Office for National Statistics

Article

Clustering local authorities against subnational indicators, England

Clustering analysis exploring similarities between local authorities in England. Experimental statistics.

Contact:
James McCrae, Ben Ford, Will Haste
subnational@ons.gov.uk
+44 1633 455135

Release date:
24 February 2023

Next release:
To be announced

Correction

11 May 2023 16:00

We have corrected an error with our coastal towns lookup, meaning an additional 11 local authorities that were not originally listed as containing coastal towns, are now listed as containing coastal towns. These are:

- Bournemouth, Christchurch and Poole
- Castle Point
- Dorset
- East Suffolk
- City of Kingston Upon Hull
- Liverpool
- Rochford
- Somerset West and Taunton
- Wealden
- West Lancashire
- Worthing

The proportion of coastal towns within each cluster have been updated in this article alongside our accompanying dataset. This happened because of human error. We apologise for any misunderstanding caused.

Table of contents

1 . Main points

- We have grouped together similar local authorities ("clusters"), using a k-means clustering method, against metrics reported on our subnational indicators explorer; our analysis allows users to understand the similarities between local authorities, while providing local authorities with control groups for investigating the impact of policy interventions.
- We have created seven different models on different themes using data available from the December edition of our subnational indicators explorer.
- Analysis of selected headline metrics across seven different themes produced four clusters of local authorities in England; the higher health and productivity, lower well-being cluster tended to be urban, densely populated areas and mostly located in London and the South East.
- In the same model, the higher health and well-being, lower connectivity cluster was made up of mostly older, rural, least deprived areas.
- The higher health and well-being, moderate educational performance cluster predominantly comprised older, low population areas.
- The higher connectivity, lower health and well-being cluster comprised mainly younger, urban, and the most deprived areas.

These are Experimental Statistics. The clustering method used is currently under development and does not control for any characteristics of place, which means that results could be subject to future revisions. We recommend caution when using the data and findings should be seen as indicative only.

The results presented are not used to influence Levelling Up policy decisions and should not be viewed as being a judgment about the performance of a local authority.

We are keen to get feedback from users to help us develop this analysis in the future. Please send any feedback to subnational@ons.gov.uk.

2 . Interactive map showing the results of cluster analysis

We have created an interactive tool (Figure 1) to present the findings for our analysis. Results for each cluster are presented in this tool across all seven themes. Local authorities of interest can be selected from the drop-down menu, and text summarising the clusters is shown.

Figure 1: Find out which local authorities share similar characteristics based on our subnational indicators explorer

Results of cluster analysis for local authorities, England

Notes:

1. Between 2 and 19 local authority (LA) areas have not been assigned to a cluster in each model (denoted by “[x]” in table 1 of the data tables) either because data were missing for at least one metric used in the model, or figures for a LA were provided using older district boundaries.

2. Medians per cluster were measured against the median score for all local authorities assigned to a cluster (regardless of cluster), the values for which can be found in our accompanying dataset. The median score across all LAs should not be treated as a national average.

3. Because of data availability and to have consistent geographical coverage across all our models, we have only included local authorities in England in our analysis.

More information is available in our data tables.

3 . Summary characteristics for each model

This section outlines the results and design of each of the clustering models presented in the map in Section 2. For each cluster, we outline where the cluster median is above or below the overall median for each metric. This is so that users can understand if each metric in each cluster are above or below the average of all local authorities included in our models.

The characteristics for the local authorities (LAs) included in each cluster are also outlined using demographic information about local authorities. These characteristics are included where there was a higher proportion per cluster than the overall proportion.

Full information about the sources we used can be found in our Clustering local authorities against subnational indicators methodology. All model results are independent from each other and comparisons between model results should not be made.

Important characteristics for each cluster are presented within this section. Full breakdowns of the characteristics, including overall proportions used for comparison, can be found in our accompanying dataset.

Headline metrics model

To better understand similarities of LAs based on multiple topic areas, we devised a model based on the seven subnational themes analysed in this article. The model produced four distinct clusters.

In our model, we chose one headline metric per theme where possible to represent each theme. The metrics included, and their subject matter date, were:

- gross value added (GVA) per hour worked (2020)
- average travel time to nearest employment centre with 500 to 4999 jobs available by public transport or walking (2019)
- gigabit capable broadband (May 2022)
- pupils at expected standards by the end of primary school (KS2) (2019)
- apprenticeships achievements per 100,000 (2020 to 21)
- healthy life expectancy (HLE) (2018 to 2020)
- life satisfaction (April 2021 to March 2022)

Further information on the metrics used and why they were chosen is available in our Clustering local authorities against subnational indicators methodology.

Higher health and well-being, lower connectivity (52 LAs)

More than half of LAs in the South West (52%) are included in this cluster, while no LAs from London, the North East and the South East are included. Examples of LAs included are Cornwall and Carlisle. Our results showed that this cluster:

- was better than the median for apprenticeships achievements, HLE and life satisfaction
- was worse than the median for GVA, public transport, broadband and KS2
- consisted of mainly rural LAs (83%) and LAs with a median age of 45 years or older (90%), and had a number of coastal towns (35%)

Higher connectivity, lower health and well-being (78 LAs)

At least one LA from all English regions are included in this cluster. Most LAs in the North East (90%) are in this cluster, and the highest proportion of North West (49%), Yorkshire and The Humber (48%) and West Midlands (40%) LAs fall into this cluster. Examples of LAs are Manchester, Liverpool, and Southampton. Our results showed that this cluster:

- was better than the median for public transport, broadband and apprenticeships achievements
- was worse than the median for GVA, KS2, HLE and life satisfaction
- consisted mainly of urban LAs (90%), had a median age of 40 years or younger (56%) and had a higher proportion of the most deprived areas (60%)

Higher health and well-being, moderate educational performance (114 LAs)

At least one LA from all English regions is included in this cluster. More than half of LAs in the East of England (62%) and the South East (59%) are included. The highest number of LAs in the East Midlands (39%) are also in this cluster. Examples of LAs are Exeter and Sheffield. Our results showed that this cluster:

- was better than the median for GVA, KS2, HLE and life satisfaction
- was worse than the median for public transport, broadband and apprenticeships achievements
- had areas with 185 to 440 people per square kilometre (38%), and had LAs with a median age of 41 to 44 years (41%)

Higher health and productivity, lower well-being (56 LAs)

Nearly all LAs in London (88%) are included in this cluster. Examples of LAs included are Westminster, Trafford and Solihull. Our results showed that this cluster:

- was better than the median for GVA, public transport, broadband, KS2 and HLE
- was worse than the median for apprenticeships achievements and life satisfaction
- consisted of mainly urban LAs (89%) and LAs where the median age was 40 years or younger (70%), and had areas with 2777 people or more per square kilometre (50%)

Economic (pay, employment and productivity) model

Our model using metrics relating to pay, employment and productivity produced four distinct clusters. The metrics included were:

- gross value added (GVA) per hour worked (2020)
- gross median weekly pay (2021)
- employment rate for people aged 16 to 64 years (July 2021 to June 2022)
- gross disposable household income (GDHI) per head (2020)

Below median on all economic metrics (103 LAs)

All of the North East's LAs (100%) are in this cluster. All regions have at least one LA in this cluster, containing over half of LAs from Yorkshire and The Humber (62%) and the North West (56%). Examples of LAs are Middlesbrough and Kingston upon Hull. Our results showed that this cluster:

- was worse than the median for all metrics
- consisted mainly of urban LAs (65%), had a higher proportion of the most deprived areas (47%), had a number of coastal towns (39%) and had areas with 441 to 1073 people per square kilometre (26%)

Above median employment rate, below median productivity (131 LAs)

All regions except the North East had a least one LA in this cluster. Over half of LAs from the West Midlands (63%), the South West (62%), the East of England (56%) and the East Midlands (51%) are in this cluster. Examples of LAs are Malvern Hills and Rutland. Our results showed that this cluster:

- was better than the median for gross median weekly pay, employment rate and gross disposable household income
- was worse than the median for gross value added
- consisted of a large number of rural LAs (39%) and LAs with a median age of 45 years or older (40%), and had areas with 185 to 440 people per square kilometre (27%)

Above median on all economic metrics (66 LAs)

This cluster is predominantly located in London (63%) and the South East (44%). Examples of LAs include Croydon and Surrey Heath. Our results showed that this cluster:

- was better than the median for all metrics
- consisted mainly of urban LAs (68%), had a higher proportion of the least deprived LAs (43%), had LAs with a median age of 40 years or younger (53%) and had areas with 2777 people or more per square kilometre (33%)

Far above median on all metrics except employment rate (7 LAs)

This cluster is almost entirely located in London, with six of the seven LAs included from London, for example Kensington and Chelsea. Our results showed that this cluster:

- was better than the median for gross value added, gross median weekly pay and gross disposable household income
- was worse than the median for employment rate
- consisted of only urban LAs (100%), had LAs with a median age of 40 or younger (71%), had a higher proportion of the least deprived LAs (29%) and had areas with 2777 people or more per square kilometre (86%)

Connectivity model

We have used indicators relating to local transport connectivity and digital infrastructure to produce a model for connectivity. The metrics included were:

- average travel time to nearest employment centre with 500 to 4999 jobs available by public transport or walking (2019)
- average travel time to nearest employment centre with 500 to 4999 jobs available by driving (2019)
- average travel time to nearest employment centre with 500 to 4999 jobs available by public transport or cycling (2019)
- gigabit capable broadband (May 2022)
- 4G coverage (May 2022)

Our model produced four distinct clusters. Although the first three clusters presented were below median for all metrics, there was variation in the distance from the medians for each cluster. This has been reflected in the cluster names.

The Isles of Scilly were removed from our model because of the long travel time to the nearest employment centre skewing the results for this area.

Far below median on all connectivity metrics (12 LAs)

This cluster included LAs from the East Midlands, the North West, the South West and Yorkshire and The Humber. Examples of LAs included are Eden and West Devon. Our results showed that this cluster:

- was worse than the median for all metrics
- consisted mainly of rural LAs (83%) and LAs with a median age of 45 or older (92%), had a number of coastal towns (33%), and had areas with fewer than 185 people per square kilometre (100%)

Below median on all connectivity metrics (39 LAs)

Of all the LAs in the South West, over a third (37%) are in this cluster. Examples of LAs are Selby and Shropshire. Our results showed that this cluster:

- was worse than the median for all metrics
- consisted mainly of rural LAs (82%) and LAs with a median age of 45 years or older (87%), contained a number of coastal towns (38%) and had areas with fewer than 185 people per square kilometre (74%)

Slightly below median on all connectivity metrics (135 LAs)

Over half of LAs in the South East (67%), the East of England (60%), the East Midlands (58%), the West Midlands (57%), and the North East (50%) are included in this cluster. Some examples of LAs included are Fareham, Oxford and Middlesbrough. Our results showed that this cluster:

- was worse than the median for all metrics
- had areas with a median age of 41 to 44 years (47%), had a higher proportion of the least deprived LAs (27%) and had areas with 185 to 440 people per square kilometre (39%)

Far above median on all connectivity metrics (115 LAs)

All LAs in London (100%) and over half of LAs in the North West (59%) are in this cluster. Examples of LAs included are Blackpool, Croydon and Peterborough. Our results showed that this cluster:

- was better than the median for all metrics
- consisted mainly of urban LAs (98%), had LAs with a median age of 40 years or younger (73%), had a higher proportion of the most deprived LAs (36%), and had areas with 2777 people or more per square kilometre (50%)

Educational attainment model

When using metrics relating to educational attainment, our model produced eight distinct clusters. This is higher than the number of clusters used in the other models, and is likely because of a greater number of metrics being included in this model. We will explore the impact of the number of metrics on the number of clusters in future articles. The metrics included were:

- pupils at expected standards by end of primary school (KS2) (2019)
- GCSEs (and equivalent) in English and maths by age 19 (2020 to 21)
- schools and nursery schools rated good or outstanding (September 2021 to August 2022)
- persistent absences for all pupils (2020 to 21)
- persistent absences for pupils eligible for free school meals (FSM) (2020 to 21)
- persistent absences for children looked after (CLA) by LAs (2020 to 21)
- children at the expected standard for communication and language by the end of the early years foundation stage (2018 to 19)
- children at the expected standard for literacy by the end of the early years foundation stage (2018 to 19)
- children at the expected standard for maths by the end of the early years foundation stage (2018 to 19)

Lower on all education metrics (18 LAs)

This cluster is focused on the North West and the West Midlands, with 21% and 17% of these regions' LAs in this cluster. Examples include Halton and Stoke on Trent. Our results showed that this cluster:

- was worse than the median for all metrics
- consisted only of urban LAs (100%), had LAs with a median age of 40 or younger (83%) and had a higher proportion of the most deprived LAs (94%)

Higher GCSEs, lower KS2 performance (51 LAs)

This cluster is focused around the South West and the East of England, with 48% and 38% of these regions' LAs in this cluster. Examples include Cornwall and Cambridge. Our results showed that this cluster:

- was better than the median for early years communication and language, early years maths, and for GCSEs performance
- was worse than the median for KS2, good or outstanding schools and CLA absences
- consisted mainly of rural LAs (65%) and LAs with a median age of 45 or older (67%) and had a number of coastal towns (35%)

Lower on all metrics, except FSM and CLA absences (40 LAs)

This cluster is Midlands focused, with 50% of the East Midlands' and 27% of the West Midlands' LAs in this cluster. Examples include Erewash and Worcester. Our results showed that this cluster:

- was better than the median for FSM and CLA absences
- was worse than the median for good or outstanding schools, early years maths and early years communications and languages
- consisted mainly of urban LAs (60%) and LAs with a median age of 45 or older (40%)

Lower foundation age performance (33 LAs)

The North West makes up 48% of this cluster, with 41% of this regions' LAs in this cluster. Examples include Preston and Fylde. Our results showed that this cluster:

- was better than the median for good or outstanding schools and FSM absences
- was worse than the median for early years communications and languages, early years maths, and early years literacy
- consisted mainly of urban LAs (73%), had LAs with a median age of 40 or younger (48%) and had a higher proportion of the most deprived LAs (46%)

Slightly above median except for FSM absences (67 LAs)

The East of England makes up 36% of this cluster, with 53% of this regions' LAs in this cluster. Examples in this cluster include South Gloucestershire and Thurrock. Our results showed that this cluster:

- was better than the median for KS2, CLA absences and early years literacy
- had a higher proportion of the least deprived LAs (25%)

Higher KS2, lower GCSE performance (32 LAs)

The South East makes up 44% of this cluster, although as the region with the most LAs, only 22% of its LAs are in this cluster. Examples in this cluster include Dover and Dartford. Our results showed that this cluster:

- was better than the median for KS2, early years literacy and good or outstanding schools
- was worse than the median for CLA and FSM absences, and GCSEs performance
- had LAs with a median age of 41 to 44 years (47%), had a higher proportion of the most deprived LAs (28%), and had a number of coastal towns (41%)

Higher against all metrics, except FSM absences (33 LAs)

The South East makes up an even larger proportion of this cluster, with 79% of LAs in this cluster being from this region, representing 41% of the region's LAs. Examples in this cluster include West Berkshire and Bracknell Forest. Our results showed that this cluster:

- was better than the median for KS2, GCSEs performance, and early years communications and languages
- was worse than the median for FSM absences
- consisted mainly of urban LAs (70%), had LAs with a median age of 41 to 44 years (52%) and had a higher proportion of the least deprived LAs (70%)

Far higher against most education metrics (29 LAs)

This cluster is focused on London, which makes up 69% of the cluster, representing 63% of its LAs. Examples in this cluster include Hackney and Tower Hamlets. Our results showed that this cluster:

- was better than the median for KS2, good or outstanding schools, and FSM absences
- was worse than the median for early years maths, and early years communications and languages
- consisted mainly of urban LAs (83%) and had LAs with a median age of 40 or younger (72%)

Skills model

Our model using metrics relating to skills training produced four distinct clusters. The metrics included were:

- apprenticeships achievements per 100,000 population (2020 to 21)
- apprenticeships starts per 100,000 population (2020 to 21)
- aged 16 to 64 years level 3 or above qualifications (2021)
- aged 19 years and over further education (FE) and skills participation per 100,000 population (2020 to 21)

Richmondshire was removed from our model because of a high number of army apprenticeships from Catterick Army Garrison skewing the results.

Above median level 3 or above qualifications, below median apprenticeships and further education (64 LAs)

At least one LA for all English regions, except the North West, were included in this cluster. Nearly all London LAs (88%) were included. Examples of LAs are Islington and York. Our results showed that this cluster:

- was better than the median for level 3 or above qualifications
- was worse than the median for apprenticeships achievements, apprenticeships starts, FE and skills participation
- consisted of mainly urban LAs (78%), had LAs with a median age of 40 years or younger (61%), had a higher proportion of the least deprived LAs (44%), and had areas with 2777 people or more per square kilometre (48%)

Slightly below median on all skills metrics (138 LAs)

There is at least one LA from all English regions included in this cluster. Over half of LAs in the East of England (76%), the East Midlands (57%), the West Midlands (57%) and the South East (52%) are in this cluster. Examples of LAs included are East Cambridgeshire, Chichester, and Nottingham. Our results showed that this cluster:

- was worse than the median for all metrics
- consisted of a notably high number of rural LAs (33%), had LAs with a median age of 45 years or older (39%) and had areas with 185 to 440 people per square kilometre (28%)

Above median apprenticeships and further education, below median level 3 or above qualifications (88 LAs)

There is at least one LA from all English regions included in this cluster. Most LAs are from the North East (83%), the North West (62%) and Yorkshire and The Humber (50%). Examples of LAs included are Salford and Teignbridge. Our results showed that this cluster:

- was better than the median for apprenticeships achievements, apprenticeships starts, FE and skills participation
- was worse than the median for level 3 or above qualifications
- consisted mainly of urban LAs (67%), had a higher proportion of the most deprived LAs (47%), had areas with 1074 to 2776 people per square kilometre (33%) and had a number of coastal towns (31%)

Above median on all skills metrics (16 LAs)

This cluster included LAs from the North West, the South East, the South West, the West Midlands, and Yorkshire and The Humber. Examples include Plymouth and Scarborough. Our results showed that this cluster:

- was better than the median for all metrics
- consisted mainly of rural LAs (60%), had LAs with a median age of 45 or older (63%), had a higher proportion of the least deprived LAs (25%), had areas with fewer than 185 people per square kilometre (63%) and had a number of coastal towns (56%)

Health model

When using metrics relating to the improvement of health, our model produced four distinct clusters. The metrics included were:

- male healthy life expectancy (2018 to 2020)
- female healthy life expectancy (2018 to 2020)
- cigarette smokers (April 2020 to December 2020)
- overweight adults (aged 18 years and over) (2019 to 20)
- overweight children at reception age (aged four to five years) (2019 to 20)
- overweight children at Year 6 age (aged 10 to 11 years) (2020 to 21)
- cardiovascular mortality considered preventable in persons aged under 75 (2017 to 2019)

Although the first two clusters presented were below median for all metrics, there was variation in the distance from the medians for each cluster. This has been reflected in the names for each cluster.

Far below median on all health metrics (52 LAs)

Nearly all LAs in the North East (83%) are included in this cluster. The highest proportion of LAs in the North West (43%) and Yorkshire and The Humber (38%) are included. Examples of these are Coventry and Doncaster. Our results showed that this cluster:

- was worse than median for all metrics
- consisted mainly of urban LAs (90%), had LAs with a median age of 40 years or younger (58%), had a higher proportion of the most deprived LAs (73%) and had areas with 2777 people or more per square kilometre (33%)

Broadly median on all health metrics (120 LAs)

Over half of LAs in the East Midlands (64%), the West Midlands (56%) and the North West (54%) are grouped in this cluster. The highest proportion of LAs in the East of England (40%) are also in this cluster. Examples of LAs included are Ipswich, Reading and Stockton-on-Tees. Our results showed that this cluster:

- was worse than the median for all metrics
- had LAs with a median age of 45 years or older (39%) and had a number of coastal towns (28%)

Lower overweight adults, higher cardiovascular mortality (19 LAs)

LAs from London and the South West are included in this cluster. Around 6 in 10 LAs in London (62%) are in this cluster. Examples of LAs included are City of Bristol and Lambeth. Our results showed that this cluster:

- was better than the median for male HLE, overweight adults, and overweight children aged 4 to 5 years
- was worse than the median for female HLE, overweight children aged 10 to 11 years, and cardiovascular mortality
- consisted of urban LAs (100%), had LAs with a median age of 40 years or younger (100%), had a higher proportion of the most deprived LAs (26%) and had areas with 2777 people or more per square kilometre (95%)

Above median on all health metrics (100 LAs)

Over half of LAs in the South West (61%), the South East (60%) and the East of England (53%) are included in this cluster. Examples of LAs included are Bromley and Dorset. Our results showed that this cluster:

- was better than the median for all metrics
- consisted of a notably high number of rural LAs (41%), had LAs with a median age of 45 years or older (40%), had a higher proportion of the least deprived LAs (46%) and had areas with fewer than 185 people per square kilometre (28%)

Well-being model

When using metrics relating to personal well-being, our model produced four distinct clusters. The metrics included were:

- life satisfaction (April 2021 to March 2022)
- anxiety (April 2021 to March 2022)
- happiness (April 2021 to March 2022)
- feeling life is worthwhile (April 2021 to March 2022)

There was variation in the relative distance from the medians for all clusters which has been reflected in the names for each cluster.

Several measures which impact on personal well-being have not been controlled for in this model including age, sex, economic activity, health status, relationship status, community well-being and mode effects. In any future analysis, we will look to control for these measures.

Far below median on all well-being metrics (74 LAs)

The majority of LAs from London (56%) are included in this cluster. All regions have at least one LA within this cluster. Examples of LAs in this cluster include Haringey and Blackpool. Our results showed that this cluster:

- was worse than the median for all metrics
- consisted mainly of urban LAs (88%), had LAs with a median age of 40 years and younger (62%) and had a higher proportion of the most deprived LAs (34%)

Better than median anxiety, worse than median on all other metrics (53 LAs)

All regions have at least one LA in this cluster. The South West and the North East had the largest proportion of their LAs in this cluster at 34% and 33%, respectively. Examples include Middlesbrough and Plymouth. Our results showed that this cluster:

- was better than the median for anxiety
- was worse than the median for life satisfaction, happiness and feeling life is worthwhile
- had LAs with a median age of 40 years and younger (51%), had a higher proportion of the most deprived LAs (27%) and had areas with 1074 to 2776 people per square kilometre (26%)

Slightly above median on all well-being metrics (103 LAs)

Nearly half of the East of England’s LAs are in this cluster (42%). Examples of LAs in this cluster include Cambridge and South Kesteven. Our results showed that this cluster:

- was better than the median for all metrics
- consisted mainly of urban LAs (50%) and had areas with 185 to 440 people per square kilometre (28%)

Far above median on all well-being metrics (77 LAs)

All regions had at least one LA in this cluster, except for the North East. Examples of LAs in this cluster include High Peak and Hambleton. Our results showed that this cluster:

- was better than the median for all metrics
- consisted of a notably high number of rural LAs (49%), had LAs with a median age of 45 years or older (53%), contained a number of coastal towns (30%) and had a higher proportion of the least deprived LAs (28%)

Conclusion

These analyses allow users to understand similarities between local authorities through the clusters they are placed in. To compare local authorities against each other, please see the subnational indicators explorer. For more information on your LA, please see our Find facts and figures about areas in England and Wales web page.

To provide any feedback on this analysis, please contact us at subnational@ons.gov.uk.

4 . Clustering local authorities against subnational indicators data

Clustering local authorities against subnational indicators, England
Dataset | Released 24 February 2023
Results of clustering analysis exploring similarities between local authorities based on subnational indicators. Experimental statistics.

5 . Glossary

Cluster Analysis

This publication uses k-means clustering as a method for identifying and grouping similar data points within a dataset together. This grouped the local authority district and unitary authorities into “k” clusters based on similarities between the geographical locations and variables. Further information on the k-means clustering method can be found in our accompanying methodology.

Gross value added (GVA)

Measures the contribution to the economy of each individual producer, industry or sector. It is an estimate of the value of the amount of goods and services that have been produced, minus the cost of all inputs and raw materials that are directly attributable to that production.

GVA can be used as a measure of how much a local economy produces over a given period.

Gross disposable household income (GDHI)

Gross disposable household income is the amount of money that all the individuals in the household sector have available for spending or saving after they have paid direct and indirect taxes and received any direct benefits. GDHI is a concept that is seen to reflect the "material welfare" of the household sector. The household sector includes residents of traditional households, as well as those living in communal establishments. GDHI also includes the business income of self-employed people.

6 . Data sources and quality

Subnational indicators explorer

All metrics included in the clustering analysis were sourced from the December 2022 update of our subnational indicators explorer. The explorer reports 40 metrics, metadata information for these metrics can be found in the data dictionary section of the explorer's dataset.

Although some metrics on the explorer were available for Scotland, Wales, and Northern Ireland, we have used England-only data in this publication to have the same consistent geographical coverage across our models. We are working with these devolved nations to improve coverage of our analysis and the data presented on our subnational indicators explorer.

Our analysis only included metrics from the explorer that reported figures at local authority district level. Metrics for UK exports, and inward and outward foreign direct investment (FDI), had to be excluded as these were reported at International Territorial Level (ITL). The metric for homicide had to similarly be excluded as results were reported at Police Force Area level. We are developing a method to include these metrics in our clustering model.

Quality

Quality and methodology information on the method, strengths, limitations, and how the data were created is available in our Clustering local authorities against subnational indicators methodology.

As the clustering method used in this release is still in the first stages of development, some results could be subject to further revisions.

Where necessary we have excluded metrics that were reported as counts (such as the number of learners achieving a funded further education and skills learning aim). This is because these metrics would not account for the population size of local authorities.

Because of older geography boundaries being used in several metrics as a result of local government restructuring , some local authorities could not be included in our analysis. Our interactive map in Section 2 only displays local authorities using 2021 boundaries. We are developing a method to try and improve our data coverage.

Metrics included in our models use the latest available figures. These can be from different dates (such as three metrics in a model could use 2022 figures but one could use 2020 figures) and are therefore contextual. For instance, coronavirus (COVID-19) measures may have had an impact on figures in 2020 and 2021.

7 . Future developments

We are developing methods to increase both the coverage of metrics and the geographic coverage within our models. We will also derive suitable metrics to include count data and incorporate more metrics.

Our analysis only covers England, but we are planning to work with devolved administrations and other stakeholders to include metrics with a UK-wide coverage.

There is scope for future analysis to explore clusters across different themes and build models using different model types.

Future analysis could look at running models at a greater geographic granularity (level of detail) to the local authority district level and incorporating the characteristics analysis undertaken here into the clustering models directly.

8 . Related links

Subnational indicators explorer
Interactive tool | Released 1 December 2022
Use our interactive tool to find out more about your local authority.

Methodology and variables
Article | Released 24 July 2018
Details of the methodology and 2011 Census statistics used for the 2011 area classifications.

Find facts and figures about areas in England and Wales
Get data about people and the communities they live in, includes population, identity, housing, people in or out of work, education and health.

Understanding Welsh Places
The Understanding Welsh Places website highlights to users the places in Wales that are most like their village, town, or community. It also helps users understand the relationships between their place and other nearby places.

NESTA mapping early years practice
A new method of clustering local authorities to help identify variations in practice that drive early years outcomes.

9 . Cite this article

Office for National Statistics (ONS), released 24 February 2023, ONS website, article, Clustering local authorities against subnational indicators, England