Modern data science and efficient transport systems

CITY SCIENCE
endless possibilities

paul.hewson@cityscience.com

What is City Science?

An ambitious, young technology company within the Oxygen House family

CITY SCIENCE
endless possibilities

What is our mission?

- To optimise complex urban systems, reducing waste and improving lives
- We want to turn the data all around us into clear, scientific insight
- Enable decision makers to understand the effect of a change before it is implemented

How will we achieve it?

- Our team of data scientists and experts are developing analytical tools for the transportation and energy sectors
- We will use multiple datasets, statistical techniques and machine learning to develop scalable insight tools
- Working closely with partner cities we use data to develop solutions

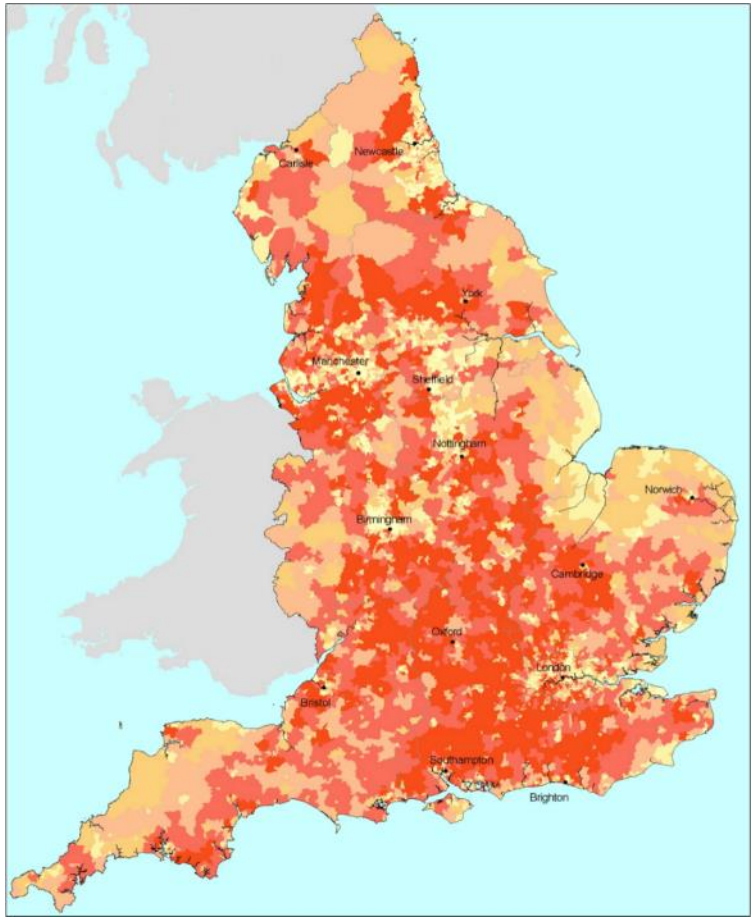
paul.hewson@cityscience.com

Ubiquity of Transport

- Economy
- Housing
- Social function
- Freedom
- Pollution
- Congestion
- Injury
- Health

paul.hewson@cityscience.com

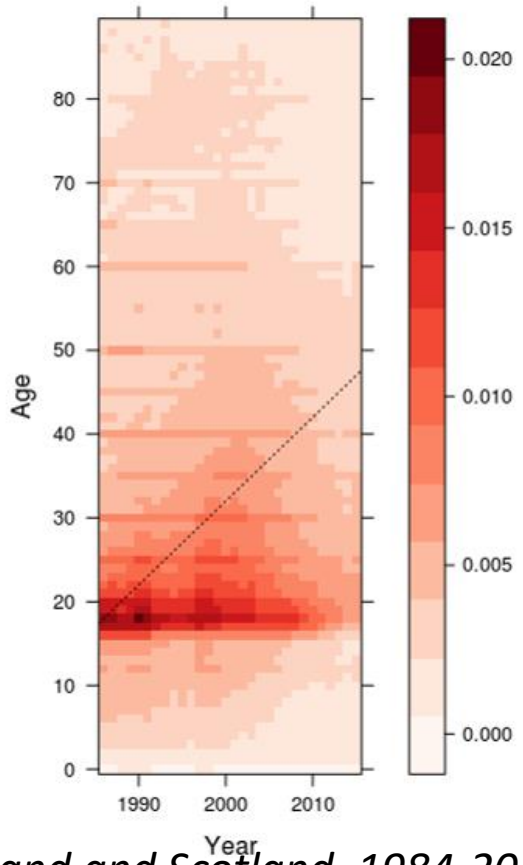
By 2035, diabetes care 'could cost the NHS £17 billion a year'



- Systematic review shows overwhelming health benefits of cycle commuting (11 fold).
- Active transport interventions in yellow shaded areas could have biggest impact on that bill.
- Small area estimates of 1 x 30 participation in Exercise among young adults (red high, yellow low)

paul.hewson@cityscience.com

Strong Cohort Effects in Road User Behaviour



England and Scotland, 1984-2014, per capita road injury rates by age

- **This (simple) tool suggests cohorts born in the 1960s had a lot of fatalities as teenagers, and continued to have a lot of fatalities as they age**
- **Even more powerful effects seen for specific groups e.g. motorbikes**
- **Highlights importance of cohort effects in road user behaviour**

paul.hewson@cityscience.com

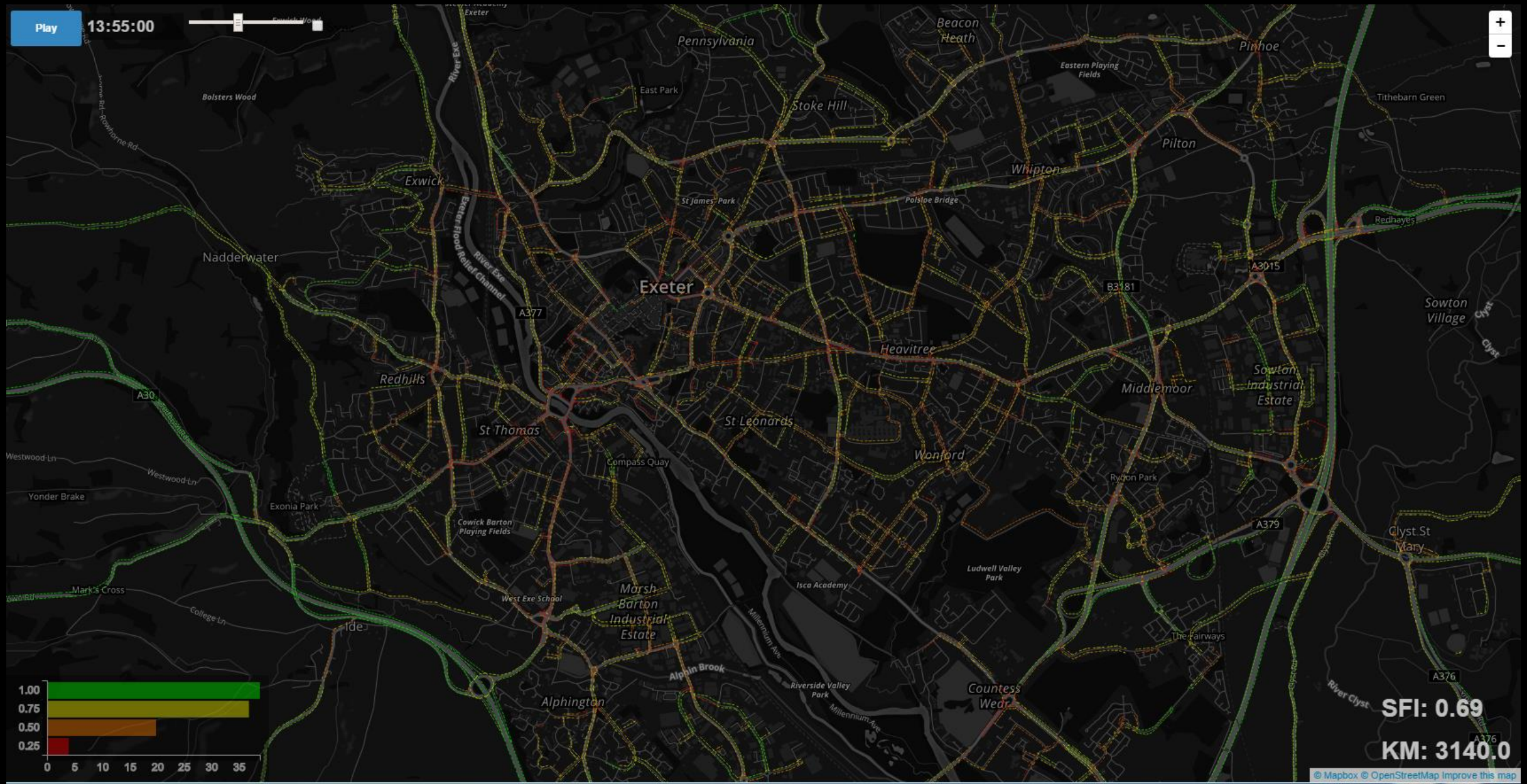
Data Sharing

- **We have been fortunate with support and data provided by local Authorities, local Businesses and other stakeholders**
- **Many other (governmental) datasets are already fully open access data; if not they are available from UK Data Archive under varying levels of restriction**
- **Ethics issues have been thought through – *i.e.*, we should use professional approaches to disclosure control rather than *ad hoc* methods**
- **Köln: https://www.offenedaten-koeln.de/dataset/verkehrskalender-der-stadt-k%C3%B6ln**

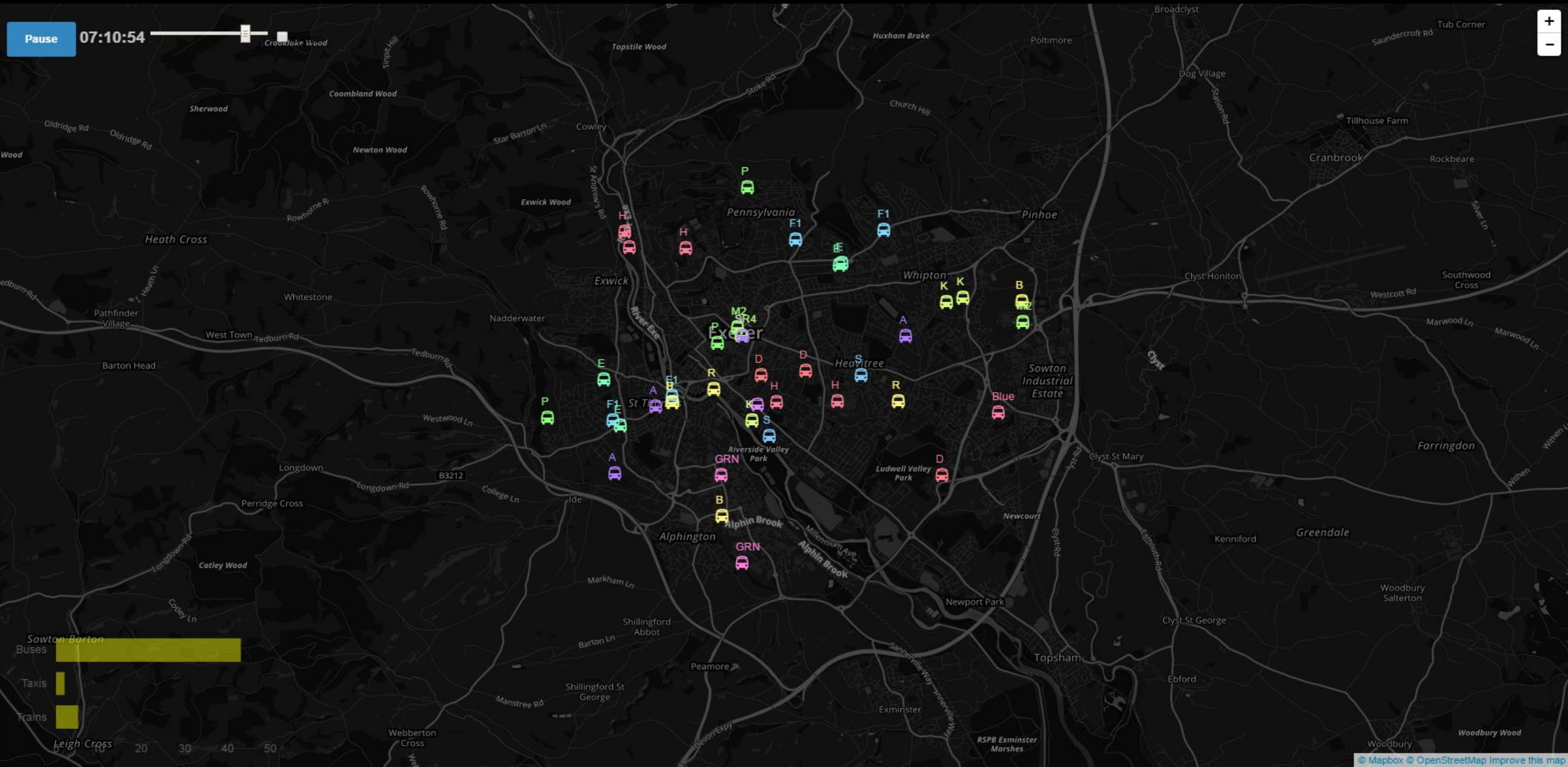
National Guidance on Disclosure Control is available

paul.hewson@cityscience.com

CITY SCIENCE
endless possibilities

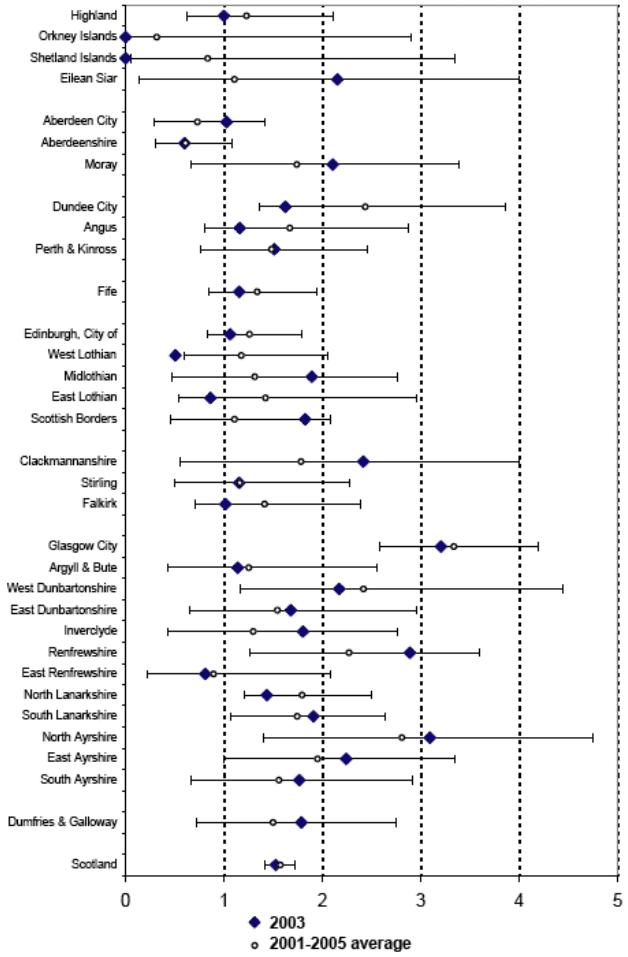


CITY SCIENCE
endless possibilities



CITY SCIENCE
endless possibilities

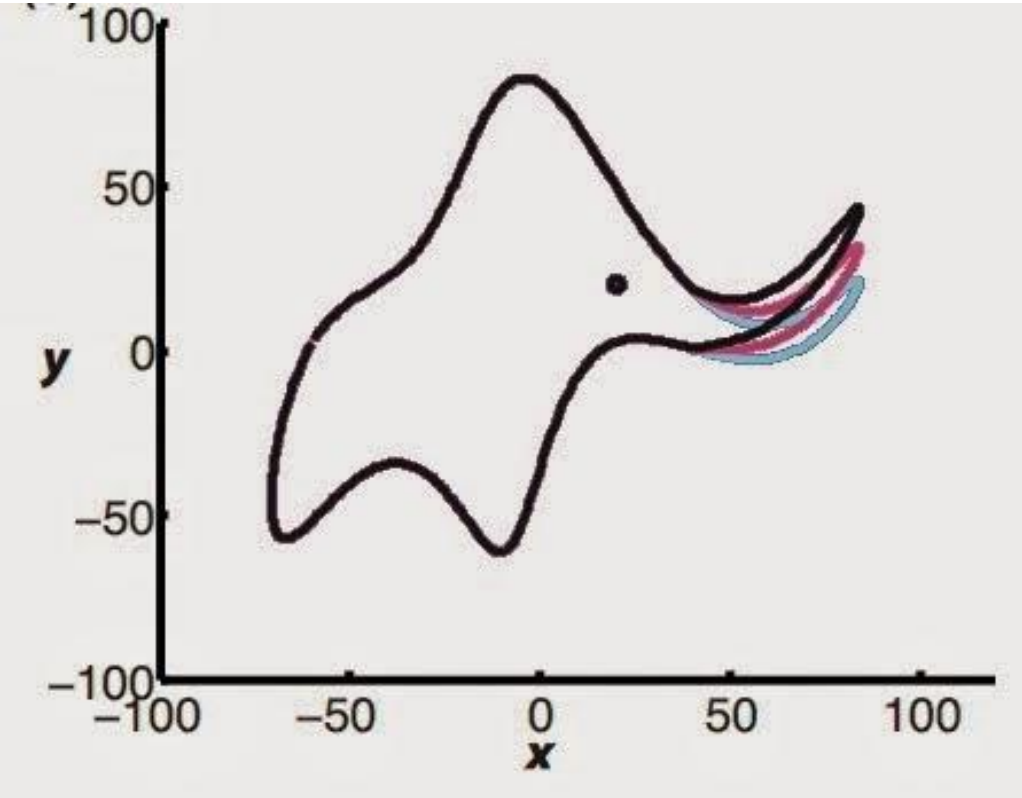
Statistical modelling and uncertainty



Scottish government estimates of child risk of fatal or serious injury as road user.

http://www.gov.scot/Publications/2009/11/23103624/66

Statistical Modelling and overfitting



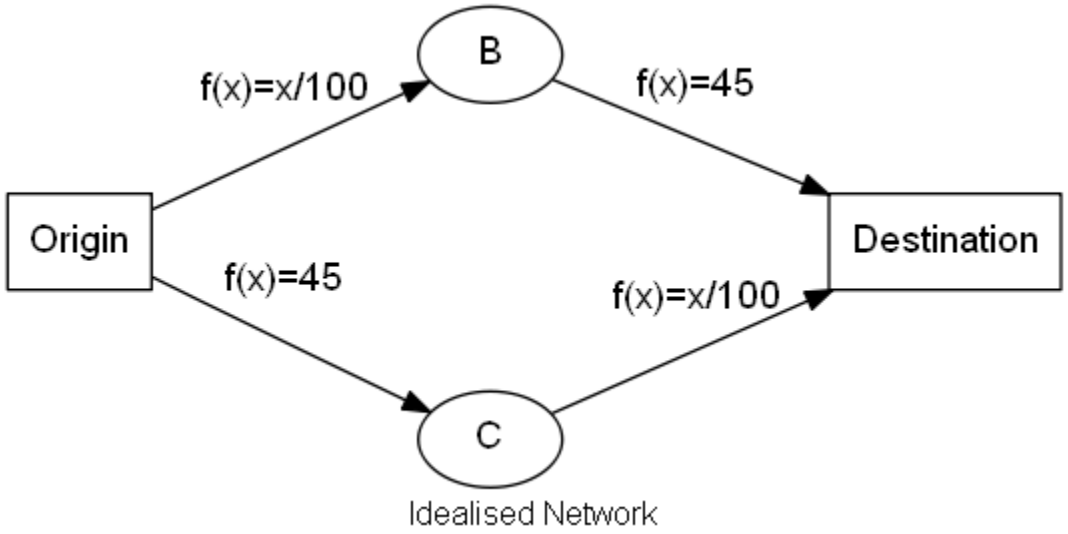
Von Neumann:
"With four parameters I can draw an elephant"

- "Drawing an elephant with four complex parameters" by Jurgen Mayer, Khaled Khairy, and Jonathon Howard, Am. J. Phys. 78, 648 (2010), DOI:10.1119/1.3254017

Assumptions to challenge

- **Are all the "parameters" (seasonal adjustments, annualization factors) sensible described as single fixed points.**
- **Transport networks used by humans (even if we all get driverless cars)**
  - **Do we always behave as economically rational mode / route selectors?**
  - **Do we all have 100% regularity in travel activity as we described in the census / that road side interview we did in 1998?**
  - **Does all our travel involve a single origin / destination?**

All models need ***ASSUMPTIONS***

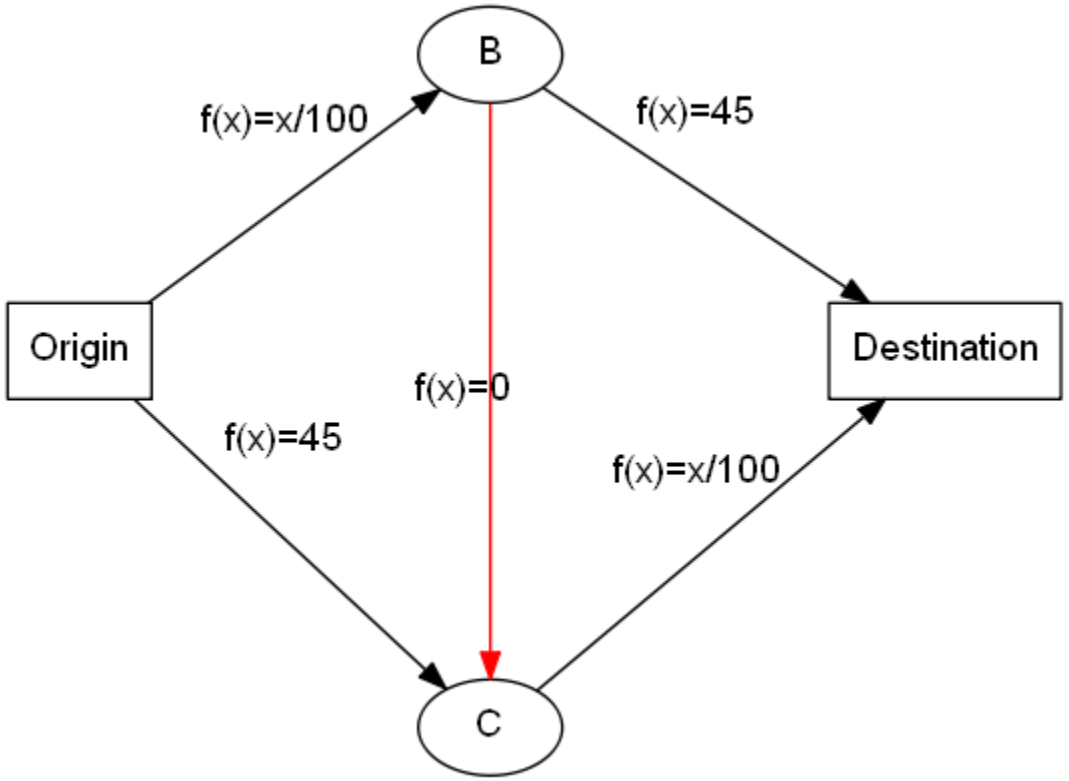


Idealised Network

For illustration - Braess paradox

- **4,000 cars travelling from Origin to Destination.**
- **A transport model would determine**
  - **equilibrium distribution has 2,000 vehicles on each route (O->B->D and O->C->D)**
  - **Optimum travel time 65 minutes.**

Mathematical modelling

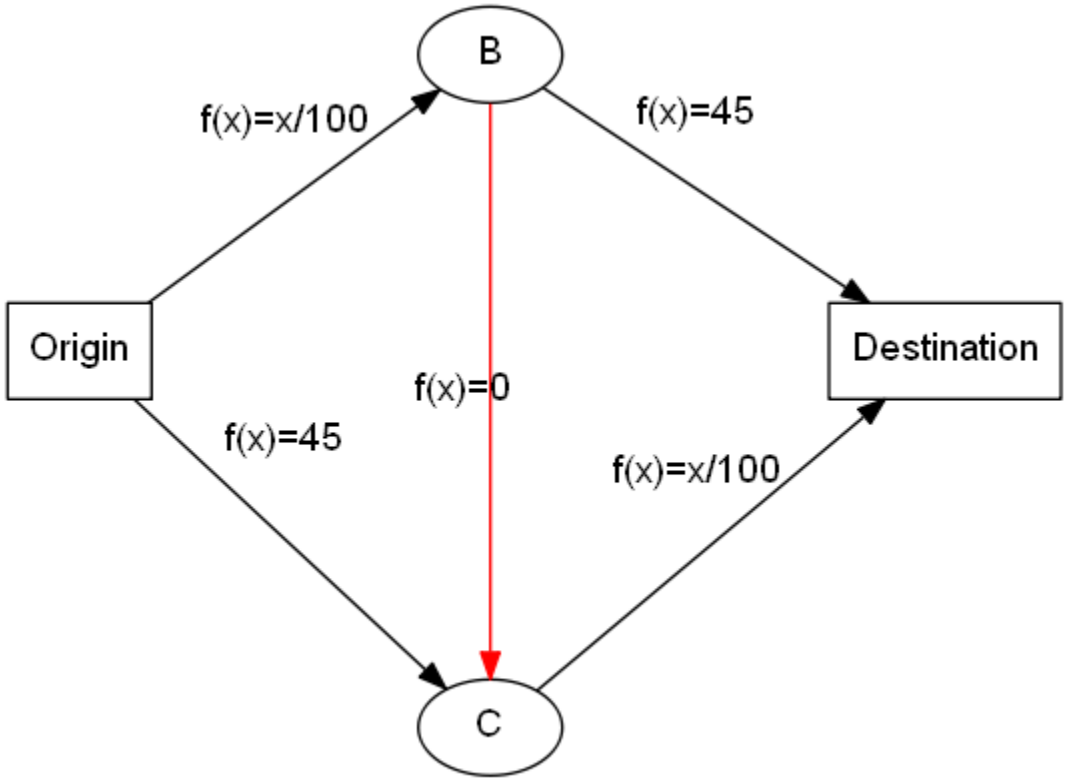


Idealised Network with Braess Paradox

Now add a link.

- **A transport model would determine**
  - **equilibrium distribution has 4,000 vehicles Origin -> B -> C -> Destination**
  - **Optimum travel time 80 minutes.**
- **There are real world examples used to argue this is a real phenomena.**

Mathematical modelling



Idealised Network with Braess Paradox

- **Filters (assumptions)**
  - **Drivers all have the same cost function (minimising time). What about avoiding congestion, difficult turns, traffic lights?**
  - **Drivers have only one purpose for a trip**
  - **What temporal and routing alternatives are available, and how do we cost them?**
  - **Actually, when evaluating public transport or road tolling we *don't* use these assumptions!**

CITY SCIENCE
endless possibilities

lightfoot Collaboration

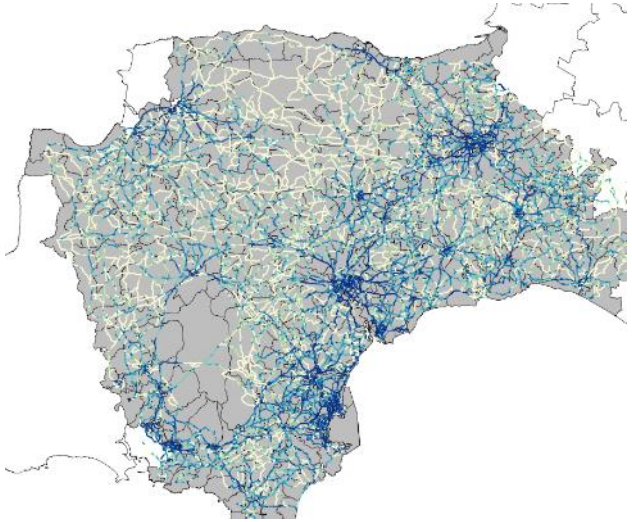
- Route choices
- Junction behaviour
- Independent validation for pre and post activation

paul.hewson@cityscience.com

CITY SCIENCE
endless possibilities

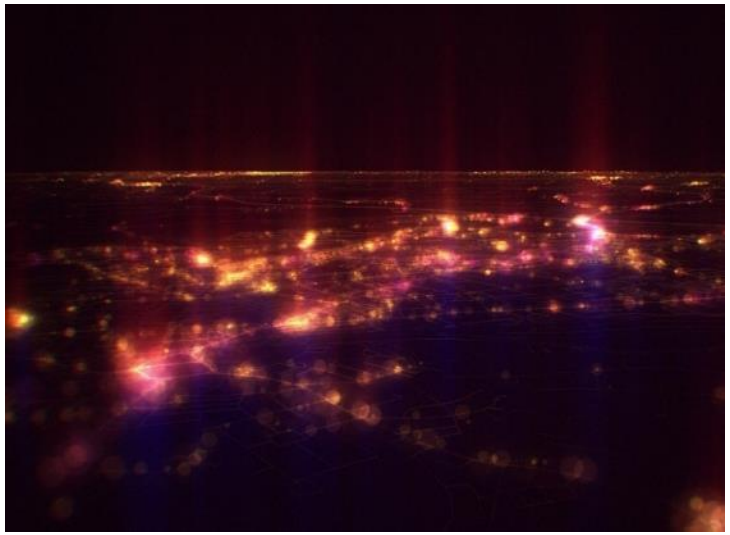
Data Collection

- New Data Collection Strategies
- Collating a wide-range of datasets in true partnership with local authorities, transport operators, private companies and beyond.



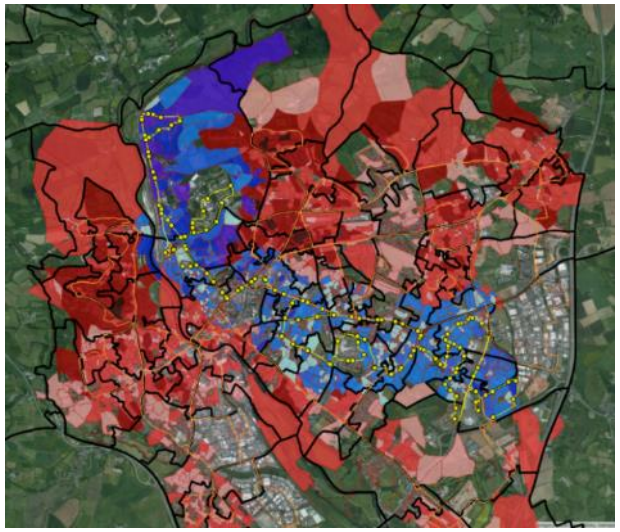
Insight

- Data-driven decision making with current state and a wider range of 'what-if' analyses.
- Using sophisticated modelling to identify pinch points and suggest low cost solutions to ease congestion – 'virtual roads'
- More analysis of how new trends will impact existing network infrastructure



Some Ideal Results

- Flexible transport plans enabling cities to adapt to the rapidly changing transportation environment.
- Integrating business cases which ensure that sustainable solutions make economic sense and can meet projected demands.
- Valuing a wider range of outcomes



A COLLABORATIVE DATA SCIENCE APPROACH TO TRANSPORTATION


CITY SCIENCE
endless possibilities
City Strategy | Data Science | Transportation | Energy
ANY QUESTIONS?
paul.hewson@cityscience.com