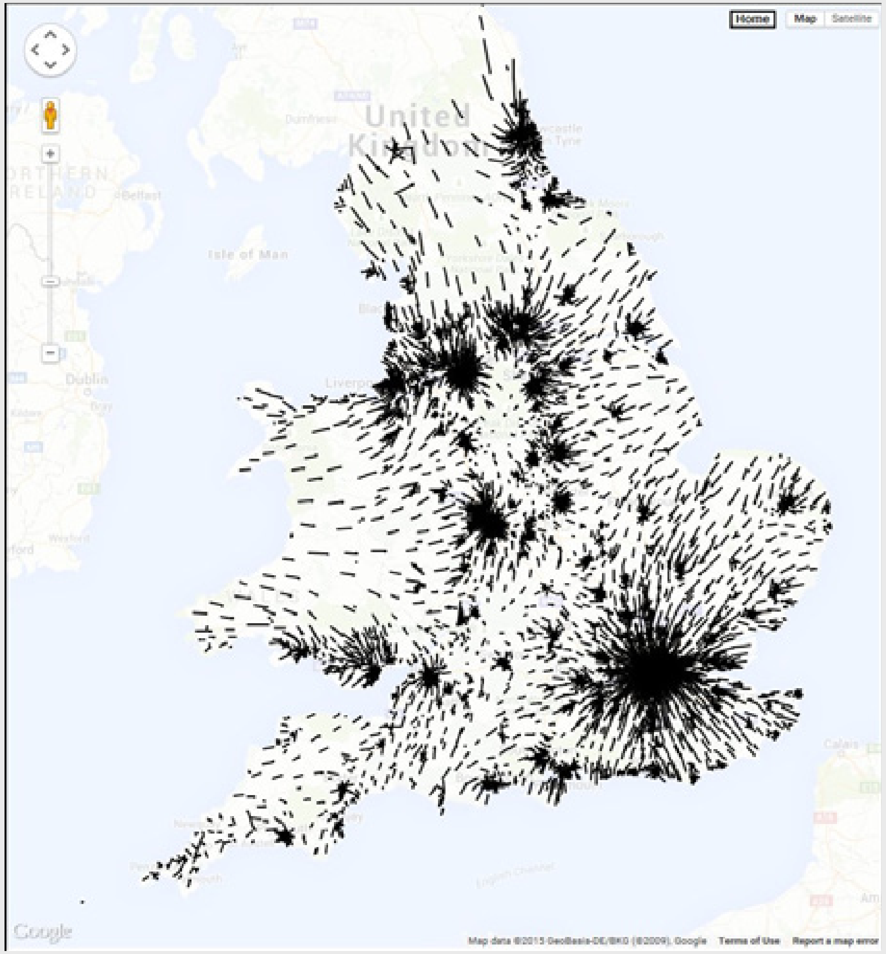
We also can't assume the client is fast enough for computation.

Frankly at this point, I am out of my depth

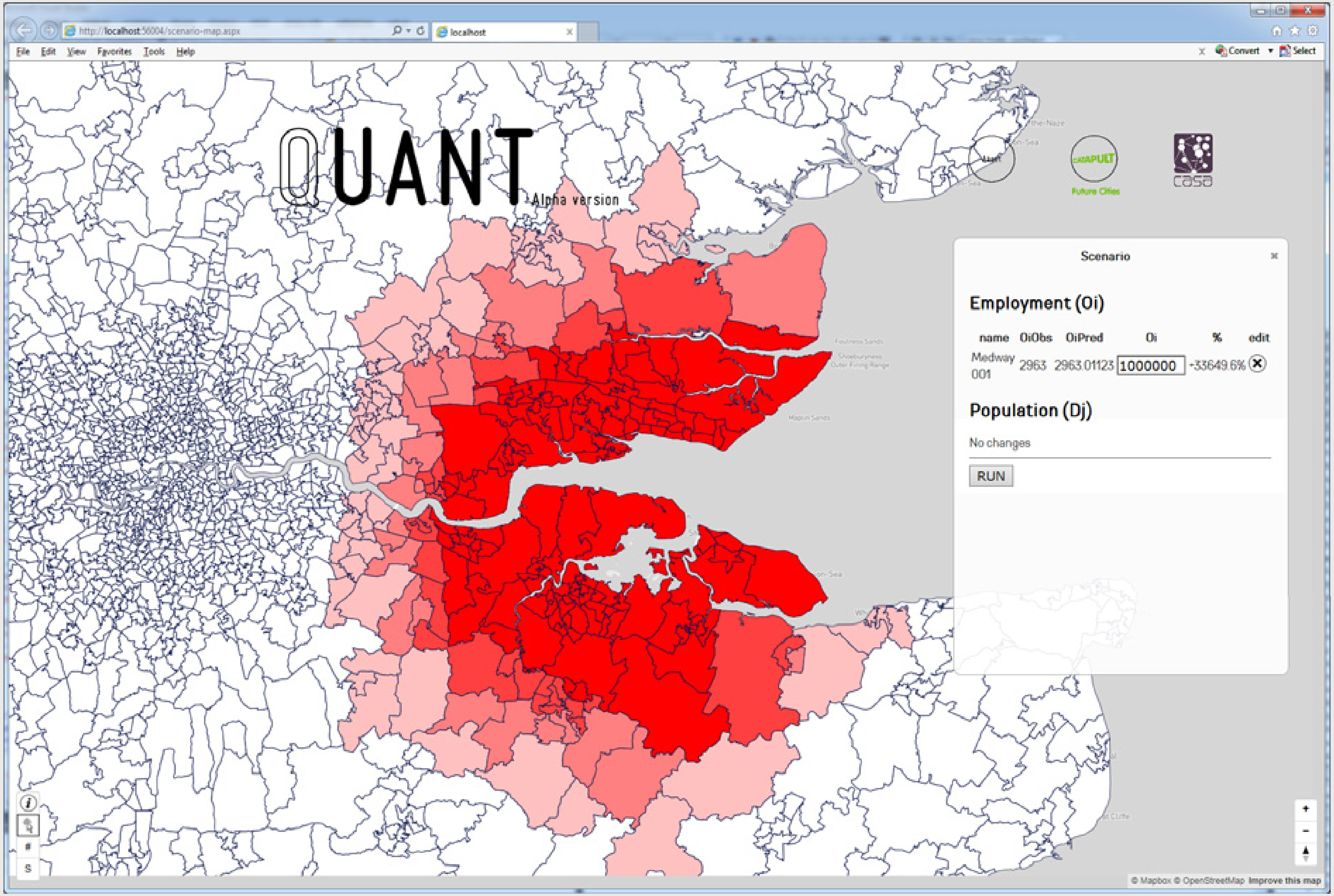


http://quant.casa.ucl.ac.uk/





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


QUANT
Alpha version
Scenario
Employment (Oi)
name OiObs OiPred Oi % edit
Medway 001 2963 2963.01123 1000000 -33649.6%
Population (Dj)
No changes
RUN
© Mapbox © OpenStreetMap Improve this map

Real-Time Streaming: What Sort of Data ?

http://www.citydashboard.org/

London Panopticon

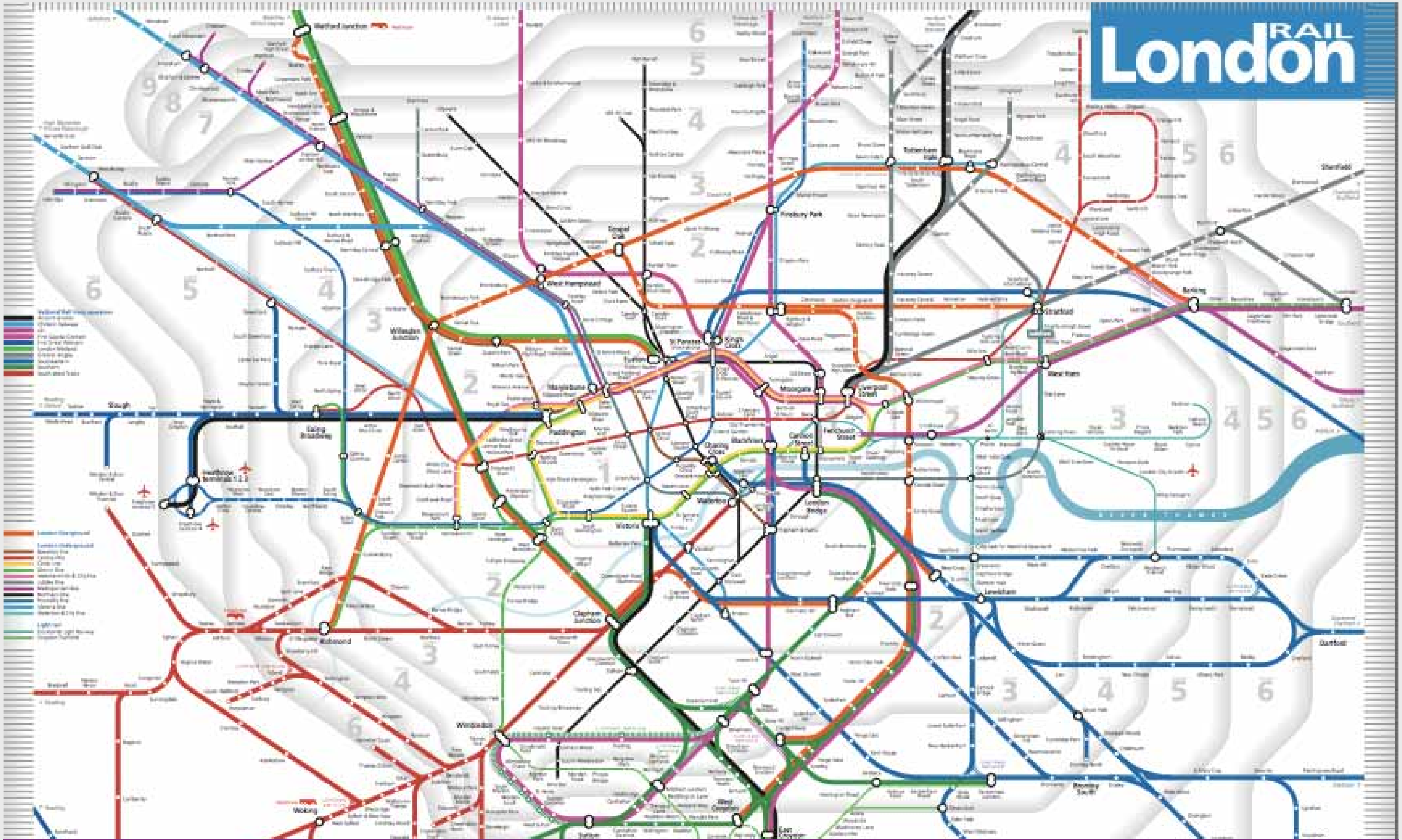
6 April 2016 London

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

Tweet Like Share 0

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

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

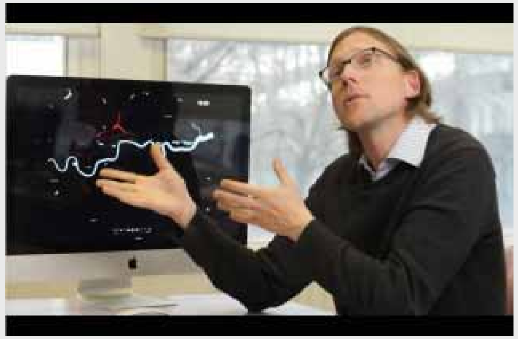
RT @markmiodownik: Am giving a ENGins seminar today for @UCLEngineering @UCLENGins all UCL engineers welcome - Roberts G06, 6:30pm. http:/...
8:58am Thu 9th October 2014

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

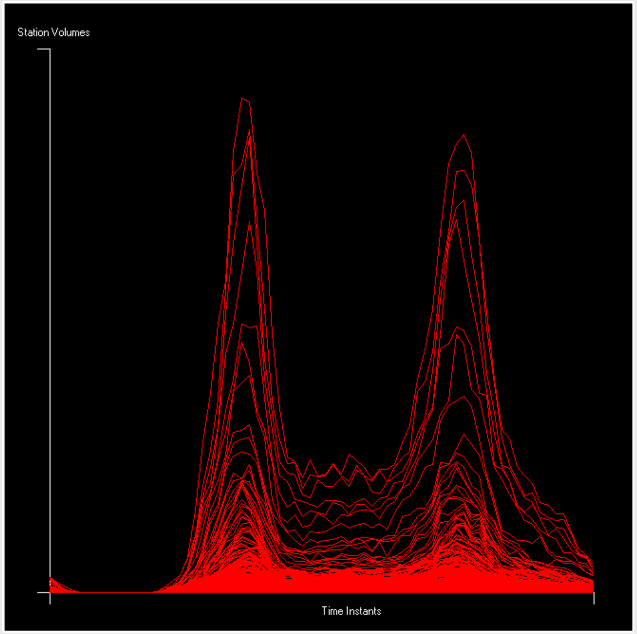
http://www.simulacra.info/

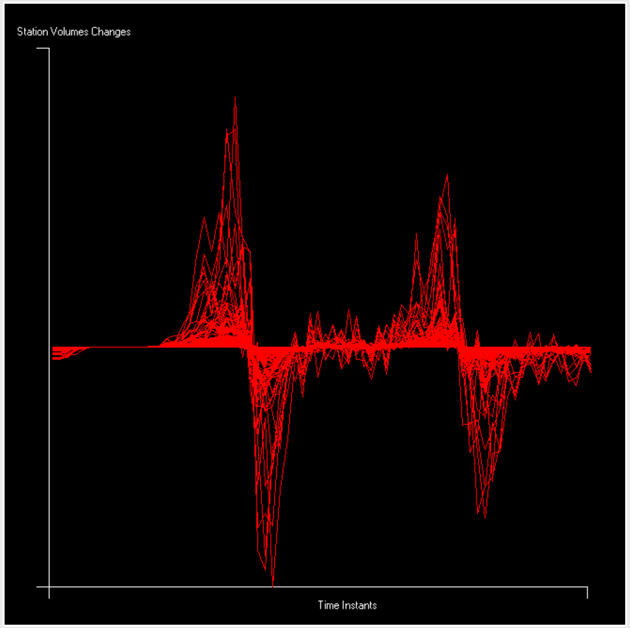


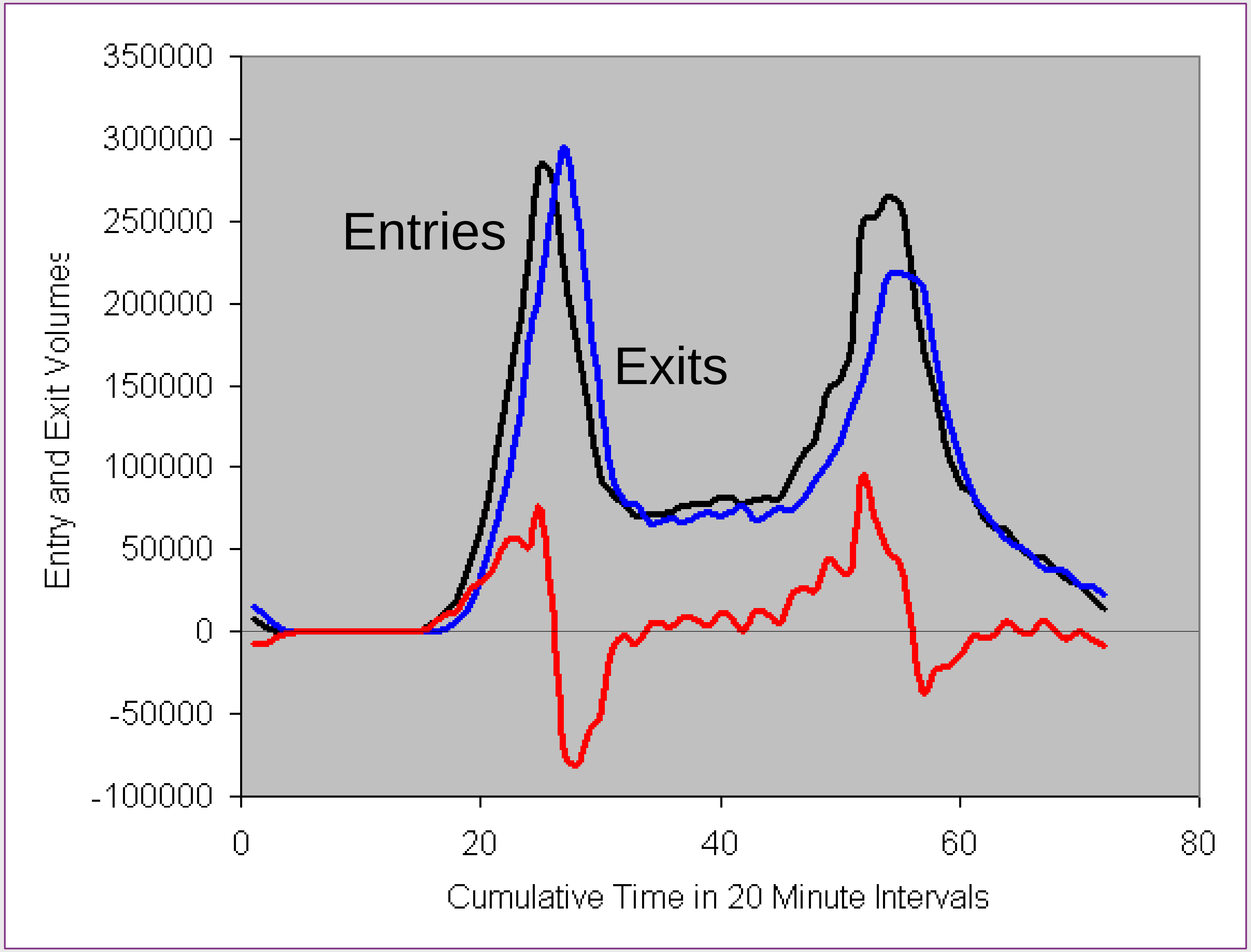
This of course was the thing that Lt Henry Harness did in Dublin in 1837 and what Minard et al. did a little later. In our LUTI models, this is an enormous problem as the scale of this assignment to networks is different

Variabilities – Heterogeneity and Travel Profiles

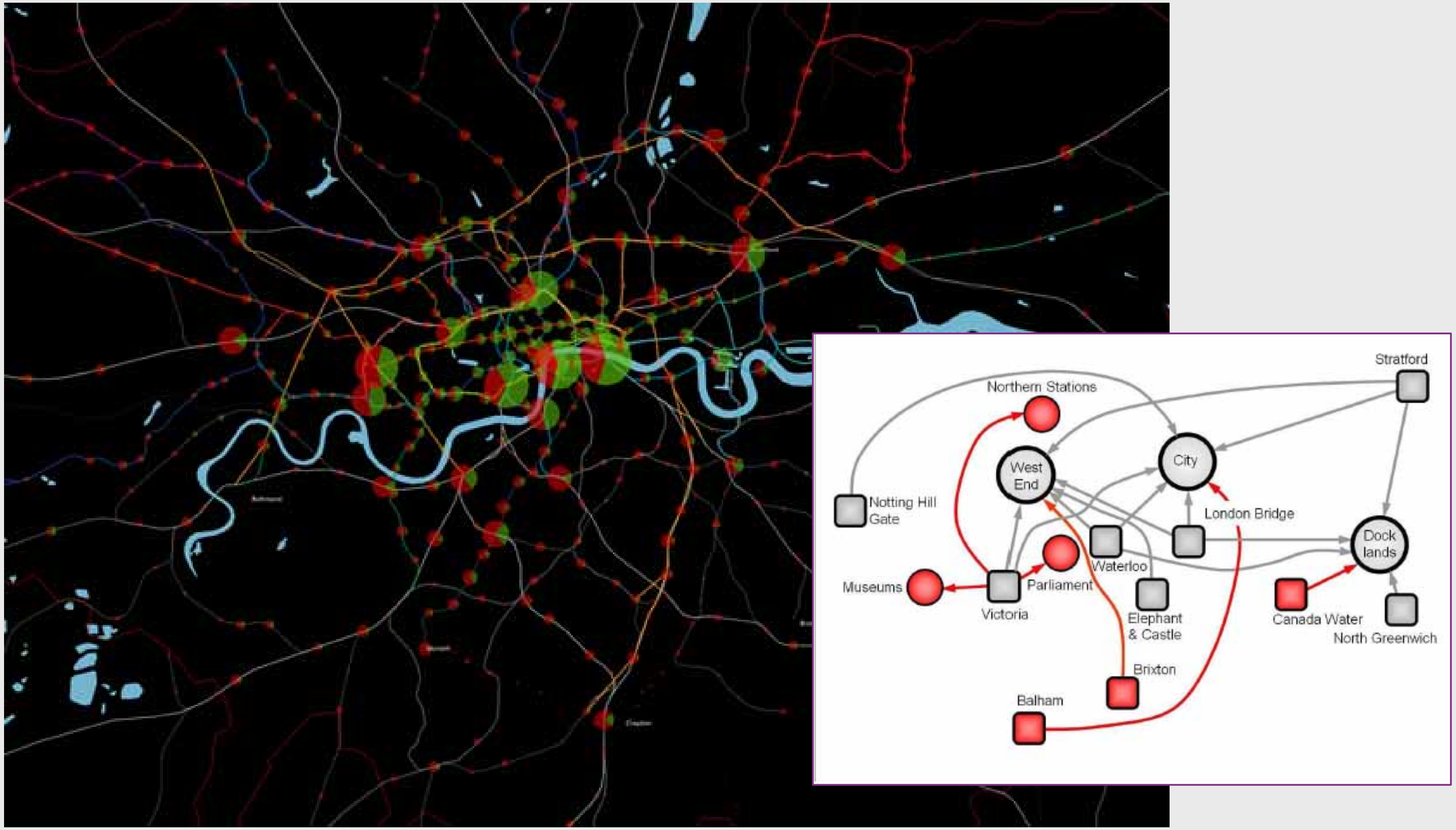
First we will look at some of the data and how it varies in terms of the diurnal flows usually morning and evening peaks, with a small blip (peak) around 10pm at night





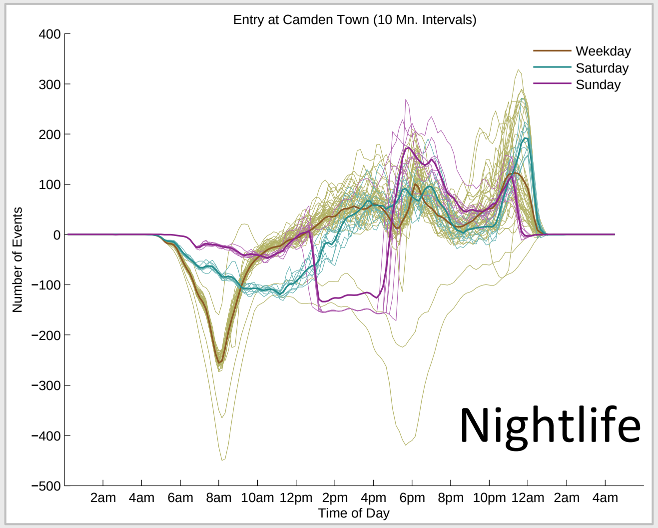

350000
300000
250000
200000
150000
100000
50000
0
-50000
-100000
Entry and Exit Volumes
Entries
Exits
0
20
40
60
80
Cumulative Time in 20 Minute Intervals

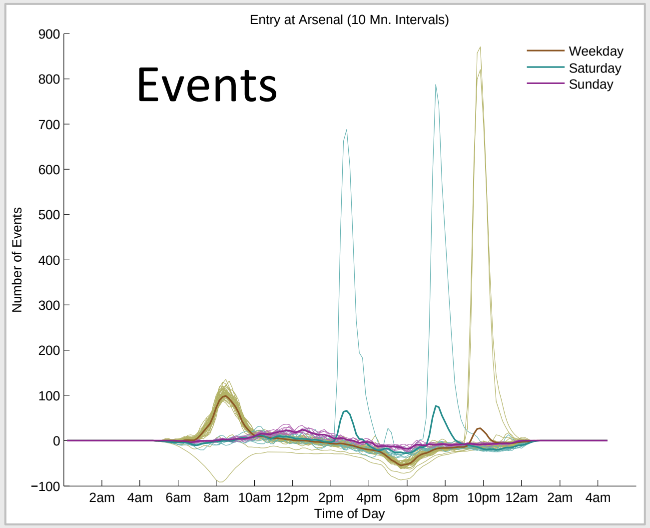
Oyster Card Data – interpreting urban structure, multitrips, etc.

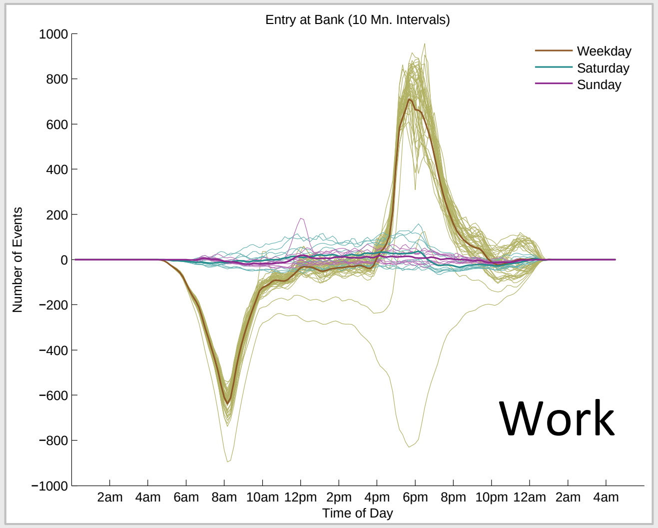


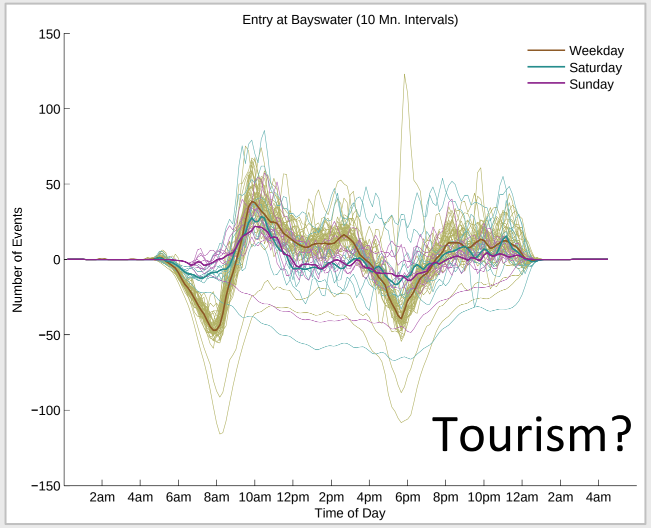
Roth C., Kang S. M., Batty, M., and Barthelemy, M. (2011) Structure of Urban Movements: Polycentric Activity and Entangled Hierarchical Flows. **PLoS ONE 6(1):** e15923. doi:10.1371/journal.pone.0015923

Particular Events: Weekdays, Saturdays and Sundays









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

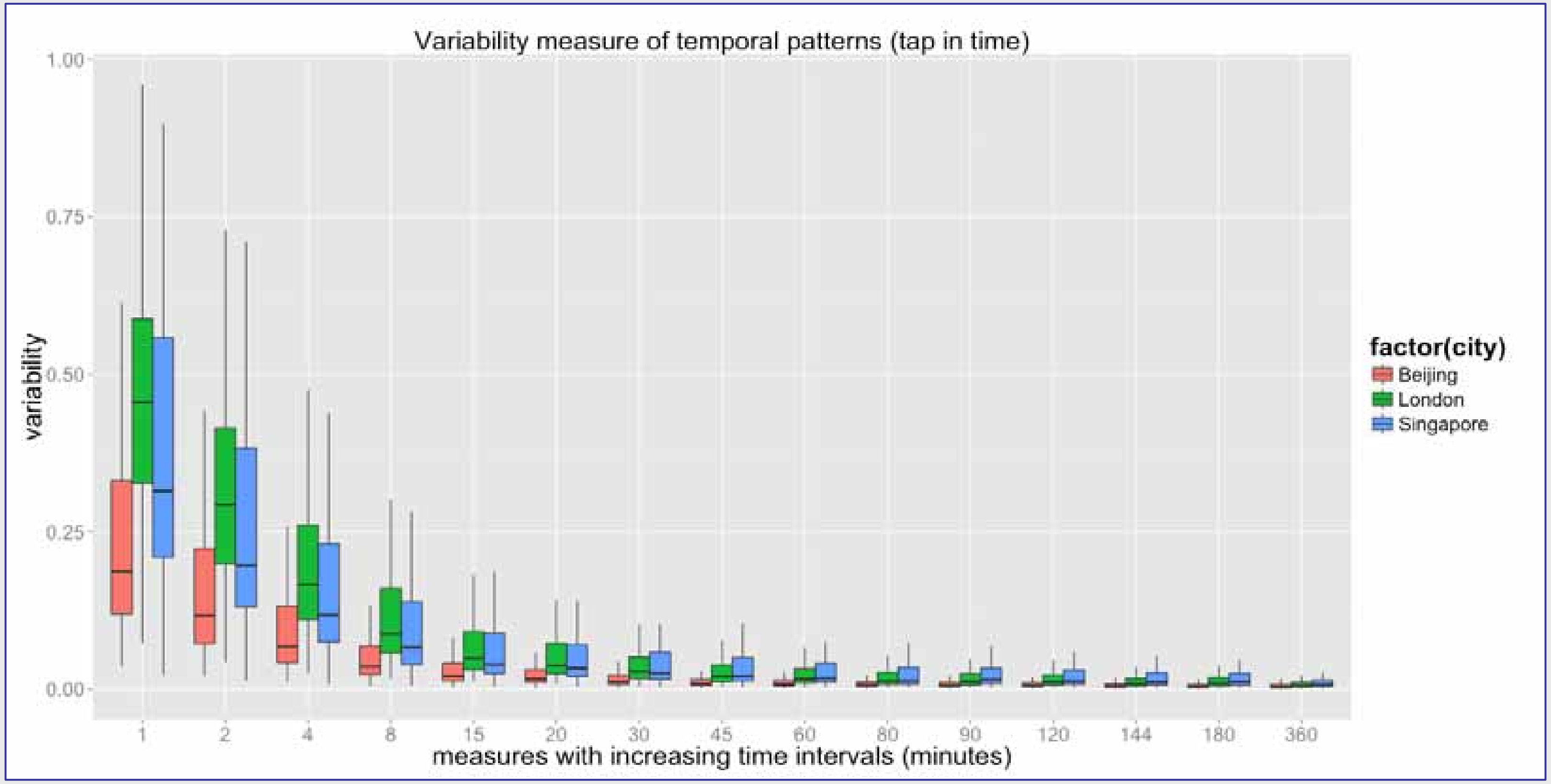
(1) Number of stations is the number of stations with smart-card records generated.

(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



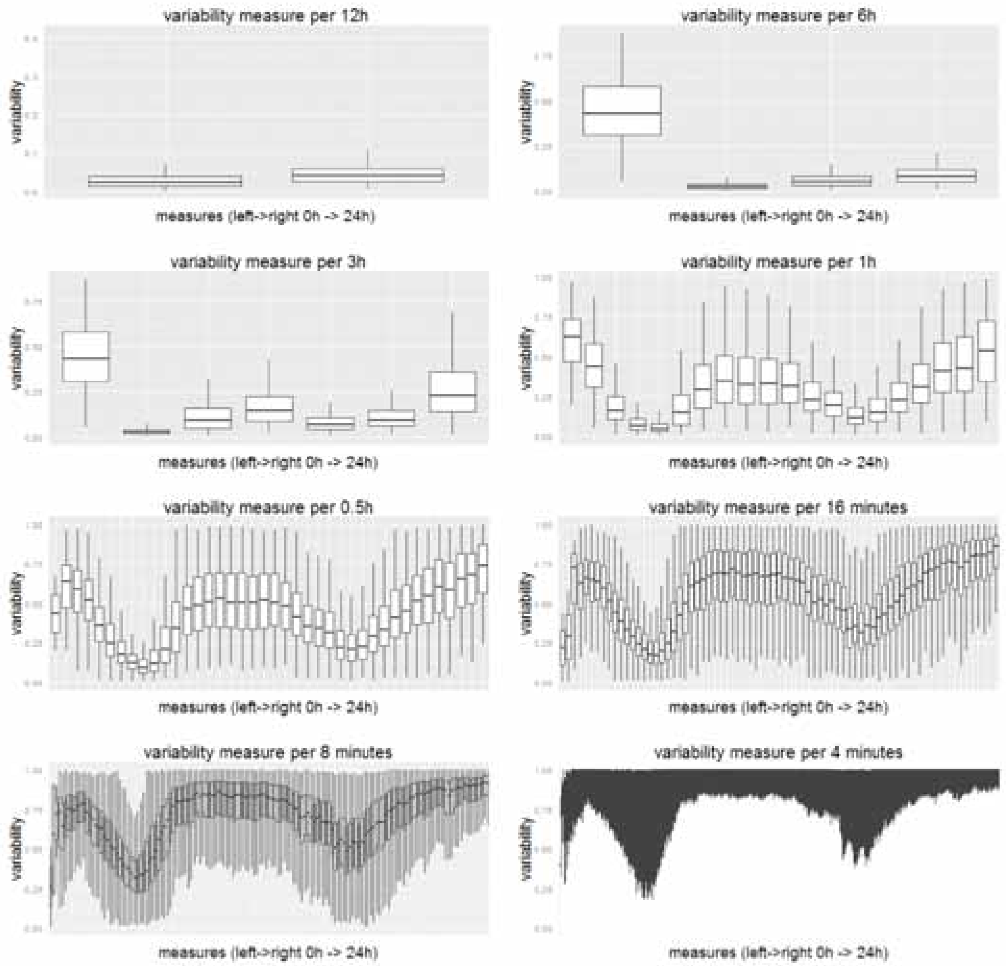
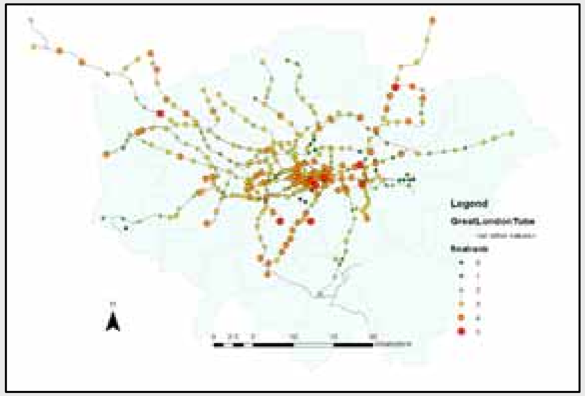


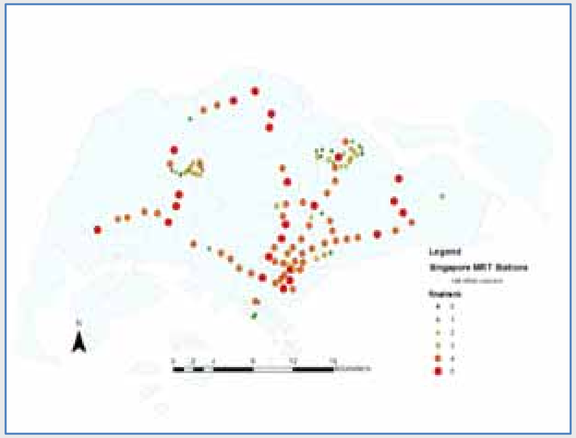
Figure 1. Variability of regularity in the trip matrix over time.

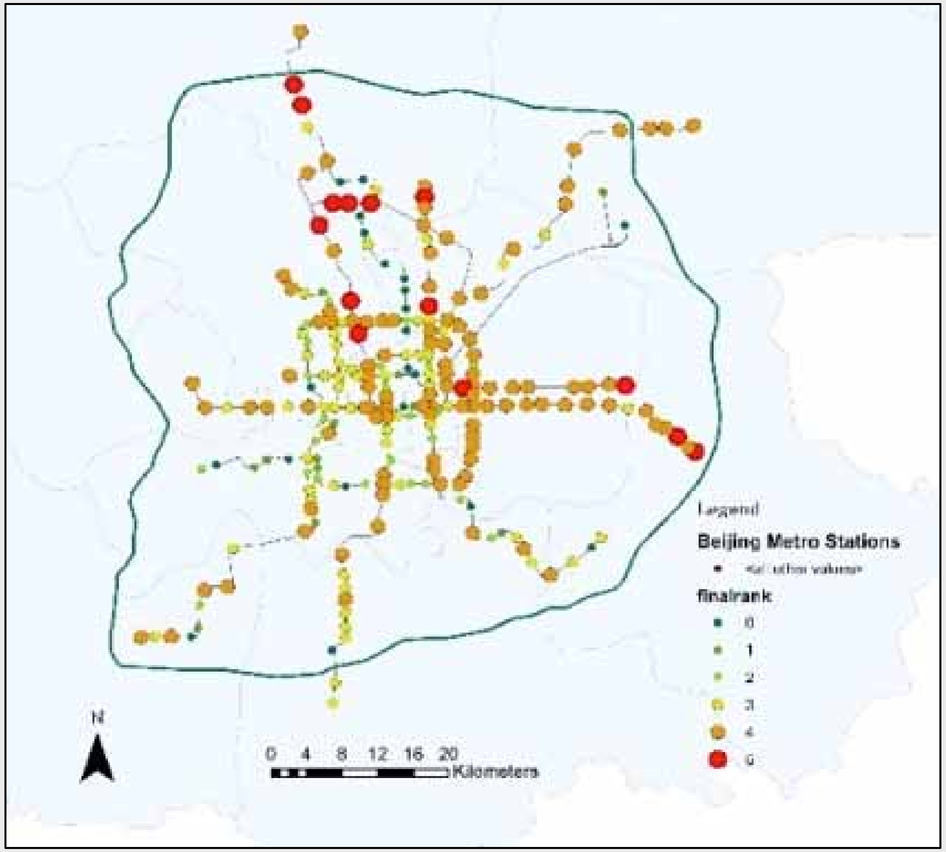
Note: Each box plot shows the variability of 400 stations over time measured at different temporal scales. Overall, eight subplots give a similar trend where lower variability appears during peak hours (around 9 am in the morning and 6pm in the evening). More details can be captured as differences of variability between each time unit are magnified as we decrease the temporal scale from 12h to 4 minutes.

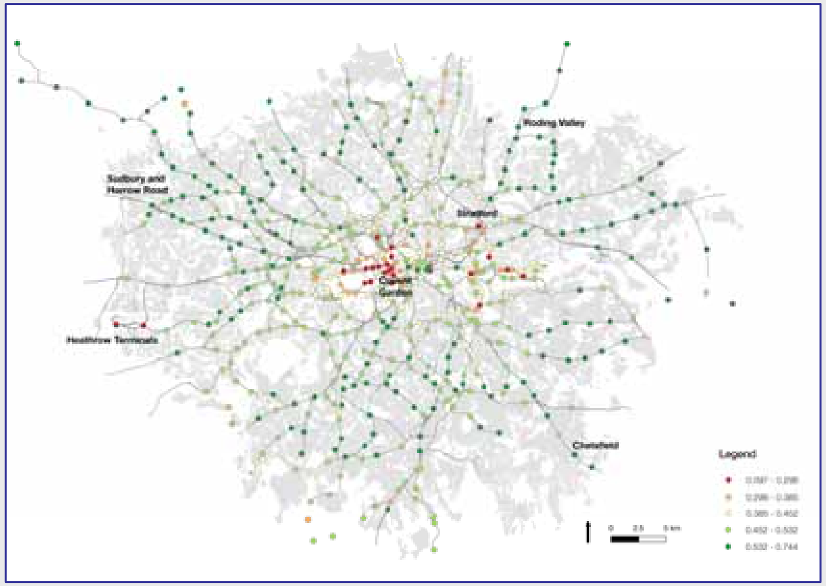
Comparing Variability for different time Intervals over the day

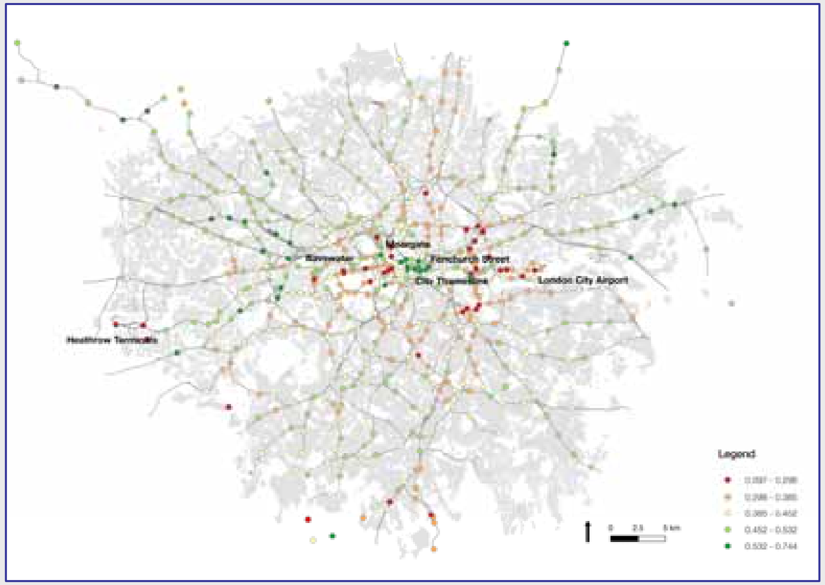
Comparing Variability for different time intervals for Three World Cities: London, Beijing and Singapore

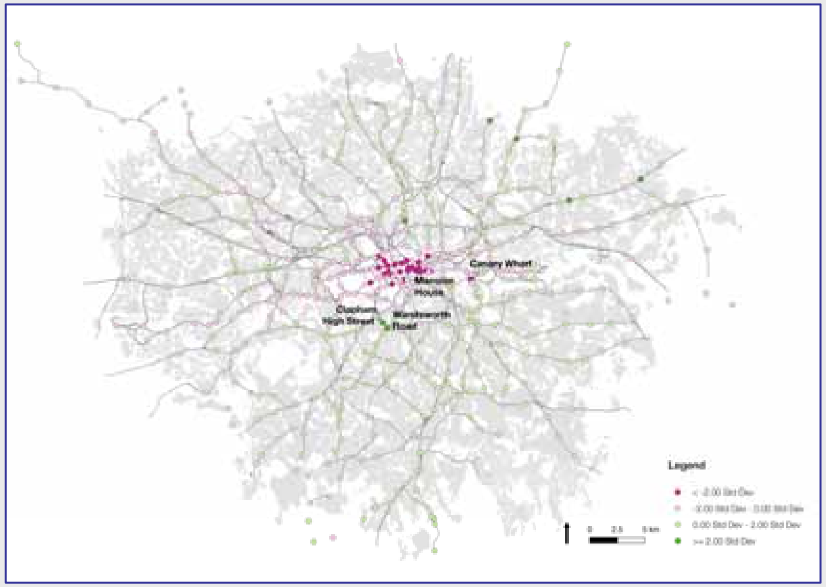












Maps of Underground and Rail stations in London visualised by the proportion of regular trips

originating at each location

ending at each location

starting and ending at each location

Disruptions – Routine Analysis of Daily Events

- Behaviours vary across network
- Different areas of network more resilient to disruption, due to available infrastructure and individual ability to change
- But areas of network are naturally closely tied through established usage patterns
- Individual-based analyses provide insight into behaviours underlying macroscopic flows

We will look at several kinds of disruption

- First hypothetical disruptions simply by examining breaks in the network

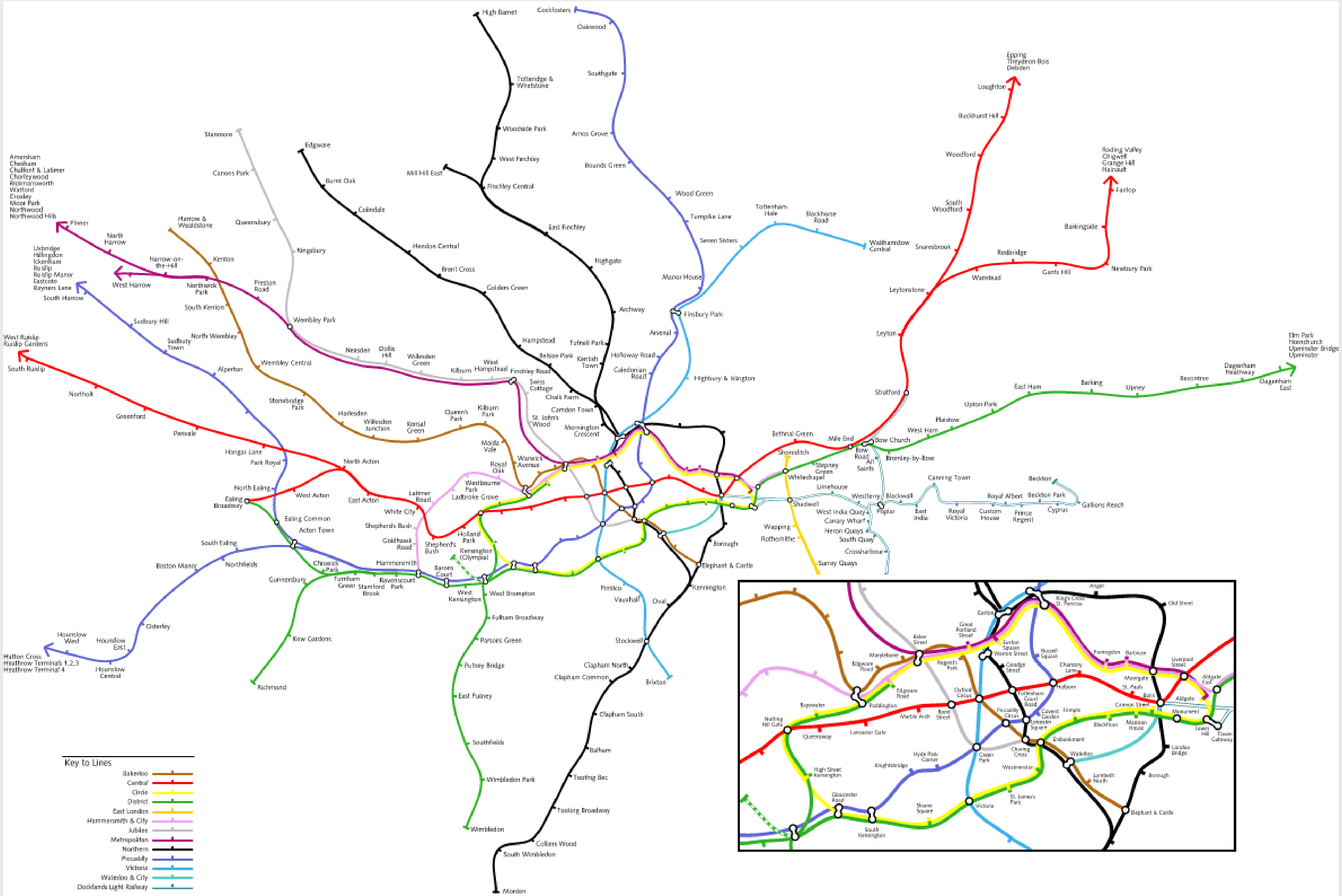
- Then an example of the Circle and District Lines which had a 4 hour stoppage on July 19th 2012

- And a Bus Strike in East London and how this shows up in the data

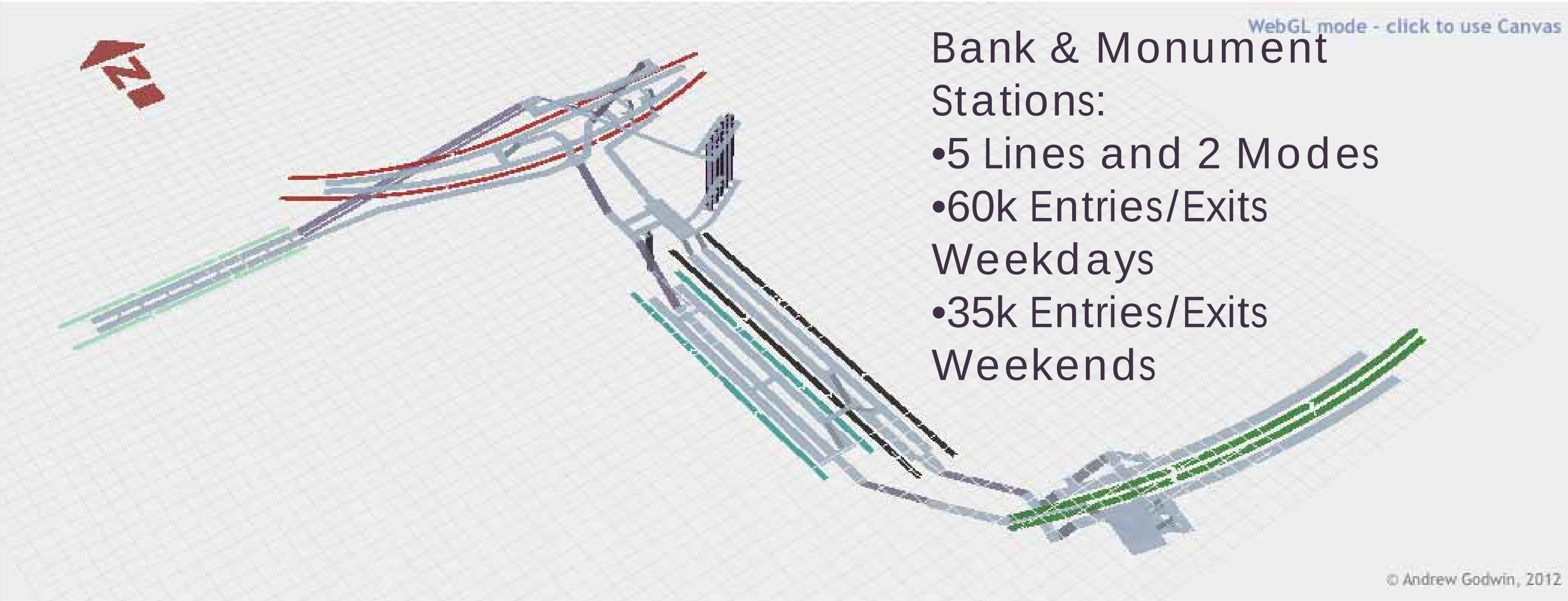
- And typical pattern of delay on all modes visualised for Greater London


Service information
Date
Time
No Service On The Entire Line
And Kings Cross
Between Baker Street
District Line
And Edgware Road
Jubilee Line
Holloway Road Station Closed


UNDERGROUND


Key to Lines
Bakerloo
Central
Circle
District
East London
Hammersmith & City
Jubilee
Metropolitan
Northern
Piccadilly
Victoria
Waterloo & City
Docklands Light Railway

Looking at Station Closures and Shifts of Travellers Using The ‘Shortest’-Paths



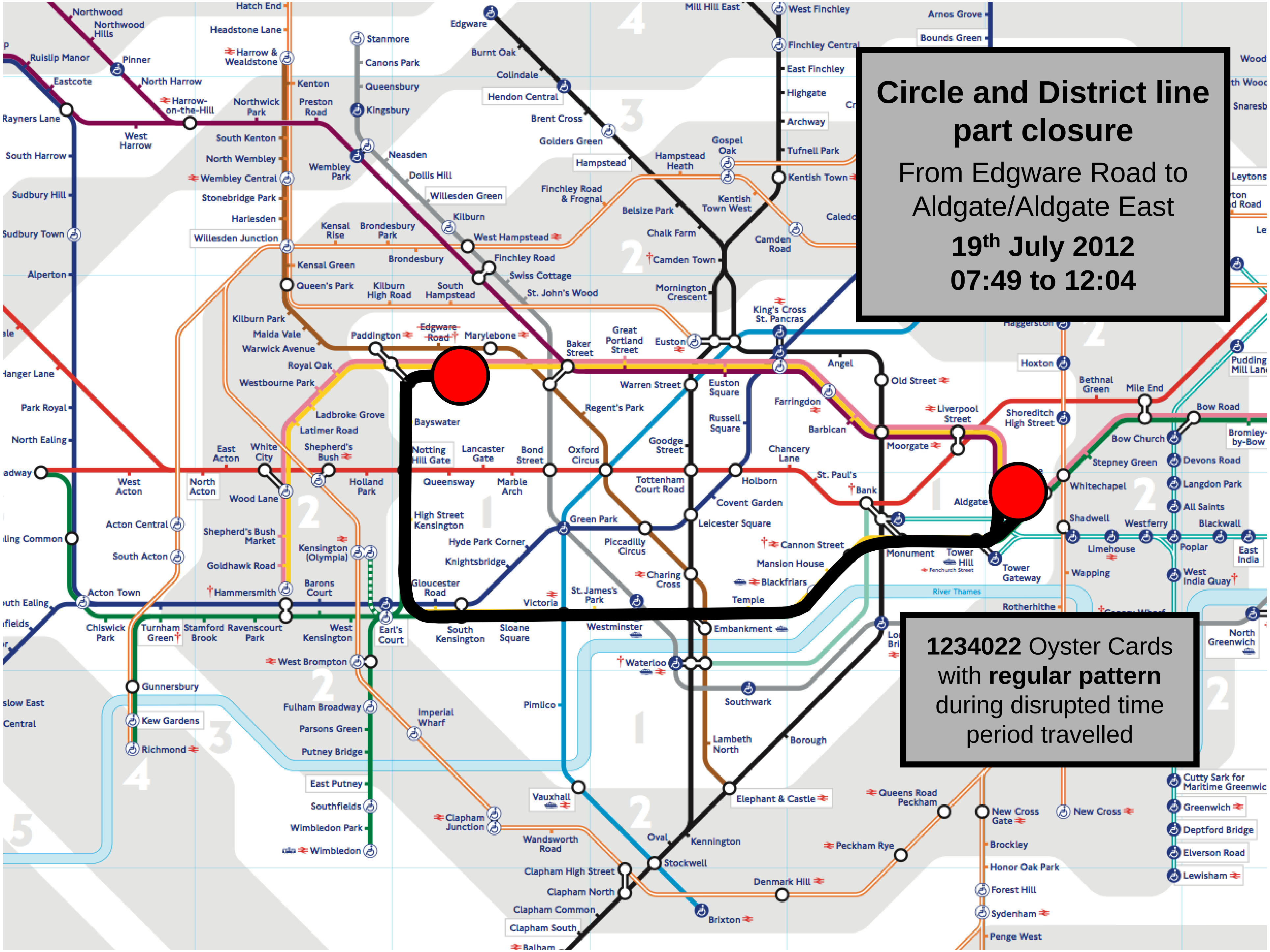
Although a simple station/line network may be sufficient for small cities, for ‘Mega-Cities’ such as London, New York, or Tokyo a much more detailed network is needed with interchanges measured down to the platform level. The ‘penalties’ for changing lines (and permitted Out-of-Station Interchanges) can be severe and should be included in a schematic network representation.

Circle and District line part closure

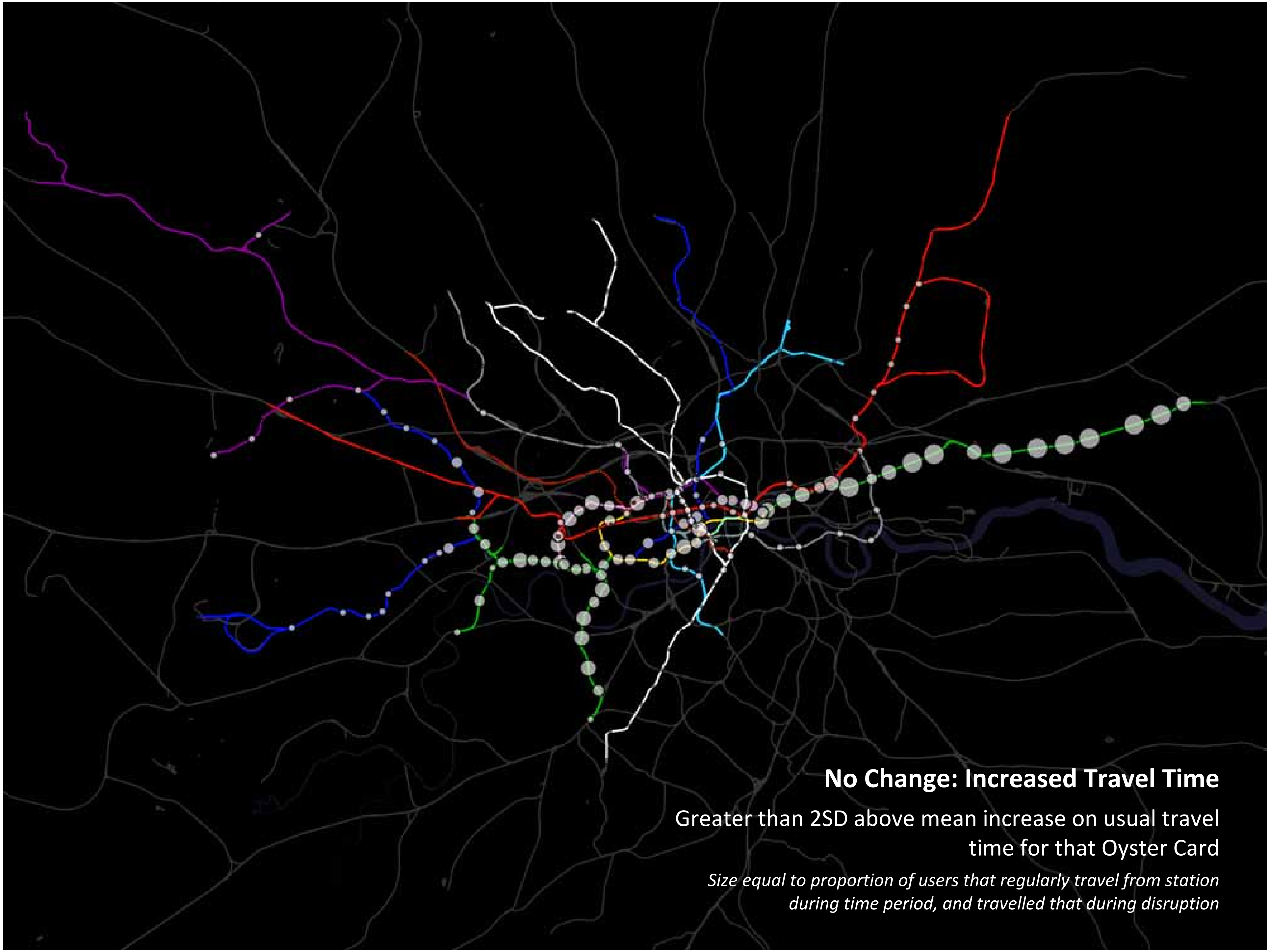
From Edgware Road to Aldgate/Aldgate East

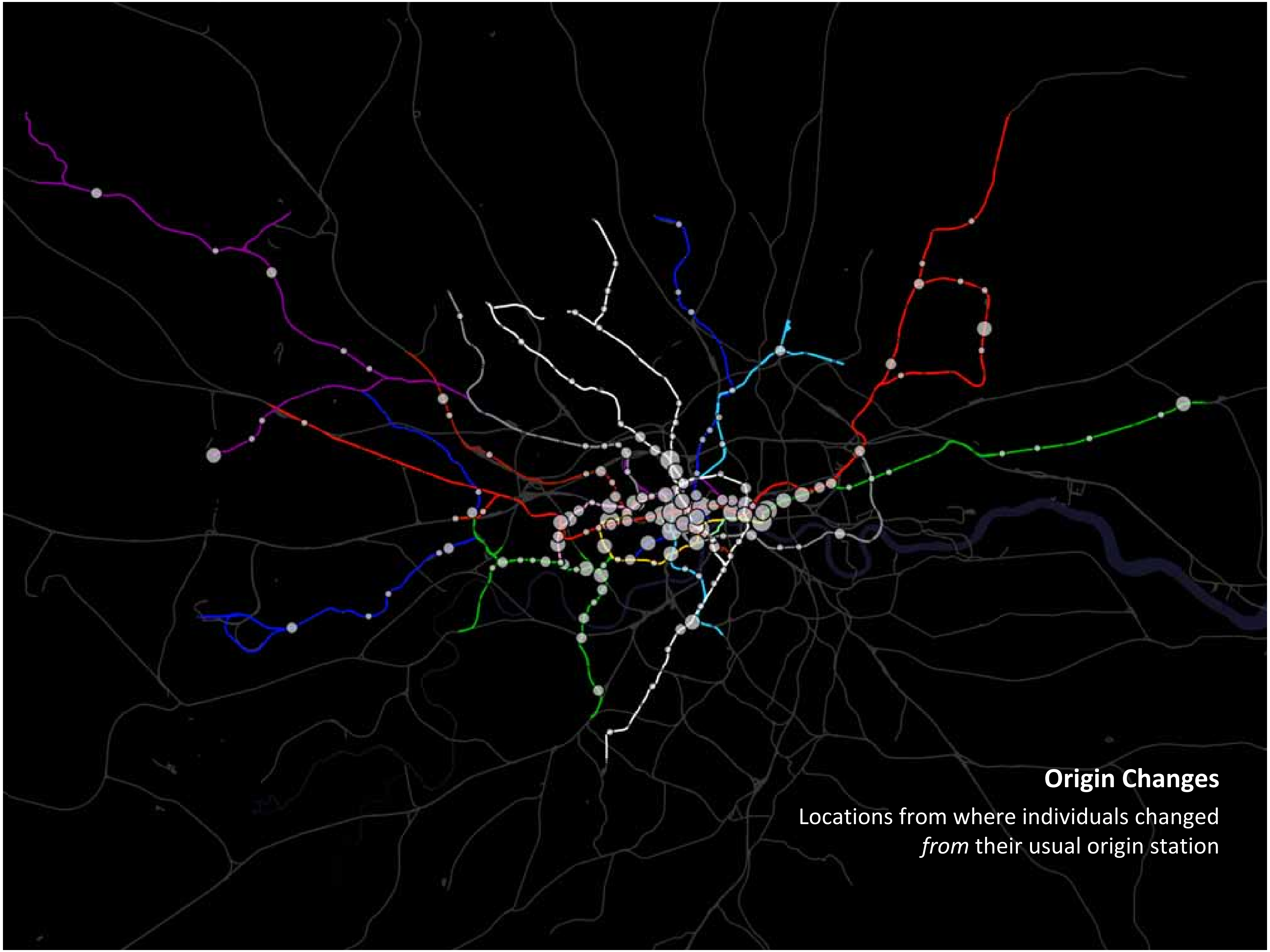
19th July 2012

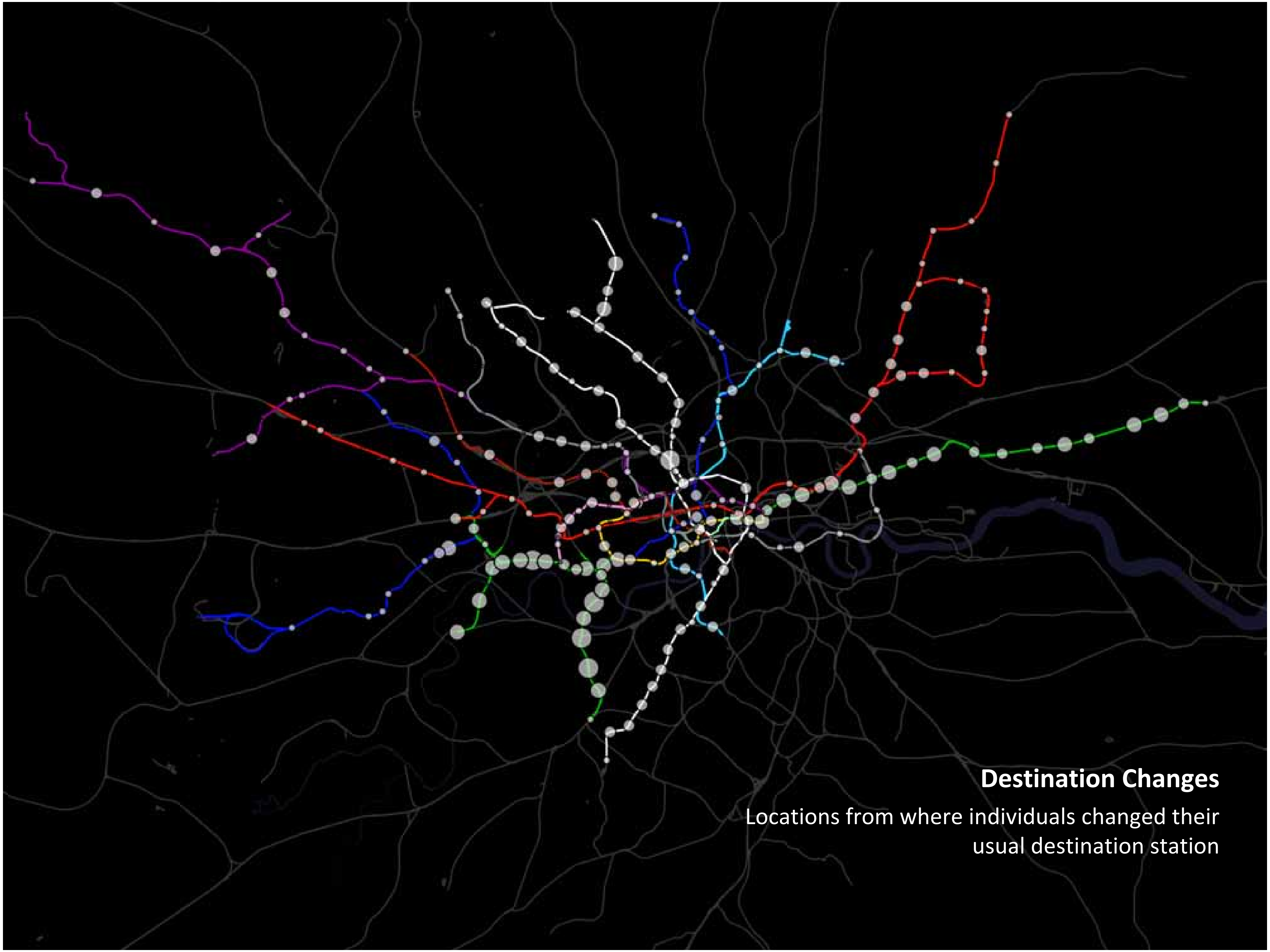
07:49 to 12:04



1234022 Oyster Cards with **regular pattern** during disrupted time period travelled


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



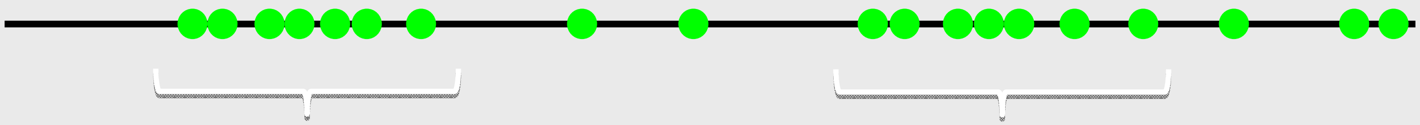
Destination Changes

Locations from where individuals changed their usual destination station

Measuring Regularity

Version 2: DBSCAN Method

Oyster Card A – Origin 747



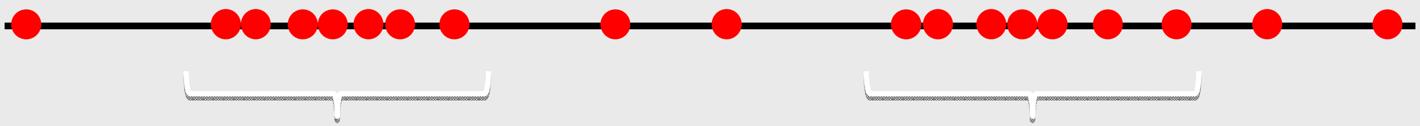
Cluster 1

Mean: 08:07; SD = 6.2
Max: 07:45; Min: 08:16
Proportion of Days = 0.8

Cluster 2

Mean: 17:34; SD = 17.1
Max: 16:58; Min: 18:57
Proportion of Days = 0.6

Oyster Card A – Destination 647



Cluster 3

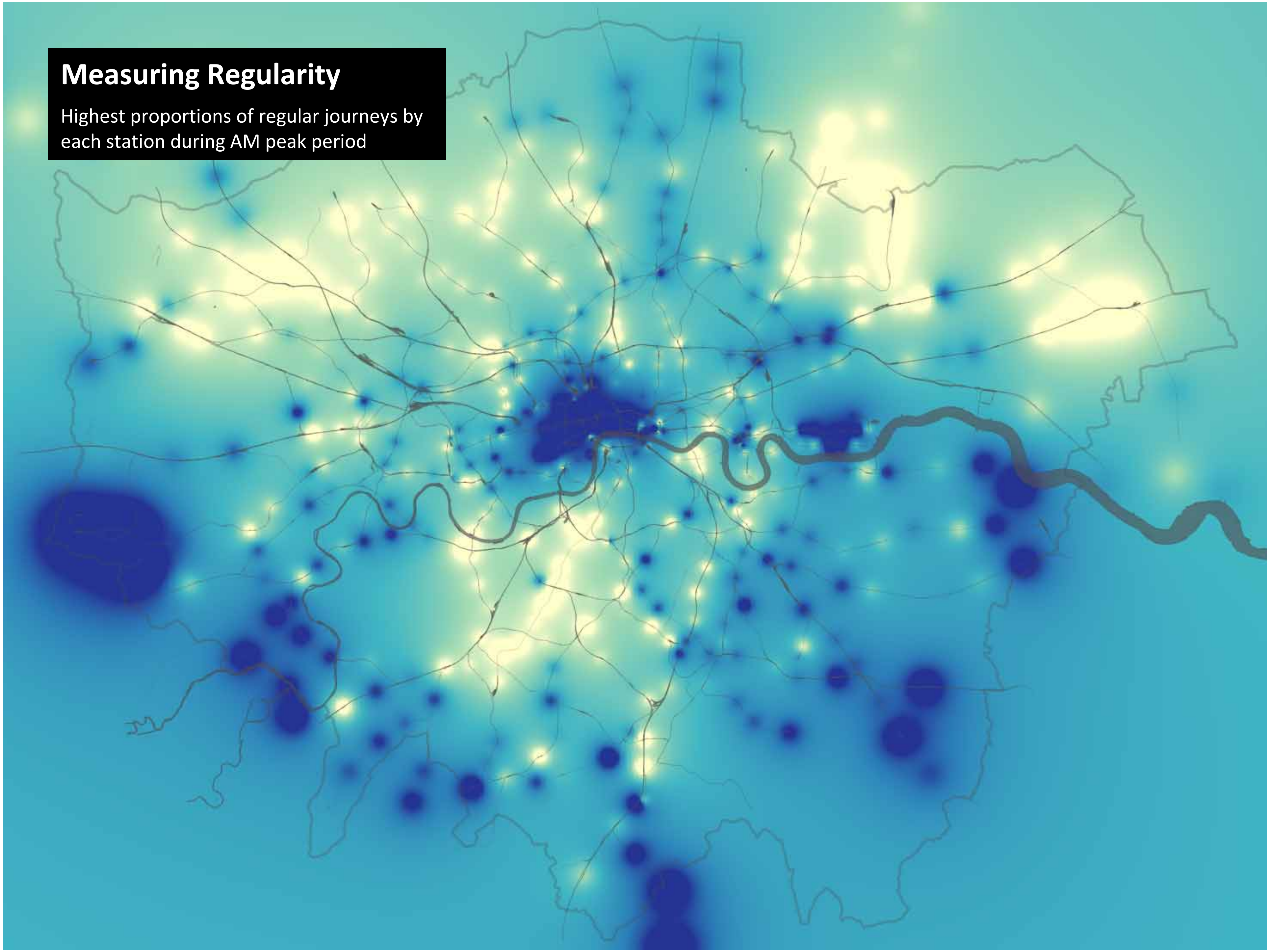
Mean: 08:37; SD = 7.4
Max: 08:13; Min: 08:48
Proportion of Days = 0.8

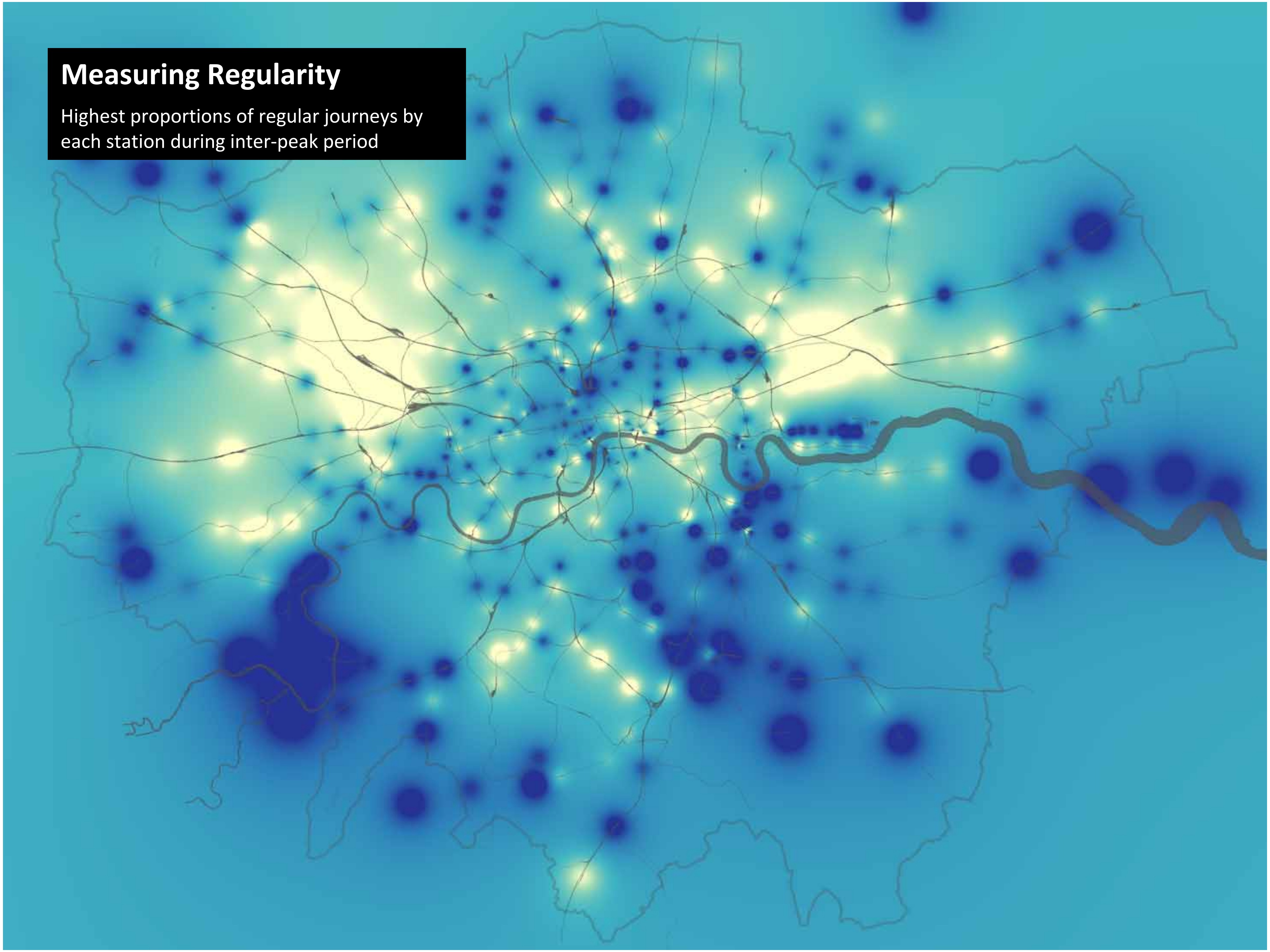
Cluster 4

Mean: 18:06; SD = 18.1
Max: 17:26; Min: 19:28
Proportion of Days = 0.6

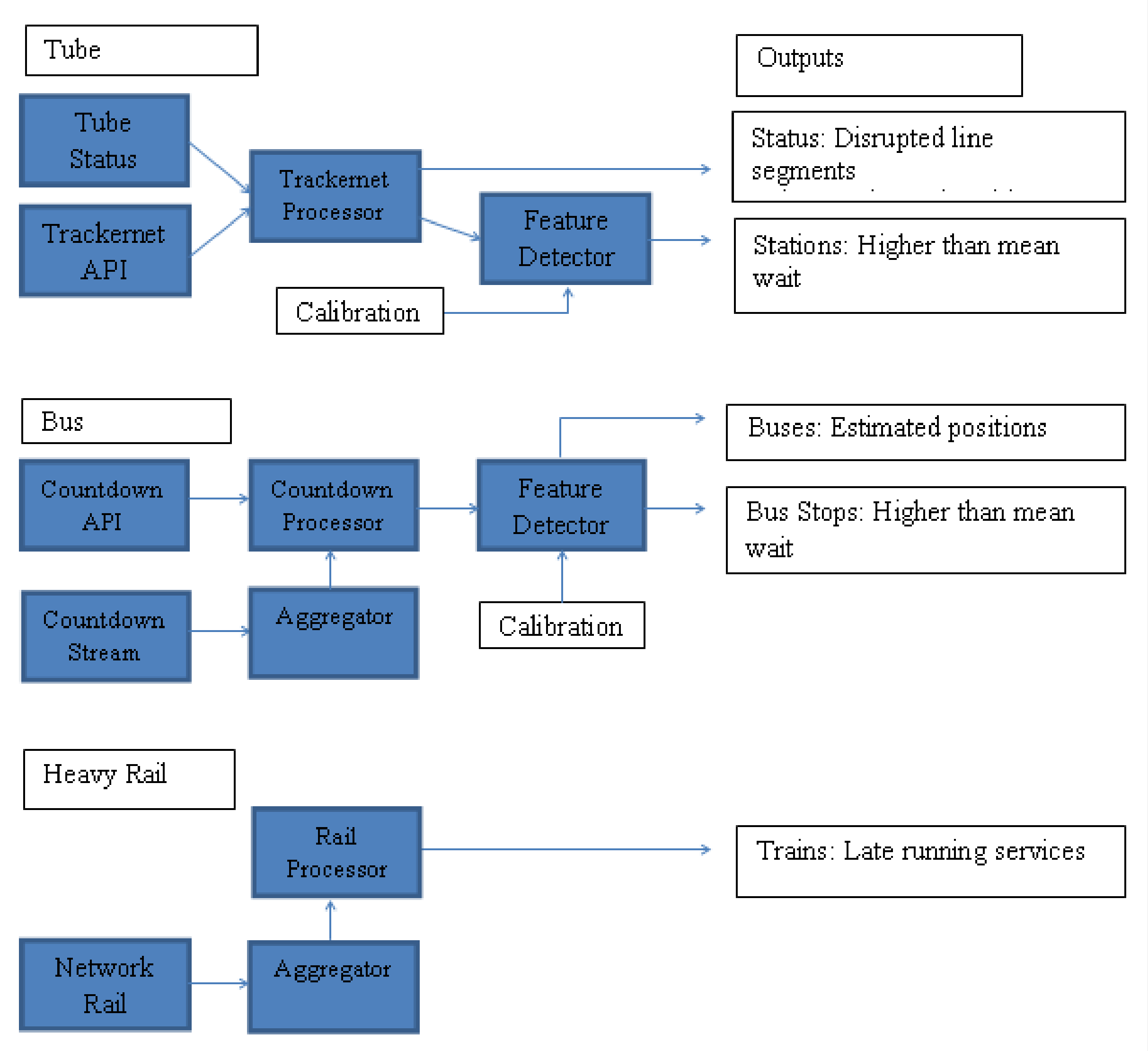
Measuring Regularity

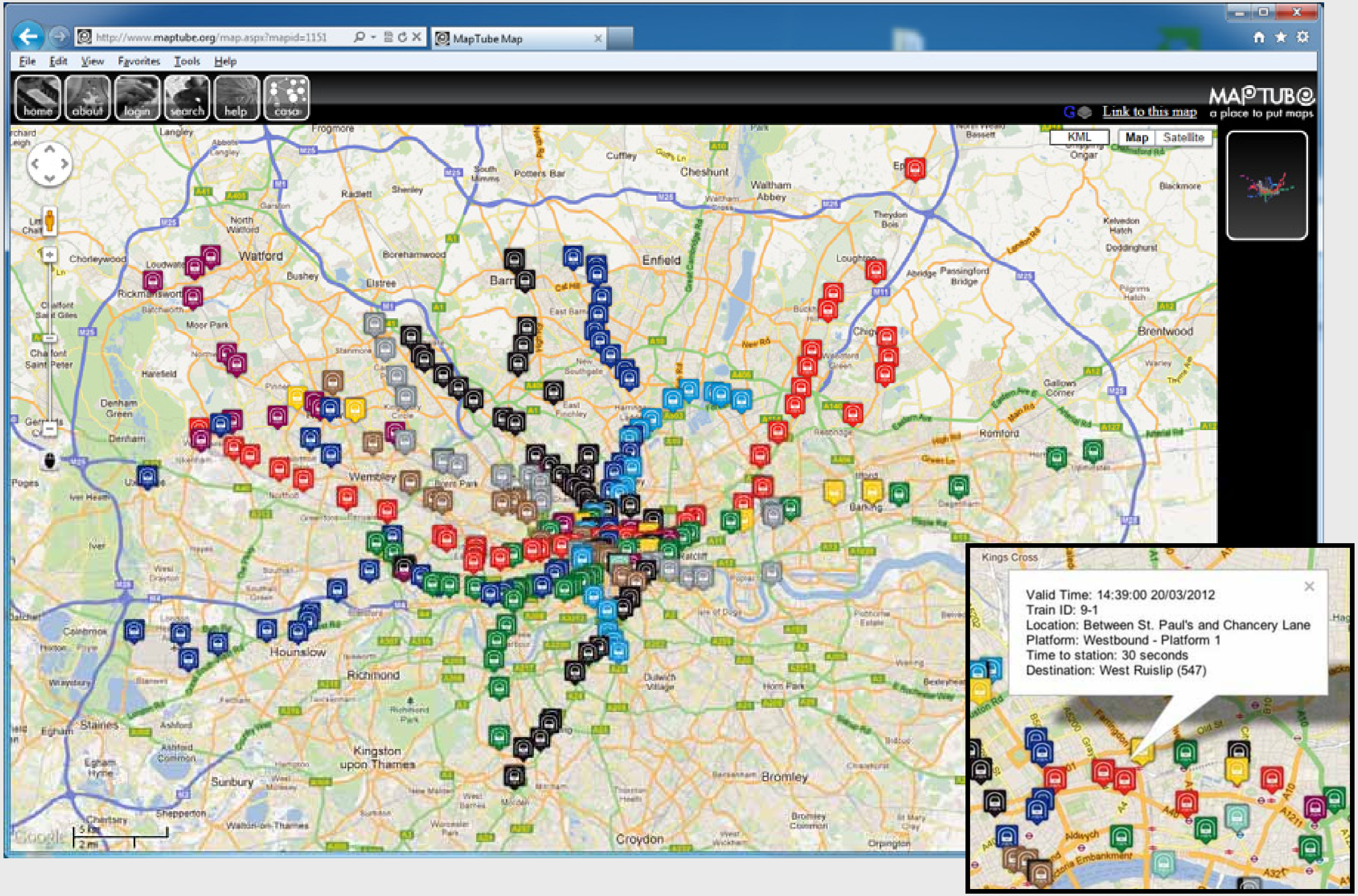
Highest proportions of regular journeys by each station during AM peak period

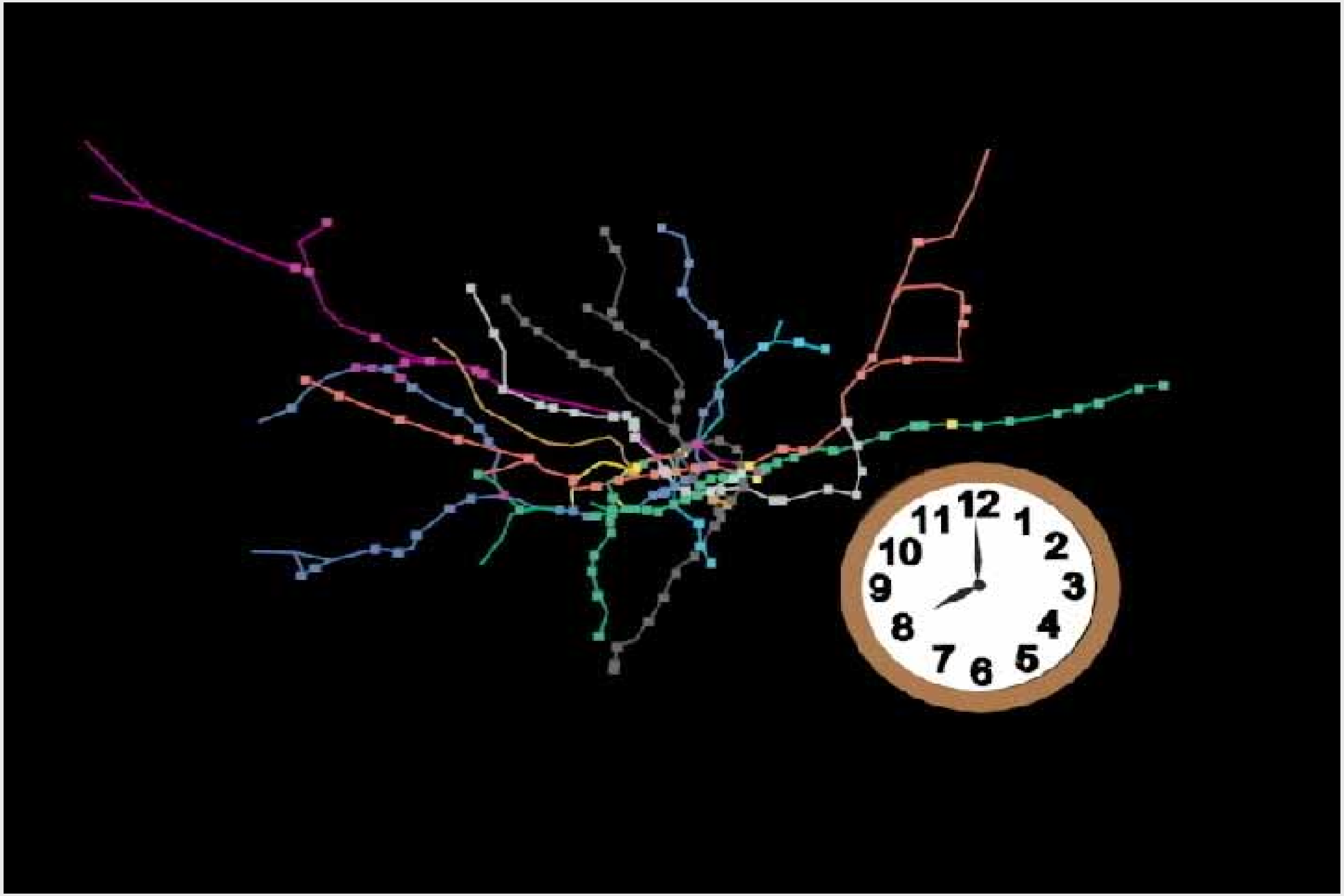



Measuring Regularity
Highest proportions of regular journeys by each station during inter-peak period

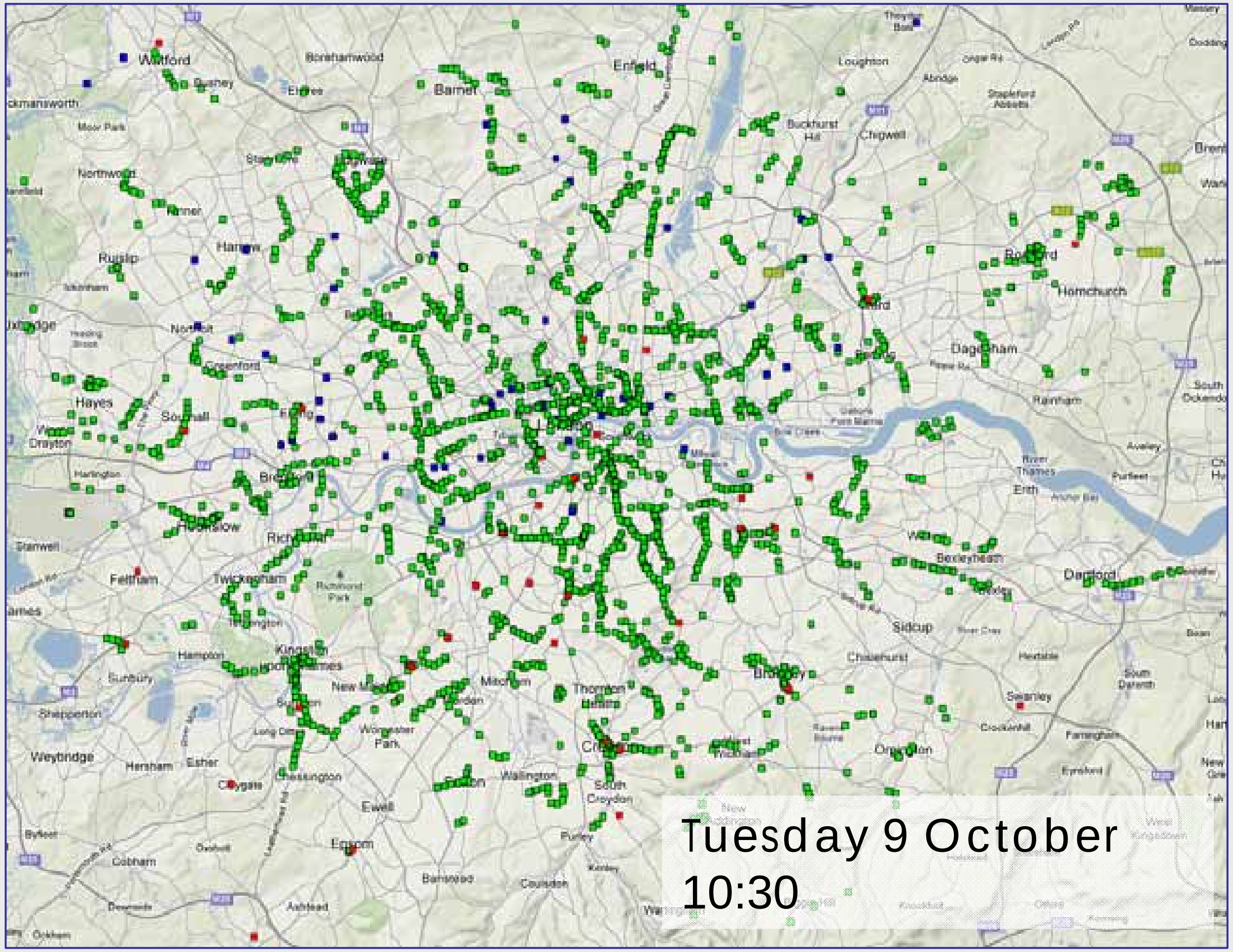
The Public Transport System in Terms of Vehicle Flows




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)


12
1
2
3
4
5
6
7
8
9
10
11

Delays from Tube, National Rail and Bus Fused



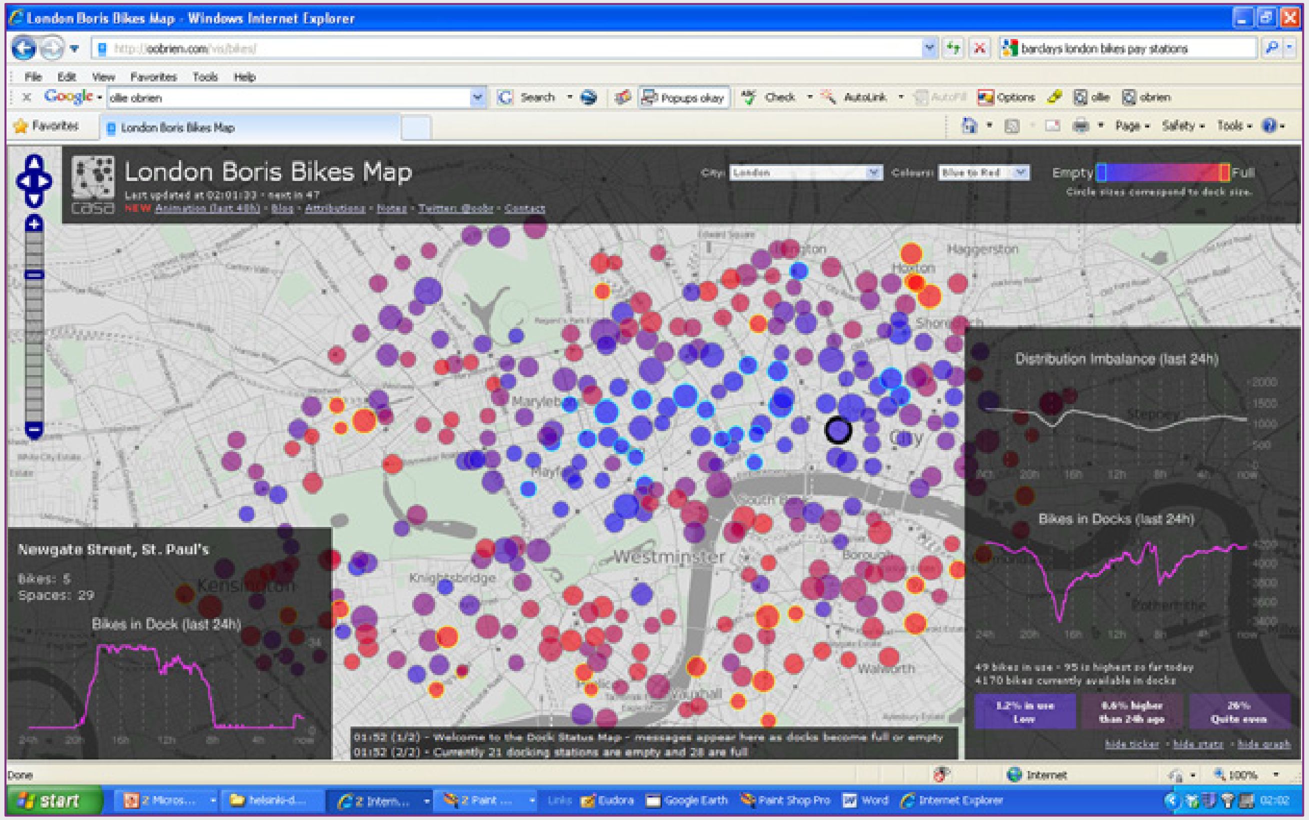
Key

- 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

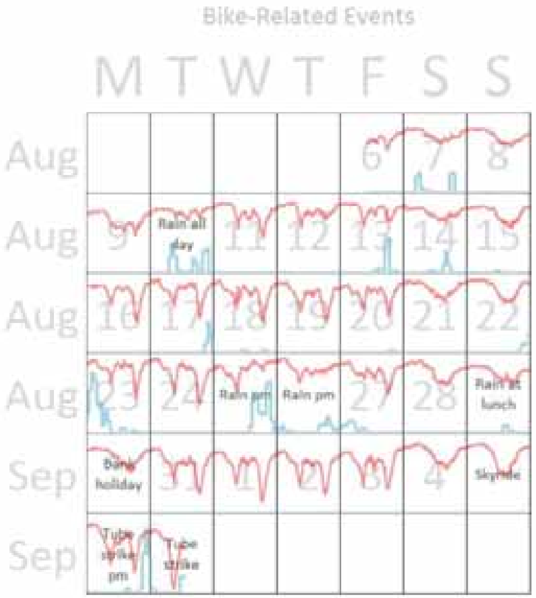


Animations of Public Bike Movements

Animations of Changes in the Bike Nodes: Docking

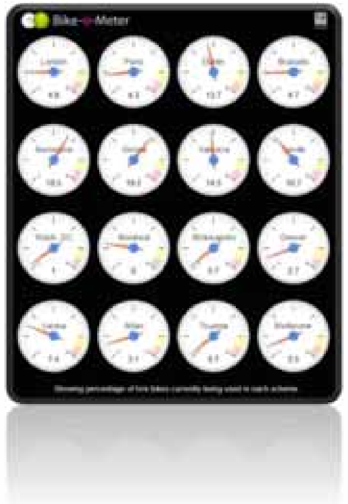
More Analysis

- **London**
- Graph shows number of bikes available to hire
- Effect of rain
  - Using the CASA weather station
- Effect of the tube strikes



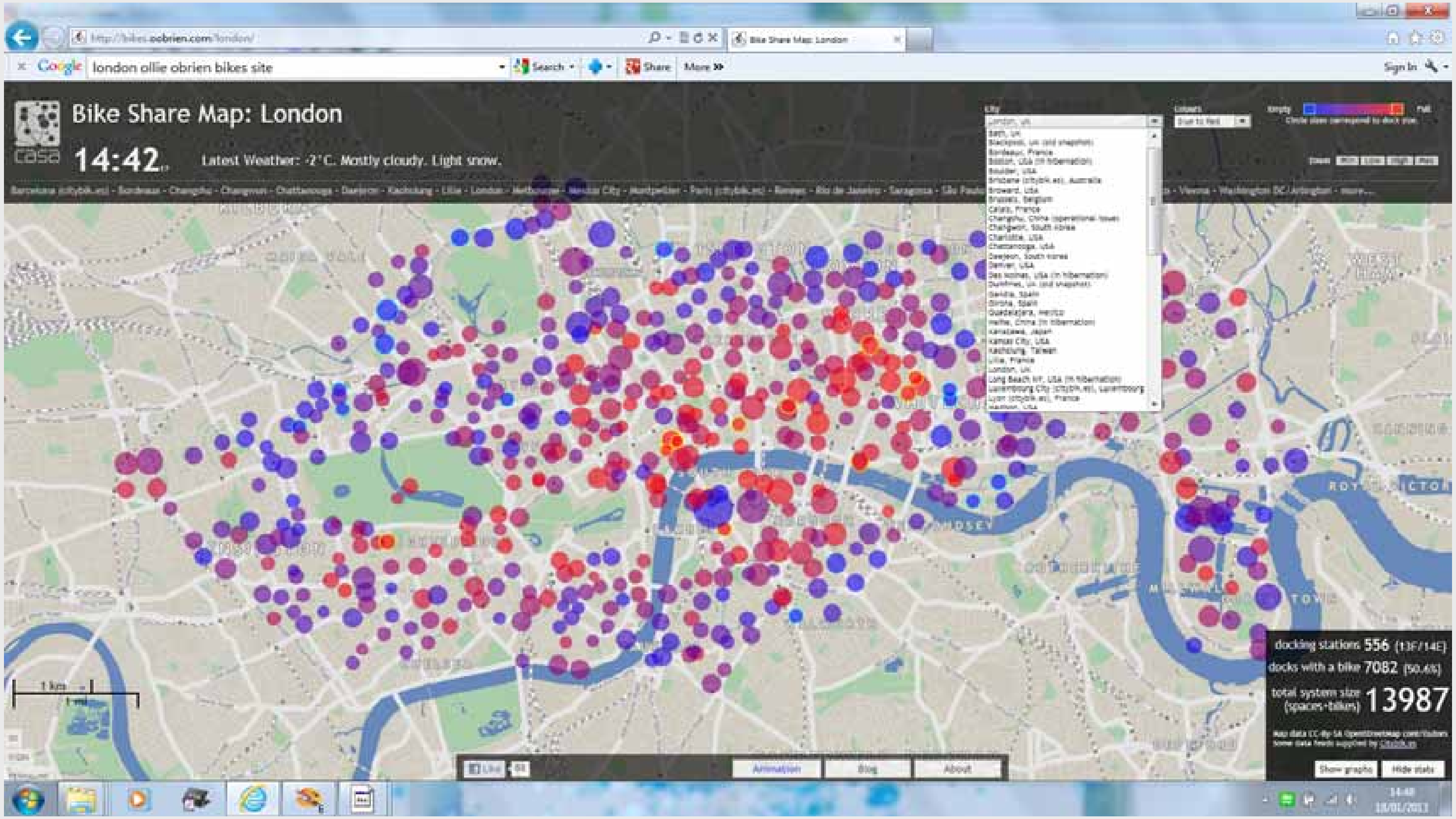
Bike-o-Meter casa.ucl.ac.uk/bom

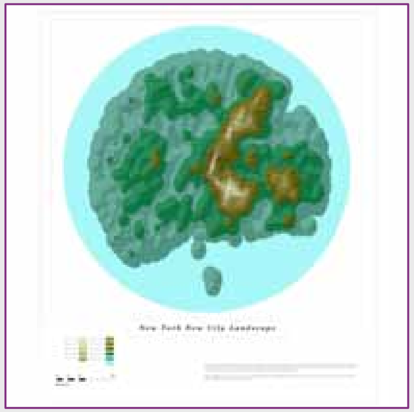
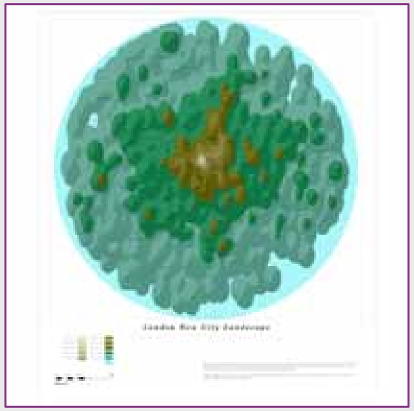
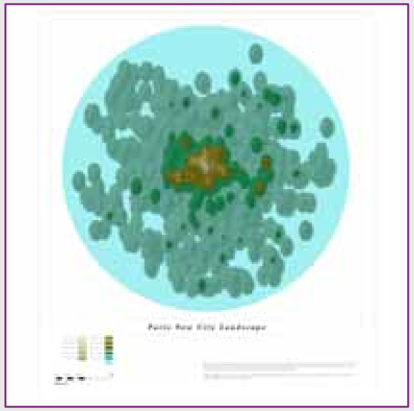
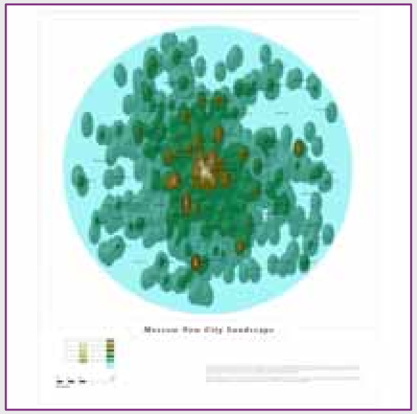
- Tweet-o-Meter for bikes
  - Steven Gray (@frogo)
  - Using Google Gauges
- See the real life Tweet-o-Meters at the new British Library "Growing Knowledge" exhibition
  - Should be easy to hack to show the Bike-o-Meters instead ☺



The Website: Real Time Visualisation of Origins and Destinations Activity

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



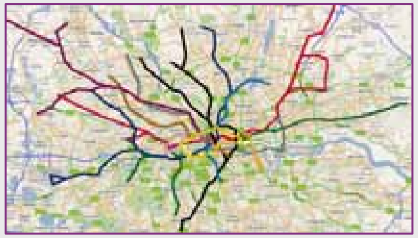
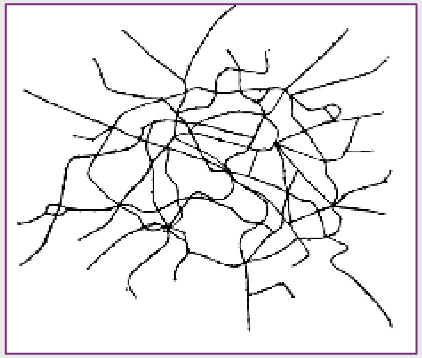
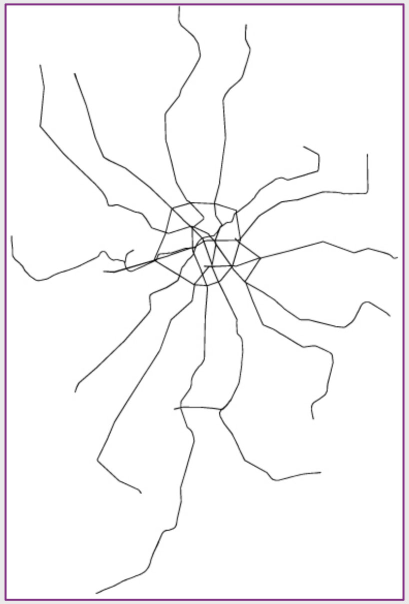





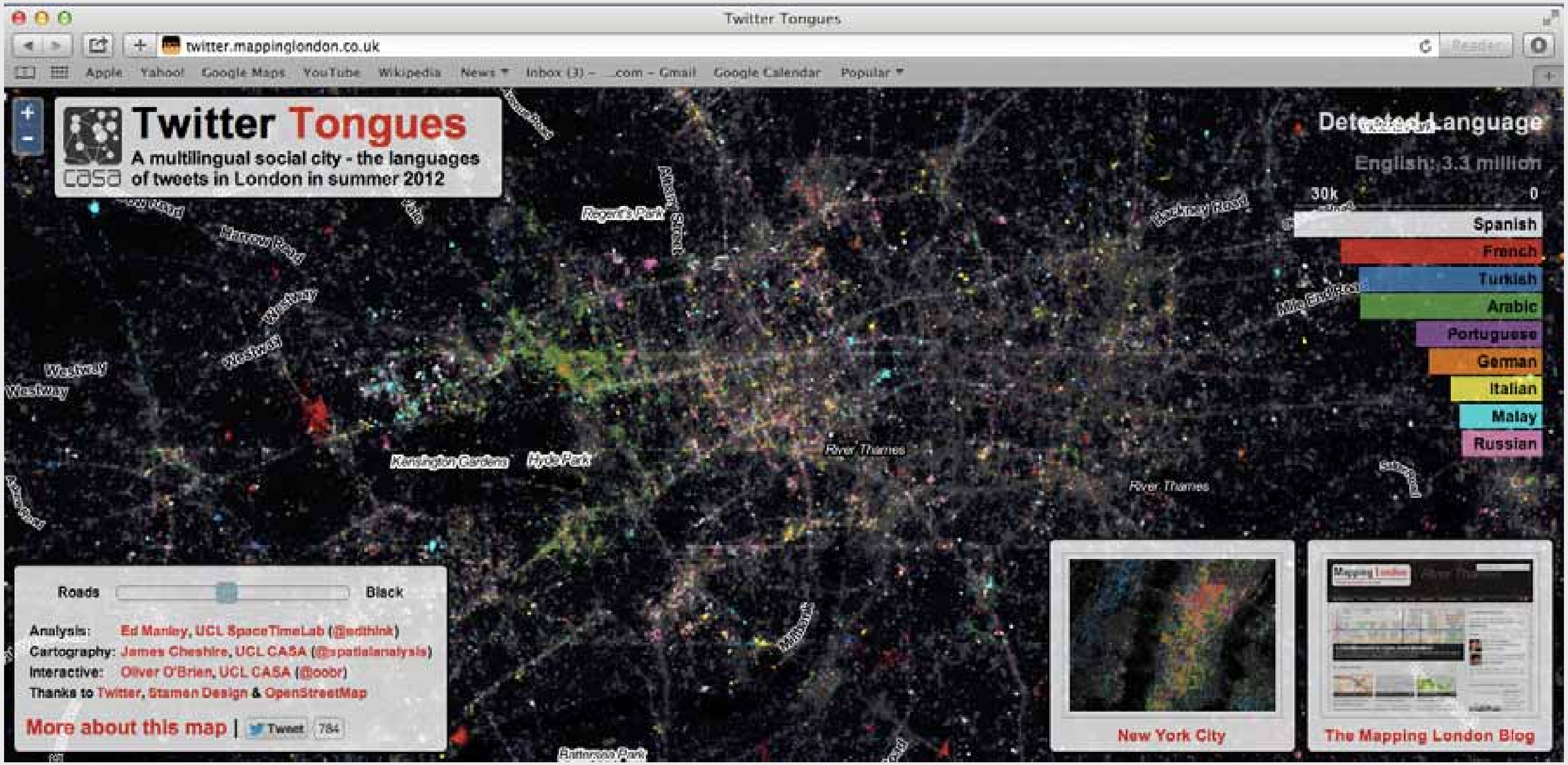
New York

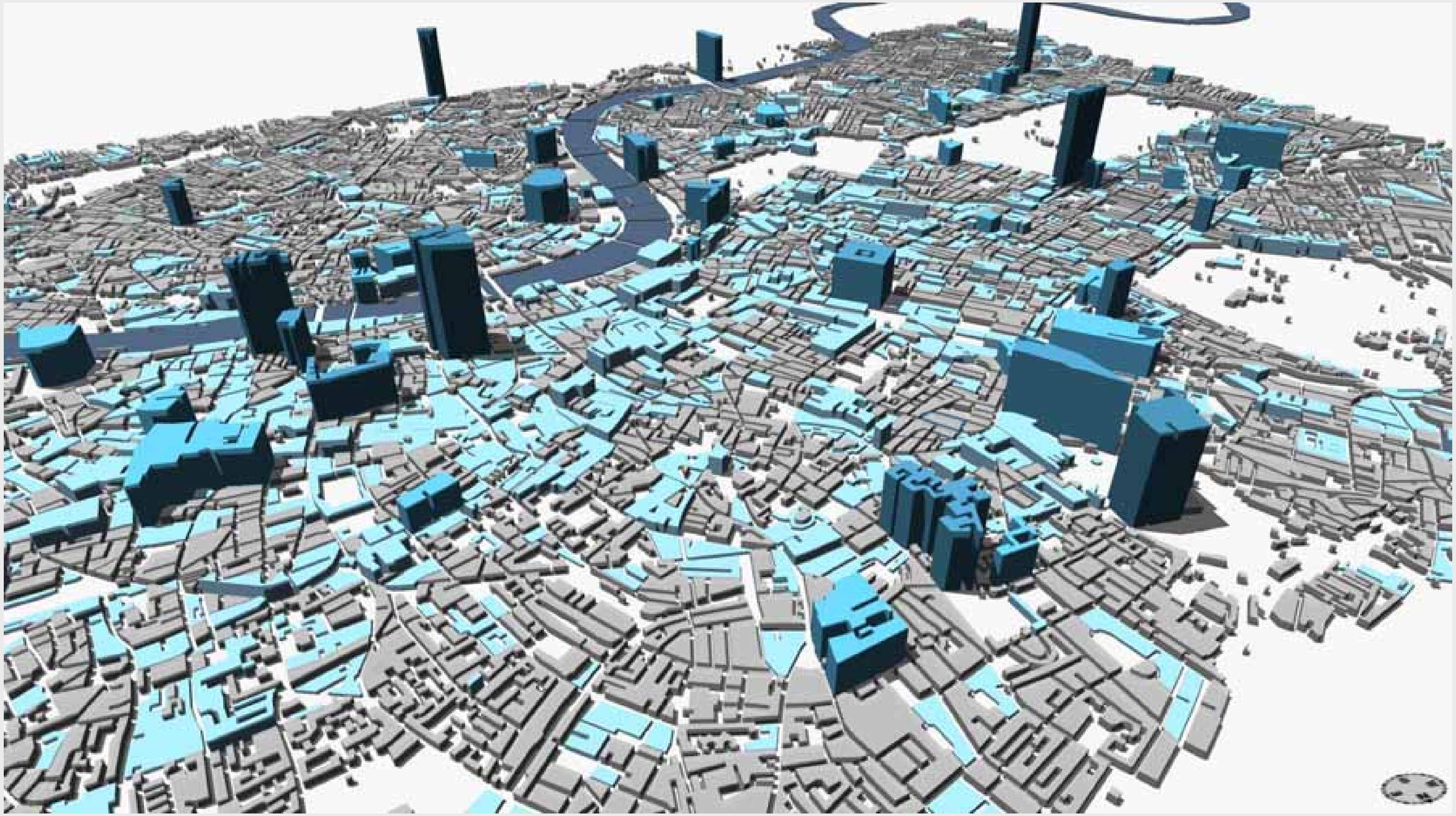
London

Paris

Moscow


Twitter Tongues
twitter.mappinglondon.co.uk
Twitter Tongues
A multilingual social city - the languages of tweets in London in summer 2012
Detected Language
English: 3.3 million
30k
0
Spanish
French
Turkish
Arabic
Portuguese
German
Italian
Malay
Russian
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



What Can We Learn: The Limits to Big Data

We need to add geo-demographics to this data – how – we barely have any possibility of doing this because of confidentiality

We only have a difference between young and old in terms of the card data

Chen Zhong my post doc, now a lecturer at Kings (KCL) has done a lot of work on this relating to extracting such data from related data sets producing synthetic results –a paper in IJGIS

International Journal of Geographical Information Science, 2014
http://dx.doi.org/10.1080/13658816.2014.914521

Taylor & Francis
Taylor & Francis Group

Detecting the dynamics of urban structure through spatial network analysis

Chen Zhong[a*], Stefan Müller Arisona[a,b], Xianfeng Huang[c], Michael Batty[d] and Gerhard Schmitt[a]

References

Manley, E., Chen, Z., and Batty, M. (2016) Spatiotemporal Variation in Travel Regularity through Transit User Profiling, to be submitted.

O'Brien, O, Cheshire, J. and Batty (2014) Mining Bicycle Sharing Data for Generating Insights in Sustainable Transport Systems, **Journal of Transport Geography**, **34**, 262–273

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

Zhong, C., Arisona, S. M., Huang, X., Schmitt, G. and Batty, M. (2014)) Detecting the Dynamics of Urban Structure through Spatial Network Analysis, **International Journal of Geographical Information Science**, http://dx.doi.org/10.1080/13658816.2014.914521

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

Zhong, C., Huang, X., Arisona, S. M., Schmitt, G., and Batty, M. (2014) Inferring building functions from a probabilistic model using public transportation data, **Computers, Environment and Urban Systems, 48**, 124–137

Zhong, C., Manley, E., Stefan Muller Arisona, S., Batty, M.,and Schmitt, G. (2015) Measuring Variability of Mobility Patterns from Multiday Smart-card Data, **Journal of Computational Science**, doi.org/doi:10.1016/j.jocs.2015.04.021

Finding Pearls in London's Oysters

JONATHAN READES, CHEN ZHONG, ED MANLEY,
RICHARD MILTON and MICHAEL BATTY

Public transport is perhaps the most significant component of the contemporary smart city currently being automated using sensor technologies that generate data about human behaviour. This is largely due to the fact that the travel associated with such transport is highly ordered. Travellers move collectively in closed vehicles between fixed stops and their entry into and from the system is unambiguous and easy to automate using smart cards. Flows can thus be easily calculated at specific station locations and bus stops and within fine temporal intervals. Here we outline work we have been doing using a remarkable big data set for public transport in Greater London generated from the Oyster Card, the smart card which has been in use for over 13 years. We explore the generic properties of the Tube and Overground rail system focusing first on the scale and distribution of the flow volumes at stations, then engaging in an analysis of temporal flows that can be decomposed into various patterns using principal components analysis (PCA) which smoothes out normal fluctuations and leaves a residual in which significant deviations can be tracked and explained. We then explore the heterogeneity in the data set with respect to how travel behaviour varies over different time intervals and suggest how we can use these ideas to detect and manage disruptions in the system.

Big Data, Automation and Smart Transit

Automation in transit systems is the most visible sign of how the city is being transformed to enhance the travel experience and efficiency of movement (Batty *et al.*, 2012). There are many ways of achieving this but one of the most significant is the use of smart cards for 'fully automatic fare collection'. These smart cards usually contain the value that the consumer has agreed to load onto the card; they meet stringent requirements for anonymity and security; and their use is such that by tapping in and out of an automated system, correct payments are ensured. Smart cards like this, in fact, go back to the late 1960s and rapid progress in their development was achieved in the 1970s and 1980s when they first made their appearance as phone cards in France. Different varieties of credit card were then emerging too, and by 1984 in places like Hong Kong, stored value cards for use on their new Mass Transit Railway (MTR) had been introduced. By the mid-1990s, contactless cards came onto the scene, first in Seoul with the UPass card, and then in Hong Kong where they introduced the Octopus card, which was then extended to other purchases in the local retail system.

Several other cities followed, but one of the most comprehensive rollouts was in London where, in 2003, the first cards were introduced on the underground ('Tube') system. These are called 'Oyster' cards – partly in tribute, it would seem, to Hong Kong's Octopus card – but the official reason is that the Oyster Card protects its 'pearl' – the stored value – in a 'hard shell'; hence, the name which we have used in the title to this paper. Our particular interest in these 'pearls' is not in their value but in the raw data that can be extracted which covers 'where' the owner of

PLOS | ONE

RESEARCH ARTICLE

Variability in Regularity: Mining Temporal Mobility Patterns in London, Singapore and Beijing Using Smart-Card Data

Chen Zhong[1]*, Michael Batty[1], Ed Manley[1], Jiaqiu Wang[1], Zijia Wang[2,3], Feng Chen[2,3], Gerhard Schmitt[4]

1 Centre for Advanced Spatial Analysis, University College London, London, United Kingdom, 2 School of Civil and Architectural Engineering, Beijing Jiaotong University, No.3 Shangyuancun, Haidian District, Beijing, P. R. China, 3 Beijing Engineering and Technology Research Centre of Rail Transit Line Safety and Disaster Prevention, No.3 Shangyuancun, Haidian District, Beijing, P. R. China, 4 Future Cities Laboratory, Department of Architecture, ETH Zurich, Zurich, Switzerland

* c.zhong@ucl.ac.uk

CrossMark

OPEN ACCESS

Citation: Zhong C, Batty M, Manley E, Wang J, Wang Z, Chen F, et al. (2016) Variability in Regularity: Mining Temporal Mobility Patterns in London, Singapore and Beijing Using Smart-Card Data. PLoS ONE 11(2): e0149222. doi:10.1371/journal.pone.0149222

Editor: Tobias Preis, University of Warwick, UNITED KINGDOM

Received: November 24, 2015

Accepted: January 28, 2016

Published: February 12, 2016

Data Availability Statement: Data are available from the Transport for London (TFL) in UK, Land Transport Authority (LTA) in Singapore and Beijing Transport Committee in China for researchers who meet the criteria for access to confidential data.

Funding: This work was co-funded by the European Research Council (https://erc.europa.eu/) under 249393-ERC-2009-AdG (PI: Michael Batty) and the National Natural Science Foundation of China(http://www.nsfc.gov.cn/) under grant number: 51408029 (PI: Feng Chen). The funders had no role in study

Abstract

To discover regularities in human mobility is of fundamental importance to our understanding of urban dynamics, and essential to city and transport planning, urban management and policymaking. Previous research has revealed universal regularities at mainly aggregated spatio-temporal scales but when we zoom into finer scales, considerable heterogeneity and diversity is observed instead. The fundamental question we address in this paper is at what scales are the regularities we detect stable, explicable, and sustainable. This paper thus proposes a basic measure of variability to assess the stability of such regularities focusing mainly on changes over a range of temporal scales. We demonstrate this by comparing regularities in the urban mobility patterns in three world cities, namely London, Singapore and Beijing using one-week of smart-card data. The results show that variations in regularity scale as non-linear functions of the temporal resolution, which we measure over a scale from 1 minute to 24 hours thus reflecting the diurnal cycle of human mobility. A particularly dramatic increase in variability occurs up to the temporal scale of about 15 minutes in all three cities and this implies that limits exist when we look forward or backward with respect to making short-term predictions. The degree of regularity varies in fact from city to city with Beijing and Singapore showing higher regularity in comparison to London across all temporal scales. A detailed discussion is provided, which relates the analysis to various characteristics of the three cities. In summary, this work contributes to a deeper understanding of regularities in patterns of transit use from variations in volumes of travellers entering subway stations, it establishes a generic analytical framework for comparative studies using urban mobility data, and it provides key points for the management of variability by policy-makers intent on for making the travel experience more amenable.

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