OLD - Mapping Museums – Notes

# People

* Fiona Candlin
* Alex P
* Jamie Larkin <j.larkin@bbk.ac.uk>
* Nick Larsson <nick@dcs.bbk.ac.uk>

# Web platform

* <http://193.61.37.198/browseproperties>
* To access the v7 data : <http://museweb.dcs.bbk.ac.uk/home>
* and you will notice the v7 in the header to the top right.
* To access v8 data: <http://museweb.dcs.bbk.ac.uk:8080/home>
* and you will notice the v8 in the header to the top right.
* user: museumuser : tlldob@bbk

# Current work

* TODO AB:
  + TODO
  + Production version 0.1: museweb.dcs.bbk.ac.uk/home (v7)
  + Get info about latest version (v8)
    - New govn data: MM\_GovernanceChange\_05.06.18.xls
    - Class2018 -> Subject matter (drop classif)
    - Duplicated ID: mm.misc.195
  + All project files: <http://193.61.37.198/showpage/meetings.html>
  + Data folder: <https://www.dropbox.com/home/Mapping%20Museums%20files/Main%20datasheet>
  + Museum closures [urgent, end of Feb]
  + [Density of visits]
  + [Find tourism]
  + Modelling for museum size based on other variables (new dataset)
  + [analysis of governance]
* Abstract for Spatial Humanities 2018 <https://docs.google.com/document/d/1NNiQoMmq-VCFAIFmqgS2ED2ZATncGB0TAz4-BhXoJU4/edit?usp=sharing>
* Meeting 30 Jan 18
  + Public engagement (social media)
* Meeting 23 Jan 18
  + Date logic – fuzzy, don’t reinvent the wheel
  + Fuzzy vs crisp time ranges
* Meeting 16 Jan 18
  + TODO
  + Location hierarcy: 4 levels
    - Country -> Region ->
    - Country -> Region ->
* Meeting 5 Dec. 17
  + Visualisation
  + Data upload: 3rd January
  + Core spreadsheet (core data) + auxiliary ones
  + Temporal sequences in governance
* Spatial humanities 2018 <http://wp.lancs.ac.uk/dighum/sh2018>
  + Deadline: 28 Feb
  + Conf: 20 Sep 2018
  + Create Google Doc for abstract
* Get dataset from Fiona (done)
* Monthly plan from offline work now to June/Jul
  + December 2017-15 Jan 2018: data cleaning / descriptive statistics / closure analysis
  + 16 Jan – Feb 2018: Demographic analysis (Census 2011/2001)
  + March 2018: Geo-demographics analysis (OAC classes) / urban-rural
  + April 2018: Voting behaviour / Brexit analysis
  + May 2018: Map production / geo-visualisation
* Extra datasets to include on the web platform
  + Granularities
    - TODO: insert hierarchy (insert into the powerpoint)
    - Output area (Census unit that includes 50 households)
    - LSOA (1500 people)
  + Data has the format
    - museum1 is\_located\_in output\_area1
    - output\_area1 has\_category cat1
    - output\_area1 has\_population yyy
    - cat1 has\_description “This category represents…”
  + To retrieve context info about a museum, join museums and datasets on the GSS
  + The interface could show the info as extra attributes, e.g., British Museum -> London cosmopolitan output area; -> urban area (low priority)
  + Search by OAC categories or ruralurban categories?
  + Datasets currently ready for insertion (provided in CSV format):
    1. Output area classification
       - OA level: museums17\_oa\_oaclassif11\_categ.tsv
       - LAD level: museums17\_lad\_oaclassif11\_categ.tsv
    2. [skip] Rural urban: museums17\_oa\_ruralurban11\_categ.tsv [England]
    3. Deprivation index: museums17\_lsoa\_deprivation15\_deciles.tsv
* Spatial analysis on museums and
  + **Closures**
  + Urban/rural
  + OAC classes (for individual museums)
  + Population (#)
  + Income
  + Education
  + Ethnic composition
  + Brexit vote
  + Census 2011, 2001 [1991]
* Think about conferences where I would like to go in 2018
* 23rd Feb workshop on closures

# Data to include in the DB

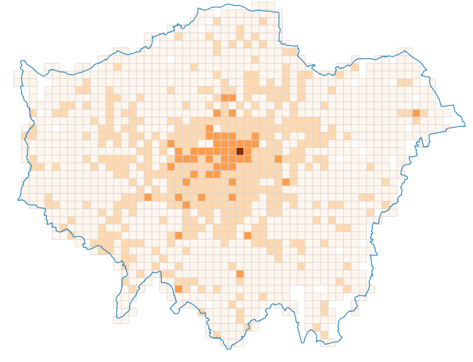
* Data folder by Andrea: <https://www.dropbox.com/sh/98x7y3epgts9g09/AACjC5k3S3JF64kDM3xq37-sa?dl=0>
* **Boundaries** (shapefiles or GeoJSON with GSS codes). What do you currently have? Here’s what’s needed for integration with Census data:
  + [ GSS codes: [http://statistics.data.gov.uk](http://statistics.data.gov.uk/) ]
  + [ <https://borders.ukdataservice.ac.uk/easy_download.html> , non-clipped are better]
  + Countries [4]
  + Regions [9]
  + Counties
  + Local authorities
  + Output areas: <https://data.gov.uk/dataset/output-areas-oa-boundaries>
  + Northern Ireland: <https://www.opendatani.gov.uk/dataset?q=boundary>
* **OS Open Names**: Ordnance Survey DB of place names in the UK, as linked data. This dataset is useful to generate custom labels on the map, selecting different types of places (villages, cities, etc.).
* **Census 2011**: tables to be downloaded from <https://www.nomisweb.co.uk/census/2011/quick_statistics>
  + Downloaded datasets:
    - Population <https://www.nomisweb.co.uk/census/2011/qs102ew>
    - Level of qualification <https://www.nomisweb.co.uk/census/2011/qs501ew>
    - Industry <https://www.nomisweb.co.uk/census/2011/qs605ew>
    - Ethnic <https://www.nomisweb.co.uk/census/2011/qs201ew>
    - Religion <https://www.nomisweb.co.uk/census/2011/qs208ew>
* **Census 2001**: tables to be downloaded from

* + <https://www.nomisweb.co.uk/home/census2001.asp>

* + <https://www.statistics.digitalresources.jisc.ac.uk/?query=2001&sort_by=search_api_relevance&sort_order=DESC>
* **Income**: median income per unit, also present in deprivation index
* **Property value**: 1995-2017
* **Output Area Classification (OAC)**: useful for a demographic summary of each area
* **Deprivation index**: summary of deprivation of each local authority
* **Urban functional areas**: Urban Audit Functional Urban Areas 2016; Urban Audit Core Cities 2016; Urban Audit Greater Cities 2016; at <http://geoportal.statistics.gov.uk>
* **Urban Rural classification** **2011**: <https://www.gov.uk/government/statistics/2011-rural-urban-classification>
* **TODO**: retrieve:
  + Income (per district if available), from <https://www.ons.gov.uk/peoplepopulationandcommunity/personalandhouseholdfinances/incomeandwealth>
  + armed forces (per county, district, from Census 2011)
  + Brexit vote

# Ideas for Visualisation

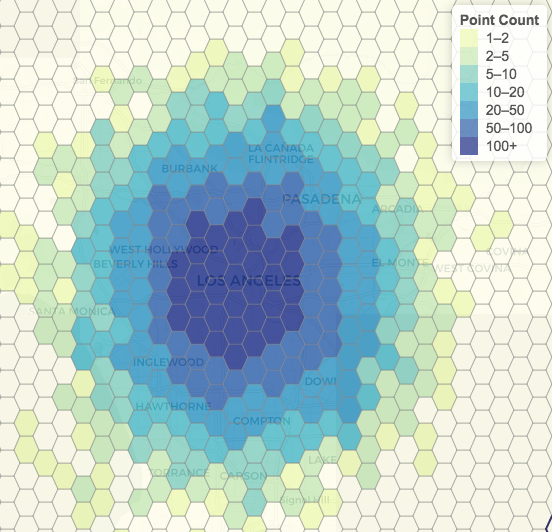
* Web mapping state of the art: <http://datashine.org.uk>
* **Base map** should be lighter, less cluttered, possibly without labels. Ideas here:
  + <http://leaflet-extras.github.io/leaflet-providers/preview/>
  + see for example *Esri.WorldGrayCanvas*
  + (Check tile licenses to ensure legal ground for use)
* **Place names** should be place above objects and never below. Layers of counties/cities/towns would also allow the user to enable/disable them when needed. Source data can be the Open Names DB, from which we can select cities/towns etc.
* To show **where museums are**, clusters of pins are not ideal, as the spatial structure is lost. I suggest a multi-scale approach: at high zoom levels (region / county), let’s use a **hexagonal or square grid**, with gradual colours, with a legend, with light colour = low count, dark colour = high count:

The data groups can be calculated with different algorithms:

* Jenks
* Quantiles
* Equal interval
* (*Quantiles* is the most understandable)

Hex grids are aesthetically pleasing, and better to show the distribution of data:

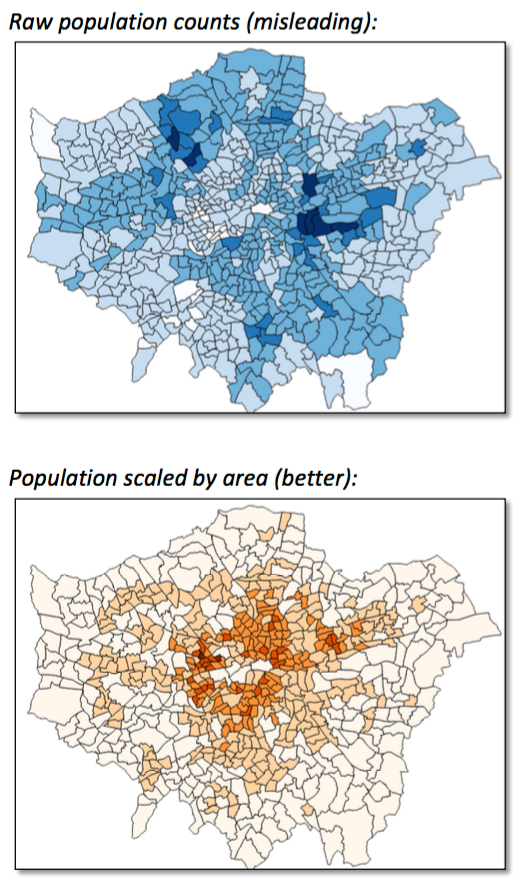


From <http://jsfiddle.net/nathansnider/ts4oyyn0>

Unlike the example above, areas without museums can be left empty, and labels should be rendered above the grid.

This approach is also good because the page loads a few hundred cells instead of thousands of museums. The cells can be pre-computed or computed on the fly. When the user zooms close enough (e.g., neighborhood or street zoom), this view can switch to a **pin/point view**, where every single museum is rendered. One can trigger the switch based on how many museums are present, e.g., if (n\_museum > 5 ) use grid, else use points.

When displaying data based on **irregular administrative boundaries**, absolute counts should **never** be used, because they can result in a misleading impression. One should always use **relative measures**, e.g., ratios, percentages, etc. For example, this is raw counts of poulation in London, vs the population density (pop per square mile):



As we are talking about a human phenomenon, **population** is one of the best measure to scale any museum metrics, for example, ( museums / population ). Other valid examples include ( museum about agriculture / tot museums in the area ), ( museum about mining / miners who live in the area ).

For non map-based visualisation, **D3** is a very powerful Javascript library: <https://github.com/d3/d3/wiki/Gallery> .

# Research questions

~4K museums

* Correlation between deindustrialisation, unemployment, tourism
* Museums as indicator of grassroots concern, lower-middle class, retired teachers, shop owners
* Factors
  + Types of museums:
    - Accredited vs non-accredited
  + Overall population
  + Economic development
  + Income per capita
  + Industrialisation / deindustrialisation
  + Rural vs urban index
  + Property prices (cost of land/facilities)
  + Level of education
  + Political orientation
  + Brexit vote as a proxy to marginalisation
* 1960-2017 time frame

# UK data

Administrative units

* This is a nice overview of the whole UK: <https://en.wikipedia.org/wiki/NUTS_statistical_regions_of_the_United_Kingdom>
* NUTS 1 (very large): <https://data.gov.uk/dataset/nuts-level-1-january-2018-super-generalised-clipped-boundaries-in-the-united-kingdom>
* NUTS 2: <https://www.europeandataportal.eu/data/en/dataset/nuts-level-2-january-2018-full-clipped-boundaries-in-the-united-kingdom>
* NUTS 3: <http://geoportal.statistics.gov.uk/datasets/473aefdcee19418da7e5dbfdeacf7b90_4>
* LAU 1: <https://www.europeandataportal.eu/data/en/dataset/local-administrative-units-level-1-january-2018-generalised-clipped-boundaries-in-united-kingdo>
* LAU 2 (very small): <https://www.ordnancesurvey.co.uk/opendatadownload/products.html#BDLINE> (I couldn't find them elsewhere)
* Alternatively, this are the NUTS boundaries for Europe in shp/GeoJSON, so you need to extract the UK from it:

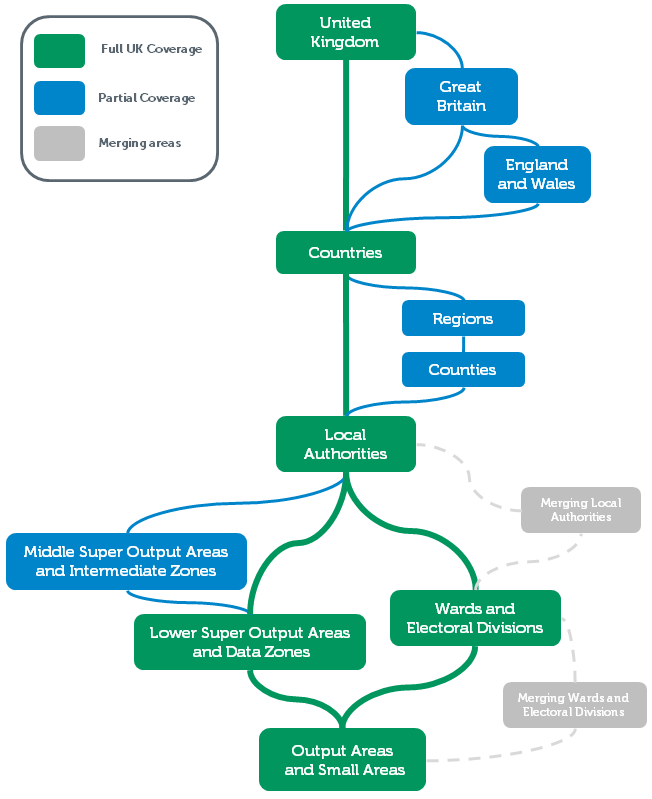
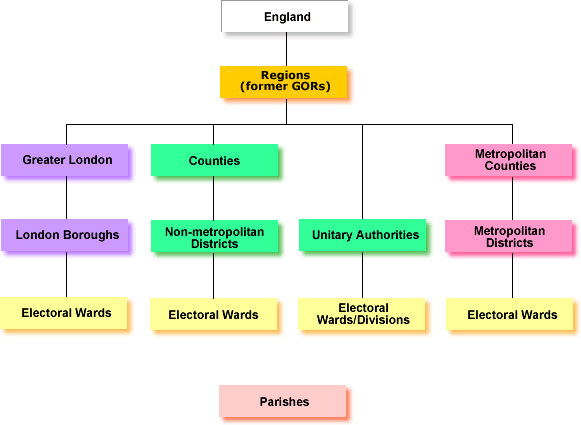
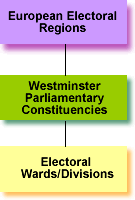
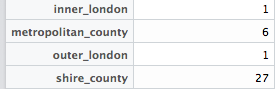
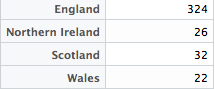
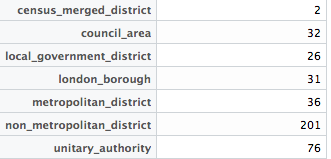
<http://ec.europa.eu/eurostat/web/gisco/geodata/reference-data/administrative-units-statistical-units/nuts>

* UK census boundaries (which are partly overlapping with NUTS):

<https://borders.ukdataservice.ac.uk/easy_download.html>

* Stefano: su NomisWeb ci sono alcuni dati (vedi link sotto, scrolla fino a “local authorities: county / unitary (as of April 2015)”)
  + http://www.nomisweb.co.uk/query/select/getdatasetbygeog.asp?cat=8&geogtype=446
  + In alternativa dovresti andare tramite InFuse, selezionando a partire da by geography (link sotto), poi sotto England check su “Counties (35 areas)” (non espandere, semplicemente check sull’elemento). Poi proseguendo, dovresti poter scaricare dati tipo esempio in allegato.
  + http://infusecp.mimas.ac.uk/infusewizgeo.aspx

# UK spatial units

* 
* 
* <https://www.ons.gov.uk/methodology/geography/ukgeographies/administrativegeography/england>
* 
* AB data (2011). All units have a unique GSS code.
  + 1 UK
  + 4 countries
  + 9 English regions
  + 35 English counties (partial): Mix of:
    - 
  + 404 UK Local authority unit. Mix of:
    - 
    - 
  + 9481 Wards (small)
  + 42143 Lower Layer Super Output Areas (very small)
* Used by Nick
  + CA: Combined authorities

UK Census

* We hold aggregate data from censuses in England, Wales and Scotland held in 1971, 1981, 1991, 2001 and 2011, as well as data for Northern Ireland from censuses in 1991, 2001 and 2011. These data are aggregated to a range of geographical levels (e.g. counties and districts) from output areas which are the smallest spatial units. <https://census.ukdataservice.ac.uk/get-data/aggregate-data>
* Tables for Census: <https://www.nomisweb.co.uk/census/2011/quick_statistics>
* Other data store for Census: <http://infuse.ukdataservice.ac.uk>
* 1971
* 1981
* 1991 <https://www.ons.gov.uk/census/2001censusandearlier/aquickguideto1991andearliercensuses>
* 2001 <https://www.ons.gov.uk/census/2001censusandearlier/aboutcensus2001>
* 2011 <https://www.ons.gov.uk/census/2011census>

Labour market stats

* <https://www.nomisweb.co.uk>

Political behaviour

* Brexit vote

Deprivation indices

* <https://census.ukdataservice.ac.uk/get-data/related/deprivation>

Tourism data:

* Tourist board: sightseer for tourism data in the UK

# Publication plan [see gdoc]

\* A geography of museums in the UK

\* A geographical history of museums in the UK

\* 1 First paper (present, museums that are open now): descriptive stats

\* 2 Second paper (history, museums in different phases): descriptive stats

\* [3 Third paper: explanatory analysis of factors]

\* Ask Rosie about informal time off for next term

\* Oct 2017 - Sep 2018

\* Regular weekly meetings

\* Release: Sept 2019

\* Research questions

\* Select datasets

# Relevant publications

* Loaded in Mendeley
* Geography/spatial analysis: This is an example of a quantitative spatial analysis of some museum-related stuff (with awful maps):
  + Widdop, P., & Cutts, D. (2012). Impact of place on museum participation. Cultural Trends, 21(1), 47–66. http://doi.org/10.1080/09548963.2012.641775
* In terms of framing of this work in geography, we can’t use terms “cultural geography” (which refers to a branch of pretty obscure anthropology), and “cultural analytics” (the analysis of large collections of cultural objects such as photos). So, we can ground the approach generically in human geography. A clear example of how geographers think about “place effects” (ie how places influence people behaviours) can be found in health geography. Would it make sense to think of how places influence museums (openings, topics, visits)?
  + Cummins, S., Curtis, S., Diez-roux, A. V, & Macintyre, S. (2007). Understanding and representing “place” in health research: A relational approach. Social Science & Medicine, 65, 1825–1838. http://doi.org/10.1016/j.socscimed.2007.05.036
* This paper presents a proper, sophisticated geographical/temporal analysis of population change in Ireland caused by the Great Famine:
  + Fotheringham, S., Kelly, M. H., & Charlton, M. (2012). The demographic impacts of the Irish famine: towards a greater geographical understanding. Transactions of the Institute of British Geographers, (Nally 2011). http://doi.org/10.1111/j.1475-5661.2012.00517.x
* Conceptual modelling/ontologies/linked data: This is a vision paper in my field (GIScience) about the impact of linked data. It might be interesting to discuss the challenges about expressing spatio-temporal / place data in the project (particularly how you modelled data uncertainty):
  + Kuhn, W., Kauppinen, T., & Janowicz, K. (2014). Linked Data - A Paradigm Shift for Geographic Information Science. In M. Duckham, E. Pebesma, K. Stewart, & A. U. Frank (Eds.), Geographic Information Science: 8th International Conference, GIScience 2014, Vienna, Austria, September 24-26, 2014. (pp. 173–186). Berlin: Springer. <http://doi.org/10.1007/978-3-319-11593-1_12>
* This is about the “spatial turn” in the humanities, which might be important to ground theoretically the spatial (and temporal) dimension of the data (see just the introduction):
  + Warf, B., & Arias, S. (Eds.). (2009). The Spatial Turn: Interdisciplinary perspectives. New York: Routledge.
* This is about historical GIS, some arguments about the impact of digital tech on history practice might be relevant to the arts too:
  + Gregory, I. N., & Healey, R. G. (2007). Historical GIS: Structuring, mapping and analysing geographies of the past. Progress in Human Geography, 31(5), 638–653. http://doi.org/10.1177/0309132507081495