

Digital & Health Literacy Index

supporting document

Last updated August 2026

Contents

1. Overview	3
2. Indicator 1: No qualifications	5
3. Indicator 2: Disability or long-term health conditions	6
4. Indicator 3: English proficiency	7
5. Indicator 4: Age (65+).....	8
6. Indicator 5: Lone living	9
7. Indicator 6: Missed appointments	10
8. Indicator 7: Overall IMD-25	11
9. Indicator 8: Digital Propensity Index (DPI)	12
10. Indicator 9: Broadband	13
11. Indicator 10: Health and fitness app usage	14

1. Overview

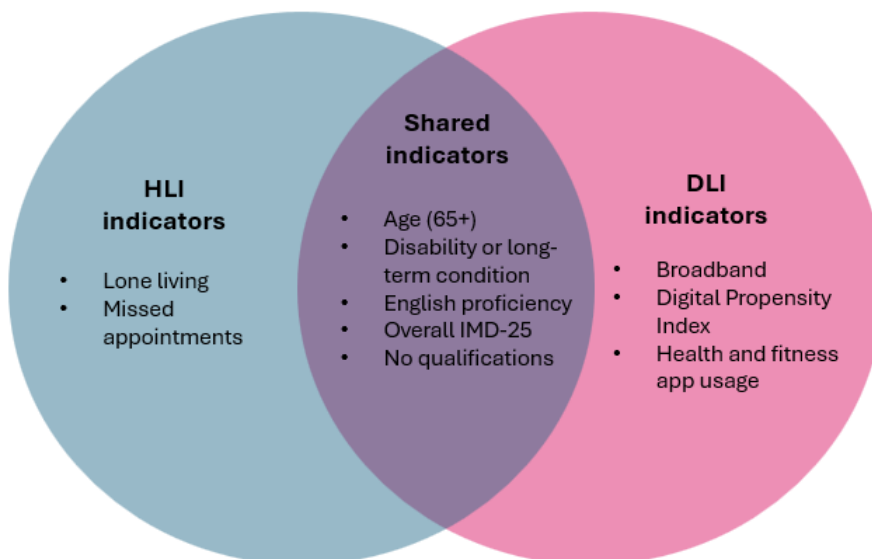
This document accompanies the Digital and Health Literacy Index (DHLI) Power BI dashboard.

There is no single, routinely available measure of digital literacy or health literacy at Lower Super Output Area (LSOA) level. As a result, the DHLI has been developed using a range of proxy indicators that are associated with digital and health literacy. The selection of these indicators was informed by evidence reviews conducted for both digital literacy and health literacy.

The index is intended to provide an indication of relative need across Southampton and should not be interpreted as a direct measure of either digital literacy or health literacy. Each indicator represents only one aspect of a complex set of behaviours, skills and circumstances. For example, missed healthcare appointments may be influenced by factors such as transport availability, caring responsibilities, work commitments or individual confidence, rather than literacy alone. Similarly, factors such as personality, motivation and social support can influence an individual's ability to access, understand and use health information but cannot be captured through routinely available datasets.

The DHLI dashboard also includes separate views of the Health Literacy Index (HLI) and Digital Literacy Index (DLI) for cases where there is a need to explore them separately. They both share some indicators. Please refer to figure 1. to see which indicators are shared.

Figure 1. Health Literacy Index (HLI) and Digital Literacy Index (DLI) indicator overlap



The Digital and Health Literacy Index (DHLI) combines all ten indicators shown above.

This document provides supporting information for each of the ten indicators included within the index, including:

- Indicator definition
- Data source

- Summary methodology
- Collinearity assessments, where applicable
- Indicator polarity (whether higher values indicate higher or lower need)
- Relevant caveats and limitations

Where indicators measure similar concepts, collinearity testing has been undertaken to assess potential overlap and support decisions on indicator inclusion.

As with any index based on proxy measures, results should be interpreted alongside local knowledge and other sources of evidence. The DHLI should therefore be considered a best available estimate of digital and health literacy need based on the data currently available.

2. Indicator 1: No qualifications

Definition: Percentage of residents aged 16 and over with no formal qualifications at LSOA level

Data source: Office of National Statistics, Census 2021 (LSOA 2021 geography)

Methodology:

1. Census 2021 data were obtained from NOMIS for residents aged 16 and over whose highest qualification level was recorded as "No qualifications".
2. Values were standardised using z-scores to enable comparison with other indicators included within the index.
3. Collinearity was assessed against the Index of Multiple Deprivation (IMD) 2025 to identify potential overlap between measures.

Collinearity:

Overall IMD 25 is included as an indicator in the DHLI. As it contains an education domain, collinearity was assessed between the Census 'no qualifications' data and the overall IMD score to evaluate overlap.

A strong correlation was identified ($R = 0.73$), with an R^2 of 0.53. This indicates that while the measures are related, they do not capture exactly the same underlying characteristics (almost half of the variance were attributed to unrelated factors).

As the measures provide both shared and unique information, the indicator was retained within the index. This decision is further supported by the evidence review, which identified educational attainment as an important factor associated with health literacy.

Indicator polarity: This indicator was treated as a risk, where higher scores indicate higher need.

Caveats:

1. Census data represents a snapshot in time. The 2021 Census may not fully reflect typical population characteristics due to the impact of COVID-19. This should be considered when interpreting results
2. To protect confidentiality, the Census includes statistical disclosure control methods such as record swapping and small adjustments to counts. These techniques may have a greater impact on estimates for smaller geographic areas.

3. Indicator 2: Disability or long-term health conditions

Definition: Proportion of residents who self-report as having a disability or long-term physical or mental health condition.

Data source: Office of National Statistics, Census 2021 (LSOA 2021 geography)

Methodology:

1. Census 2021 data were obtained from NOMIS using the following categories:
 - Disabled under the Equality Act: Day-to-day activities limited a lot
 - Disabled under the Equality Act: Day-to-day activities limited a little
 - Not disabled under the Equality Act: Has long term physical or mental health condition but day-to-day activities are not limited
2. The three categories were combined to create a single measure of residents reporting a disability or long-term health condition.
3. Values were standardised using z-scores to enable comparison with other indicators included within the index.

Indicator polarity: This indicator was treated as a risk, where higher scores indicate higher levels of need.

Caveats:

1. This indicator is based on self-reported information and may be influenced by how individuals perceive and report their health, disability status, or limitations to daily activities. Responses may vary between individuals, potentially leading to some under-reporting or over-reporting.
2. The Census was conducted during the COVID-19 pandemic, which may have influenced how some residents reported their health and day-to-day activities.
3. To protect confidentiality, the Census includes statistical disclosure control methods such as record swapping and small adjustments to counts. These techniques may have a greater impact on estimates for smaller geographic areas.

4. Indicator 3: English proficiency

Definition: Proportion of residents who report that they cannot speak English well or at all.

Data source: Office of National Statistics, Census 2021 (LSOA 2021 geography)

Methodology:

1. Census 2021 data were obtained from NOMIS using the following categories:
 - Main language is not English (English or Welsh in Wales): Cannot speak English well
 - Main language is not English (English or Welsh in Wales): Cannot speak English
2. The percentages for both categories were combined to create a single measure of limited English proficiency within each LSOA.
3. Values were standardised using z-scores to enable comparison with other indicators included within the index.

Indicator polarity: Low English proficiency was treated as a risk factor. Higher scores indicate higher levels of need.

Caveats:

1. This indicator is based on self-reported Census data and may be affected by differences in how individuals assess and report their English language proficiency.
2. Census data represent a snapshot in time and may not reflect changes in population composition or language proficiency since 2021.
3. To protect confidentiality, the Census includes statistical disclosure control methods such as record swapping and small adjustments to counts. These techniques may have a greater impact on estimates for smaller geographic areas.

5. Indicator 4: Age (65+)

Definition: Proportion of residents aged 65 and over.

Data source: Office for National Statistics Mid-year population estimates (LSOA 2021 geography)

Methodology:

1. Mid-year population estimates were obtained from NOMIS for residents aged 65 years and over.
2. The percentage of residents aged 65 years and over was calculated for each LSOA.
3. Values were standardised using z-scores to enable comparison with other indicators included within the index.

Indicator polarity: This indicator was treated as a risk factor, where higher scores indicate higher level of need.

Caveats: Mid-year population estimates are modelled data rather than direct population counts. As such, it is possible that there may be some inaccuracy, particularly at smaller geographic levels such as LSOAs.

6. Indicator 5: Lone living

Definition: Proportion of self-reported one-person households.

Data source: Office of National Statistics, Census 2021 (LSOA 2021 geography)

Methodology:

1. Census 2021 data were obtained from NOMIS using the following categories:
 - One-person household: Aged 66 years and over
 - One-person household: Other
2. The percentages for both categories were combined to create a single measure of lone living within each LSOA.
3. Values were standardised using z-scores to enable comparison with other indicators included within the index.

Indicator polarity: This indicator was treated as a risk factor, where higher scores indicate higher need.

Caveats:

1. Living alone does not necessarily indicate social isolation or a lack of support. Individuals living in one-person households may still have strong social networks and access to support.
2. Census data represent a snapshot in time (March 2021) and do not capture changes in household composition or living arrangements over time.
3. Census responses are self-reported and may be affected by how individuals interpret their household circumstances, particularly in cases of temporary or shared living arrangements.
4. To protect confidentiality, the Census includes statistical disclosure control methods such as record swapping and small adjustments to counts. These techniques may have a greater impact on estimates for smaller geographic areas.

7. Indicator 6: Missed appointments

Definition: Percentage of missed GP appointments per registered patient population (weighted to LSOA level based on GP practice catchment).

Data source: NHS England Digital

Methodology:

1. GP appointment data and registered patient data were linked using a geographic lookup to enable analysis at LSOA level.
2. Data were cleaned prior to analysis:
 - Patients registered to GP practices outside Southampton LSOAs were excluded.
 - Appointments with an unknown attendance status were excluded.
3. GP practice-level missed appointment rates were calculated and weighted using the distribution of registered patients across LSOAs to estimate a missed appointment rate for each LSOA.
4. Values were standardised using z-scores to enable comparison with other indicators included within the index.

Indicator polarity: This indicator was treated as a risk factor, where higher scores indicate higher level of need.

Caveats:

1. Appointment attendance information is not available directly at LSOA level. Rates are estimated using GP practice attendance data and patient registration patterns and should therefore be considered an approximation rather than a direct measure.
2. This indicator is based on registered patient populations rather than residents. As a result, some individuals registered to Southampton LSOAs may not reside within the area, and some residents may be registered with practices outside Southampton.
3. Missed appointments can occur for a range of reasons, including access barriers, transport issues, work commitments, caring responsibilities, forgetfulness, or ill health. Therefore, higher rates should not be interpreted as a direct measure of individual behaviour.

8. Indicator 7: Overall IMD-25

Definition: Overall level of deprivation per LSOA, based on the Index of Multiple Deprivation (IMD) 2025.

Data source: The Ministry of Housing, Communities and Local Government (MHCLG)

Methodology:

1. Data was taken from 'IMD2025 All Domains Quintiles and Deciles' dataset.
2. The overall IMD-25 score was extracted for each of the relevant Southampton LSOAs.
3. Values were standardised using z-scores to enable comparison with other indicators included within the index.

Indicator polarity: Higher levels of deprivation were treated as a risk factor, where higher scores indicate higher level of need.

Caveats:

1. IMD 2025 combines information from multiple deprivation domains. Where overlap with other indicators was identified, collinearity was assessed to support decisions on indicator inclusion.
2. The IMD 2025 dashboard map and the deprivation map in the DHLI dashboard are not directly comparable. While both use MHCLG LSOA-level data, deciles within the DHLI were calculated using Southampton LSOAs only, whereas IMD deciles are based on the national distribution of LSOAs across England.

9. Indicator 8: Digital Propensity Index (DPI)

Definition: Estimated likelihood of households engaging with online services, based on observed and modelled online Census 2021 response behaviour, reported at LSOA level.

Data sources: Office for National Statistics (ONS) Digital Propensity Index for England and Wales LSOAs

Methodology:

1. As the dataset uses 2011 LSOA codes, a best-fit lookup was applied to map scores to 2021 LSOAs.
2. Where 2011 LSOAs were split for 2021 LSOAs, the same DPI score was applied to each associated 2021 LSOA.
3. Values were standardised using z-scores to enable comparison with other indicators included within the index.

Collinearity:

Collinearity was assessed between the Digital Propensity Index (DPI) and Health and Fitness App Usage, as both indicators relate to digital engagement and may capture similar behaviours.

A moderate correlation ($r=0.66$) was observed, with an R^2 value of 0.44. This suggests that approximately 56% of variance is attributable to other factors. This indicates that, while related, they do not capture the same underlying characteristics.

The indicators are considered to reflect different aspects of digital engagement, providing complementary information supporting the decision to retain both indicators in the DHLI.

Indicator polarity: This indicator was treated as protective for digital literacy. A lower score indicates higher level of need/risk of digital exclusion.

Caveat:

1. The conversion from 2011 to 2021 LSOAs assumes that there is even distribution in those areas.
2. The DPI is derived from Census 2021 data and reflects population characteristics at that point in time. Demographic and digital behaviours may have changed since the data were collected.
3. The Digital Propensity Index is a modelled measure and should be considered an estimate rather than a direct measure of digital engagement.

10. Indicator 9: Broadband

Definition: Proportion of premises with access to superfast broadband (defined as speeds of 30 Mbps or above), based on May 2022 data, aggregated to LSOA level using population weighting.

Data source: Geographic Data Service (derived from Ofcom broadband availability data).

Methodology:

1. The most recent available data was exported (May 2022) from the multi-year dataset. The indicator represents the percentage of premises with access to superfast broadband (≥ 30 Mbps) at Output Area (OA) level.
2. Broadband availability data were aggregated from 2011 OA's to 2021 LSOAs to ensure consistency with other DHLI indicators. To account for variation in population distribution, a population-weighted approach was applied whereby OA level broadband availability values were weighted by the OA population prior to LSOA aggregation.
3. Values were standardised using z-scores to enable comparison with other indicators included within the index.
4. Unlike other indicators, a modified decile approach was applied due to a large number of LSOAs sharing the highest broadband availability score. All LSOAs with the maximum observed value were assigned to Decile 10, while the remaining LSOAs were distributed across Deciles 1–9. This ensured that areas with identical broadband availability received the same classification while retaining a complete 1–10 decile scale. This approach was necessary because a standard rank-based decile calculation did not produce a full 1–10 decile range due to the high number of tied values.

Indicator polarity: This indicator was treated as protective for digital literacy. A lower score indicates higher level need/risk of digital exclusion.

Caveat:

1. Some Southampton OAs were missing from the source dataset. This may be related to the supplier's use of best-fit geographic allocation between OA boundaries, which could affect the accuracy of some estimates.
2. Due to missing OA data, no broadband value was available for LSOA E01035446. This LSOA was assigned the same value as E01035447, as both areas were previously represented within the same 2011 geography.
3. Broadband availability reflects the presence of infrastructure rather than actual broadband uptake or usage. Therefore, it should be considered a proxy measure of digital access rather than digital engagement.
4. The modified decile approach used for this indicator differs from that applied to other indicators in the index. This was necessary to ensure that areas with identical broadband availability scores were assigned the same decile classification.

11. Indicator 10: Health and fitness app usage

Definition: Estimated percentage of residents who use health and fitness apps, based on Mosaic modelled propensities and aggregated to LSOA level.

Data sources:

- Experian Mosaic UK 8 (persona-level propensity for health and fitness app usage).
- MosaicHouseholdExtract_3 (LSOA-level distribution of Mosaic personas)
- NOMIS population estimates

Methodology:

1. Health and fitness app usage data were derived from Experian Mosaic UK 8 propensity data. The selected measure reflects the likelihood of different Mosaic persona types using mobile applications for health and fitness purposes.
2. An estimated level of health and fitness app usage was calculated for each LSOA:
 - a. The distribution of Mosaic persona types within each LSOA was calculated.
 - b. Population weighting was applied to estimate the percentage of residents likely to use health and fitness apps within each LSOA.
3. Collinearity was assessed against the Digital Propensity Index (DPI) to identify potential overlap between measures.
4. Values were standardised using z-scores to enable comparison with other indicators included within the index.

Collinearity: See the collinearity assessment reported within the DPI indicator section.

Indicator polarity: This indicator was treated as a protective for digital literacy. A lower score indicates higher level of need/risk of digital exclusion.

Caveats

1. This indicator represents a modelled estimate of health and fitness app usage rather than actual observed usage at LSOA level.
2. Estimates are based on Experian Mosaic segmentation assumptions and may not fully reflect the behaviour of all residents within an area.
3. Population weighting was based on residents aged 18 years and over, reflecting the age range modelled within the Mosaic dataset.