Local Digital Capital Index Methodological Note 2023

Constructing the index that measures local digital capital (LDC) – How do we move from the model to a unit-free index?

Many socioeconomic indexes, such as the Human Development Index, convert a raw input figure into a unit-free index (typically on a scale from 0-1 or 0-100), using the formula below.

x the raw figure in the data
a the "minimum" figure for that variable among the nations and regions
b the "maximum" figure for that variable among the nations and regions

$$x \text{ index} = \frac{x - a}{b - a}$$

But to ensure that the Local Digital Capital Index's within component scores are reflective of a nation's/region's true performance and not just their comparative performance, we decided to set *a=0.* As such, a nation or region will only score zero on any individual LDC measure if the data for that measure is indeed zero.

Where there is a natural maximum of 100%, such as for infrastructure variables like superfast broadband and gigabit coverage we set *b=1*. This allows us to benchmark the nations/regions against one another and against an achievable long-term benchmark of full coverage. For other variables that are measured in percentage terms we have taken a more flexible approach.

Where the maximum benchmark was less clear and achieving 100% on a measure is an implausible target, such as percentage employment in tech jobs, we set *b* equal to the actual maximum data point for the nations/regions.

For raw data, such as the value of R&D spending (scaled by labour hours) or SME bank lending (scaled by labour hours), we set *b* to the raw maximum of the nations and regions. The individual *within* component index score for a nation/region is therefore essentially a percentage compared to the leading nation/region.

Variables that needed to be scaled in per capita or per hours worked terms?

The economic variable of interest used in our econometric model was GVA per hour worked, which is essentially a measure of labour productivity and very highly correlated to GDP per capita (and therefore average living standards) since labour hours per capita vary little between nations/regions. Therefore, the data added to the model needed to be scaled in per capita or in per hours worked in order for the index to make economic sense.

Data that is as a percentage of the population should strictly be rescaled in per hours terms (rather than per capita). However, since the model is set up in "log-log functional form", the economic effects of a percentage change in each independent variable (such as percentage of the properties that have superfast broadband coverage) on labour productivity are unaffected by the multiplicative scaling difference between per capita and per labour hours worked. As such, percentage data is not scaled.

'Between component' weightings

The econometric model helped us understand which components of LDC had a greater positive impact on labour productivity. Consistent with previous empirical work in other similarly developed countries, we found a larger economic impact of digital skills and a slightly lower impact for the other components and therefore weighted skills highly (30% of the overall index). This also reflects the findings of the European Commission's Digital Economy and Society Index (DESI), which weights 'human capital' (including multiple measures of digital skills) joint highest of their components.

The variables 'research and innovation', 'finance and investment' and 'infrastructure' are weighted at 20% each, and 'adoption' is weighted at 10%. The 'trade support' component is currently weighted at 0% and does not contribute towards the overall Index results. This is because 'trade support' as conceptualised by techUK's original Local Digital Capital Report cannot be easily quantified. What we are measuring instead are raw export figures for both NUTS1 and NUTS2 areas. While this does tell us how successful a given area is in terms of exports, we cannot infer that this translates to the amount of 'trade support' available to businesses (tech businesses in particular), in a given area. As such, the 'trade support' results are presented in isolation, and do not contribute to the overall Index results. This may change in subsequent years if new data becomes available.

In 2023 we introduced a new component 'data ecosystems'. This looks to measure the existence and availability of open, accessible and interoperable data. For this component, we took the same approach to weighting as we did for 'trade support.' This is because measuring the regional availability of open, accessible and interoperable data cannot easily be done and there currently is no well-tested methodology. We have used several proxies, such as the number of employees with data-focused jobs and measures from the DCMS Business Data Survey, such as the percentage of businesses that share data with public bodies and charities and non-profit organisations. This tells us how many people might be working in data in each region, which will be linked to the amount of data available and data skills, and how much data-sharing is taking place. However, we cannot infer that this directly measures 'data ecosystems' as defined by techUK. Therefore, the 'data ecosystems' results are presented in isolation and do not contribute to overall results. As with trade support, this may change in subsequent years if new data and methodologies become available.

'Within component' weightings

A higher weighting was generally given to the LDC measure within each component that produced the most precise (i.e. lower standard deviation and result robust to different exact specification of the model) results, as we have greater confidence that variation in measure of LDC is driving differences in labour productivity between the nations/regions. Where various within component measures produced similar model results, we took a balanced within component weighting (such as 50-50, or 1/3-1/3-1/3).

The next challenge was to translate NUTS1 within component weightings to the NUTS2 level. For the components 'infrastructure', 'trade support' and 'research and innovation', all the variables used to generate results at the NUTS1 level were also reported at the NUTS2 level, thus the within component weightings remained the same.

However, for 'adoption', 'skills' and 'finance and investment', fewer data inputs were reported at the NUTS2 level. This means that 'adoption' for example, is measured using six different variables at the NUTS1 level, but only four of these are used at the NUTS2 level. For these three components, instead of creating new weightings at the NUTS2 level, weightings were made proportional to the regional analysis for the variables which carried over.

For the new 2023 variable introduction in the 'skills' component, we still placed a greater emphasis on 'internet users' as the best measure and shared the remaining 0.3 weighting between the other six variables. We took a similar approach to the addition 5G coverage into the 'infrastructure' component, continuing to place a larger emphasis on mobile coverage compared to broadband.

As new data becomes available, the ambition for the Index will be to utilise the same variables across both the NUTS1 and NUTS2 levels across all components.

Local Digital Capital Results: 2022 v 2023

This year's Local Digital Index largely reflects the work of last year. The vast majority of the variables used last year were still available and were often updated. There was only one case of a data set being discontinued and a replacement was easily found.

The first six components: skills, digital adoption, digital infrastructure, finance and investment, research and development and trade support, remained largely unchanged. The substantial change in this year's Index was the introduction of 'data ecosystems', though this has a limited impact on the overall Index result as the component is discounted from the Index weighting.

As outlined last year, the purpose of an Index is to track changes year-on-year, which necessitates a consistency in approach and inputs used; to utilise the same data sets and same methodology. This must be balanced with the wish to ensure the Index is composed of the most up to date and genuinely reflective data. For this iteration, most components have remained largely unchanged. Components have been strengthened where possible, such as a new measure in the 'data skills' component, and we have looked for updated data and for NUTS2 to be added where possible.

Given the consistency in approach and data inputs from 2022 to 2023, it is possible to draw comparisons between the rankings seen in 2022 to those seen in 2023. While it is necessary exercise some caution in drawing conclusions about whether the Index indicates an improvement or decline in performance in local digital capital over a, relatively, short period of time, the repeated data sets mean that comparison is more possible than 2022 to 2021.

The overall position of London, the South East, East Midlands, East of England, North West and West Midlands remained the same as in 2022. Some areas have changed, including the North East moving up slightly from 9th to 10th position, Wales from 12th to 11th and Scotland from 6th to 4th, while Northern Ireland moved down from 11th to 12th position, South West from 4th to 6th and Yorkshire and Humber down from 9th to 10th.

Clearly, the 2023 results are broadly consistent with the results from 2022. This is reflected at the NUTS1 and NUTS2 level with small changes accounted for by marginal changes in data.

Component breakdown

Digital skills

Public Services Use: This data is from the Ofcom Technology Tracker and measures the percentage of the population who access local or national government services online, e.g., to find information, complete processes such as tax returns or to contact a local MP. This is benchmarked against a maximum figure of 100%. This is measured at the NUTS1 level only.

Internet Users: This data is the ONS Internet Users data, and measures the percentage of the population who have used the internet in the previous three months (most frequent measure). This is benchmarked against a maximum figure of 100%. This measure is used across both the NUTS1 and NUTS2 level.

Higher Education Digital Study: This data is from the Higher Education Statistics Agency (HESA) and measures the number students in a region studying computing, and engineering and technology, per 1,000 population. The figure is measured by region of educational provider and benchmarked against the maximum score achieved by an individual region. This is measured at the NUTS1 level only.

NVQ4+: This data is from the Annual Population Survey and measures the number of people of working age (between 16 and 64), with at NVQ4 qualifications or higher. This is benchmarked against the maximum score achieved by any individual nation or region. This is measured across both the NUTS1 and NUTS2 levels.

Downloading Information: This data is from the Ofcom Technology Tracker and measures the percentage of people who use the internet to download information for study or work. This is benchmarked against a maximum figure of 100% and measured at the NUTS1 level only.

Higher Education Number of Students: This data is from the Higher Education Statistics Agency (HESA) and it measures the total number of higher education students (all subjects and levels) by student domicile (i.e., the total number of higher education students who come from a given area). This is benchmarked against the maximum score achieved by an individual region and measured across the NUTS1 and NUTS2 levels.

Business Data Skills: This data is from DCMS Business Data Survey, and it measures the percentage of businesses that 'strongly agree' or 'tend to agree' that their business has sufficient data skills to meet its needs. This is benchmarked against the maximum score achieved by any individual nation or region and is measured at the NUTS1 level.

Digital adoption

Jobs in ICT: This data is ONS Workforce Jobs by Region data and measures the percentage of jobs in 'information and communication' (SIC2007 sector J). This is benchmarked against the maximum score reached by an individual region and measured at the NUTS1 level only.

UK Business Count: This data is from the UK Business Count and it measures the count of businesses in the in the 'digital sector' according to the DCMS 4-digit SIC2007 definition. This is benchmarked against the maximum score achieved by an individual region and measured across both the NUTS1 and NUTS2 levels.

Digital Employment Share: This data is from the Business Register and Employment Survey and it measures the total number of jobs accounted for by the 'digital sector' according to the DCMS 4-digit

SIC2007 definition. This is benchmarked against the maximum score achieved by an individual region and measured across both the NUTS1 and NUTS2 levels.

Digital Occupations: This data is from the Annual Population Survey and it measures the total number of jobs accounted for by digital and associated occupations. This is benchmarked across the maximum score achieved by an individual region and is measured across both the NUTS1 and NUTS2 level. The definition includes two SOC2010 sub-major occupation groups:

- Science, Research, Technology and Engineering Professionals
- Science, Research, Technology and Engineering Associated Professionals

Investment in ICT: This data is ONS Experimental Regional Gross Fixed Capital Formation (GFCF) by asset type and industry. This measures annual GFCF in information and communication assets across all industries, in GBP millions. This is benchmarked across the maximum score achieved by an individual region and is measured across both the NUTS1 and NUTS2 level.

Digital Infrastructure

Superfast Broadband: This data is from the Ofcom Connected Nations reports, and it measures the percentage of premises with superfast (30Mb-300Mb) broadband coverage (the available coverage, not the take-up). This data is benchmarked against a maximum figure of 100% coverage and measured across both the NUTS1 and NUTS2 level.

Ultrafast Broadband: This data is from the Ofcom Connected Nations reports, and it measures the percentage of premises with ultrafast (300Mb-1Gb) broadband coverage (the available coverage, not the take-up). This data is benchmarked against a maximum figure of 100% coverage and measured across both the NUTS1 and NUTS2 level.

Gigabit: This data is from the Ofcom Connected Nations reports, and it measures the percentage of premises with gigabit broadband coverage (the available coverage, not the take-up). This data is benchmarked against a maximum figure of 100% coverage and is measured across both the NUTS1 and NUTS2 level.

Mobile Coverage 4G: This data is from the Ofcom Connected Nations reports, and it measures the percentage of premises with indoor 4G services from four or more providers. This is benchmarked against a maximum figure of 100% coverage and measured across both the NUTS1 and NUTS2 levels.

Mobile Coverage 5G: This data is from the Ofcom Connected Nations, Spring Update. It measures the percentage of premises where there is 5G coverage from at least one provider. This is benchmarked against a maximum figure of 100% coverage and measures across both the NUTS1 and NUTS2 levels.

Research and Innovation

R&D Spend: This is ONS data on regional R&D spend and it measures the total amount of R&D spending in a region by all sectors (Government, Higher Education, businesses and non-profits) in GBP millions. This is benchmarked against the maximum figure achieved by any individual nation or region and measured across both the NUTS1 and NUTS2 level.

Innovate UK Grants: This is Innovate UK data on the value of Innovate UK Grants awarded in the AI and Data Economy sector, measured in GBP. This is benchmarked against the maximum figure achieved by any nation or region and measured across both the NUTS1 and NUTS2 level.

HMRC R&D Tax Credits: This is HMRC data on the total cost of R&D tax credits awarded to businesses located in a given area, in GBP millions. This is benchmarked against the maximum figure achieved by any nation or region and measured across both the NUTS1 and NUTS2 level.

Finance and Investment

Equity Finance: This is data is from the British Business Bank Equity Tracker and it measures the percentage of equity finance value invested into each nation or region. This is benchmarked against the highest percentage achieved in any nation or region and measured at the NUTS1 level only.

VC in Tech: This data is published by Dealroom and measures the total value of venture capital investment in tech companies in a region, in USD millions. This is benchmarked against the highest figure achieved in any nation or region and measures at the NUTS1 level only.

SME Lending: This data is published by UK Finance and it measures the total value of bank lending (loans and overdrafts) to SMEs in a region. We have divided this number by the count of SMEs in each region, in order to see the approximate value of lending by SME in GBP. This is benchmarked against the highest figure achieved in any nation or region and measured at the NUTS1 level only.

High Growth: This is ONS Business Demography data, and it measures the number of high growth businesses in a given area. This is benchmarked against the highest figure achieved in any nation or region and measured across both the NUTS1 and NUTS2 level.

Investment in ICT Sector: This is ONS experimental Regional Gross Fixed Capital Formation (GFCF) by asset type and industry data. This measures the amount of investment in all assets in the 'Information and Communication' sector in GBP millions. This is benchmarked across the highest figure achieved in any nation or region and it is measured at both the NUTS1 and NUTS2 level.

Inward Investment in ICT: This is ONS Foreign Direct Investment (FDI) data and it measures the value of inward FDI in the information and communication sector in GBP millions. This is benchmarked against the highest figured achieved in any nation or region and measured at both the NUTS1 and NUTS2 level.

Trade support

Goods exports: This is HMRC Regional Trade in Goods data, and it measures the total value of exports in goods in nations and regions on an annual and quarterly basis in GBP millions. This is benchmarked against the highest figure achieved in any nation or region and measured across both the NUTS1 and NUTS2 level. The data for the NUTS2 level is found here.

Services Exports: This is experimental ONS data, and it measures the value of service exports by subnational areas of the UK in GBP millions. This is benchmarked against the highest figure achieved in any nation or region and measured across both the NUTS1 and NUTS2 level.

Data Ecosystems

Data Acquisition: This is taken from the DCMS Business Data Survey and measures the percentage of businesses that responds 'yes' to 'acquires or collects any data.' This is benchmarked against the highest figure achieved in any nation or region and is measured across the NUTS1 level only.

Data Sharing: This is taken from the DCMS Business Data Survey and measures the percentage of businesses that responds 'yes' to 'shares any data.' This is benchmarked against the highest figure achieved in any nation or region and is measured across the NUTS1 level only.

Data Recipients: This is taken from the DCMS Business Data Survey and measures the percentage of businesses that receive personal data from 'public bodies' and 'charities or non-profit organisations.' This is benchmarked against the highest figure achieved in any nation or region and is measured across the NUTS1 level only. We used the 2021 Survey for this measure due to lack of regional data in 2022.

Data Availability: This is taken from the DCMS Business Data Survey and measures the percentage of businesses that responded 'a great deal' or 'a fair amount' to the belief that data from outside their business has become more readily available in the last ten years. This is benchmarked against the highest figure achieved in any nation or region and is measured across the NUTS1 level only. We used the 2021 Survey for this measure due to lack of regional data in 2022.

Data Roles: This data is taken from NOMIS Annual Population Survey, regional occupation by employment type, and measures the number of people in the roles: 'database administrators and web content technicians', 'data analysts' and 'data entry administrators.' This is benchmarked against the highest figure achieved in any nation or region and is measured across the NUTS1 level only.