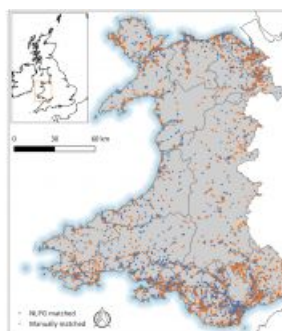
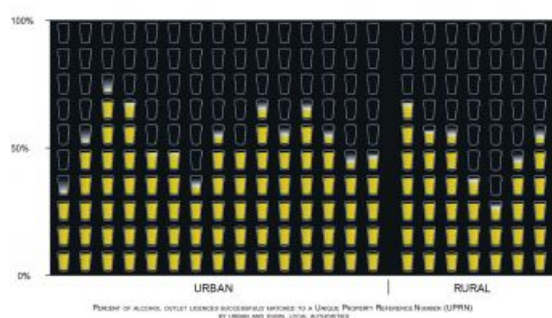


JOINING UP THE DOTS

Realising the Potential of Routinely Collected Administrative Data



The use of routinely collected administrative data for purposes other than those it was collected for has become an essential data resource for approved health and social science research.

Initiatives such as the Data.gov.uk and the Open Government Licence underpinned by the Freedom

of Information Act, has meant that publicly funded bodies have made their data available to researchers and the public. This shift has led to a number of data focussed research centres across the UK including the Farr Institute (www.farrinstitute.org), Administrative Data Research Network (www.adrn.ac.uk) and the Consumer Data Research Centre (www.cdrc.ac.uk) focusing research attention on facilitating and conducting approved research using health informatics, administrative data and consumer generated data respectively.

CHALICE, a recently published piece of research [1], examined the impacts of spatio-temporal alcohol availability on crime and health, utilised routinely collected alcohol outlet licence data collated from the 22 Unitary Authorities in Wales for a five-year period (2006-2011). Our research team assumed that this would be a relatively straightforward process as there is a legal obligation to keep a register of licensed premises within each authority as detailed by the Licensing Act 2003 (Part 2-8).

We found that, on average, only 51% of licence addresses could be precisely located using the address supplied. Although most Unitary Authorities could provide a list of outlets over the study period, despite a standard licence application form, the quality and details contained within each record varied greatly. None of the premise address details supplied had been verified against the Local Land and Property Gazetteer, thus licence data needed to be geocoded using the unverified address details supplied in the licence register. This resulted in a wide range of match rates (28-72%) of alcohol outlets successfully geocoded against AddressBase Premium.

Through innovative data linkage techniques developed at these centres, it is now common practice to link disparate data together, down to individual levels, in secure research platforms such as the Secure Anonymised Information Linkage (SAIL) platform housed at the Swansea University Medical School. Data linkage requires a common reference framework to be successful. The Local and National Land and Property Gazetteer (LLPG and NPLG) and related Ordnance Survey AddressBase products provide a common reference framework for address level data, in the form of the Unique Property Reference Number (UPRN) but in our experience, it is being underutilised in all levels of administrative data.

Currently, government and academic researchers alike are missing out on valuable insights into society and health through linked data research because of data collection issues. The use of a standardised method for collecting and verifying address data against LLPG and NPLG registers has the potential to facilitate world leading longitudinal research in local and national government departments and UK research institutes. As GIS professionals, we must encourage best practice and use of geographic information across all levels of government, business, education and research so that the wider societal benefits can be realised.

A more in depth analysis of this work has recently been published in an open access journal Applied Spatial Analysis and Policy (<http://link.springer.com/article/10.1007/s12061-016-9184-4>).

References

[1] www.journalslibrary.nihr.ac.uk/phr/phr04030#/abstract

This article was published in GIS Professional June 2017

<https://www.gim-international.com/content/article/realising-the-potential-of-routinely-collected-administrative-data>
