

Designing with Data

Weeks 1 and 2

James Stewart, Ewan Klein

Weeks 1 and 2

Sources and Forms

Uses and Abuses

- ◆ Find or create data
- ◆ Critically evaluate it : value, validity and life story
- ◆ Understand and handle it

Why do you need to critically evaluate data?

To judge it's value and validity in helping you meet your objectives.

Validity:

- The degree a measurement measures what it is supposed to measure.
- That a measurement is well-founded and corresponds to the real world

Value:

- The importance, or worth or usefulness

Quality and representativeness

Design Goal

- Change the world (for the better) e.g.
 - tool for action
 - Evidence for decisions or policy
- Unsuitable, invalid data will mean your design is created on the wrong basis
 - Wrong assumptions
 - Wrong hypothesis, theory or model

Error shows up when:

- non-use,
- Challenged
- Something worse

Source of Data and Form

- Motivation for collection
- Instrument of collection
- Control/comparator
- Structure
- Ownership/Ethics

Motivation

- *Who* commissioned data, *why*, how much did it *cost*, *how important* was it
- What were they trying to do?
 - What questions were being asked?
 - Are they collecting data for indicators?
 - Were they conducting an experiment?
 - Were they testing a model?
 - Were they trying to prove causality?
 - Doing it on the cheap?

Instrument

- *Means* E.g.

Survey: what questions, when, how, who

Machine: data logs, physical sensors?

Sample or whole population data set?

- *Question the:*

Design of instrument

Design of Administration/deployment

Testing and validity of instrument: calibration

e.g. validity of questions

Cost

Units – standardised or adhoc

Expected quantified errors,

unexpected potential human/environmental error

Comparator/ Control

- What data items can be used for comparison and/or control
- E.g. age, ethnicity against census data, school roll data, etc
- Time series data

Data structure

A continuum

- Unstructured data: free text, prompted or unprompted
- Semi-structured – text answers to tested questions
- Structured - controlled answers
- Numerical structured

Ownership and Ethics

- Are you allowed to use it?
- Under what conditions?
- Why? Licence, Law.
- Who might be affected by you using the data?
- Was consent obtained for your use?
- Does linking with other data potentially go against licence, law or consent?
- Example
- <http://vartree.blogspot.co.uk/2014/04/i-know-where-you-were-last-summer.html>

Questions

- Explore different techniques to used to establish:
- Validity of data
- Representativeness
- What makes good quality quantitative data?
- What makes good qualitative data?

GPS Links

- Open Streetmap GPS traces

- https://wiki.openstreetmap.org/wiki/API_v0.6#GPS_traces
- Visualisation: <https://www.mapbox.com/blog/openstreetmap-gps-layer/>
 - OSM mapping data: <http://www.openstreetmap.org> ; <https://www.mapbox.com/data-platform/>

- Endomondo

- <https://www.endomondo.com/>
- My Example: <https://www.endomondo.com/workouts/412993269/11812342>

- Mapmyride

- <http://www.mapmyride.com/workout/749547209>

- GPSIES

- http://www.gpsies.com/#11_55.94999694824219_-3.1999969482421875_cycle

- Wikiloc

- <http://wikiloc.com/>

- OS to GPS convertor

- <http://www.ordnancesurvey.co.uk/gps/transformation>
- <http://www.ordnancesurvey.co.uk/business-and-government/help-and-support/navigation-technology/os-net/coordinate-spreadsheet.html>

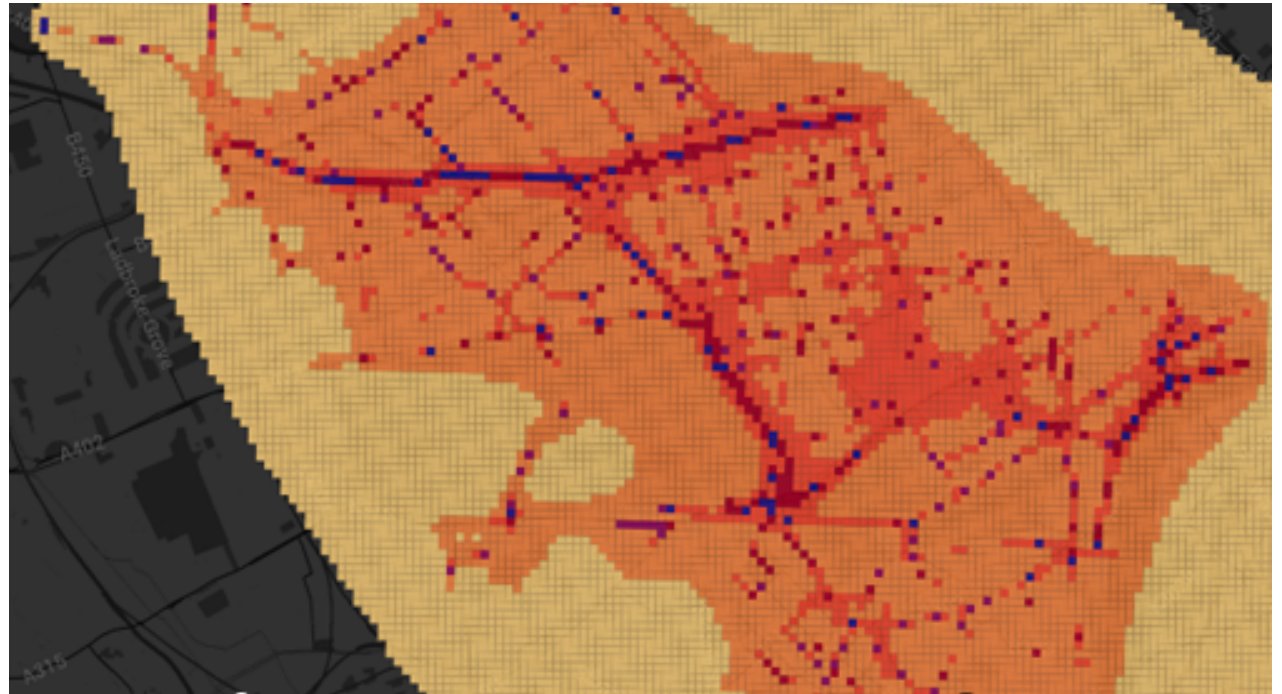
NO2 survey

Greenpeace competition

<https://comp-soc.com/ilwhack/page/challenge-foe>

FoE Advocacy page

<http://www.foe-scotland.org.uk/node/1740>



CiMo (Cambridge Uni) <http://www.eeci.cam.ac.uk/cimo/index.shtml>

Odd things

- <http://mashable.com/2010/08/06/twitter-mapping-tools/>
- <http://spatial.ly/2012/10/londons-twitter-languages/>

Article: Mapping the global Twitter heartbeat: The geography of Twitter, by Kalev H. Leetaru, Shaowen Wang, Guofeng Cao, Anand Padmanabhan, and Eric Shook. *First Monday*, Volume 18, Number 5 - 6 May 2013, <http://firstmonday.org/article/view/4366/3654>

- A tweet:
https://github.com/edinburghlivinglab/cyclehack/blob/master/twitter_rw/tweet.json
- Cycle counters details
<http://edinburgh.stv.tv/articles/279922-sustrans-explain-technology-behind-new-union-canal-bike-counter/>
- CiMo (Cambridge Uni) <http://www.eeci.cam.ac.uk/cimo/index.shtml>
- UK OpenGov
 - <http://data.gov.uk/>