Use of geospatial data, analysis and tools in regional and local government: a scoping project by Northumbria University

Project Summary Report

19 August 2022

Authors:

Josephine Ellis

Dr Paul Greenhalgh

Dr Helen King

Contents

List of Figures

List of Tables

Glossary of acronyms

AONB	Area of Outstanding Natural Beauty
ASNW	Ancient Semi-Natural Woodland
EPC	Energy Performance Certificate
GIS	Geographic Information Systems
HELAA	Housing and Economic Land Availability Assessment
IMD	Index of Multiple Deprivation
LEP	Local Enterprise Partnership
LNR	Local Nature Reserve
LSOA	Lower Super Output Area
MSOA	Middle-layer Super Output Area
NNDR List	National Non-Domestic Ratings List
NNR	National Nature Reserve
NPPF	National Planning Policy Framework
NTCA	North of Tyne Combined Authority
PPG	Planning Practice Guidance
SATs	Standard Assessment Tests
SCat codes	Special Category codes
SIC codes	Standard Industrial Classification codes
SSSI	Site of Special Scientific Interest
TVCA	Tees Valley Combined Authority
VOA	Valuation Office Agency

1   Introduction

Funding was granted in 2022 for a six-month project, seeking to explore the potential for greater use of geospatial data in local and regional government across the North East.

The project bid identifies different practices in the use and implementation of geospatial data in local and regional government, particularly with regard to the built and natural environment, to fulfil statutory duties of local planning and economic development, to make best use of built assets, deliver future sustainable development and protect economic wellbeing of citizens. It also identified recognised deficits in terms of skills and access to data, leading to mixed practices in decision-making and to the excessive use of private consultants to produce single-use analyses which may be out of date before they are even completed.

The outcome of this scoping project will feed into a broader project, the objectives of which are to:

1) Work with Local Government, Local Enterprise Partnerships and other public sector stakeholders to identify spatial data availability and needs at local, city and sub-regional scales
2) Identify accuracy, interoperability and coverage issues to inform formatting, filtering and validation of datasets to be represented spatially
3) Develop process/protocol for creating a common GIS platform on which to share identified spatial datasets
4) Create Geographic Information Systems platform using ESRI ArcGIS to portray spatial data and test efficacy with stakeholders to refine and improve as fit purpose as a data sharing and querying platform

This report discusses the extent to which these four broad objectives have been explored and makes recommendations for further activities.

2   Objective 1: Work with Local Government, LEPs and other public-sector stakeholders

Over the project period a series of meetings and other interactions took place, based around meetings with the two combined authorities in the North East: the North of Tyne Combined Authority (TVCA), and the Tees Valley Combined Authority (TVCA). Meetings also took place with the North East Local Economic Partnership and with representatives of some local authorities. These are listed at Appendix 5.

It would be outside the scope of this stage of the project to list all the uses to which geospatial data is being put in local government currently. The increased use of open-source GIS and the increased availability and awareness of geospatial data means that officers are increasingly able to use geospatial data and tools in the course of their work without this being recorded by any single officer or department within an organisation. At the same time, an increasing governmental requirement to make use of geospatial data and tools (for example, in land-use planning) means that it has become a matter of course in some contexts. For example, as we discuss below, local authorities are required to publish data on brownfield sites in the form of a map. The requirement for local authorities to make use of geospatial data and tools is likely to increase in coming years as a consequence of initiatives such as the UK's Geospatial Strategy (2020) and the recommendations of the Geospatial Commission's Planning and Housing Landscape Review (2020).

As we discuss later, however, there is evidence of a great deal of variability in terms of data use and data sharing.

Combined Authority representatives were asked to identify their priorities for development of GIS and use of geospatial data, giving their reasons for choosing these priorities, what data they already have, and what additional data is needed, using the matrix at Appendix 1, Table 1.

2.1 Priorities for further use of geospatial data and tools (TVCA; NTCA)

2.1.1 Brownfield Land Sites (TVCA)

An analysis of existing brownfield land - not just the land suitable for residential development, but all previously-developed land - is needed to inform strategy and pipeline of future brownfield development. The project could entail the collation of brownfield land data from each local authority, in a geospatial format, including the location of sites, details of the age and construction of properties, and land value data. TVCA also seek to develop a Tees Valley-wide approach that overlays mapped data sets, such as: environmental data / natural assets, local economic data on deprivation, income, employment, affordability and data to assess impact on infrastructure and local services. This approach could also be used to assess all potential development sites, particularly those considered within a local authority's Housing and Employment Land Availability Assessment (HELAA).

2.1.2 Natural Assets (TVCA)

Collation of data on natural assets – including all protected wildlife sites such as nature reserves and Sites of Special Scientific Interest (SSSIs), together with specific habitats such as ancient semi-natural woodland, and valued landscapes such as Areas of Outstanding Natural Beauty (AONBs) and those identified in local Landscape Character Assessments - is needed to inform Net Zero and Local Nature Recovery Strategies. It could link to the Brownfield Land analysis. The same is true for heritage assets such as listed buildings and protected parks and gardens; and flood risk zones, which, though they are not a natural asset as such, entail similar restrictions on development. Much of this data, as we discuss below, is already available in geospatial format. The added value of this project is to explore how that data could be more productively used.

2.1.3 Broadband services (NTCA)

Collation and display of data on gigabit capable broadband, and 5G connectivity, is required to support an appropriate investment strategy to enable economic growth and equitable access to opportunity, and to measure progress against the Government's target for 100% ultra-fast broadband coverage by December 2026. NTCA request that the data on broadband connectivity be mapped alongside demographic and economic data such as: population density, economic activity and employment rates, and deprivation. It needs also be looked at with reference to commercial delivery plans. Ideally, the data would be mapped at property level, with the ability to analyse and filter data by the connection type available – e.g., super-fast broadband/ ultra-fast broadband, etc. NTCA also request a version showing the proportion of property connected to ultra-fast broadband within a given LSOA boundary.

2.1.4 Schools (NTCA)

Develop a map of schools which provides insight into pupil performance and the social/ economic characteristics of the region. This could be used to support applications for funding via new funding streams in relation to education improvement and Child Poverty Prevention. The school location data itself should be mapped and analysed with reference to the school population characteristics (Free School Meals, Attendance, Grades etc); indices of multiple deprivation; and the local claimant count, at LSOA or MSOA level.

2.1.5 Economic clusters and low-carbon businesses (NTCA)

Analysis of economic clusters in relation to specific sectors and employment density. The purpose is to understand the implications of sector investment strategies on places, and the potential displacement caused by significant changes to economic conditions; to identify growth opportunities for targeted economic areas as defined by NTCA; and, in particular, to map high and low carbon sectors and businesses to understand Just Transition implications for an area. There is the potential to add existing investment activities and data to this analysis.

2.1.6 Business start-ups and deaths (NTCA)

Understand where businesses are being created, how they are distributed, and what SIC category they fall under. This could be considered with reference to the CA's economic growth objectives. Ideally, the analysis would establish, for each LSOA, the number of new businesses (raw), number of new businesses as proportion of existing, and the number of new businesses as proportion of local SICs. It would also give users the option to filter by time.

3 Objective 2: Identify accuracy, interoperability and coverage issues

3.1 Data types

The first stage in this process is to consider *which* datasets ought to be represented spatially, and at what level. It would be unrealistic to try to compile a list of all the geospatial datasets used in all of the local and sub-regional bodies across the region, as we say above. Many of these datasets are, in any case, used only to make decisions at a local authority level, and have little relevance outside the boundaries of that area – for example, data on the condition and age of trees and shrubs in local authority-owned open spaces.

It falls within the scope of this study to consider when data might be relevant at a to regional authorities. These data, perhaps, fall into two broad categories:

3.1.1 Data that can be used in regional-level decision-making.

This type of data is necessary to inform regional decision-making on matters which operate at a greater than local scale. Examples of these matters are:

- *Transport:* this operates at a variety of scales from the very local, to the travel-to-work area, to the regional and national. Datasets might include traffic capacity and traffic flows, the location of public transport infrastructure and the highway network, and the location of residential areas and important destinations.
- *Flood risk and flood management:* this operates at a variety of levels but is highly dependent upon topography – particularly the geography of river basins and their component parts.
- *Green infrastructure networks*: green infrastructure planning, by definition, is multi-scalar. It might cover matters such as access to green space within walking distance of homes (neighbourhood level); urban heat islands (settlement level); or equitable distribution of green space across an area (local authority level). At a regional level, it entails, inter alia, consideration of habitat networks and flood resilience.
- *Location of development:* although spatial planning is the responsibility of local authorities and there is no supra-local level, regional government must have regard to existing centres of population, large employers and industrial areas, and potential growth areas, in order to direct funding or support towards the places that most need it.

3.1.2 Data that should be used in a similar way across the region.

In order for regional government to operate properly, local government needs to operate consistently. In order for this to happen, there needs to be similar decision-making mechanisms across a region, and this requires a similar attitude to the use and application of data.

A very significant aspect of this is land-use planning. We have conducted an exercise in which we list datasets which are used or referenced, directly or otherwise, in planning practice guidance (PPG) or the National Planning Policy Framework. We then state whether these datasets exist at present, and, if so, where they can be found. The findings of this exercise are presented in Appendix 2.

Some of these are *inputs* to the planning process – such as the boundaries of nature reserves and heritage sites. Many, although by no means all, of these datasets are publicly available and are created by governmental bodies. Others - such as open spaces and brownfield sites - are created or collated by the local authorities themselves. The collation of these datasets may be a straightforward matter of observation – identifying points or areas relating to a single quality or feature ("original datasets", e.g., open spaces); or it may require some sort of analysis to create a dataset derived from another ("derived datasets", e.g., brownfield sites suitable for development).

Some of the most important of these (taken from Appendix 2. are:

> ***Original datasets***
> Groundwater or soil contamination
> Public open space
> Public rights of way
> Rights of access
> Utilities infrastructure
> Quality and quantity of healthcare, education, sports, recreation, places of worship
> Location of hazardous establishments
>
> ***Derived datasets***
> Brownfield sites (deemed suitable or unsuitable for residential? development)
> Buildings in need of repair
> Buildings in need of demolition
> Sites whose topography makes them undevelopable
> Land locked sites
> Sites whose development would have unacceptable traffic management impacts
> Viable and non-viable sites
> Sites which are well-related to existing settlements
> Sites which are or can be made sustainable [in transport terms]
> Town centres
> Nature Recovery Networks/ Ecological Networks
> Landscapes

A caveat applies here. Although PPG recommends the use of all of these datasets, it might not be the case that all of them are necessarily useful at a metropolitan or sub-regional level. The dataset of buildings in need of repair or demolition is, perhaps, unlikely to be useful outside the local authority – and in any case, its compilation entails a degree of subjective judgement which may be difficult to standardise. Meanwhile, viability assessments, depending as they do upon current values for land, commodities and housing, are liable to have a short shelf life in relation to currency of data.

Another set of datasets are those which are *outputs* of the planning system. The most significant of these are probably those which form part of planning policy – the boundaries of development sites, green belts, town centres and so on. Other datasets might include those which manifest the consequences of planning policy, such as the areas where development has taken place.

3.2   Interoperability and coverage issues

These issues primarily apply, in this context, when datasets which are relevant at a supra-local level are compiled – if they are compiled at all – at a local level. In those circumstances it *should* be possible for a region-wide dataset to be compiled, on the basis of data provided by local authorities.

Particularly with regard to derived datasets, it *should* be reasonable to assume that the same methodology is being followed across a region to determine, for example, which brownfield sites are suitable for development, or where the boundaries of town centres should lie.

We cannot claim with confidence that these aspirations are being met at the moment.

We requested the dataset of brownfield sites suitable for residential development from all of the local authorities in the region (see 2.1.1). This dataset is unusual among those listed above insofar as national standards exist for its publication and format. As mentioned above, local authorities must produce a CSV file, listing all of the brownfield sites suitable for development within its jurisdiction, and providing, in standardised fields, specific information about each site in a specific format. There must also be a link to an online location where a map of each site can be found.

Therefore, local authorities *must* provide the data in a standard format, and they *must* present it in geospatial format. Even so, we were unable to obtain it in a geospatial format (as opposed to being able to view it on an online map or PDF) for the entire region. Out of the 12 local authorities in the region, two published the data on a regional data-sharing website, "Data Mill North"; five provided us with the data on request; five, despite several attempts to make contact, did not provide it at all. The data could be merged to give partial coverage, but required some reformatting to do so (mostly, to ensure consistent field names).

Furthermore, we cannot be entirely sure what methodology has been applied to determine that the sites within the dataset are suitable for development (the only guidance is in the Town and Country Planning (Brownfield Land Register) Regulations 2017 and is not extensive).

Since this is the case with regard to a particularly tightly-defined dataset, we have little confidence that other spatial datasets – open spaces, for example, or town centres – would be consistent and interoperable, or would be made public.

4   Objective 3: Develop process/protocol for creating a common GIS platform

As we discuss above, it should be possible for local authorities and other governmental bodies to publish a very wide range of geospatial data. Some of the LAs in our region already publish a certain amount of data on the regional data-sharing website, "Data Mill North". This website, however, currently does not provide a means for viewing geospatial data; it can only be downloaded and then opened up in another application.

An example of how this might work is Defra's "MAGIC" website. This contains an interface by which users can select map layers to be displayed upon an interactive map. These map layers consist of datasets provided by various partner organisations responsible for aspects of the natural or built environment, such as Natural England, Historic England, the Forestry Commission, and so on. The data can also be downloaded - either directly from the website, or via a link to the provider's website.

Similarly, the Coal Authority produce an [interactive map](#) showing where there are constraints to development associated with mining activities, and DEFRA have an interactive [data services platform](#) enabling users to view and download spatial and other datasets.

The process or protocol for establishing this type of interface would entail:

- Constructing a website with an interactive map, into which registered users could upload data.
- Establishing which datasets should be published to this website, and how they should be compiled and formatted.
- Monitoring and managing the website to ensure that datasets are up to date and are in line with the standards.

5 Objective 4: Create a GIS platform to portray, share and query spatial data

The project began by investigating the use of ArcGIS Online products, in particular, the "Dashboard" function. This interface allows users to upload datasets and present them within an interactive map which carries out selected analyses upon the data; the user can direct the application to analyse the entire dataset, or the part of it visible within the part of the interactive map being displayed, or the part which falls within an area defined by a separate polygon dataset. (so, for example, the application could be directed to analyse only data falling within a given local authority area).

The following sections discuss, in turn, the areas of work suggested by our regional partners in section 1 of this report. They consider:

- Whether ArcGIS Online Web Mapping Applications or Dashboards might be a useful tool in these areas of work, or in displaying our findings
- What sort of geospatial analysis might be carried out in desktop GIS (ArcGIS or QGIS)
- What features an "ideal" interactive mapping portal would have, to make it usable by local authority and combined authority partners, and to ensure that the analysis provides good answers to the right questions, including ability to carry out spatial queries with multiple layers.

5.1 Brownfield Land

Data on brownfield land suitable for development, as discussed above, is published by local authorities in a format specified by national government.

As we discuss below with regard to natural assets, it is currently quite simple to produce an online map showing where brownfield or other sites are situated with regard to any mappable constraints or material considerations. This allows us, not only to assess a site's suitability for development in relation to historic or natural environment constraints, but also to establish characteristics of the local environment in the vicinity of a given site – such as levels of deprivation, incomes and employment, and affordability of accommodation. An interactive online mapping interface map depicting indices of multiple deprivation (by decile, by Lower Super Output Area (LSOA)) and brownfield sites can be seen [here](#) a screenshot of example content is provided in Figure 1.
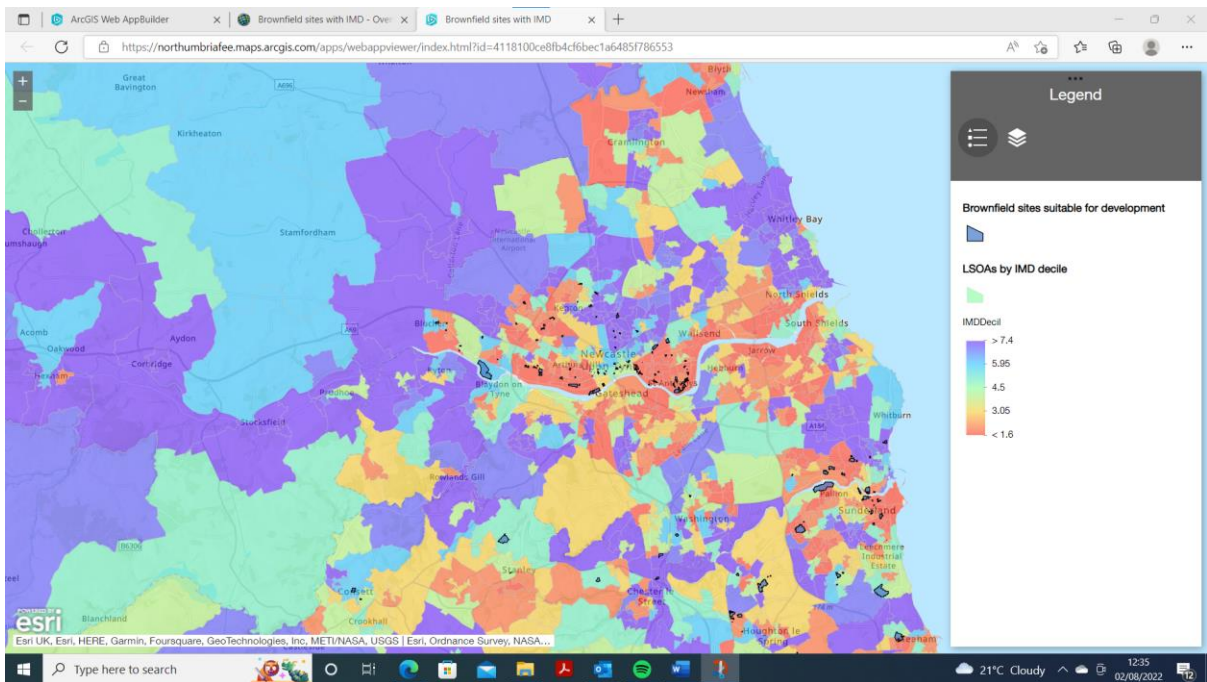
Figure 1: Online interactive map of brownfield sites, superimposed upon a map of IMD deciles. Brownfield land data courtesy of the following local authorities: Darlington Borough Council, Durham County Council, Hartlepool Borough Council, Middlesbrough Council, Newcastle City Council, Redcar and Cleveland Borough Council, and Sunderland City Council.   IMD data courtesy of the Ministry of Housing, Communities and Local Government.

ArcGIS Online allows us to carry out analysis relating to the distribution and type of brownfield sites across the area, and then display these findings on an interactive "dashboard". This is useful in terms of quantifying the amount of brownfield land across the region, displaying statistics relating to the information required by government, such as planning status and ownership; and comparing the situation in different local authority areas (
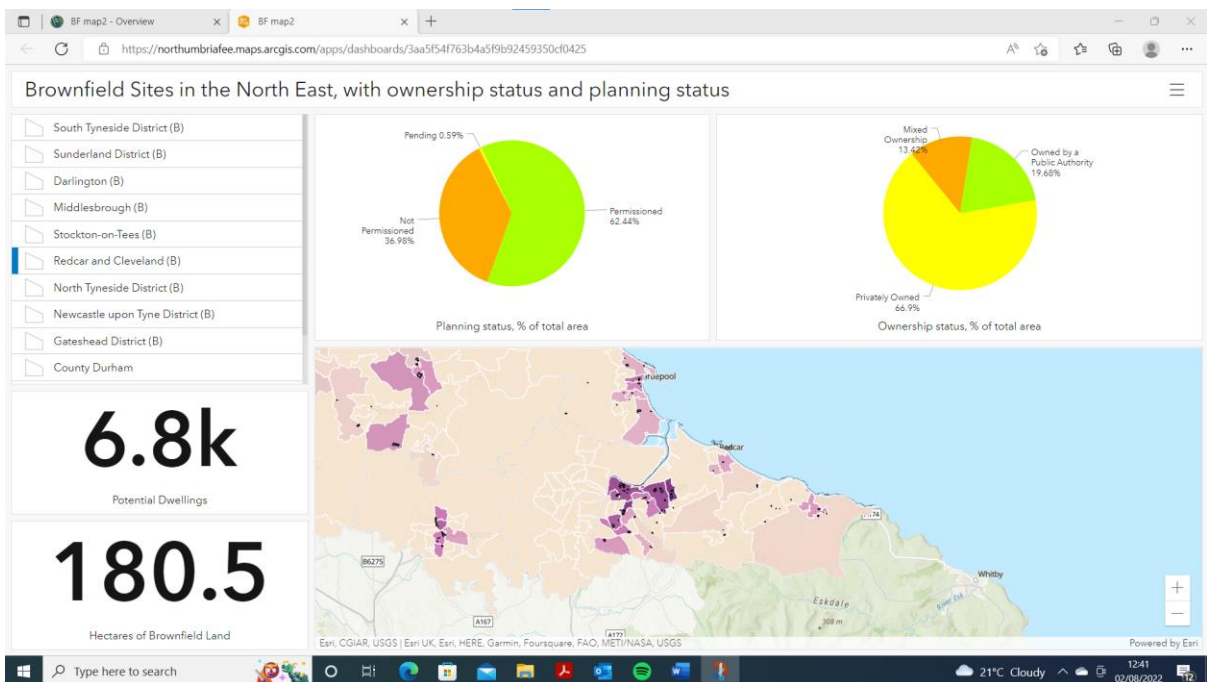

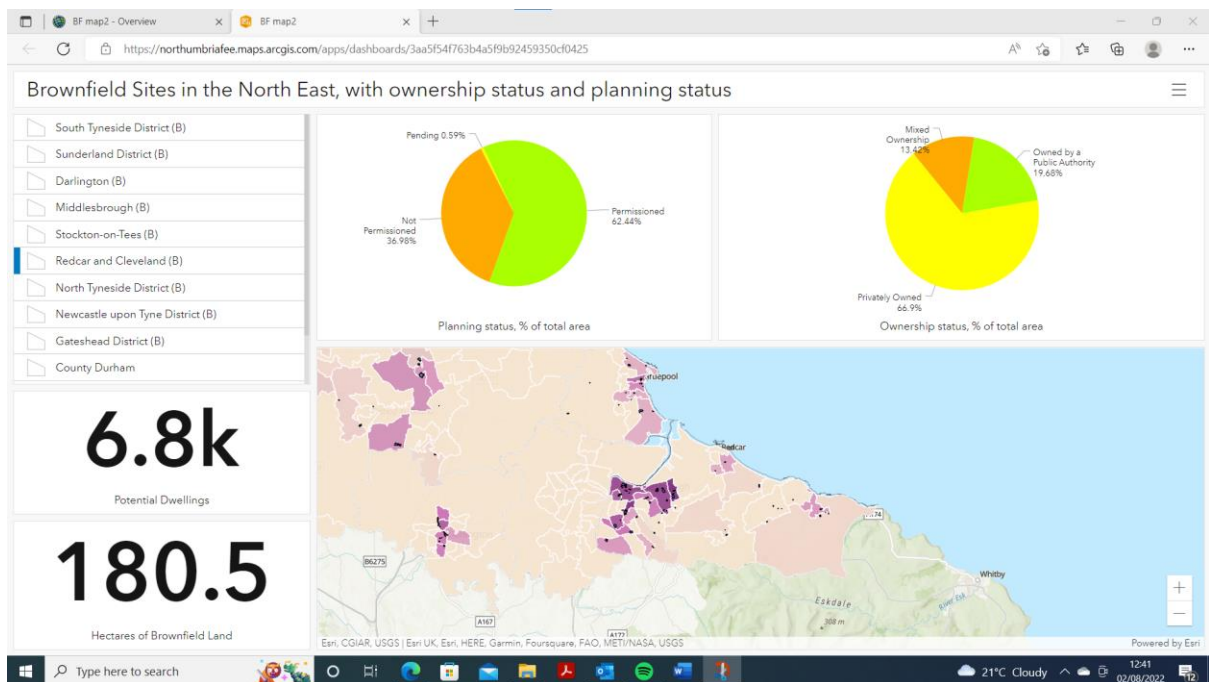Figure 2). An exemplar dashboard can be viewed here.

Figure 2: Dashboard of brownfield sites with planning status and ownership. Brownfield land data courtesy of local authorities, as in Figure 1.

A very significant caveat about this work is that the information required is not freely available. As discussed above, although local authorities must publish a CSV with brownfield land data in it, they have no obligation to publish polygon data relating to it and, in fact, out of twelve authorities in the region, only seven did so routinely or on request. Further, there were some difficulties in combining the datasets to make even a partial region-wide dataset. In some cases, the sites listed within the CSV were not the same as those in the polygon dataset. Elsewhere, the metadata relating to each site was formatted or titled differently, so that it was difficult to merge different datasets into one. The way the data is shared requires an individual to compile the data each time it is to be used which also links into data currency issues, potential to introduce error and upkeep and maintenance of datasets. Good analysis requires good data.

5.2   Natural Assets

Much of the data on natural assets - such as national and local nature reserves (NNRs and LNRs), sites of special scientific interest (SSSIs) and ancient semi-natural woodlands (ASNWs) - is already made available via ArcGIS Online, which contains a repository of datasets provided by contributor organisations, including UK government quangos and departments. This means that it is very easy to establish a web map displaying the natural assets which might be considered with reference to the suitability of potential development sites. Such a map can be viewed here.
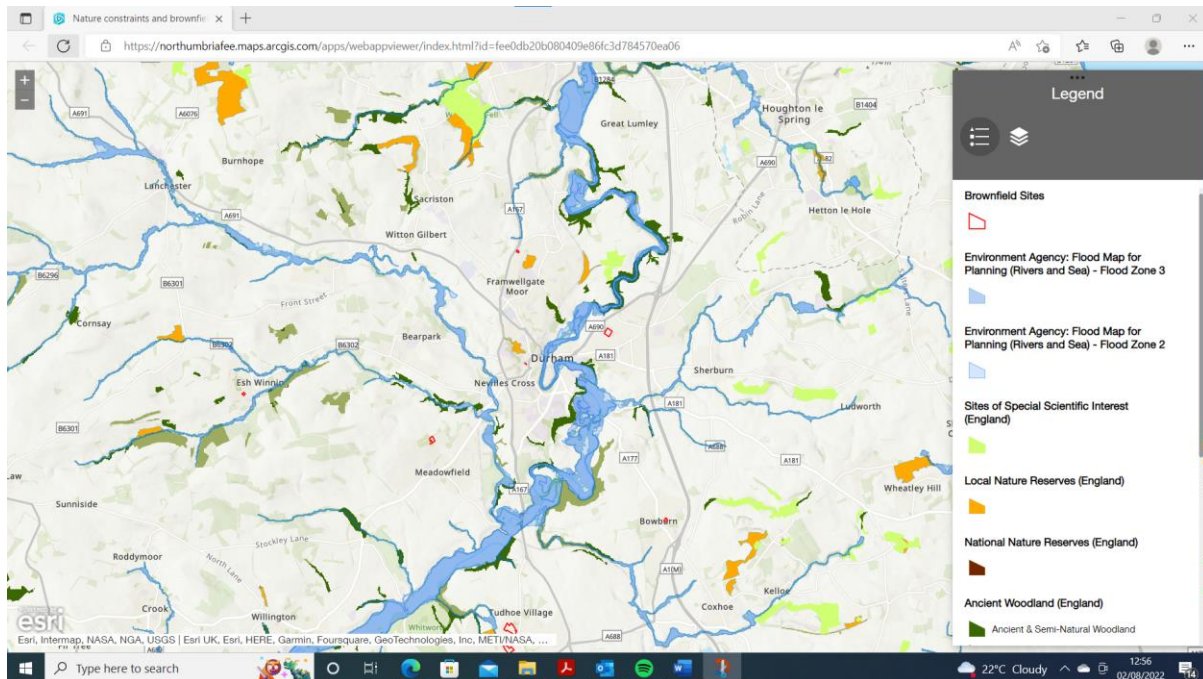
Figure 3: Online interactive map of potential development sites and natural assets. Data on natural assets provided by HM Government and made available via ESRI.

However, it could be argued that this doesn't provide a very easy way of comparing potential development sites (brownfield or otherwise) against natural assets. Local authorities are required to carry out such an assessment as part of their Housing Land Availability Assessment (HELAA). In order to do so, they need to be able to state whether each site in their portfolio intersects with a very wide range of constraints, including, but not confined to, natural assets; it should also involve consideration of heritage assets, flood zones, contaminated land and so on. Displaying these constraints upon a map allows the user to *see* whether potential development sites intersect with them, but transferring this knowledge into a written HELAA site assessment is still a tedious and time-consuming process. What is more, it is not always enough to know that a potential development site doesn't *intersect* with a mapped constraint; it is sometimes also necessary to know whether these constraints can be found nearby.

An alternative (but not web-based) approach can be found by using ArcGIS Desktop's "Model Builder" function (Figure 4). This allows the user to prescribe a number of operations, one after the other, upon the same input dataset. In this case, the input dataset would be the set of potential development sites, and the operations to be carried out would be a series of queries comparing the site boundaries with the constraint boundaries. The output would be a new dataset, depicting the site polygons as in the input dataset, but with the addition of new attributes stating whether these sites intersected with any of the relevant constraints.

We produced an indicative version, using only natural environment assets - to be of optimum use within the HELAA process, it might be better to include all mappable constraints.

In the schematic diagram of the process as shown in Figure 4, the input datasets (local nature reserves, national nature reserves, etc. – and the original set of brownfield sites) are shown as blue circles. The yellow oblongs represent a series of operations carried out upon the input datasets, and the green circles are the outputs of each operation. In this case, the output of the first operation (a spatial join) is a dataset including information on whether each brownfield site intersects with a LNR, which then

is one of the inputs for the next spatial join, this time involving NNRs – and so on. (See glossary for acronyms.)
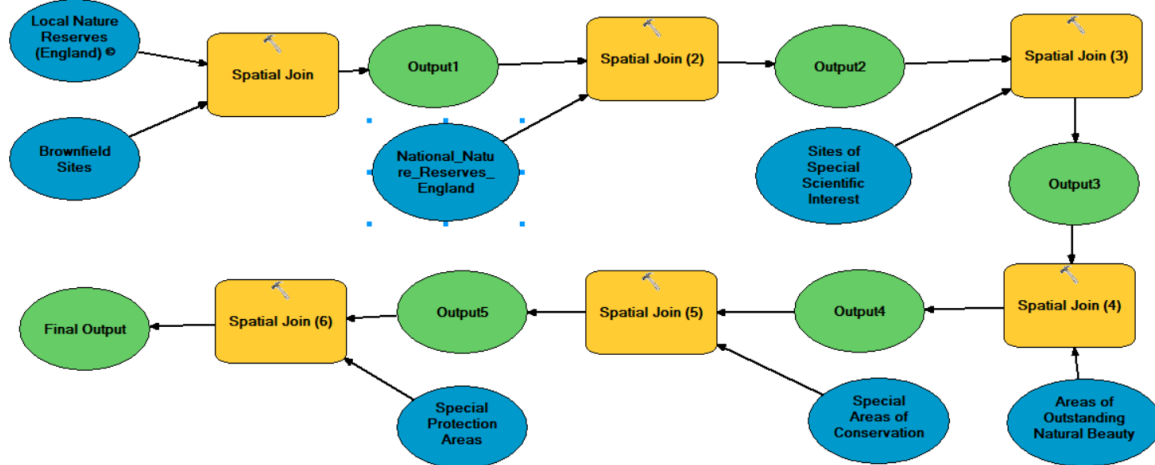

Figure 4: Diagram of the process determining whether polygons in an input dataset (Brownfield Sites) intersect with various natural environment assets.

In fact, when this analysis was carried out, we found that, by their nature, brownfield sites rarely intersect with protected sites and habitats (though sometimes brownfield sites can support unusual assemblages of species, so it is not impossible). The analysis, however, failed to capture proximity to protected sites. Therefore, the exercise was repeated. This time, instead of only comparing sites to the natural environment assets themselves, in some cases we compared them to buffers of the sites. The buffers were generally 800m around each natural environment asset, but in the case of Special Areas of Conservation and Special Protection Areas, which are more stringently protected, the buffers were 3000m.

This is a perfectly good way of standardising part of the HELAA process, which could easily be extended to cover other types of constraint or material consideration, such as heritage sites and properties, or flood zones. It does not, of course, include site-specific ecological assessments or any other analyses that could not be carried out on the basis of geospatial datasets.

Currently, though, this model is not publicly accessible or usable. This would make it harder for the analysis to be presented as a standard to be used across the region. The ideal would be for this analysis to be embedded within an online interface into which datasets of brownfield or HELAA sites could be uploaded, and which would return the analysis.

5.3   Broadband services

Data on broadband services is available online via Ofcom's Connected Nations project. It can be found here. It is available at various spatial levels, from postcode area (the smallest) to local authority area (the largest).

The data is provided in the form of CSV files which can be linked to polygon datasets relating to the relevant spatial division.

In this case, broadband data was mapped at an Output Area level. It was then possible to depict this in an interactive web map which colour-codes areas depending upon their level of access to ultra-fast

broadband. Other details relating to each Output Area can be accessed by clicking on it. This map can be viewed [here](#).
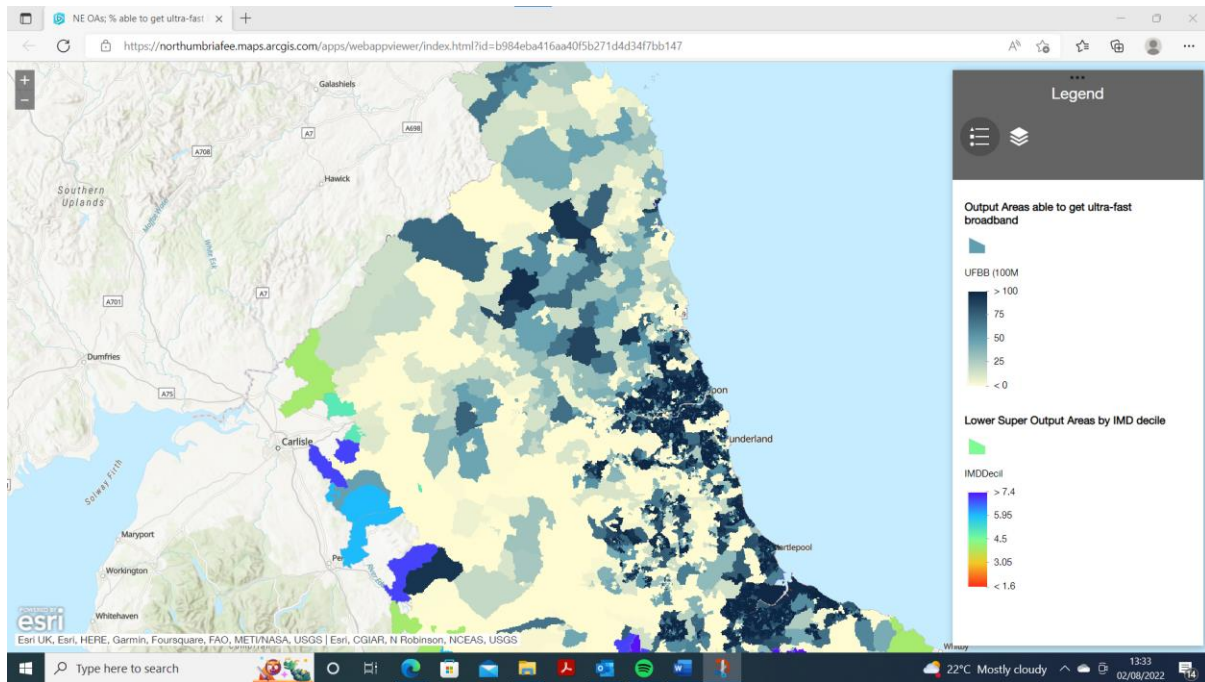

Figure 5: Online interactive map of broadband provision, with IMD decile. Broadband accessibility data courtesy of Ofcom. IMD data courtesy of the Ministry of Housing, Communities and Local Government.
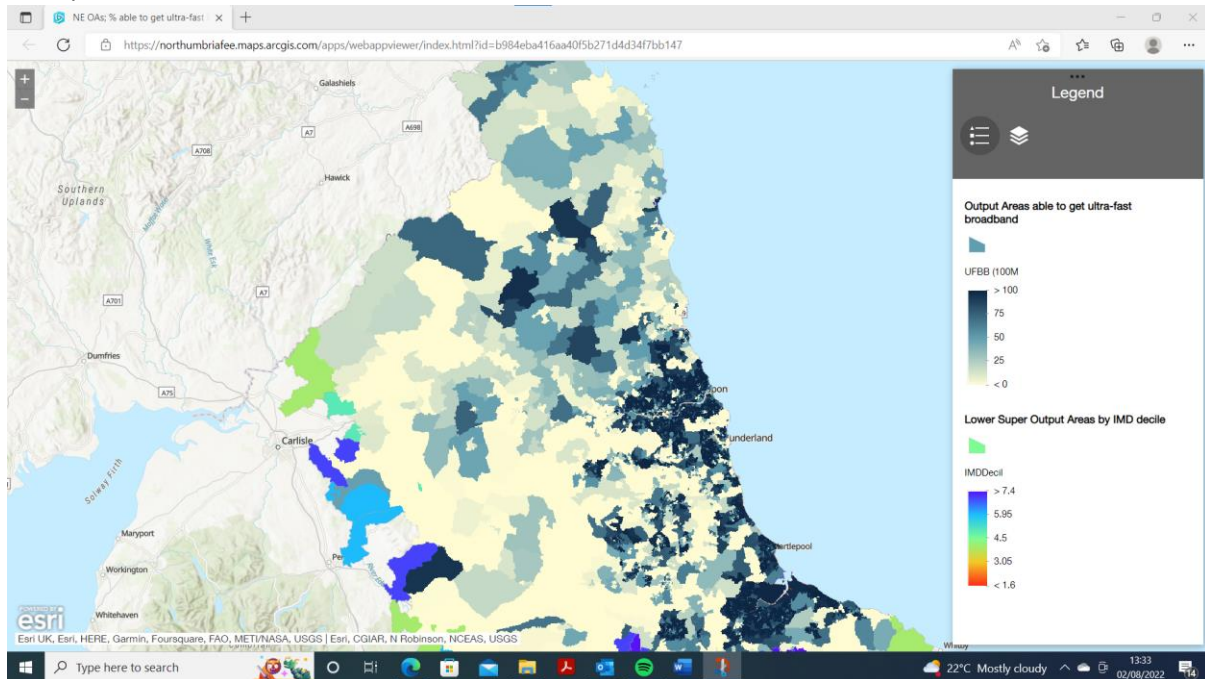
Data presented in

Figure 5 provides a good way of checking an individual area's broadband access, and also clearly depicts the "gaps" in broadband provision.

The map also depicts the index of multiple deprivation (IMD) decile for each LSOA. The problem with this, however, is that only one characteristic (broadband access, or deprivation) can be shown at one time. Another option might be to append IMD ratings (and other geospatial statistics relating to social and economic characteristics) to the broadband dataset, so that they appear within the same pop-up in the web map.

Some more detailed research could, however, be done, to establish what relationships exist between broadband accessibility and economic indicators, including deprivation, economic activity and employment rates; or whether discrepancies in broadband accessibility can simply be explained with reference to population density.

5.4 Schools data

The Department for Education publish details of all of the schools within the UK, as a CSV file containing eastings and northings, so that they can be mapped as points. The file contains numbers and percentages of children taking free school meals.

A polygon dataset of LSOAs with Index of Multiple Deprivation ratings can be downloaded from https://data-communities.opendata.arcgis.com/. Claimant data relating to Lower Super Output Areas can be downloaded from and downloaded as a query from NOMIS, https://www.nomisweb.co.uk/.

The schools dataset was filtered to select only open, state-funded, primary schools. Voronoi polygons were then created using QGIS to create an indicative catchment area for these schools. Schools in England and Wales do not have a specified catchment area; parents can apply for their children to join any eligible school in their local authority area, but proximity to the school may be used as a selection criterion.)
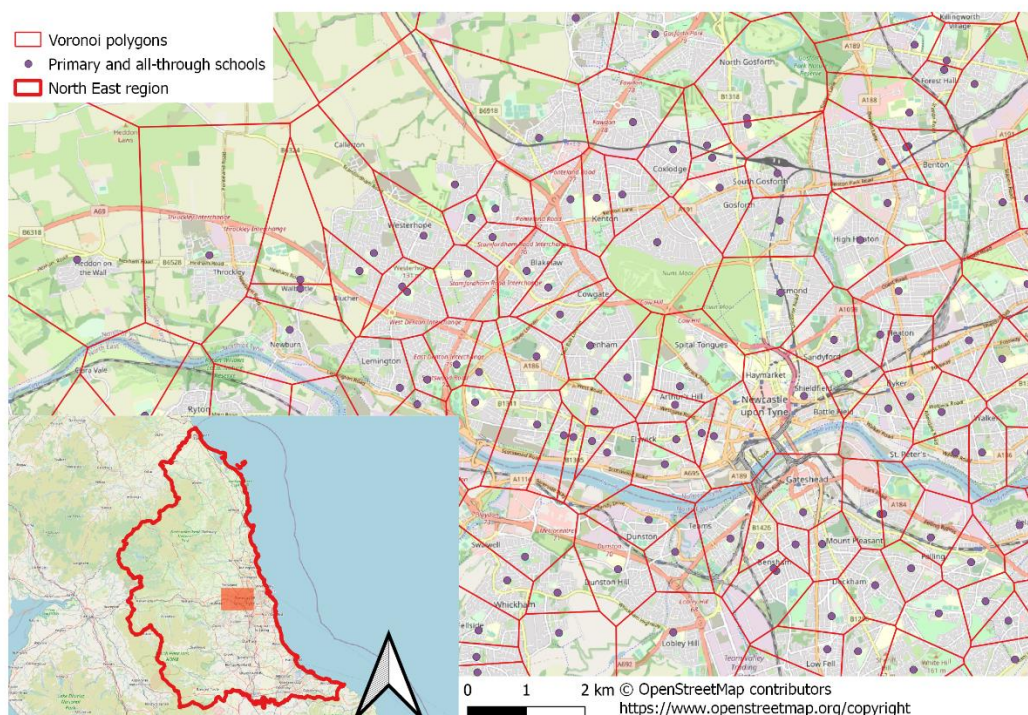

Figure 6: Voronoi polygons around primary schools on a map showing part of Newcastle upon Tyne. Schools data courtesy of the Department for Education.

Average figures for IMD and for the proportion of the population claiming benefits were calculated for each Voronoi polygon. In the case of IMD, the average IMD decile was calculated on the basis of LSOAs intersecting with each Voronoi polygons. In the case of benefits, the total number of benefit claimants, and the total number of residents, in LSOAs intersecting with each Voronoi polygon, was calculated, and then a percentage figure was calculated on the basis of these totals.

These data were appended to the schools records, so that they feature as attribute data. The schools could then be mapped on a publicly-accessible web map, in which schools' size and the proportion of children taking free school meals are depicted symbolically, and these and other metadata can be obtained as a pop-up by clicking on each point (
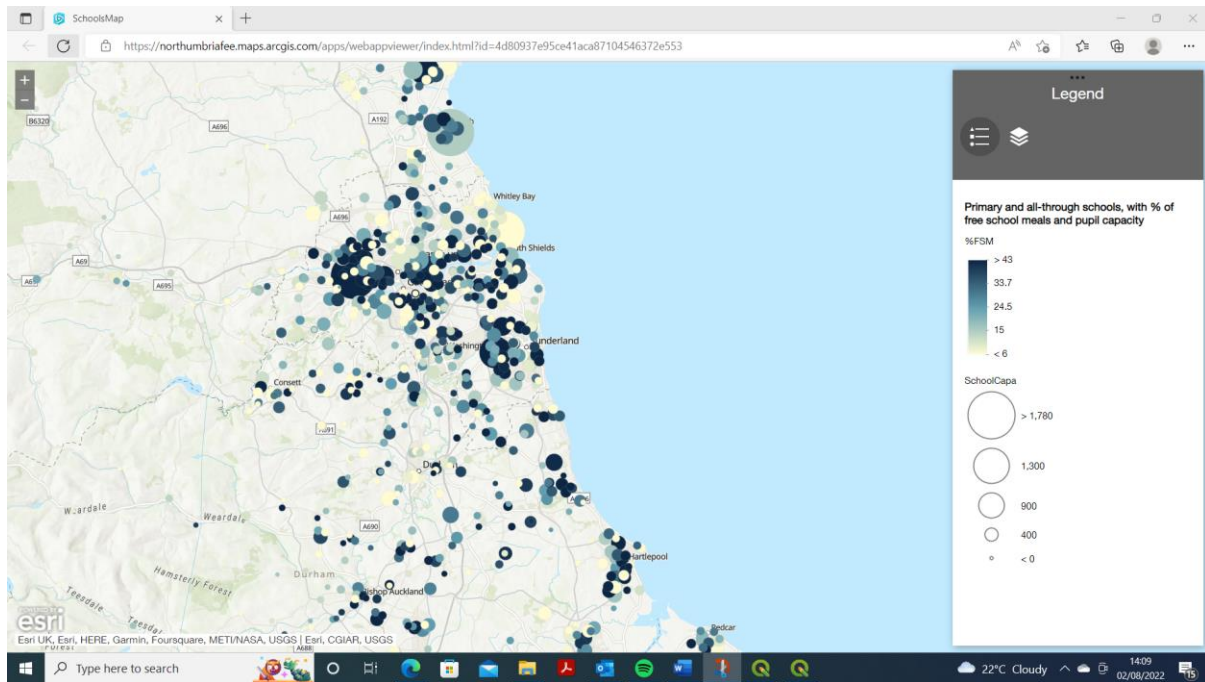

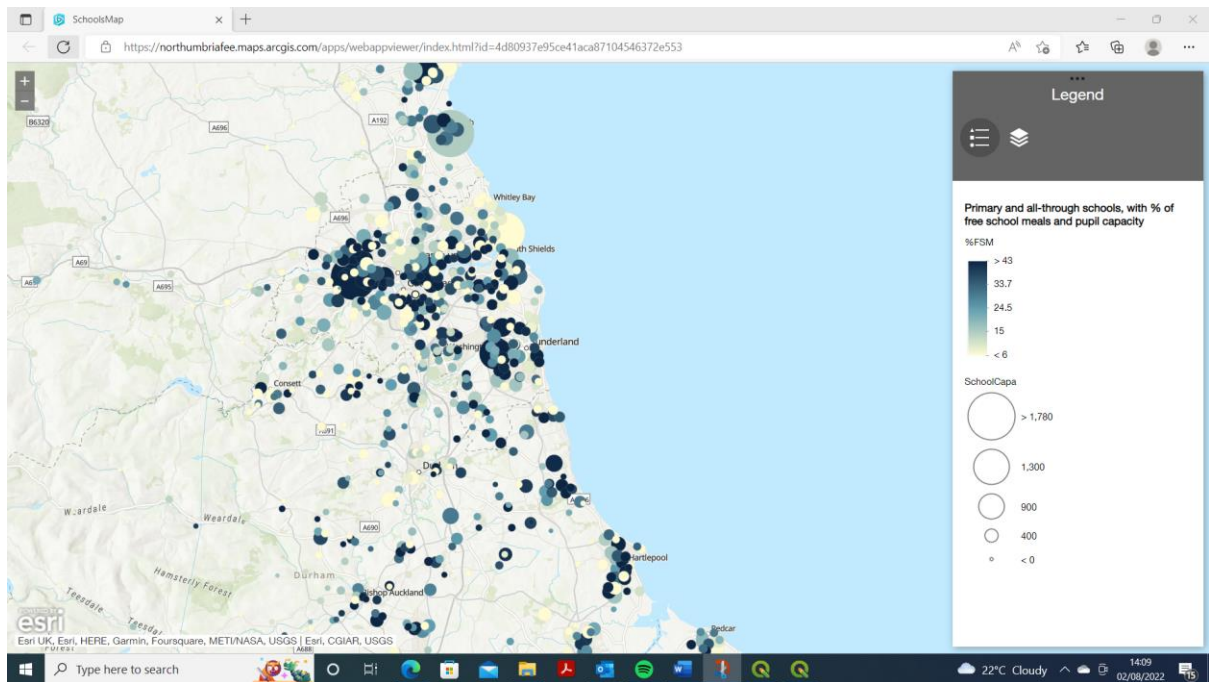Figure 7). This can be viewed [here](#).

Figure 7: Online interactive map of primary schools, indicating size of school and proportion of free school meals.

We could not find an up-to-date dataset of SATs results or other indicators of pupil attainment at primary level for each school. However, the dataset does indicate the results of each school's latest OFSTED inspection. This map is a reasonably good way of providing information on the socio-demographic situation relating to each school across the region. However, again, it doesn't provide an analysis of the relationship in between sociodemographic characteristics, or between these characteristics and pupil attainment. Another caveat is that the "catchment areas" represented by Voronoi polygons are very approximate, since, in practice, pupils do not always attend their nearest school.

A further piece of work could be carried out, using a multivariate analysis or similar process to investigate the relationship between sociodemographic characteristics, and pupil attainment, and to identify the establishments which are doing much better, or much worse, than expected.

5.5 Low-carbon (and other) enterprises

Data on all businesses across the country, the type of activities they pursue, and their rateable value can be obtained from the VOA in the form of the National Non-Domestic Ratings List (NNDR) and mapped (by postcode).

This data was used to create an ArcGIS dashboard (
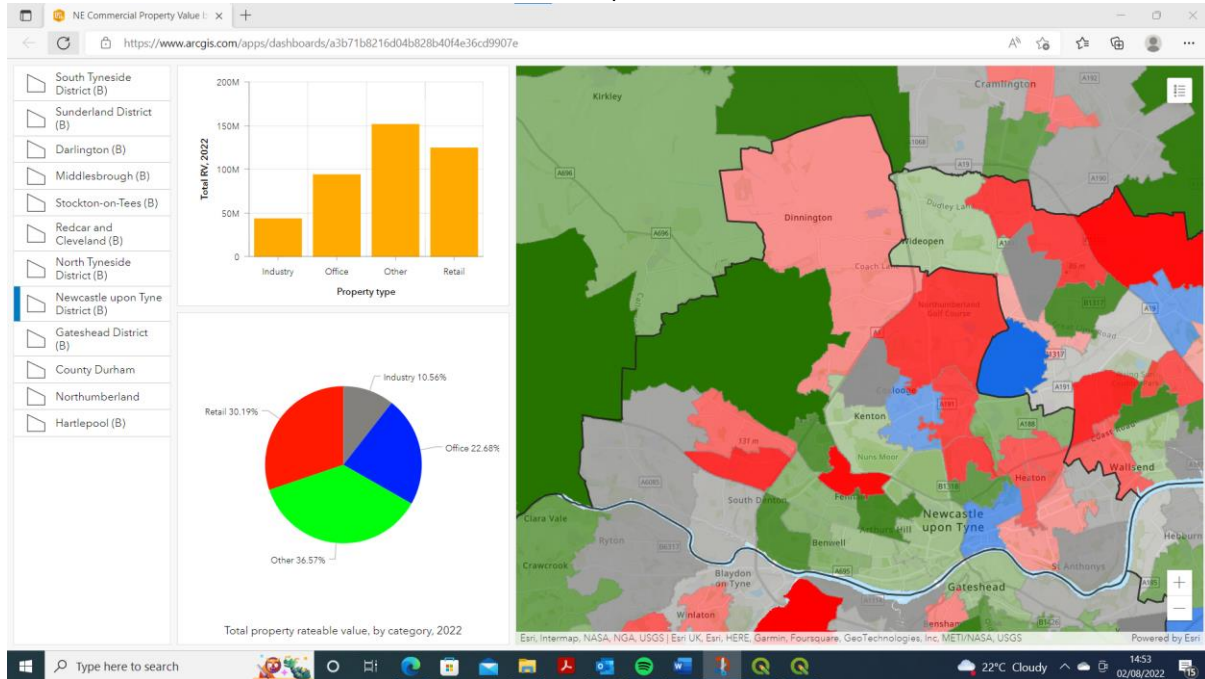
Figure 8) which provides information about the amount of rateable value within each local authority area and which sector is predominant. This can be viewed here.
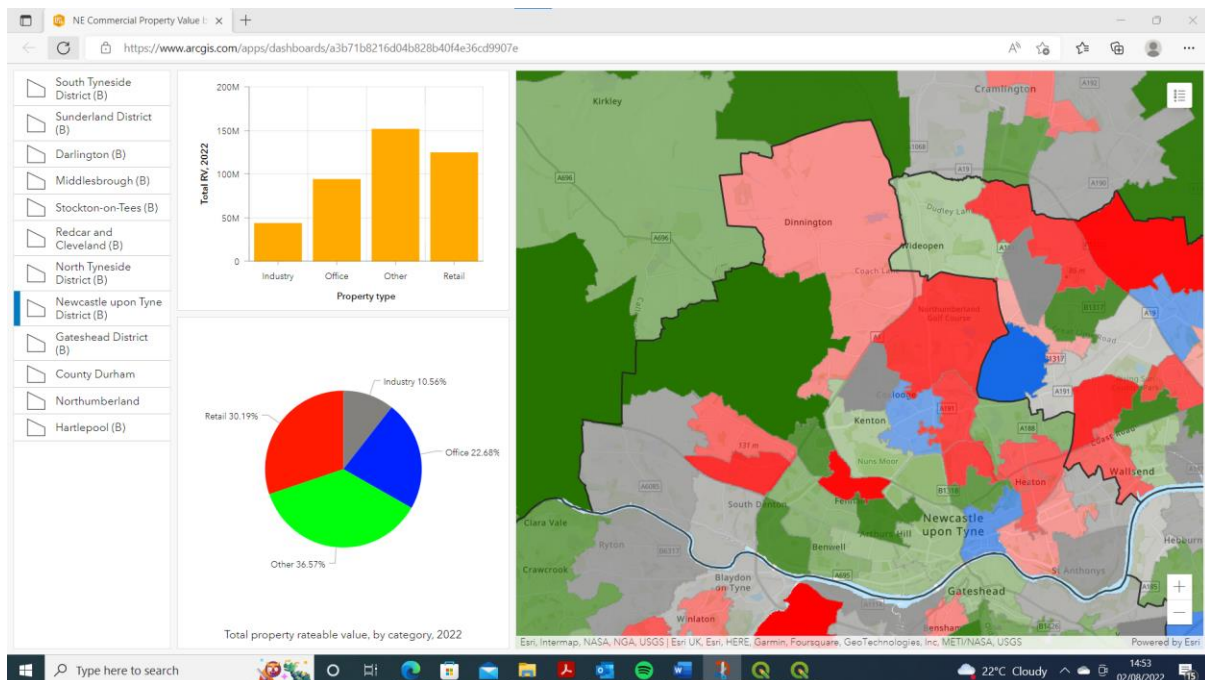
Figure 8: Dashboard showing rateable value of economic sectors within each local authority area. Commercial property data courtesy of the Valuation Office Agency.

Another map was created specifically with regard to low- or high-carbon industries. One constraint to this is that we do not have data on the carbon emissions of each business. The only guide we have in the NNDR is the SCat code attached to each record. There are several hundred SCat codes, each relating to a different sort of business, ranging from the very broad, e.g ,"offices and computer

centres" to the very specific, e.g. "fish farms". "High carbon" businesses were therefore selected on the basis of these SCat codes. A list of these can be found in Appendix 3, Table 2.

These selected businesses were depicted in an interactive web map which colour-codes them by activity (waste, energy, minerals processing, metal processing) and indicates their size in terms of rateable value (

Figure 9). This map can be viewed here.

Figure 9: Interactive map of enterprises in high-carbon sectors, indicating sector and size. Commercial property data courtesy of the Valuation Office Agency.

This is a reasonable way of indicating where clusters of economic activity in given sectors can be found. However, in this form it is only partially useful: as discussed above, the NNDR does not provide information on the carbon 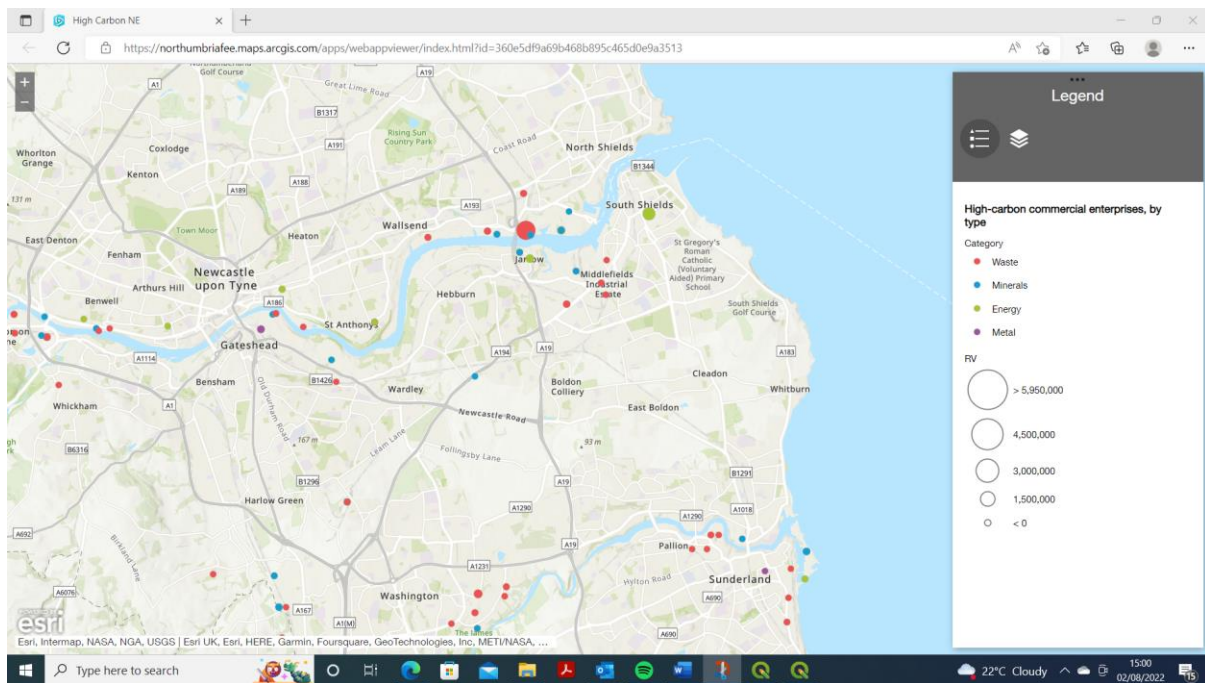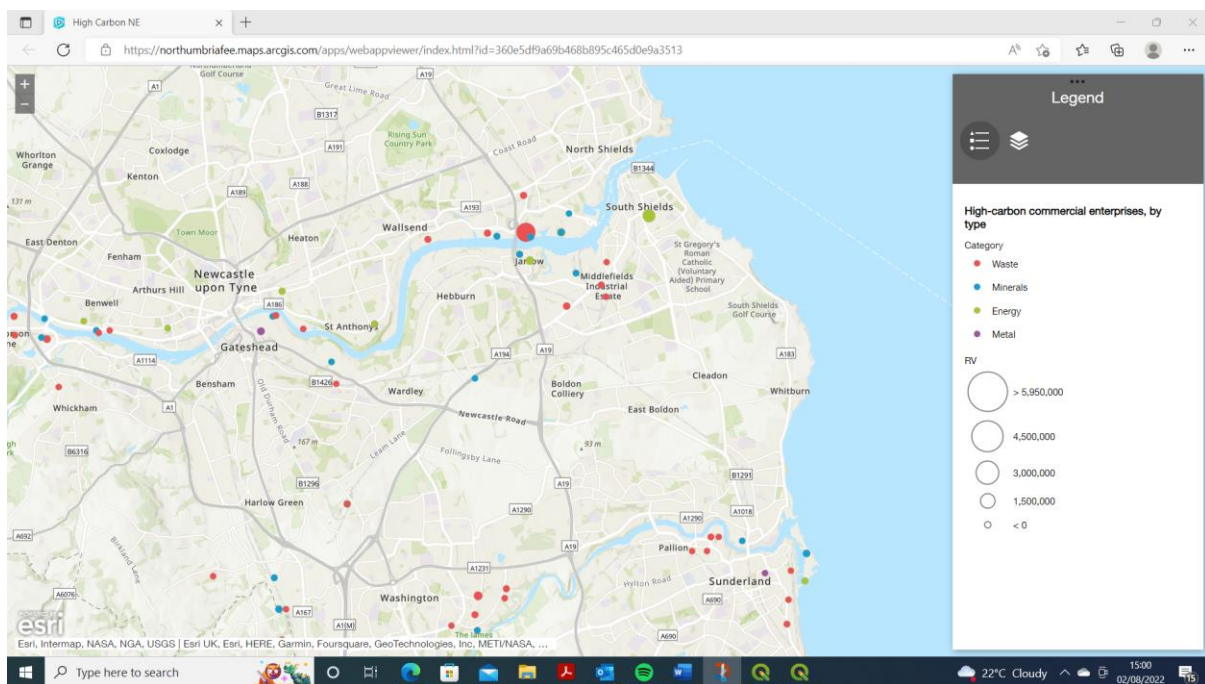emissions of enterprises or offsets. Some clues, with regard to some types of enterprise, might be obtained by cross-referencing the NNDR with Energy Performance Certificate (EPC) data, which is publicly available. However, although EPC data, including historic records, now contain the Ordnance Survey's Unique Property Reference Numbers (UPRNs), the NNDR does not. This makes it difficult to join the two datasets. In any case, EPC data does not, of course, give information on the carbon emissions associated with industrial activities.

Further research on this topic is needed to determine the location of the most significant carbon emitters and sectors.

5.6   Business start-ups and deaths.

The NNDR list is fully updated every few years; the last two updates were 2017 and 2010. In between these dates, the current NNDR is updated on a regular basis, as new businesses are created and others cease to exist. This up-to-date version of the NNDR can be filtered to include only businesses in existence at present, whereas older versions can be filtered in order to include only those in existence at the time the list was established. Three versions of the NNDR can therefore be obtained, dating from 2010, 2017 and 2022. Sadly, it does not seem to be possible to obtain NNDR lists for intermediate points.

All hereditaments in the NNDR have a Unique Address Reference Number (UARN). This means that we can establish which businesses ceased to exist, and which have been created, between 2010 and 2022 by comparing the UARNs in each list: UARNs found in 2010 but not in 2022 represent business deaths, whereas UARNs found in 2022 but not in 2010 represent business births. (The same principle applies to discovering business births and deaths between 2010 and 2017, and between 2017 and 2022.)

There is a caveat, however: sometimes the same enterprise on the same site may have different UARNs at different points in time – for example, when hereditaments are split or merged. This means that the proportion of births and deaths may be exaggerated.

It is possible to sub-divide commercial properties by broad sector, sub-sector and SCat (i.e. specific business class). This means that patterns of change in different sectors of the economy can also be shown. There is another caveat, however: the VOA categorise businesses (or rather, hereditaments) by SCat rather than by SIC codes, and there does not seem to be a good way of determining SIC from SCat because there are more of the former than the latter.

Some indicative analysis of trends in business births and deaths is provided at Appendix 4.

A web map has been created to indicate the percentage of businesses in 2022 (
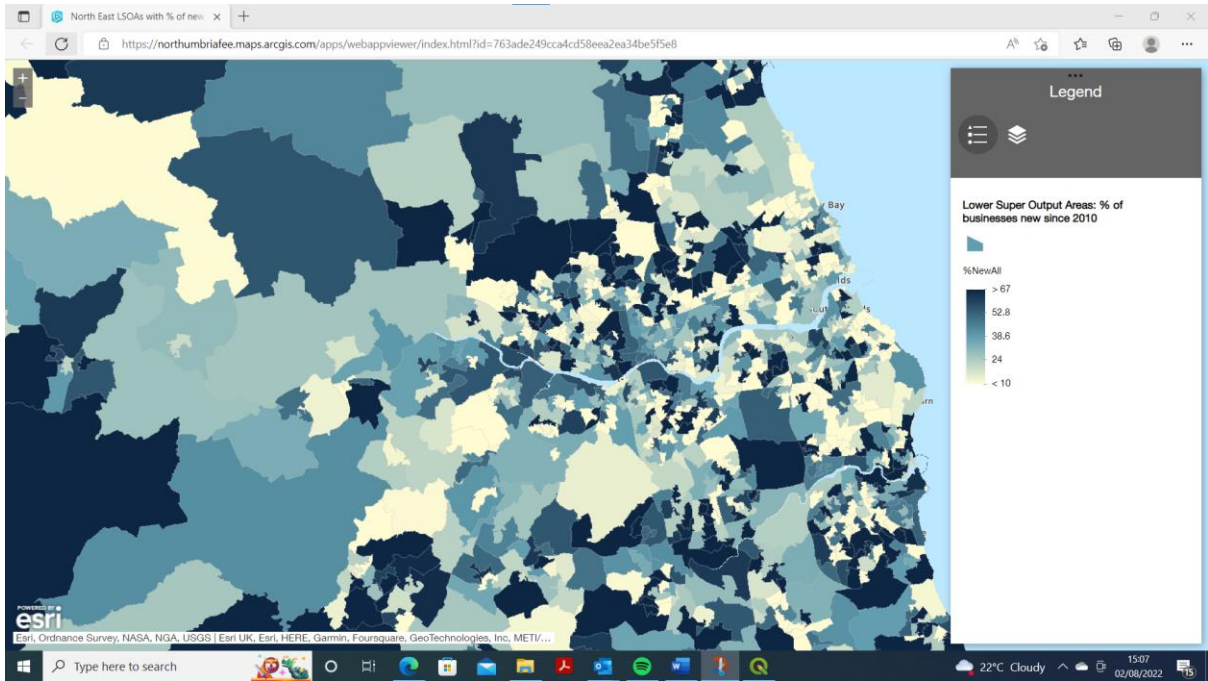
Figure 10), in each LSOA, which have been "born" since 2010; this can be viewed here. Further information about births and deaths in each sector, in each LSOA, can be found by clicking on the LSOA in question.
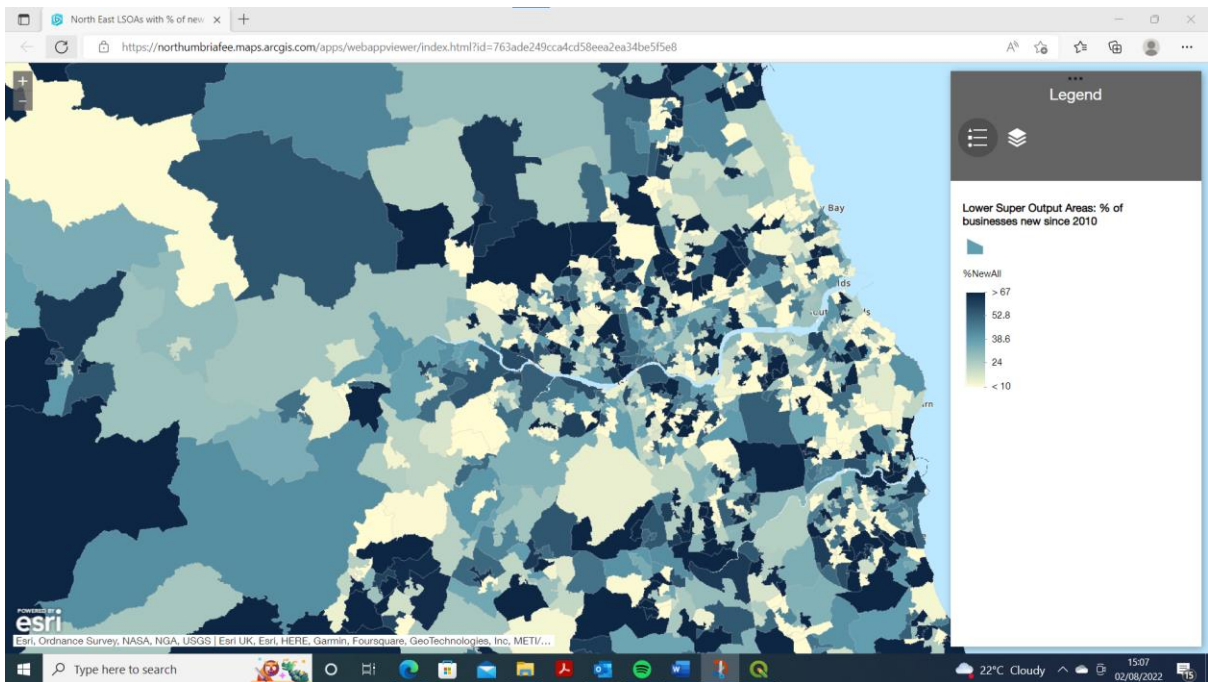
Figure 10: Interactive map of percentage of businesses in each LSOA which have been created since 2010. Commercial property data courtesy of the Valuation Office Agency.

6   Recommendations

A key recommendation is that a region-wide Geospatial Standardisation Group (GSP) (covering the wider North East Region including Tees Valley, County Durham, South and North of Tyne) be

established to decide on the standards for the production and use of geospatial data and tools across the region. It should contain representatives from local authorities, the combined authorities, the LEP and the region's universities.

The GSP should conduct an audit to determine which geospatial datasets local authorities are required or expected to compile. A regional protocol for these datasets should be established, to standardise:

- The methodology by which these datasets are obtained, whether by collation of primary data, or by analysis of other datasets
- The format by which the data should be collated, including type and format of metadata
- A standard means of publication, working on the assumption that geospatial data should be made freely available online unless there are legal or other barriers to this.

A second recommendation is that a regional GIS platform for the display of geospatial data should be created, in which users could select one or more geospatial datasets, provided by local authorities or other governmental bodies, and view them on an interactive map. A web-based interface should be established for the analysis of potential development sites. This site would contain all of the datasets which, in the opinion of the GSP, should be taken into consideration in the assessment of potential development sites. These would include natural and built environment assets and buffers thereof, and datasets relating to access to facilities and public and active transport. The interface would allow users to upload datasets of sites and would then produce a new geospatial dataset or matrix detailing, for each site, whether it was affected by any of the constraints or assets listed.

Further research should be conducted, specifically to establish the impacts of "gaps" in broadband accessibility. This research would include consideration of the characteristics of the population not able to access good broadband services, and whether poor connectivity exacerbates other inequalities, or is having an effect upon economic performance. Further research should also be carried out on the carbon emissions of businesses and their spatial distribution.

The regional GSP should make representations to national government, requesting greater interoperability between datasets. In particular, UPRNs, and co-ordinate information, should be added to the NNDR.

Other suggested activities include:

- Testing and refining of protocol
- Comprehensive audit of spatial analysis, tools, data, software etc. being deployed by LAs, CAs, LEP and others
- Sub group to pursue common themes linked to policy areas

For further information or conversation about the project please contact:

Dr Paul Greenhalgh
Principal Investigator
Professor of Real Estate and Regeneration
Department of Architecture and Built Environment
Northumbria University
paul.greenhalgh@northumbria.ac.uk

Dr Helen King
GIS Lead
Department of Geography and
Environmental Science
Northumbria University
helen.m.king@northumbria.ac.uk

Appendix 1: Example Proforma/Matrix

Table 1: Example proforma/matrix to capture themes and priorities to be completed by project partners.

Theme	Task	Why data is needed	What data do we have	Data needed	Geography Level
Place					
Place					
People					
People					
Business					
Business					

Key:
Priority
Day to day ongoing mapping

Appendix 2: List of datasets which are used or referenced in PPG or the NPPFdirectly or otherwise

Set 1: Polygon or point datasets relating to a single quality or feature

Name	Subset	Location	Required/ invoked in PPG	Collated by/ notes
Groundwater contamination	Contamination	n/a	PPG	Local authority
Soil contamination	Contamination	n/a	PPG	Local authority
Former mines	Contamination	*http://mapapps2.bgs.ac.uk/coalauthority/ home.html*	PPG	Coal Authority
Minerals safeguarding areas	Resource management	n/a	Minerals	Local Authority
Protected species	Natural environment	n/a	PPG requires "an assessment of existing and potential components of ecological networks, biodiversity resources and landscapes" and mentions "Habitats Sites" in the context of HRA (and therefore only	Information isn't publicly available. Can be purchased from ERIC North East.
Priority habitats	Natural environment	On the ArcGIS Online interface as "Living Atlas" layers and also available at https://environment.data.gov.uk/.		Natural England
Ancient woodland	Natural environment	On the ArcGIS Online interface as "Living Atlas" layers and also available at https://environment.data.gov.uk/.		Natural England
Local Wildlife Sites	Natural environment	On the ArcGIS Online interface as "Living Atlas" layers and also available at https://environment.data.gov.uk/.		Natural England
Local Nature Reserves	Natural environment	On the ArcGIS Online interface as "Living Atlas" layers and also available at https://environment.data.gov.uk/.		Natural England

National Nature Reserves	Natural environment	On the ArcGIS Online interface as "Living Atlas" layers and also available at https://environment.data.gov.uk/.	relevant to European protected sites)	Natural England
Sites of Special Scientific Interest	Natural environment	On the ArcGIS Online interface as "Living Atlas" layers and also available at https://environment.data.gov.uk/.		Natural England
Special Protection Areas	Natural environment	On the ArcGIS Online interface as "Living Atlas" layers and also available at https://environment.data.gov.uk/.		Natural England
Special Areas of Conservation	Natural environment	On the ArcGIS Online interface as "Living Atlas" layers and also available at https://environment.data.gov.uk/.		Natural England
Ramsar sites	Natural environment	On the ArcGIS Online interface as "Living Atlas" layers and also available at https://environment.data.gov.uk/.		Natural England
Tree Protection Order trees/ woodland	Natural environment/ heritage	n/a		Local authorities
Flood risk from rivers and sea (Zones 2 and 3)	Natural environment	On the ArcGIS Online interface as "Living Atlas" layers and also available at https://environment.data.gov.uk/.	PPG/ NPPF	Defra. Need for "flood and water management" invoked in PPG
Surface water flood risk	Natural environment	On the ArcGIS Online interface as "Living Atlas" layers and also available at https://environment.data.gov.uk/.	PPG /NPPF (indirectly)	Defra. Many different datasets available. Need for "flood and water management" invoked in PPG
Listed buildings	Heritage	https://historicengland.org.uk/listing/the-list/data-downloads/	PPG mentions "the relevant historic environment record [and] the	Historic England
Scheduled Ancient Monuments	Heritage	https://historicengland.org.uk/listing/the-list/data-downloads/		Historic England

"Listed" Parks and Gardens	Heritage	https://historicengland.org.uk/listing/the-list/data-downloads/	National Heritage List for England"	Historic England
Historic battlefields	Heritage	https://historicengland.org.uk/listing/the-list/data-downloads/		Historic England
World Heritage Sites	Heritage	https://historicengland.org.uk/listing/the-list/data-downloads/		Historic England
Protected wrecks	Heritage	https://historicengland.org.uk/listing/the-list/data-downloads/		Historic England
Buildings of local conservation interest	Heritage	n/a		Local authorities
Conservation areas	Heritage	https://historicengland.org.uk/listing/the-list/data-downloads/		Historic England dataset is suspect: it says there's no conservation areas in Newcastle or Gateshead, which there certainly are.
Greenbelt	Planning designations	https://data.gov.uk/dataset/ccb505e0-67a8-4ace-b294-19a3cbff4861/english-local-authority-green-belt-dataset	No (but to be fair this is an output of planning , not an input)	Ministry of Housing, Communities and Local Government
Public open space	Land-use designations	n/a	PPG mentions "quality and quantity of GI, sports, recreation.."	Local authorities (Mastermap layer also exists but may be less accurate/ precise)
Public rights of way	Land-use designations/ infrastructure	n/a		Local authorities
Rights of access	Legal	n/a	PPG	Not publicly available
Land Ownership	Legal	n/a	Not directly	Not publicly available. Can be purchased from Land Registry Office

Utilities infrastructure	Infrastructure	n/a	PPG references achievability of sites	Not sure
Communications infrastructure	Infrastructure	n/a	PPG references achievability of sites	Not sure
Local highways infrastructure	Infrastructure	n/a	PPG references achievability of sites	Ordnance Survey has datasets but local authorities themselves may have more detailed knowledge of their own highways
Watercourses and water bodies	Topography	https://www.ordnancesurvey.co.uk/business-government/products/open-map-rivers	PPG – indirectly, wrt flood and water management	Ordnance Survey. Whether free or not depends on the level of information required
Locations of deprivation	Demographics	https://www.gov.uk/guidance/english-indices-of-deprivation-2019-mapping-resources	PPG	Department for Levelling Up, Housing and Communities and Ministry of Housing, Communities & Local Government
Best and most versatile agricultural land	Resources	https://environment.data.gov.uk/DefraDataDownload/?mapService=NE/AgriculturalLandClassificationProvisionalEngland&Mode=spatial	PPG	Defra
Quality and quantity of healthcare	Resources	n/a	PPG	Local authorities/ NHS
Education, sports, recreation, places of worship	Resources	n/a	PPG	Local authorities

Location of hazardous establishments	Contamination	n/a	PPG	Local authorities
Air Quality Management Areas	Contamination	https://uk-air.defra.gov.uk/aqma/	PPG	Defra

Set 2: Datasets derived from analysis of other datasets to aid the categorisation of sites or premises in planning terms

Note: PPG and NPPF repeatedly mention the suitability and achievability of housing and employment land in general terms;this table only covers specific aspects thereof.

Name	Subset	Required/ invoked in PPG/ NPPF	Derived from
Buildings in need of repair	Built environment management	No	Site-specific analysis
Buildings in need of demolition	Built environment management	No	Site-specific analysis
Sites whose topography makes them undevelopable	Site suitability	PPG references achievability of sites	Filtering by gradient
Land locked sites	Site suitability	PPG references achievability of sites	Existing urban fabric and highway network
Sites whose development would have unacceptable traffic management impacts	Site suitability	NPPF. PPG references achievability of sites	Existing highway network

Viable and non-viable sites	Site suitability	NPPF and PPG	Land prices, cost of development, local property prices
Housing, employment, retail and leisure land requirements	Built environment	NPPF and PPG	Demographic predictions, economic growth predictions, assessments of density (note: these assessments yield a required quantity of land, not a geospatial dataset)
Sites which are well-related to existing settlements	Site suitability	Yes - NPPF	Existing urban fabric and built-up areas dataset
Sites which are or can be made sustainable [in transport terms]	Built environment	NPPF	Built-up areas dataset, highway network, public transport route maps
Housing Market Area	Built environment	PPG	Built-up areas dataset, travel-to-work areas, data on relocation
"The most appropriate geography to prepare policies for employment"	Built environment	PPG	Built-up areas dataset, travel-to-work areas, predictions of economic growth, data on movement of companies
Suitability and achievability of housing and employment land	Infrastructure/ Built environment	PPG	Proposed sites, constraints as discussed below
Town centres	Infrastructure/ Built environment	PPG	Existing locations of commercial property
Existing and future supply of land	Built environment	PPG	Site-specific inquiries or analyses
Nature Recovery Networks/ Ecological Networks / Need for green infrastructure	Natural environment	PPG	Existing green infrastructure, proposed development or predicted population change in the future
Landscapes	Natural environment	PPG	Landscape character assessment

Need for education	Infrastructre	PPG	Existing educational establishments, proposed development or predicted populsation change in the future
Need for transport infrastructure	Infrastructure	PPG	Existing transport infrastructure and its capacity, likely impact of proposed development or population change under different scenarios
Need for flood and water management	Natural environment/ Resources/ Infrastructure	PPG	Flood risk areas, centres of population and other receptors, existing and planned flood defences; different scenarios for green infrastructure in and around urban areas
Need for digital infrastructure	Infrastructure	PPG	Existing digital infrastructure, centres of population
(Centres of) population	Demographics	PPG	Dataset of built up areas exists; most dense areas could be obtained by downloading dataset of (L or M)SOAs and selecting the most populous for their size

PPG also invokes detailed assessments compiled either by local authorities, or by other parties, which are likely to require geospatial data as an input, and may generate geospatial data as an output: Sustainability Appraisal and, possibly, Habitats Regulations Assessment (that's only required where sites protected at European level may be affected – but there are several in our region); River Basin Management Plans, AONB Management Plans, Green Infrastructure Plans, Tree and Woodland Strategies, Landscape Character Assessments, Water Resource Management Plans, Catchment Management Plans, Shared Nitrogen Action Plans, and conservation area management plans and appraisals.

Appendix 3: High-carbon businesses.

Table 2: Attributes of high carbon businesses.

SCATnocode	Sector	Sub-sector	SCat_Desc	Category
9	Industry	OTHER - INDUSTRIAL	Aluminium Smelting Works	Metal
17	Industry	OTHER - INDUSTRIAL	Asphalt Plants	Minerals
32	Industry	OTHER - INDUSTRIAL	Brickworks (Traditional) Clay Tile/Pipe Works	Minerals
33	Industry	OTHER - INDUSTRIAL	Bulk Cement Storage Depots	Minerals
51	Industry	OTHER - INDUSTRIAL	Cement Tile Works	Minerals
52	Industry	OTHER - INDUSTRIAL	Cement Works	Minerals
58	Other	OTHER	Civic Amenity Sites	Waste
63	Industry	OTHER - INDUSTRIAL	Coking & Carbonising Plants	Minerals
71	Industry	OTHER - INDUSTRIAL	Concrete Batching Plants	Minerals
72	Industry	OTHER - INDUSTRIAL	Concrete Block Works	Minerals
73	Industry	OTHER - INDUSTRIAL	Concrete Product Works	Minerals
88	Other	UTILITIES	District Heating Undertakings & Networks	Energy
90	Other	UTILITIES	Domestic Fuel Installations	Energy
94	Other	UTILITIES	Electricity Undertakings (Non-Statutory)	Energy
110	Industry	GENERAL - INDUSTRIAL	Foundries	Metal
115	Other	UTILITIES	Gas Processing Plants	Energy
142	Industry	OTHER - INDUSTRIAL	Iron and/or Steel Works	Metal
149	Other	UTILITIES	Landfill Gas Generator Sites	Waste
150	Other	OTHER	Land used for Waste Composting	Waste
157	Other	STORAGE & DISTRIBUTION	Liquid Bulk Storage (Incl Petrol & Oil) (National Scheme)	Energy
166	Industry	OTHER - INDUSTRIAL	Mineral Depot & Premises	Minerals
167	Industry	OTHER - INDUSTRIAL	Mineral Producing Hereditament - Blockstone	Minerals
168	Industry	OTHER - INDUSTRIAL	Mineral Producing Hereditament - Brine	Minerals
169	Industry	OTHER - INDUSTRIAL	Mineral Producing Hereditament - Chalk	Minerals
170	Industry	OTHER - INDUSTRIAL	Mineral Producing Hereditament - China Clay	Minerals

171	Industry	OTHER - INDUSTRIAL	Mineral Producing Hereditament - Clay	Minerals
172	Industry	OTHER - INDUSTRIAL	Mineral Producing Hereditament - Coal	Minerals
173	Industry	OTHER - INDUSTRIAL	Mineral Producing Hereditament - Fluorspar	Minerals
174	Industry	OTHER - INDUSTRIAL	Mineral Producing Hereditament - Gas	Minerals
175	Industry	OTHER - INDUSTRIAL	Mineral Producing Hereditament - Hardrock	Minerals
176	Industry	OTHER - INDUSTRIAL	Mineral Producing Hereditament - Inert	Minerals
177	Industry	OTHER - INDUSTRIAL	Mineral Producing Hereditament - Oil	Minerals
178	Industry	OTHER - INDUSTRIAL	Mineral Producing Hereditament - Other Mineral Category	Minerals
179	Industry	OTHER - INDUSTRIAL	Mineral Producing Hereditament - Putrescible	Minerals
180	Industry	OTHER - INDUSTRIAL	Mineral Producing Hereditament - Sand & Gravel	Minerals
181	Industry	OTHER - INDUSTRIAL	Mineral Producing Hereditament - Sand	Minerals
183	Industry	OTHER - INDUSTRIAL	Mineral Producing Hereditament - Slate	Minerals
200	Industry	OTHER - INDUSTRIAL	Nuclear Establishments	Energy
205	Industry	OTHER - INDUSTRIAL	Oil Refineries	Energy
218	Industry	OTHER - INDUSTRIAL	Potteries	Minerals
219	Other	UTILITIES	Power Generators	Energy
221	Industry	OTHER - INDUSTRIAL	Property used for Secondary Aggregate Processing	Minerals
233	Industry	OTHER - INDUSTRIAL	Refuse Destructor Plants/Disposal Sites	Waste
244	Industry	OTHER - INDUSTRIAL	Scrap Metal/Breakers Yard	Waste
245	Industry	OTHER - INDUSTRIAL	Sea Dredged Aggregate Processing Plants & Depots	Minerals
246	Other	UTILITIES	Sewage Works (National Scheme)	Waste
255	Industry	OTHER - INDUSTRIAL	Spoil Heap Workings	Minerals
297	Industry	OTHER - INDUSTRIAL	Waste Incinerator Plants	Waste
298	Other	OTHER	Waste Recycling Plants	Waste
299	Industry	OTHER - INDUSTRIAL	Waste Transfer Stations	Waste

Appendix 4: Indicative analysis of patterns in change in commercial property from 2010 to 2022

The brief for this element of the project was provided by the North of Tyne CA.

"Using VAT registration data, is there a way we can map new businesses (business registrations) attached to a postcode? Across UK SPF and some of our other economic growth objectives, we'd like to understand where businesses are being created, how they are distributed, and what SIC category they fall under. Again pins of postcodes and addresses would be useful but an aggregated and proportionate view by LSOA area would be useful (eg. Number of new businesses (raw), number of new businesses as proportion of existing, number of new businesses as proportion of local SIC), and an option to filter by time."

First steps:

The starting point is three datasets of commercial property, taken from the National Non-Domestic Ratings List. This dataset gives the rateable value for all commercial properties, together with the type of business using them. Businesses are categorised by Special Category codes (SCats) which can be grouped into sub-sectors (insert table) and then into the four broad sectors (otherwise known as bulk classes) of retail, office, industrial and "other". In this study, only the records pertaining to hereditaments in the North East region were selected.

Each record, corresponding to a hereditament, is given an unique address reference number (UARN) in the NNDR list.

Methodology

The records were geocoded by postcode, using a dataset of postcode centroids. This meant that it was also possible to identify, for each hereditament, which Lower Super Output Area it fell into.

Business start-ups and deaths - that is, new and "disappearing" businesses, were selected by comparing the UARNs in one dataset with the UARNs in another. Those which were found in a later dataset but not in an earlier one were assumed to be businesses formed in the interim (henceforth "start-ups". Those which were found in an earlier dataset but not a later one were assumed to have ceased to exist (henceforth "business deaths"). This gave us four datasets: start-ups from 2010-2017 and from 2017-22; business deaths from 2010-2017 and from 2017-22.

There is a caveat, however, insofar as it is possible for a hereditament to be given a new UARN under certain circumstances, especially if hereditaments are combined, split or altered.

The records were grouped into sub-sectors, using the VOA classification. Then the total amount of RV in start-ups and business deaths was calculated. This was then compared to the total amount of RV in the NNDR lists as a whole to show, firstly, what percentage of RV in the 2017 and 2022 lists was accounted for by start-ups (i.e. entries in the 2017 list not found in 2010, and in the 2022 list not found in 2017); and secondly, what percentage of RV in the 2010 and 2017 lists was accounted for by business deaths - businesses that subsequently ceased to exist. This was calculated for each sub-sector and for the NNDR lists as a whole. To establish, roughly, the amount of change per year, these figures were divided by seven (for the period between 2010 and 2017) and five (for 2017 – 22).

Findings

The most significant finding is that there is a great deal of churn in commercial property. In total, 37% of UARNs in 2017 were not present in 2010. The corresponding figure between 2017 and 2022 is 23%, indicating a similar rate of change per year. That this is actually, for the most part, churn, rather than growth, is indicated by the fact that business deaths over the same periods were 33.7% and 21.7% respectively.

The next thing to note is that some sub-sectors have more churn in them than others – as distinct from decline or growth. In fact, although the total amount of RV in commercial property in the North East increased by about 1.77 % in between 2010 and 2017, different sectors behaved differently. As Figures 1 and 2 below show, slight increases in assembly and leisure, health, education, industry, and storage and distribution, were balanced by decreases in offices, shops, utilities and the "other-industrial" category, which includes minerals and metal extraction and processing sites, chemical works and garages.

Churn, where this was not explained by growth or decline, was generally comparable with the overall figure. It was slightly lower with regard to non-residential and residential institutions, and financial and professional services; and slightly higher with regard to education, retail and offices – and extremely high with regard to transport. In the latter case, though, when the "start-ups" and "deaths" are mapped, many of them are in exactly the same place, indicating that the "dead" and "new" properties may be the same ones, but with different UARNs.
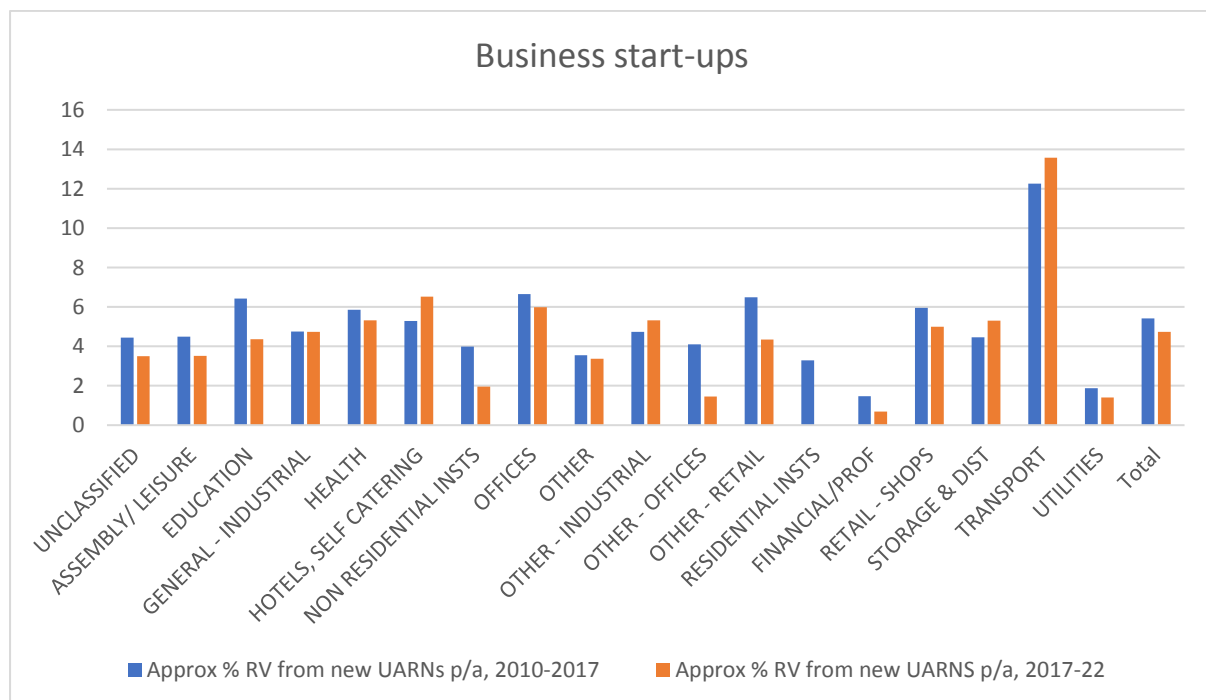

Figure 11: Business start-ups, by sub-sector, 2010 – 2022.
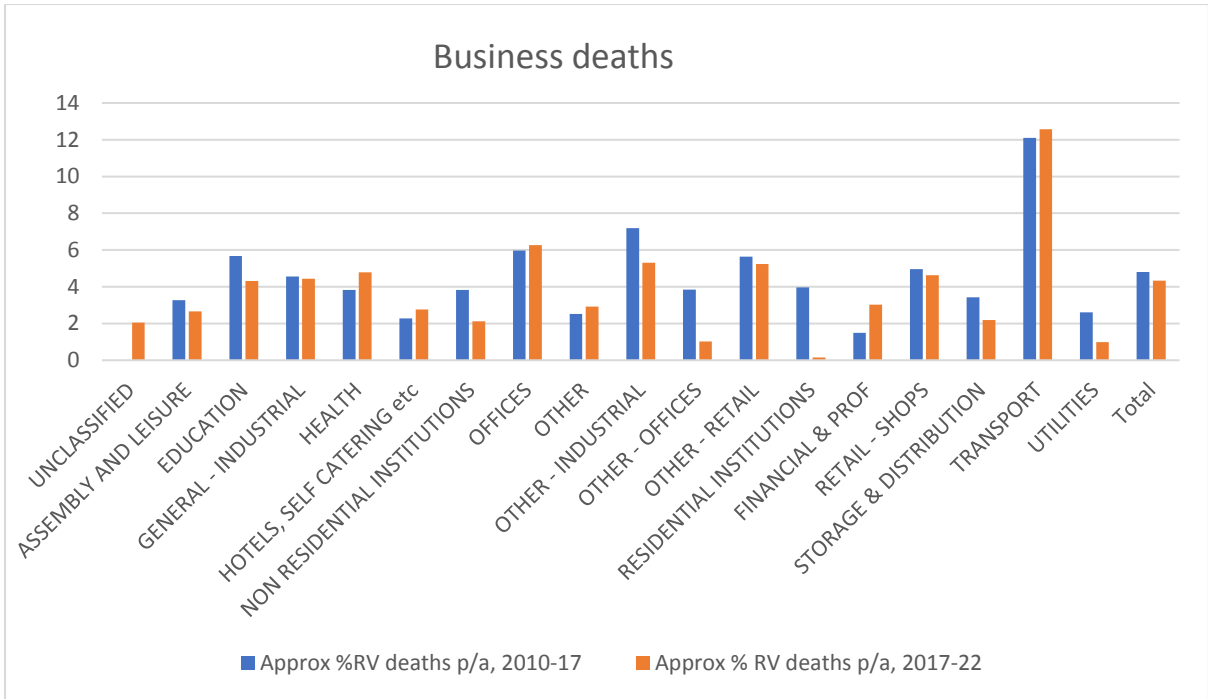
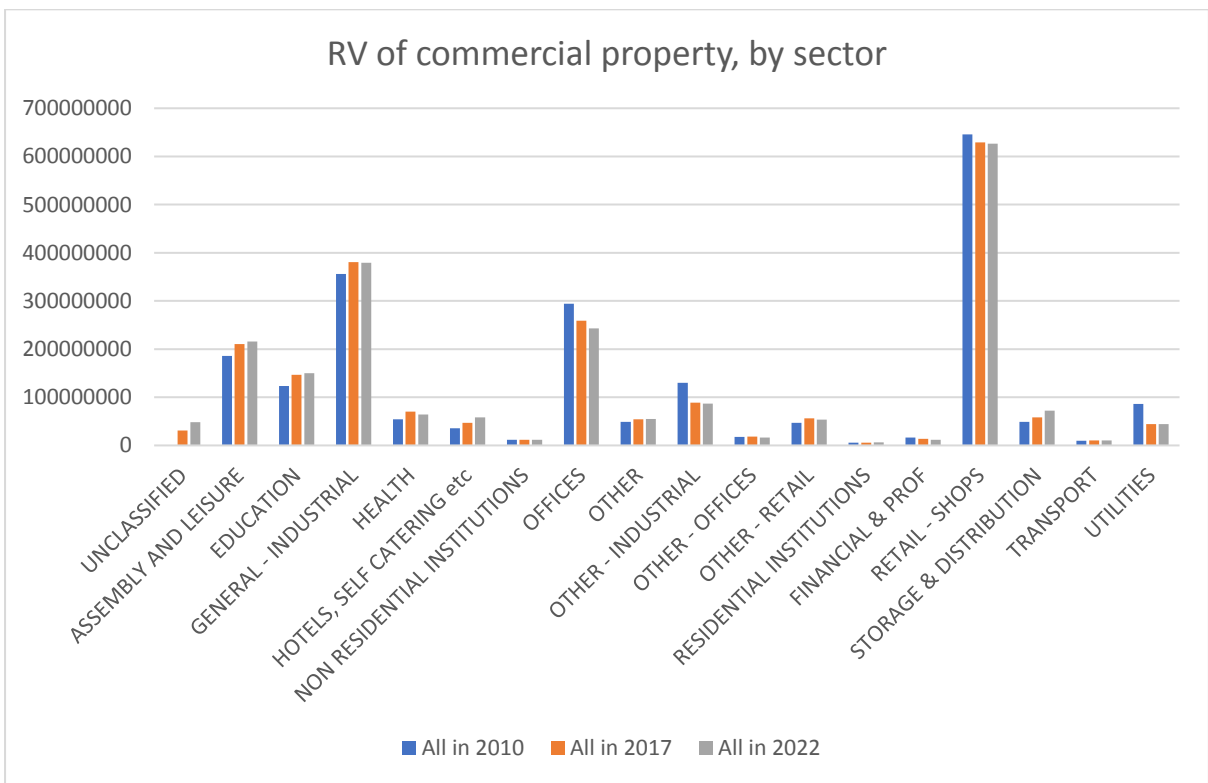Figure 12: Business deaths, by sub-sector, 2010 - 2022.

Figure 13: Total commercial property value, 2012-2022.

Appendix 5: Schedule of meetings with project partners

Meetings took place via Teams, apart from those on 08/06/22 and 13/06/22.

Northumbria University representatives were Paul Greenhalgh, Josephine Ellis and Helen King.

Table 3: Schedule of meetings with project partners.

Date	Organisation	Representatives
24/02/22	North East Local Economic Partnership	Richard Baker Paul Corbett
02/03/22	North of Tyne Combined Authority	Martin Wood Alex Black Aliyah Farooq
04/03/22	Tees Valley Combined Authority	Geraldine Brown
14/03/22	North of Tyne Combined Authority	Alex Black Aliyah Farooq Janet Ross
22/03/22	Tees Valley Combined Authority	Geraldine Brown Natasha Wilson Clare Ross Julie McCartney
29/03/22	North of Tyne Combined Authority	Alex Black Aliyah Farooq
30/03/22	Stockton Borough Council	Viv Lattimer David Bage Johanne Parker Wedad Mitoubsi
20/04/22	Tees Valley Combined Authority	Geraldine Brown Natasha Wilson
03/05/22	Darlington Borough Council	Philippa Rayner
08/06/22	North of Tyne Combined Authority	Alex Black Aliyah Farooq Other colleagues via Teams
13/06/22	Tees Valley Combined Authority	Natasha Wilson Geraldine Brown Julie McCartney
19/07/22	Tees Valley Combined Authority	Natasha Wilson Geraldine Brown
19/07/22	North of Tyne Combined Authority	Aliyah Farooq