Codebook for `Regional Human Capital Database`

Jane Gingrich, Dan McArthur, with Mihnea Cuibus

December 18, 2025

1 Dataset Overview

- **Dataset name:** Regional Human Capital Database (RHCD), Regional Elections Database (RED), Regional Migration Database (RMD)
- **Source:** Data sources described in country notes
- **Unit of observation:** LAU, or other statistical units, 2018 boundaries (V1) and 2021/2024 (V2). RMD only available for V1.
- **Time period:** 1980s–2020s
- **Number of observations:** 416,712 (RHCD), 737,309 (RED)
- **Number of variables:** Varies by country

2 Variable Definitions

Table 1: Variable Codebook

Variable	Label / Description	Type	Notes	Dataset
Ccode	Unique country identifier	ISO country code		All
Year	Year	Census/Election year of observation	See notes; range varies by country	All
Decade	Decade of observation	Census/Election decade of observation		All
Version	2018 or 2021/2024 LAU or equivalent		V2 updates boundaries for AT, CH, ES, FR, IT, DE. See notes on Norway	All

Continued on next page

Variable	Label / Description	Type	Notes	Dataset
LAU	Local identifier	See country notes; generally local area unit	No missing values	All but codes vary for Germany, NZ, Portugal across datasets (see note)
Launame	Local identifier name	As above	No missing values	As above
Laumatch	Match code for RHCD-RED datasets	As above	As above	All
NUTS3CODE	Regional identifier	NUTS3	Only Europe	RHCD
Macro-region	Regional identifier	Higher spatial units for AUS, CAN, USA, JPN, NZ	Excludes Europe	RHCD
TL3	Regional identifier	NUTS3 (Europe), OECD-TL3 for non-European	Missing NZ	RED
Fuacode	Functional Urban Area Code	OECD - 2011 boundaries	Missing NZ; areas not in FUA	All
Fuaname	Functional Urban Area Name	As above	As above	All
Urban_area_type	Ordinal categorization of urban areas	See notes	Missing NZ; partial for Japan	All
Area_km_sq	Size of LAU in km^2	LAU or equivalent		All
Distance_major	Distance to the center of a major FUA in km	Only available for areas in a major FUA	Missing for Ireland and NZ	RED
Distance_*city	Distance to specific cities in km	Varies by country		RED
East	Dummy for former GDR (Germany)			RED
Population (gender_age)	Total population	Subset by gender (m/f), age bins	Age breakdown unavailable for some units (see notes)	RHCD

Continued on next page

Variable	Label / Description	Type	Notes	Dataset
Population_adj (gender_age)	Total population minus missing observations	Denominator used to calculate adjusted shares	Australia, Ireland, NZ, Sweden, Switzerland	RHCD
Ed_isced_0to2 (gender_age)	Total number with ISCED 0–2	Subset by gender (m/f), age	Age breakdown unavailable for some units; missing in UK (earlier years), Germany	RHCD
Ed_isced_3to5 (gender_age)	Total number with ISCED 3–5	Subset by gender (m/f), age	Age breakdown unavailable for some units; missing in UK (earlier years), Germany	RHCD
Ed_isced_6to8 (gender_age)	Total number with ISCED 6–8	Subset by gender (m/f), age	Age breakdown unavailable for some units	RHCD
Ed_isced_adj (gender_age)	Total number with ISCED by group (0–2, 3–5, 6–8) with missing removed	Subset by gender (m/f), age	Age breakdown unavailable for some units	RHCD
Ed_isced_pct (gender_age)	Share of population 15+ with ISCED by group (0–2, 3–5, 6–8)	Subset by gender (m/f), age	Age breakdown unavailable for some units; 0–2 and 3–5 missing for some years in UK and Germany	RHCD
Ed_isced_pct_adj (gender_age)	Share of population 15+ with ISCED by group (0–2, 3–5, 6–8) with missing removed	Subset by gender (m/f), age	Only some units	RHCD
Total_votes	Total votes cast	Votes		RED
Total_voters	Total voters in unit	Voters	Missing Aus, NZ, USA	RED
To1	Turnout	Share	Missing Aus, NZ, USA	RED
Weight	Weight variable	Share of national votes in area		RED
party001–party0096	Votes for party using codebook party classification	Total local votes		RED
share_party001–party0096	Share for party using codebook party classification	Share of local votes		RED
pf_*	Votes for party using Party Facts ID	Total local votes		RED
vote_(family)	Votes by party family	Total local votes	See codebook	RED

Continued on next page

Variable	Label / Description	Type	Notes	Dataset
pct_(family)	Share by party family	Share for party family	See codebook	RED
den_pctile	Density decile	0–9 deciles in national election-specific density distribution		RED
ed_pctile	Education decile	0–9 deciles in national ISCED 6–8 (university) education distribution from RHCD		RED
Flag	Flag for units that are not fully matched across time series		Only for those with missing data	RED
population_foreign_born	Total population foreign born		See country notes	RMD
population_foreign_born_pc	Share of population foreign born		See country notes	RMD
population_noncitizens	Total population non-citizen		See country notes	RMD
population_noncitizens_pc	Share of population non-citizen		See country notes	RMD
population_foreign_combined	Foreign born/non-citizen combined		See country notes	RMD
population_foreign_combined_pc	Share foreign born/non-citizen combined		See country notes	RMD

3 Value labels and Notes

CCODE

- 36 = Australia
- 40 = Austria
- 56 = Belgium
- 124 = Canada
- 208 = Denmark
- 246 = Finland
- 250 = France
- 276 = Germany
- 300 = Greece
- 372 = Ireland
- 380 = Italy
- 392 = Japan
- 528 = Netherlands
- 554 = New Zealand
- 578 = Norway
- 620 = Portugal

- 724 = Spain
- 752 = Sweden
- 756 = Switzerland
- 826 = UK
- 840 = USA

Year

- Year of observation. Note: election years and census years generally do not align.

Decade

- Decade of observation. Note: election years and census years generally do not align, but decades do. France does not have any RHCD observations in the 2000s.

LAU

- The LAU label is used to identify the local units. For most countries, the LAU is the local area unit.
- In V1 this refers to LAU-2018, and in V2 the LAU-2021/2024 boundaries.
- In all datasets, the LAU for Greece and Ireland is NUTS3, not the local area unit (see notes).
- There are differences between the RHCD and election data for Germany, Portugal and New Zealand. In the RHCD data the unit for Germany is NUTS3; for New Zealand it is the territorial authority. See country notes.

LAUNAME

- English name of the above unit.

Laumatch

- Code for matching the units across RHCD and RED datasets. Because years differ across datasets, matching requires interpolating or collapsing observations by decade.
- Unmatched: Australia (no pre-1990s elections, 100–150 SA2 per year; see country notes); Austria (no pre-1980 elections, 10–70 per decade unmatched); Belgium (11 obs. total unmatched); Canada (no pre-1990s elections, 50–60 per decade unmatched); Denmark (no pre-1990s RHCD); France (no census observation in the 2000s, no elections pre-1993, 950 unmatched observations in the 2010s of very small areas < 50 population); Germany (unmatched GDR pre-1990); Greece; Ireland; Italy (40–50 unmatched per decade; Aosta not included in elections); New Zealand (no pre-1990s elections, 1 per decade unmatched); Norway (no pre-1980 elections); Spain (300–700 missing per decade; these are small units suppressed in census data); Sweden (several municipalities missing RHCD data before 2020).

NUTS3CODE

- European NUTS3 units.

Macroregion

- Australia: states; Canada: provinces; US: states; NZ: regional authorities; Japan: prefectures.

TL3

- Territorial level 3 unit code. For Europe this is NUTS3. For Australia, Canada, and the US this follows OECD TL3 definitions.

FUACODE

- Functional Urban Areas, 2012 boundaries. OECD.

3.1 Fuaname

- As above, by name.

3.2 Urban_area_type

- 5 - FUA major city, core area
- 4 - FUA major city, non-core area
- 3 - FUA metropolitan area
- 2 - Towns (Eurostat tercet, or national equivalent)
- 1 - Rural (Eurostat tercet, or national equivalent)

3.3 Area_km_sq

- Area in square km of the LAU or equivalent unit.

3.4 Distance_major

- Estimated distance in kilometers for units in a major FUA to the core city center.

3.5 Distance_city

- Estimated distance in kilometers for all units to a given city.

4 RHCD

4.1 Population - by gender and age

- Total population, subset by male (m)/female (f) and age groups where available.

4.2 Population - by gender and age _adj

- Total population, subset by male (m)/female (f) and age groups where available, which removes respondents with missing educational information.

4.3 ed_isced_0to2 - ed_isced_6to8 by age, sex

- Total number of people age 15+ with ISCED 0–2 / 3–5 / 6–8 qualification subset by male (m)/female (f) and age groups where available. See country notes for qualification types and adjustments.

4.4 ed_isced_0to2pc - ed_isced_6to8pc by age, sex

- Total share of people age 15+ with ISCED 0–2 / 3–5 / 6–8 qualification subset by male (m)/female (f) and age groups where available.

4.5 ed_isced_0to2pc_adj - ed_isced_6to8pc_adj by age, sex

- As above, but missing removed from 0–2 and the population baseline to calculate shares.

5 RED

5.1 Total_votes

- Total votes cast in the LAU.

5.2 Total_voters

- Total eligible voters in the LAU (not available for Australia, NZ, USA).

5.3 To1

- Total votes / total voters.

5.4 Weight

- Votes in the LAU as a share of all votes cast in country-year. Used as a weighting variable for unit size.

5.5 party001–party0096

- Votes in LAU for a specific party; see codebook for party names.

5.6 share_party001–share_party0096

- Share of votes in LAU for a specific party; see codebook for party names.

5.7 pf_*

- Party votes linked to Party Facts ID code (or V-PARTY for USA and UK).

5.8 vote_(family)

- Votes in LAU aggregated by party family. See codebook for family classifications.

5.9 pct_(family)

- Share of votes in LAU aggregated by party family. See codebook for family classifications.

5.10 den_pctile

- Density based on voters (votes where missing) per square kilometer in country-year.

5.11 ed_pctile

- Education deciles based on decade-average decile of the share of ISCED 6–8 attainment (from RHCD).

5.12 flag

- Flag = 1 where total LAU units are fewer than the full time series due to missing data or unmatched merges/splits. Generally affects a small number of cases.

6 RMD

6.1 population_foreignborn

- Total population in local area not born in country.

6.2 population_foreignborn_pc

- Share of population in local area not born in country.

6.3 population_noncitizens

- Total population in local area not a citizen.

6.4 population_noncitizens_pc

- Share of population in local area not a citizen.

7 Country Information

- Describe any data cleaning steps.
- Mention derived variables.
- Cite related papers or documentation.

Australia

Flags

- There are substantial missing data in the education series. Individuals with missing information are included in the ISCED 0–2 category. The adjusted (_adj) variables exclude the missing group from both the numerator and denominator and provide an alternative measure of the population skill distribution, which reduces the share of the lowest category.
- Matched election data should be used with caution. Not all SA2 units can be matched to polling booths. While results are generally reliable, aggregation to higher-level units yields fully matched results (see notes below).

Human Capital and Population Data Sources

- Census of Population 2006–2021 time series profiles contains information on the highest qualification among individuals *with* a reported qualification. However, this indicator alone does not capture individuals who have not completed secondary school. We therefore combine the highest qualification variable with a measure of the age left school to construct an educational qualification measure.
- For the 1986, 1996, and 2006 censuses, we submitted a special request to the Australian Bureau of Statistics (ABS) for detailed information matchable to the 2016 SA2 boundaries and qualification structures.
- In some statistical regions, data on highest qualification are missing. In approximately 1% of statistical units, around one quarter of individuals have missing qualification data.

Qualifications

- **ISCED 6–8**: Postgraduate degree level; graduate diploma and graduate certificate level; bachelor's degree level.
- **ISCED 5**: Advanced diploma; diploma level.
- **ISCED 3–4**: Individuals who have completed Year 11, or who (in 1986/1996) left school at age 17 or above, or who have completed a Level 3–4 certificate.
- **ISCED 1–2**: Individuals who left school before Year 11 or before age 17 (for 1986/1996), as well as individuals with missing qualification information.

Age Ranges and Adjustments to Education Estimates

- Applies to the population aged 15 and over.

Core Geography

- The base spatial unit is Statistical Area Level 2 (SA2), using 2016 boundaries. For 1986, 1996, 2006, and 2016, the Australian Bureau of Statistics provided concordances to this level of aggregation (with small amounts of random error). For 2021, data are matched backward to the 2016 geography.

Migration Data Sources

- Data for 2006, 2016, and 2021 are publicly available from the Australian Bureau of Statistics; data for 1986 and 1996 were obtained through a special request to the ABS.
- Data are available for foreign-born individuals.

Electoral Data Sources

- Australian Electoral Commission.
- First-preference ballots are used.
- Legislative (House of Representatives) elections: 3 October 1998; 10 November 2001; 9 October 2004; 24 November 2007; 21 August 2010; 7 September 2013; 2 July 2016; 18 May 2019; 21 May 2022.

Electoral Geography

- For 2011, 2016, 2019, and 2022, the longitude and latitude of polling booths are available. This allows exact matching of polling booths to lower-level geographic units but does not define catchment boundaries (i.e., a booth located in one SA2 may draw voters from another SA2).
- For years without coordinates, many polling booths are matched using coordinates from 2011–2016 where booth names and addresses coincide, or by postcode, although postcodes may span multiple SA2s.
- Large parts of rural Australia use postal balloting, including geographically extensive regions (e.g., East Pilbara in Western Australia), with approximately 6,700 residents and an area of about 372,000 square kilometers.

    1. In total, approximately 350,000 voters in 1998 and 200,000 voters in 2001 and 2004 lack geographic information.
    2. Not all SA2s match to polling booths. This partly reflects the fact that individual polling booths may serve multiple SA2s and that booth and SA2 boundaries do not align precisely. There are 178 SA2s (out of 2,159) that do not match; most have zero or very small populations (< 50), although a small number are urban SA2s within large conurbations (e.g., Brisbane) with populations exceeding 7,000.
    3. Collapsing to higher-level units (SA3) improves spatial accuracy but substantially reduces geographic precision (SA3 populations range from 30,000 to 130,000 across 358 units). We therefore retain the SA2 classification, noting that full precision requires aggregation to SA3. The unmatched units, while covering a substantial geographic area, account for only a small fraction of the voting population.
    4. Mail-in votes, hospital votes, prison votes, and unmatched booths are assigned to SA2s proportionally based on their share within the electoral division.

Austria

Human Capital and Population Data Sources

- National censuses: 1971, 1981, 1991, 2001, and 2011. Detailed data were procured through Statistik Austria STATcube.
- Register-based statistics from 2018 and the 2021 Census.

Qualifications

- **ISCED 6–8**: University degrees; universities of applied sciences.
- **ISCED 4**: Post-secondary colleges (1971–2001); pedagogical academies, social academies, and health system academies (2011–2018). The latter represent a form of higher vocational education targeted at individuals who already hold a university entrance degree and include higher schools for teacher training.
- **ISCED 4**: Higher technical and vocational schools; five-year secondary vocational programs including advanced training, as well as colleges and special courses.
- **ISCED 3**: Upper secondary general education (grammar schools); graduation from the AHS lower cycle is not regarded as completion of upper secondary education.
- **ISCED 3**: Apprenticeships, including completion of apprenticeship training, preliminary and final examinations, and skilled worker examinations.
- **ISCED 3**: Intermediate technical and vocational schools (1971–2001); one- to four-year vocational schools that do not confer a certificate qualifying for university admission (known as *Fachschulen* or technical schools), including intermediate medical and health-care schools and master craftsman examinations (2011, 2018).
- **ISCED 2**: Compulsory schooling, including individuals with a mandatory school-leaving qualification who have not completed upper secondary education, as well as individuals who provided no information.

Age Ranges and Adjustments to Education Estimates

- Applies to the population aged 15 and over.
- Age breakdowns are available.

Core Geography

- LAU boundaries: 2018 (2,100 units; V1) and 2021 (2,091 units; V2).
- Municipal mergers are tracked using official files (*Auflösungen bzw. Vereinigungen von Gemeinden ab 1945*).
- Some manual coding was required for Styria, which substantially restructured its municipalities.
- For 1983 and 1986, 12 municipalities split and cannot be meaningfully matched forward. This affects 53 contemporary municipalities and approximately 53,000 voters in contemporary terms.

Migration Data Sources

- All migration data are publicly available from Statistik Austria.
- Data are available for individuals with a foreign nationality (non-citizens) for the entire time period. Information on foreign-born individuals is additionally available for 2011 and 2018.

Electoral Data Sources

- Bundesministerium des Innern historical series.
- Legislative (Nationalratswahlen) elections: 24 April 1983; 23 November 1986; 7 October 1990; 9 October 1994; 17 December 1995; 3 October 1999; 24 November 2002; 1 October 2006; 28 September 2013; 15 October 2017; 29 September 2019; 29 September 2024.

Electoral Geography

- See the core geography section above.
- Postal votes are aggregated at the province (NUTS1) level and assigned proportionally by population to individual units within each province.
- Minor discrepancies (approximately 10,000–20,000 voters) occur from the 1990 elections onward, with larger discrepancies in 1983 and 1986 when postal votes were not included (approximately 120,000 votes not regionally attributable in each election).

Belgium

Flags

- Electoral and RHCD geographies differ. Some electoral data are missing; see notes below.

Human Capital and Population Data Sources

- Population census data for 1981, 1991, 2001, 2011, and 2017.
- Data for 2011 and 2017 are publicly available from Statbel.
- Data for 1981, 1991, and 2001 were obtained by request from Interface Demography, with additional data provided directly by Statbel.

Educational Qualifications

- **ISCED 6–8**: 1981 – universitair of daarmee gelijkgesteld onderwijs (lange type of 3e graad). 1991 – ander niet-universitair onderwijs van het korte type; universitair en gelijkgesteld. 2001 – higher academic cycle 1*; higher academic cycle 2; advanced academic (doctorate). 2011 – enseignement supérieur* (graduat/bachelier; licence/master); doctorat. 2017 – higher*.
- **ISCED 6–8 (adjusted)**: See below. Adjustments are applied to higher academic cycle 1 in 2001 and 2011, and to the higher category in 2017.
- **ISCED 3–5**: 1981 – hoger secundair algemeen; hoger secundair technisch; hoger secundair kunst; hoger secundair beroeps; hoger secundair met beperkt uunooster; kleuteronderwijzer(es); onderwijzer(es) lager onderwijs; regent(es) of geaggregeerde van het onderwijs van de lagere secundaire graad; ander hoger onderwijs van het korte type (1e en 2e graad); hoger technisch of kunstonderwijs met beperkt uunooster. 1991 – hoger secundair beroeps; hoger secundair algemeen vormend + kunst; pedagogisch van het korte type; hoger secundair technisch; ander niet-universitair onderwijs. 2001 – higher secondary professional education; higher secondary technical education; higher secondary education; post-secondary education; higher non-academic education. 2011 – enseignement secondaire supérieur; enseignement post-secondaire non supérieur. 2017 – medium.
- **ISCED 0–2**: 1981 – lager onderwijs; lager secundair algemeen; lager secundair technisch; lager secundair kunst; lager secundair beroeps; lager secundair met beperkt uunooster; education followed up to age 14 without a diploma; education followed above age 14 without a diploma; no diploma and no formal education. 1991 – no (Belgian or foreign) diploma; primary education; lower secondary vocational education; lower secondary general + arts; lower secondary technical education. 2001 – no formal education; primary education; lower secondary professional education; lower secondary technical education; lower secondary education. 2011 – pas de diplôme ou de certificat; enseignement primaire. 2017 – low.

Age Ranges and Adjustments to Education Estimates

- Although age breakdowns are available for many years, they are not included in the final dataset. In some years (e.g. 2011), tables with age breakdowns report less granular qualification categories, making reconciliation to ISCED classifications more difficult.

- Four major adjustments are applied to the Belgian education data. First, a large share of the population does not respond to the census qualification question, with over 10% missing in each wave. Second, foreign qualifications are treated differently over time and are excluded from the main qualification totals in 2001. Third, the higher category combines ISCED 5 (short-cycle tertiary) and ISCED 6 (long-cycle tertiary). Finally, in some waves, students are excluded. Because of these differences across waves, we apply the following adjustment strategies without constructing an additional "adjusted" variable.

    1. We benchmark educational distributions against nationally published aggregates from the Belgian Labour Force Survey (LFS) for 1987–2020.
    2. Missing and foreign qualifications are assigned based on the mismatch between the observed distribution and national totals. This approach likely undercounts highly skilled foreigners in Brussels; caution should therefore be exercised when interpreting Brussels data for 2001, where foreign qualifications are excluded. We assume that the breakdown of "missing" and "foreign" categories in 2001 is similar across geographic areas, with local variation driven by baseline levels.
    3. Where students are excluded (2001), we use Statbel data on educational attainment for the population aged 18+, assigning all students aged 18+ to ISCED 3–5. We then use data on 15–17-year-olds and assign them to ISCED 0–2, as they are unlikely to have completed upper secondary education. This may misclassify a small number of 18-year-olds still completing upper secondary or post-secondary qualifications, but the effect is likely minor.
    4. To separate long- and short-cycle tertiary education within ISCED 6, we apply an 80–20 split. We assign 80% of higher academic cycle 1 qualifications (enseignement supérieur non-universitaire de type court; graduat (A1); baccalauréat professionnalisant / hoger onderwijs) to ISCED 3–5 and 20% to ISCED 6–8. This yields national distributions consistent with disaggregated totals. For 2017, we further adjust categories using locally specific disaggregated distributions from 2011.

Core Geography

- LAU, 2019 boundaries.
- 589 municipalities. Data for 1991, 2001, and 2011 are consistent with 2019 boundaries and matched using municipality codes. Post-2020 municipal mergers can be incorporated by collapsing to new boundaries, for which codes are available.

Migration Data Sources

- Migration observations in the early 1990s are from 1992, while education observations are from 1991.
- Coverage is limited to 1981, 2001, and 2011.
- 1981 census data are available as a PDF from Statbel and were manually digitized; later register data are publicly available from Statbel.
- Data are available for individuals with a foreign nationality (non-citizens).

Electoral Data Sources

- Elections in 2003, 2007, 2010, 2014, and 2019 from the Interior Ministry.
- Historical series for pre-2003 elections.
- Turnout data are not available for 1981, 2003, or 2007.
- Federal legislative elections (Chambre/Kamer): 8 November 1981; 13 October 1985; 13 December 1987; 11 November 1991; 21 May 1995; 13 June 1999; 18 May 2003; 10 June 2007; 13 June 2010; 25 May 2014; 26 May 2019.

Electoral Geography

- Election results are reported at the level of the **electoral district (canton)** (211 units), each of which contains multiple municipalities (589 units).
- To match results to urban types, electoral districts are allocated to municipalities within each canton. Data are presented at the LAU level, with district votes assigned proportionally by population shares. **Note:** these LAU-level results are imputed; accurate results require collapsing back to the canton level.
- Rhode-Saint-Genèse (RSG) is a new canton formed in 2014, previously part of Brussels–Hal–Vilvorde. The split affects Sint-Genesius-Rode, Kraainem, Linkebeek, Wemmel, Drogenbos, and Wezembeek-Oppem. Municipal attachments to Hal and Zaventem therefore change over time and are not fully consistent pre- and post-2014. RSG itself cannot be reliably traced backward. Post-2014 voting structures also vary, as RSG electors may vote using multiple ballots.
- Zaventem canton data are unavailable prior to 1991.
- Caution should therefore be exercised when interpreting time-series results for municipalities in the Zaventem, Hal, and RSG areas.

Canada

Human Capital and Population Data Sources

- National censuses: 1986, 1991, 1996, 2001, 2006, 2016, and 2021.
- Data are publicly available from Statistics Canada and the University of Toronto.

Educational Qualifications

- **ISCED 6–8**: University with a bachelor's degree or higher (1981, 1986); bachelor's degree (1991–2006); university certificate or diploma above the bachelor level (1991–2006); medical degree (1991–2006); master's degree (1991–2006); earned doctorate (1991–2006); university certificate, diploma, or degree at the bachelor level or above (2011); university certificate, diploma, or degree at the bachelor level or above (2016); bachelor's degree (2016); university certificate or diploma above the bachelor level (2016); degree in medicine, dentistry, veterinary medicine, or optometry (2016); master's degree (2016); earned doctorate (2016).
- **ISCED 4–5**: University with a university or non-university certificate or diploma (1981); university without a degree (1986); other non-university certificate or diploma (1991–2001); university without a certificate, diploma, or degree (1981); university certificate or diploma below the bachelor level (1991–2016); college, CEGEP, or other non-university certificate or diploma (2006–2016); certificate of apprenticeship or certificate of qualification (2016).
- **ISCED 4–5 (alternative classification)**: Trades certificate or diploma (1981–2001); some postsecondary education (2001); apprenticeship or trades certificate or diploma (2006–2016); postsecondary certificate, diploma, or degree (2016).
- **ISCED 3**: Grades 9–13 with a secondary school graduation certificate (1981, 1986); secondary graduation certificate (1991–2001); high school diploma or equivalent (2006–2016).
- **ISCED 1–2**: Less than Grade 9 (1981, 1986); Grades 9–13 without a secondary certificate or diploma (1981, 1986); no degree, certificate, or diploma (1991, 1996, 2006, 2011, 2016); less than high school graduation certificate (2001).

Age Ranges and Adjustments to Education Estimates

- Applies to the population aged 15 and over.
- Disaggregated age groups are available.

Core Geography

- The base spatial unit is the 2016 Census Consolidated Subdivision (CCS). CCS units group census subdivisions into larger areas that preserve urban diversity while consolidating lower-density regions.
- Data are typically collected at finer spatial scales (generally census subdivisions) and aggregated to CCS boundaries.
- Areas are matched to 2016 boundaries using the degree of spatial overlap. Shapefiles are provided by the University of Toronto Libraries.

Migration Data Sources

- Migration data are publicly available from Statistics Canada.
- Areas are matched to 2016 boundaries using the degree of spatial overlap. Shapefiles are provided by the University of Toronto Libraries.
- Data are available for individuals with a foreign nationality (non-citizens).

Electoral Data Sources

- Election results are drawn from the Elections Canada Researchers File.
- Federal elections held on: 2 June 1997; 27 November 2000; 28 June 2004; 23 January 2006; 14 October 2008; 2 May 2011; 19 October 2015; 21 October 2019; 20 September 2021.

Electoral Geography

- Raw voting data are reported by polling booths and broader voting categories (e.g. mail ballots).
- Polling booths are matched, where possible, to Census Consolidated Subdivisions (CCSs). Where a polling booth spans multiple CCS units, votes are allocated proportionally based on the physical area of overlap.
- Polling booths that cannot be matched, advance polls (with larger catchment areas), and mail ballots are assigned proportionally to CCS units based on their share within each electoral constituency. Caution should be exercised when interpreting 2021 results, as many voters used advance polls and mail ballots due to the COVID-19 pandemic.
- Geographic boundary matching introduces some spatial imprecision.
- Matching to polling booths uses shapefiles provided by the University of Toronto Libraries.
- The share of geographically assignable votes is higher in 2015 and 2019 than in earlier elections. In previous years, approximately 300,000 votes (around 2–2.5%) are not geographically assignable; these missing votes are relatively evenly distributed across parties.
- For 1997 and 2000, raw vote totals do not fully match published aggregates. Some votes in Quebec are misallocated between the NDP and Bloc Québécois (17,043 votes in 1997 and 20,213 votes in 2000). These inconsistencies are present in the official researcher files.

Denmark

Human Capital and Population Data Sources

- Register-based statistics covering 1991–2024.
- Series: HFUDD11: Befolkningens højst fuldførte uddannelse (15–69 år) for 2006–2024.
- Series: HFU1: Befolkningens højeste fuldførte uddannelse (15–69 år) efter område, herkomst, uddannelse, alder og køn for 1991–2006.
- All data are publicly available.

Educational Qualifications

- **ISCED 6–8**: Long-cycle higher education (65); bachelor (60); medium-cycle higher education (50) (1991–2005); PhD programs (H80); master's programs (H70); bachelor's programs (H60); vocational bachelor's programs (H50) (2006–2018).
- **ISCED 5**: Short-cycle higher education (40/H40) (1991–2005 and 2006–2018).
- **ISCED 3**: Vocational education (35) (1991–2005); qualifying educational programs (H35) (2006–2018); vocational education and training (VET) (H30) (2006–2018).
- **ISCED 3**: Vocational upper secondary school (25) (1991–2005); general upper secondary school (20) (1991–2005); upper secondary education (H20).
- **ISCED 1–2**: Basic school, Grades 8–10 (10) (1991–2005); primary education (H10) (2006–2018).

Age Ranges and Adjustments to Education Estimates

- Applies to the population aged 15–69.
- We adjust the upper threshold by averaging education levels in the two oldest five-year age groups and assuming a linear trend across educational categories for the excluded population aged 69+. We then estimate educational shares based on the municipality-level population size of this group and add these estimates to aggregate ISCED totals.
- Disaggregated age groups are available.

Core Geography

- LAU, 2018 boundaries (99 units).
- Denmark's system of local government underwent a major reorganization in 2007. Municipal mergers are tracked using official merger files.
- Pre-2007 elections are manually matched, which requires aggregating disaggregated Copenhagen data and selected other municipalities.

Migration Data Sources

- Coverage is limited to data from 1995, 2000, 2005, and 2008–2019.
- Data are publicly available from Statistics Denmark.
- Data are available for individuals with a foreign nationality (non-citizens).

Electoral Data Sources

- From 2007 onward, election results are available from Elections 2007–2022.
- For elections before 2007, results were digitized from statistical yearbooks and manually added; these data are now available from the Danish Election Database.
- Legislative elections (Folketing): 10 January 1984; 8 September 1987; 12 December 1990; 20 November 1991; 21 September 1994; 11 March 1998; 15 February 2005; 13 November 2007; 15 September 2011; 18 June 2015; 5 June 2019; 1 November 2022.

Electoral Geography

- See the core geography section above.

Finland

Human Capital and Population Data Sources

- Register-based statistics covering 1970, 1975, 1980, 1985, and annually from 1987 to 2019.
- Data are publicly available from Statistics Finland.

Educational Qualifications

- **ISCED 6–8**: Bachelor's or equivalent; master's or equivalent; doctoral or equivalent.
- **ISCED 5**: Short-cycle tertiary education.
- **ISCED 4**: Post-secondary non-tertiary education.
- **ISCED 3**: Upper secondary education.
- **ISCED 2**: Basic education qualification.

Age Ranges and Adjustments to Education Estimates

- Applies to the population aged 15 and over.
- Disaggregated age groups are available.

Core Geography

- LAU, 2018 boundaries (308 units).
- A small number of boundary changes are matched manually.

Migration Data Sources

- Coverage is limited to annual data from 1980 to 2019.
- Data are publicly available from Statistics Finland.
- Data are available for individuals with a foreign nationality (non-citizens) for the entire 1980–2019 period. Information on foreign-born individuals is additionally available for 1990–2019.

Electoral Data Sources

- Time-series data on party support and turnout are publicly available.
- Legislative elections (Eduskunta): 20 March 1983; 15 March 1987; 17 March 1991; 19 March 1995; 21 March 1999; 16 March 2003; 18 March 2007; 11 April 2011; 19 April 2015; 14 April 2019; 2 April 2023.

Electoral Geography

- See the core geography section above.

France

Flags

- Migration data are available at the department level, not the LAU.
- See notes on imputation of some educational qualifications.

Human Capital and Population Data Sources

- Census: 1975, 1982, 1990, 1999, 2010, and 2020.
- Series: INSEE harmonized census data.

Educational Qualifications

- **ISCED 7–8**: Diplôme universitaire de 2ème ou 3ème cycle (includes doctorate — ISCED 8 — and masters — ISCED 7).
- **ISCED 5–6**: Diplôme universitaire de 1er cycle. This combines university and other post-secondary short-cycle qualifications (Bac+3 and Bac+2). The Licence (ISCED 6), BTS/DUT or equivalent (ISCED 5), and paramedical or social qualifications are included. A very small share of level 4 is present in this group; overall the category combines short- and long-cycle higher education (ISCED 5 and 6).
- **ISCED 6 (corrected)**: Adjustment applied for 1982 and 2020 to the ISCED 5–6 category based on national aggregates (see dataset paper and appendix)
- **ISCED 3**: Diplôme de niveau Baccalauréat (général, technologique, professionnel) — ISCED 3A.
- **ISCED 3 (certificat)**: Diplôme de niveau CAP, BEP — ISCED 3C.
- **ISCED 0–2**: No diploma above the BEPC / brevet des collèges / DNB. Some brevets map to 3C, but DNB or lower is treated as ISCED 1–2 combined.

Age Ranges and Adjustments to Education Estimates

- The original data are restricted to individuals who have completed their education (age 16+; *Individus âgés de 16 ans ou plus, et ayant terminé leurs études*).
- We expand the lower threshold to include 15-year-olds by adding 15-year-old population counts into the ISCED 0–2 category, using census population estimates.
- Because students are not included in the original tables, we allocate the difference between the total 16–24 population and the 16–24 population with reported educational qualifications to ISCED 1–2 and ISCED 3–4 in a 1:2 ratio.

Core Geography

- LAU: 2018 boundaries (V1) and LAU 2021 boundaries (V2).
- Municipal merges are tracked using official INSEE files. A very small number of splits cannot be matched.

Migration Data Sources

- Coverage: NUTS3 (department) level only.

Electoral Data Sources

- Legislative election results (first round) are available from data.gouv.fr for 1993–2022.
- Before 1993, only large towns and certain circumscriptions can be matched; departments can be matched from 1958.
- Legislative elections (Assemblée Nationale) — first round: 21 March 1993; 25 May 1997; 9 June 2002; 10 June 2007; 10 June 2012; 11 June 2017; 12 June 2022.

Electoral Geography

- LAU: 2018 boundaries (35,281 units) and LAU 2021 boundaries (34,966 units).
- Municipal merges are tracked with official INSEE files.
- Arrondissements of Paris, Lyon and Marseille cannot be linked consistently over time; neighborhoods in these cities are therefore treated as single units for time-series comparability.
- Between 2002 and 2017 a very small number of tiny towns (fewer than 10) do not match.
- For 1997, 15 towns (combined population circa 8,000) are unmatched in the election data; additionally Beuvry (1997), Tournefeuille (1993), and 20–30 small towns are missing from the raw data.
- In 1993, approximately 75,000 voters in Marne cannot be matched due to inconsistent naming practices (aggregate groups, e.g. Commune 701–713); use Marne with caution for 1993.
- A very small number (about 25) of municipal splits cannot be matched.
- Vote totals reported here are lower than national aggregates because we omit votes from Réunion, Mayotte, Guadeloupe, and other French overseas constituencies. These overseas voters account for close to 900,000 voters in 1993 and upwards of 3,000,000 voters in 2017. Overseas voters are omitted because they vote in special constituencies rather than via mail-in ballots attached to mainland areas.

Germany

Flags

- RHCD and RED are available at different levels of aggregation.
- The V2 version uses the most recent LAU match files, but some German municipal mergers are not captured. Because these mostly involve smaller towns joining existing centers, a small number (fewer than 50) of observations may include merged areas in 2025 without a flag variable. Caution should therefore be used when interpreting changes between the 2021 and 2025 elections in small areas.

Human Capital and Population Data Sources

- 1987 Census of Population (Federal Republic of Germany only). Available from GESIS.
- 2011 and 2021 Censuses of Population. Available from the Zensus Datenbank.
- Longer time series are not available due to data constraints. Germany did not conduct a population census between 1987 and 2011. Regional microcensus data below NUTS2 are not publicly available at lower levels of aggregation. Other sources, such as labor force surveys, can provide more localized snapshots of the working-age population.
- Lower-level estimates can be produced by combining 2011 census data with higher-level microcensus aggregates; these imputations are available upon request.

Educational Qualifications

- Only ISCED 6–8 qualifications are reliably available.
- **ISCED 6–8**: University degrees (1987); *Abschluss einer Fachakademie oder Berufsakademie*, *Fachhochschulabschluss*, and *Hochschulabschluss Promotion* (2011, 2021).
- Qualifications below ISCED 6 cannot be distinguished from one another. Tables in both 1987 and 2011 report the total number of qualifications held rather than individuals' highest qualification. From these tables, the number of individuals holding ISCED 6–8 qualifications can be inferred.

Age Ranges and Adjustments to Education Estimates

- 1987: residents aged 15–64. Educational attainment among those aged 65+ is estimated using attainment rates among those aged 48+ in the 1970 census, aggregated at the NUTS1 level.
- 2011: residents aged 15 and over.

Core Geography

- Human capital data are available only at the NUTS3 level (Kreise/Districts), using 2018 boundaries.
- Areas are combined to approximate 2011 boundaries in Mecklenburg–Vorpommern; to account for 2016 boundary changes in Göttingen (Lower Saxony); and to reflect the 2001 merger of *Kreisfreie Stadt Hannover* and *Region Hannover*.
- Data for 1987 are available for areas in the Federal Republic of Germany only, excluding West Berlin.

Migration Data Sources

- 1987 migration data are available from GESIS; 2011 migration data are available from the Zensus Datenbank.
- Data are available for individuals with a foreign nationality (non-citizens).

Electoral Data Sources

- Election results are available from the Federal Returning Officer (Bundeswahlleiter).
- Federal legislative elections (Bundestag), **second vote** (*Zweitstimme*): 6 March 1983; 25 January 1987; 16 October 1994; 27 September 1998; 22 September 2002; 18 September 2005; 27 September 2009; 22 September 2013; 24 September 2017; 26 September 2021; 23 February 2025.

Electoral Geography

- Electoral data are reported at the LAU (Gemeinde) level rather than at the county (Kreis) level, using 2018 LAU boundaries (11,072 units) for V1.
- V2 uses 2024 LAU boundaries (10,980 units). A small number of boundary changes occurring after 2024 cannot be matched.
- A sizeable share of municipalities (approximately 10%) required manual matching due to multiple boundary changes over time.
- Matching was conducted using the Destatis Gemeindeverzeichnis.
- Some municipalities could not be matched: 10 in 2017; 2–4 per election between 1998 and 2013 (typically fewer than 500 voters); and 23 municipalities in 1994 (approximately 9,000 voters) due to major boundary changes.
- For 1983–1987, no data are available for Berlin. In addition, 22 contemporary municipalities (approximately 20,000 residents in current population terms) and 8 municipalities in the raw data (approximately 7,000 voters) could not be forward matched.
- Some cases involve municipal splits that cannot be fully reconciled (e.g. Thurm, Neudorf, Reichenberg, Radefeld, Groß Gladebrügge); these are matched to the largest successor unit.
- Postal votes are assigned proportionally at the *Wahlkreis* level. Not all *Wahlkreise* report postal votes.

Greece

Human Capital and Population Data Sources

- Ten-percent microdata from the 1971–2011 population censuses obtained from IPUMS-International.
- Full census data at the regional level are available from ELSTAT; however, educational qualifications are not provided with the same degree of granularity.

Educational Qualifications

- **ISCED 6–8**: University completed.
- **ISCED 5**: Some college completed.
- **ISCED 4**: Post-secondary technical education.
- **ISCED 3**: Secondary education completed (general track); secondary education completed (technical track).
- **ISCED 2**: Lower secondary education completed.
- **ISCED 1**: Primary education (6 years) completed.
- **ISCED 0**: Some primary education completed; no schooling; less than primary education completed (not stated).

Age Ranges and Adjustments to Education Estimates

- No age adjustments are necessary; estimates are derived directly from microdata.

Core Geography

- Geographic boundaries are pre-harmonized by IPUMS-I to 2010 boundaries, i.e. prior to the Kallikratis Reforms of 2011.
- Data are available at the 2010 NUTS3 (department) level.
- IPUMS microdata allow the identification of larger municipalities within NUTS3 regions containing Athens and Thessaloniki.
- Within the Attiki NUTS3 region, we distinguish 46 municipalities, separating central Athens, Piraeus, and many suburban areas. Four additional residual categories cover the remainder of the Athens, East Attiki, Piraeus, and West Attiki prefectures.
- Within the Thessaloniki department, we distinguish 12 municipalities covering most of the Thessaloniki urban area, along with a residual unit covering the remainder of the department.
- This yields a total of 107 geographic units, with a median population of 67,380 in 2011.

Migration Data Sources

- Migration data are estimated using 10% census microdata from the 1971–2011 censuses obtained from IPUMS-International.
- Data are available for individuals with a foreign nationality (non-citizens).

Electoral Data Sources

- Although highly localized election results are available, they cannot be consistently downloaded and reconciled to stable geographic boundaries over time.
- We therefore rely on CLEA data matched to stable NUTS3 boundaries. This approach yields less geographically granular data for Greece than for most other countries, and matching to urban areas is imperfect.
- Legislative elections (Hellenic Parliament): 18 October 1981; 2 June 1985; 18 June 1989; 5 November 1989; 8 April 1990; 10 October 1993; 22 September 1996; 9 April 2000; 7 March 2004; 16 September 2007; 4 October 2009; 6 May 2012; 17 June 2012; 25 January 2015; 20 September 2015; 7 July 2019.

Electoral Geography

- 2018 NUTS3 boundaries (46 units for Greece).
- To approximate urban exposure, NUTS3 regions are matched to LAUs, and an average urban-type measure is constructed for each NUTS3 unit.
- This geographic approach prevents separate identification of suburbs of the major city (Athens) and the mid-sized city (Thessaloniki) in the electoral data.

Ireland

Flags

- There are more granular geographic RHCD data than RED data.
- An adjusted variable is available to address substantial missing education data.

Human Capital and Population Data Sources

- Population censuses: 1981, 1991, 1996, 2002, 2006, 2011, 2016, and 2022.
- Data are publicly available from the Central Statistics Office (CSO).
- Substantial adjustments are required to estimate qualification breakdowns due to highly aggregated education categories in the 1981 and 1991 censuses.

Educational Qualifications

- **ISCED 4–8**: University (1981); third-level education (1991).
- **ISCED 6–8**: Third-level degree or higher (1996, 2002, 2006); ordinary bachelor's degree or professional qualification (2011); honours bachelor's degree or professional qualification (2011); postgraduate diploma or degree, doctorate (Ph.D.) (2011); ordinary bachelor's degree or professional qualification (2016); honours bachelor's degree or professional qualification (2016); postgraduate diploma or degree, doctorate (Ph.D.) (2016).
- **ISCED 4–5**: Third-level non-degree (1996, 2002, 2006); technical/vocational (2011); advanced certificate or completed apprenticeship (2011); higher certificate (2011); technical/vocational (2016); advanced certificate or completed apprenticeship (2016); higher certificate (2016).
- **ISCED 6–8 (corrected)**: Adjustment applied in 1981 and 1991 to estimate the share of individuals with ISCED 4–8 qualifications attaining ISCED 6–8, using IPUMS-I microdata at the NUTS3 level.
- **ISCED 2–3**: Secondary, vocational, or secondary and vocational education (1981); second level – first stage and second level – second stage (1991).
- **ISCED 3**: Upper secondary education (1996, 2002, 2006, 2011, 2016).
- **ISCED 2**: Lower secondary education (1996, 2002, 2006, 2011, 2016).
- **ISCED 3–5 (corrected)**: Adjustment applied in 1981 and 1991 to estimate the share of individuals with ISCED 4–8 qualifications attaining ISCED 4–5, and those with ISCED 2–3 attaining ISCED 3, using IPUMS-I microdata at the NUTS3 level.
- **ISCED 0–1**: Primary education, including not stated (1981); primary education, including no formal education (1991, 1996, 2002, 2006); no formal education and primary education (2011, 2016).
- **ISCED 0–2 (corrected)**: Adjustment applied in 1981 and 1991 to estimate the share of individuals with ISCED 2–3 qualifications attaining ISCED 2, using IPUMS-I microdata at the NUTS3 level.

Age Ranges and Adjustments to Education Estimates

- Students are assumed to belong to ISCED 0–2 if below age 19 and to ISCED 3–5 if aged 19 or above.
- There is substantial missing information on educational attainment. The adjusted variable excludes missing respondents from both the numerator and denominator.

Core Geography

- Geographic classification is based on County and City Councils, which are the primary units of local government in Ireland.
- The following adjustments were made to harmonize boundaries over time:
  - Dún Laoghaire–Rathdown, Fingal, and South Dublin are combined into a single "Dublin County" unit to match 1981 boundaries.
  - Galway County and Galway City are combined into a single "Galway" unit to match 1981 boundaries.
  - Cork City and Cork County are combined into a single "Cork" unit to adjust for boundary changes in 2019.
- The resulting classification contains 27 units, compared to 31 units in 2019.
- Units are nested within NUTS3 regions but are larger than LAUs.
- In 2016, units have a mean population of 176,365, ranging from 32,044 (Leitrim) to 792,805 (Dublin County).

Migration Data Sources

- Migration data are publicly available from the Irish Central Statistics Office.
- Data are available for foreign-born individuals.

Electoral Data Sources

- Although highly localized election results are available, they cannot be consistently reconciled to stable geographic boundaries over time.
- We therefore rely on CLEA data matched to stable NUTS3 boundaries. This approach yields less geographically granular data for Ireland than for most other countries, and matching to urban areas is imperfect.
- Legislative elections (Dáil): 18 February 1982; 24 November 1982; 17 February 1987; 15 June 1989; 25 November 1992; 6 June 1997; 17 May 2002; 24 May 2007; 25 February 2011; 26 February 2016; 8 February 2020.

Electoral Geography

- 2018 NUTS3 boundaries (8 units for Ireland).
- To approximate urban exposure, NUTS3 regions are matched to LAUs, and an average urban-type measure is constructed for each NUTS3 unit.
- This geographic approach prevents separate identification of Ireland's mid-sized cities (Cork, Galway, Limerick, and Waterford) in the electoral data.

Italy

Human Capital and Population Data Sources

- Censuses of Population and Housing: 1981, 1991, 2001, 2011, and the 2021 Permanent Census.
- Series: ISTAT census data.
- All data are publicly available. The 1981 census was scanned and digitized; data for later years are available online.
- We apply a small correction to the ISCED 5–6 category in 2001 and 2011 using IPUMS data, as well as a correction to age ranges.

Educational Qualifications

- Census classifications vary over time, and the most disaggregated categories are not available at the municipal level. We therefore apply small corrections to adjust for combined qualifications.
- **ISCED 6–8**: *Laurea* (1981–1991)[1]; *laurea* or *diploma universitario o terziario di tipo non universitario* (2001); *titoli universitari* (2011). In 2021, the university category corresponds to the "primo livello" of tertiary education.
- The 2001 and 2011 censuses include ISCED 5 qualifications within the university category; we therefore apply small corrections for these years. In 2011, *titoli universitari* include short-cycle university degrees (*diploma universitario*, 2–3 years). In 2001, the tertiary category includes ISCED 6 (*Laurea*) and *diploma universitario o terziario di tipo non universitario*. We estimate the share of non-university tertiary diplomas (2001) and short-cycle university degrees (2011) by NUTS2 region and gender using IPUMS microdata, adjust the ISCED 6 category downward, and reallocate the remainder to ISCED 3–5. Since this group represents a small share of the overall tertiary population, the adjustment amounts to a shift of approximately 4–8% in 2011 and 8–18% in 2001.
- In the 2021 Permanent Census, the university category includes the ITS qualification (ISCED 5). This qualification has only existed since 2011 and includes a very small number of students (16,772 in 2021, compared to over one million students in higher education). Given its short duration and limited size, we adjust this category downward. Changes in diploma structures after 2004 limit direct comparability with earlier adjustments; however, the 2021 categories separately identify most short-cycle post-secondary education, making the series broadly comparable for bachelor's degrees and above.
- **ISCED 3–5**: *Diploma* (1981)[2]; upper secondary education diploma (2001–2011); post-graduate non-university diplomas and AFAM diplomas (2001–2011); secondary and professional education (3–4 years) and *certificato di specializzazione tecnica superiore* (IFTS) (2021).
- **ISCED 2**: *Licenza media inferiore* (1981)[3]; *licenza di scuola media inferiore o di avviamento professionale* (2001–2011, 2021).

[1] 1981: coloro che, avendo seguito un corso completo di studi universitari (della durata minima di quattro anni), hanno conseguito un diploma di laurea

[2] coloro che hanno conseguito un diploma di maturità o di abilitazione di scuola secondaria di secondo grado

[3] coloro che hanno conseguito la licenza di scuola secondaria di primo grado

- **ISCED 1**: *Licenza elementare* (1981, 1991)[4]; *licenza di scuola elementare* (2001, 2011, 2021).
- **ISCED 0**: No formal qualification (1981, 1991)[5]; no degree, including illiterate individuals and those literate without formal qualifications (2001–2021).

Age Ranges and Adjustments to Education Estimates

- Original data for 1981–2011 include residents aged 6 and above (*Popolazione residente in età da 6 anni in poi*); for 2021, residents aged 9 and above.
- The population aged 7–14 (or 9–14 in 2021) is removed from the ISCED 0–2 category to obtain a 15+ population. This is done using single-year population estimates, assuming all individuals in these age ranges fall within ISCED 0–2.
- ISCED 5 is estimated as described above.

Core Geography

- LAU, 2018 boundaries (V1), 2021 boundaries (V2)
- Municipal mergers are tracked manually. A very small number of splits cannot be matched directly. Several splits (e.g. the Roman suburb of Fiumicino) are back-estimated. For 1981 and 1991, population ratios from 2011 are used to estimate historical shares and construct a consistent time series.
- In V1, post-2018 geography is matched back to 2018 boundaries by applying population shares to split or absorbed municipalities. This creates a synthetic version of merged municipalities for 2021. Between 2018 and 2021, 90 units experienced a merger or absorption (out of 7,960 stable units). V2 uses 2021 boundaries.

Migration Data Sources

- Coverage is limited to 1991, 2001, 2011, and 2021.
- Data are publicly available from ISTAT and the Permanent Census.
- Data are available for individuals with a foreign nationality (non-citizens).

Electoral Data Sources

- Election results are available from the Interior Ministry.
- Legislative elections (Camera): 26 June 1983; 14 June 1987; 5 April 1992; 21 April 1996; 13 May 2005; 9 April 2006; 13 April 2008; 24 February 2013; 4 March 2018; 25 September 2022.

[4] coloro che hanno conseguito la licenza elementare

[5] coloro che hanno dichiarato di saper leggere e scrivere pur non avendo conseguito la licenza elementare

Electoral Geography

- LAUs are matched by name from election results to LAU codes, introducing some scope for error. Municipalities with duplicate names were double-checked.
- Municipal mergers are tracked manually using government records and local sources.
- Data from Aosta are missing prior to 2006.
- Geographically disaggregated results for 2018 exhibit an undercount (1,258,487 voters) due to differences in the geolocation of postal and absentee ballots. This issue does not arise in other years; approximately 3% of voters in 2018 are unassigned.

Japan

Human Capital and Population Data Sources

- Census data at the prefectural level are available for 1980–2020; municipal-level data are available for 1990–2020.
- Data are publicly available from Statistics Japan and digitized statistical yearbooks.
- Educational qualification data by gender are available for 1990, 2010, and 2020; age-specific breakdowns are available only for 2020.
- The original census tables include classifications by population density within prefectures; however, these units cannot be consistently identified over time and are therefore not used.

Educational Qualifications

- **ISCED 6–8**: University graduates (卒業者　大学・大学院); in the 2020 census this category is further disaggregated into university and graduate school.
- **ISCED 3–5**: High school graduates and graduates of technical colleges (卒業者　短大・高専). These categories are separately identified from 2000 onward. Prior to 2020, the high school category also includes middle school graduates from the pre-1944 system (高校・旧中).
- **ISCED 0–2**: Elementary, middle school, and higher elementary education (小学校・中学校・高小), including former youth schools in the 1980 census; elementary and middle school (小学校・中学校) in the 1990 census; and primary education (初等教育) from 2000 onward.

Age Ranges and Adjustments to Education Estimates

- Available data include all residents aged 15 and over.
- **Students**: All students aged 20 and above are assigned to ISCED 3–5, along with one-fifth of students aged 15–19. The share of students aged 15–19 is approximately 80% in 1980–2000 and declines to about 70% in 2010–2020. In the overlapping year 2020, this allocation closely matches more detailed census tabulations. Accordingly, four-fifths of students aged 15–19 are assigned to the ISCED 0–2 category. While some older students may ultimately attain ISCED 6–8 qualifications, this group constitutes a small share of total students.

Core Geography

- Municipalities, using 2022 boundaries.
- A substantial number of municipal mergers occurred during the study period. From 2000 onward, these are tracked using official Statistics Japan crosswalk files.
- For municipal boundary changes prior to 2000, we adopt the harmonization approach developed by Keisuke Kondo.

Migration Data Sources

- Coverage is limited to the period 2000–2010.
- Data are available for foreign-born individuals.

Electoral Data Sources

- Election results are consistently available only at the prefectural level over time. We therefore use CLEA data matched to prefectural boundaries. This approach yields substantially less geographic detail for Japan than for most other countries, and matching to urban areas is imperfect.
- More geographically disaggregated election results are available for the period 2017–2024.

Electoral Geography

- See the data limitations described above.

Netherlands

Flags

- RHCD data are not strictly comparable to those of other countries. Coverage and quality are more limited and the data should therefore be used with caution.

Human Capital and Population Data Sources

- Data for the Netherlands span 1996–2013 using CBS Series 71883 – Beroepsbevolking; gemeenten 1996–2013.
- Data for later years are drawn from Series 84773.

Educational Qualifications

- **ISCED 6–8**: *Hoger*. Includes Levels 5–7 of the Dutch *Standaard onderwijsindeling*, covering university education and higher vocational education (four-year HBO). A small number of ISCED 5 HBO two-year programs are included, but only a small fraction of students attend these programs. The structure and role of higher-level HBO have changed over time, which may lead this measure to overestimate the ISCED 6–8 population.
- **ISCED 3–5**: *Middelbaar*. Includes Level 4 of the Dutch *Standaard onderwijsindeling*, covering upper secondary education such as VWO and HAVO.
- **ISCED 0–2**: *Lager*. Includes Levels 1–3 of the Dutch *Standaard onderwijsindeling*, covering primary education and the first phase of secondary education: LBO/VBO/VMBO, MULO/MAVO, and the first three years of HAVO/VWO.

Age Ranges and Adjustments to Education Estimates

- Data for 1996–2013 cover the population aged 15–65, while data from 2013 onward cover the population aged 15–75.
- Educational attainment for older age groups is estimated using IPUMS Dutch national microdata for 2001 and 2011, which allow assignment of broad educational distributions to the 65+ and 75+ populations. These distributions are then adjusted using relative differences observed in the disaggregated data. For example, if a municipality has an ISCED 6–8 share 10% above the national average, the estimated share for the older population is increased by 10%.

Core Geography

- Municipal code changes are matched using CBS match files, supplemented with limited manual coding of municipal mergers.
- Only one municipality cannot be matched in the pre-2000 period (Zeewolde).

Electoral Data Sources

- Election results are provided by the Dutch Government.
- Legislative elections (Tweede Kamer): 26 May 1981; 8 September 1982; 21 May 1986; 6 September 1989; 3 May 1994; 6 May 1998; 15 May 2002; 22 January 2003; 22 November 2006; 9 June 2010; 12 September 2012; 15 March 2017; 17 March 2021; 22 November 2023.

Electoral Geography

- Electoral data are matched backward from 2021 boundaries to 2018 municipal codes for 355 units (V1), and directly to 2021 boundaries for V2.
- See the core geography section above for details on boundary matching.

New Zealand

Human Capital and Population Data Sources

- Population census data for 1981, 1991, 2001, 2006, 2013, and 2018.
- Data from 2001 onward are publicly available from Statistics New Zealand census data.
- Data for 1981 and 1991 were purchased from Statistics New Zealand.

Educational Qualifications

- **ISCED 6–8**: Bachelor‘‘s degree and Level 7 qualifications; postgraduate and honours degrees; master‘‘s degree; doctoral degree (2006, 2013, 2018); bachelor‘‘s degree and higher degree (2001); bachelor‘‘s degree and postgraduate qualification (1991); bachelor‘‘s degree, postgraduate diploma, master‘‘s degree, and doctorate (1981).
- **ISCED 4–5**: Level 4 certificate; Level 5 diploma; Level 6 diploma (2006, 2013, 2018); basic, skilled, intermediate, and advanced vocational qualifications (2001); trade or New Zealand certificate; teacher, nursing, and technician certificates; university below bachelor‘‘s level (1991); non-university qualification and undergraduate diploma or certificate (1981).
- **ISCED 3**: No qualification; Level 1, Level 2, and Level 3 certificates (2006, 2013, 2018); sixth form qualification and higher school qualification (2001); school qualification equivalent to Level 3 (1991); seventh form, sixth form, UE, or scholarship (1981).
- **ISCED 0–2**: No qualification; Level 1 certificate (2006, 2013, 2018); no qualification and fifth form qualification (2001); no qualifications and school qualification equivalent to Level 2 (1991); no secondary education, third, fourth, or fifth form (1981). Individuals with missing qualification information are included in this category.

Age Ranges and Adjustments to Education Estimates

- Includes all residents aged 15 and over.
- Disaggregated age breakdowns are available.
- An adjusted variable excluding missing qualifications from both the numerator and denominator is provided.

Core Geography

- The primary geographic unit is the Territorial Authority, using 2018 boundaries; this is the main unit of local government in New Zealand.
- Auckland Territorial Authority contains approximately one-third of the national population. We therefore use its 13 wards to provide additional spatial granularity.
- The Chatham Islands Territory is excluded.
- Data for 2006, 2013, and 2018 are available from Statistics New Zealand at 2018 boundaries.
- Data for 2001 are aggregated to 2018 boundaries from 2001 Area Units (approximately 1,800 units) using areal-weighted spatial interpolation.
- Data for 1981 and 1991 were aggregated to 2018 boundaries by Statistics New Zealand.

Migration Data Sources

- Coverage is limited to census years 2001, 2006, 2013, and 2018.
- Data are publicly available from Stats NZ.
- Data are available for foreign-born individuals.

Electoral Data Sources

- Party vote shares for elections from 1999–2020 are matched by polling booth. Data are provided by Election New Zealand.
- Legislative elections (House of Representatives): 27 November 1999; 27 July 2002; 17 September 2005; 8 November 2008; 26 November 2011; 20 September 2014; 23 September 2017; 17 October 2020.

Electoral Geography

- The most disaggregated electoral data are reported at the polling-booth level, which is nested within electoral districts.
- Polling booths can be matched to Statistical Areas 2018 (SA2018; 1,633 units). However, approximately 400 SA2018 units cannot be consistently matched to any polling booths, resulting in units with no assigned voters. To ensure fully reliable results, electoral data are therefore collapsed to the ward level (243 units), where all polling booths are reliably matched.
- For 2017 and 2020, polling booths are matched using latitude and longitude. In earlier elections, where polling booths are located at the same sites, the 2017 and 2020 coordinates are reused.
- Polling booths that cannot be matched via coordinates are matched manually using address and state suburb information.

Norway

Flags

- Norwegian municipal reorganization means that post-2017 boundaries are not yet matched

Human Capital and Population Data Sources

- Register-based data covering 1970, 1980, and annually from 1986 to 2019.
- Data are publicly available from Statistics Norway (StatBank).
- Several adjustments are required to distinguish ISCED 5 and ISCED 6 qualifications, as boundaries between these categories cannot be directly observed. Additional adjustments are also made to harmonize age ranges.

Educational Qualifications

- **ISCED 7–8**: Tertiary education (long cycle), comprising higher education lasting more than four years.
- **ISCED 5–6**: Tertiary education (short cycle), comprising higher education lasting up to four years. The Norwegian NUS classification combines ISCED 5, 6, and 7 into a single category. More detailed data allowing full separation of ISCED 5 from ISCED 6 and 7 are not widely available. We therefore assume that one-third of individuals in this category hold ISCED 5 qualifications and that the remaining two-thirds hold ISCED 6 or higher qualifications in order to estimate the ISCED 6+ population.
- **ISCED 3–4**: Intermediate-level courses based on completed upper secondary education that are not accredited as higher education.
- **ISCED 3–4**: Upper secondary education. Prior to 2016, tertiary vocational education was included in this category. Tertiary vocational education comprises intermediate-level courses based on completed upper secondary education that are not accredited as higher education.
- **ISCED 2**: Basic school level.

Age Ranges and Adjustments to Education Estimates

- Available data include all residents aged 16 and over. We adjust to a 15+ reference population by assigning 15-year-olds to the ISCED 0–2 category.
- Statistics Norway notes: "For many immigrants, SSB has no information about their level of education. From the 2014 figures, level of education includes estimated levels of education for missing values for immigrants. For more information, see 'About the statistics'."

Core Geography

- LAU, 2018 boundaries.
- Municipal mergers are tracked manually. Some municipalities cannot be matched in 1970, while others are merged forward over time.
- Ongoing municipal consolidation in Norway has resulted in additional post-2018 boundary changes, with further changes in 2023. At present, we have not matched the new boundaries back, but future iterations of this dataset will do so.

Migration Data Sources

- Coverage is limited to 1970, 1980, and annually from 1986 to 2019.
- Migration data are also available at post-2018 municipal boundaries; these can only be matched using the corresponding education file (see Statistics Norway).
- Data are available for foreign-born individuals.

Electoral Data Sources

- Statistics Norway provides election results and turnout data.
- Legislative elections (Storting): 14 September 1981; 9 September 1985; 11 September 1989; 13 September 1993; 15 September 1997; 10 September 2001; 12 September 2005; 14 September 2009; 9 September 2013; 11 September 2017.
- Legislative elections (Storting): 13 September 2021 are available in a separate file, but cannot be fully matched (see above).

Electoral Geography

- See the core geography discussion above.

Portugal

Flags

- RHCD and RED at different units. RED fully nested.

Human Capital and Population Data Sources

- Population census data for 1981, 1991, 2001, 2011, and 2021 obtained from Pordata.
- The 1991 census is not available online and was obtained directly from INE Portugal.

Educational Qualifications

- **ISCED 6–8**: Higher education.
- **ISCED 4–5**: Middle-level education.
- **ISCED 3**: Upper secondary education.
- **ISCED 2**: Compulsory education, third cycle.
- **ISCED 1**: Compulsory education, first and second cycles.
- **ISCED 0**: No educational qualifications, including individuals unable to read or write.

Age Ranges and Adjustments to Education Estimates

- Data are available for all residents aged 15 and over.

Core Geography

- Human capital data are available at the *concelho* level.
- Boundary changes, including the creation of new *concelhos*, are addressed by matching *freguesias* over time. *Concelhos* with substantial boundary changes (Odivelas, Trofa, and Vizela) are assigned shares of their predecessor *concelhos* based on the population shares of constituent *freguesias*.

Migration Data Sources

- Migration data for 2001 and 2011 are publicly available from INE Portugal; data for 1981 and 1991 were obtained through direct requests to INE Portugal.
- Data are available for foreign-born individuals.

Electoral Data Sources

- Election results are obtained from the Interior Ministry.
- Legislative elections (Assembly of the Republic): 25 April 1983; 6 October 1985; 19 July 1987; 6 October 1991; 1 October 1995; 10 October 1999; 17 March 2002; 20 February 2005; 27 September 2009; 5 June 2011; 4 October 2015; 6 October 2019; 30 January 2022.

Electoral Geography

- Electoral data are reported at the *freguesia* level, which is more geographically disaggregated than the human capital data at the *concelho* level. *Freguesias* are fully nested within *concelhos*.
- Boundary splits are addressed by assigning population-weighted fractions based on historical *freguesia* populations, while mergers are manually traced over time using official match files.

Spain

Human Capital and Population Data Sources

- Population census data for 1981, 1991, 2001, 2011, and 2021.
- Series: INE Census.
- Series: INE 2021 Census.
- Publicly available data cover 1991, 2001, 2011, and 2021.
- The 1981 census was provided separately by INE. These data are not fully reliable due to a substantial population undercount; the counted population is disproportionately more educated. Results for 1981 should therefore be used with caution.
- Census education data are reported for individuals living in family units.

Educational Qualifications

- Reporting changes in the 2021 census follow the CNED-2014 classification.
- **ISCED 6–8**: *Tercer grado.* Includes the following qualifications across censuses: *arquitecto e ingeniero técnico y diplomado* (approved completion of the third year) from higher technical schools; *diplomado* from university schools and faculties; *arquitecto o ingeniero superior*; *licenciado universitario*; higher non-university degrees; doctoral degrees; postgraduate and specialization degrees for licenciados; *diplomatura universitaria*, technical architecture, technical engineering or equivalent; university bachelor‘‘s degree or equivalent; official master‘‘s degree (from 2006 onward); medical specialties or equivalent; doctorate.
- **ISCED 4–5**: *Formación profesional* (second level); industrial mastery; *bachiller superior*; higher-level vocational training.
- **ISCED 2**: Completion of the final year of ESO, EGB, or elementary baccalaureate; possession of a school-leaving or primary studies certificate; elementary baccalaureate; *graduado escolar*; completed EGB or equivalent; first-level vocational training; industrial official qualification; other intermediate qualifications.
- **ISCED 2**: Primary education or completion of five years of EGB or equivalent; attendance at school for five or more years without completing the final year of ESO, EGB, or elementary baccalaureate (2011).
- **ISCED 0–1**: Illiterate individuals and those with no formal education.

Age Ranges and Adjustments to Education Estimates

- Data include all residents aged 16 and over.
- We adjust to a 15+ reference population by assigning 15-year-olds to the ISCED 0–2 category using census population estimates.

Core Geography

- LAU, 2018 boundaries (V1), 2021 boundaries (V2)
- Due to data suppression in small municipalities, detailed population–education tabulations undercount the population by approximately 1% in 2011 (an undercount of 397,000 individuals aged 15+). Undercounting is more severe in Galicia and Asturias, where approximately 4.5% of the population is missing.
- In 2021, education totals are unavailable for 568 small municipalities (fewer than 50 residents), and education-by-gender data are unavailable for 1,356 municipalities, affecting approximately 80,000 individuals.
- Numerous municipalities split between 1986 and 1993. To harmonize data forward, original municipalities are divided proportionally among successor units. Pre-1996 data should be used with caution for the following municipalities: Albacete (2003), Pozo Cañada (2901), Almoradí (3015), Orihuela (3099), Pilar de la Horadada (3902), Los Montesinos (3903), Berja (4029), Balanegra (4904), Badajoz (6015), Valdelacalzada (6901), Pueblonuevo del Guadiana (6902), Cervelló (8068), Cerdanyola del Vallès (8266), Badia del Vallès (8904), La Palma de Cervelló (8905), Jerez de la Frontera (11020), Medina-Sidonia (11023), Benalup-Casas Viejas (11901), San José del Valle (11902), Vilafamés (12128), Sant Joan de Moró (12902), Ortigueira (15061), Cariño (15901), Guadalajara (19130), Marchamalo (19171), Donostia (20069), Astigarraga (20903), Beas de Segura (23012), Arroyo del Ojanco (23905), Cervo (27013), Burela (27902), Colmenar Viejo (28045), Tres Cantos (28903), Antequera (29015), Málaga (29067), Torremolinos (29901), Villanueva de la Concepción (29902), Ansoáin (31016), Arakil (31025), Cizur (31076), Galar (31109), Cendea de Olza (31193), Berrioplano (31902), Berriozar (31903), Irurtzun (31904), Beriáin (31905), Orkoien (31906), Zizur Mayor (31907), Vilanova de Arousa (36061), A Illa de Arousa (36901), Palazuelos de Eresma (40155), San Cristóbal de Segovia (40906), Lebrija (41053), La Luisiana (41056), La Puebla del Río (41079), Cañada Rosal (41901), Isla Mayor (41902), El Cuervo de Sevilla (41903), El Perelló (43104), Tarragona (43148), Vila-seca (43171), Salou (43905), L'Ampolla (43906), La Canonja (43907), Paterna (46190), San Antonio de Benagéber (46903), Barakaldo (48013), Durango (48027), Iurreta (48910), Alonsotegi (48912), Zaragoza (50297), Villamayor de Gállego (50903).

Migration Data Sources

- Coverage is limited to 1991, 2001, and 2011.
- Data are publicly available from INE Spain.
- Data are available for individuals with a foreign nationality (non-citizens).
- In 2011, information on foreign nationals is missing for 3,796 municipalities (46%), as INE suppresses counts below five individuals. These municipalities are predominantly small rural areas, with an average population of 560 (compared to a national average of 5,195).

Electoral Data Sources

- Election results are obtained from the Ministry of the Interior.
- Due to extensive boundary changes during the 1980s, electoral data collection begins in 1986.

- Legislative elections (Cortes Generales): 22 June 1986; 29 October 1989; 6 June 1993; 3 March 1996; 12 March 2000; 14 March 2004; 9 March 2008; 20 November 2011; 20 December 2015; 26 June 2016; 28 April 2019; 23 July 2023.

Electoral Geography

- See the core geography discussion above.

Sweden

Human Capital and Population Data Sources

- Register-based data are available from 1985 to 2024.
- Data are publicly available from Statistics Sweden.
- Adjustments to age ranges are required; these rely on additional population-based data.

Educational Qualifications

- **ISCED 6–8**: Postgraduate education (ISCED97-6); post-secondary education of three years or more (ISCED97-5A).
- **ISCED 5**: Post-secondary education of three years or less (ISCED97-4A and 5B).
- **ISCED 3**: Upper secondary education (ISCED97-3A); upper secondary education of two years or less (ISCED97-3B).[6]
- **ISCED 2**: Primary and lower secondary education (ISCED97-2).
- **ISCED 1**: Less than nine years of education.

Age Ranges and Adjustments to Education Estimates

- The original data include all residents aged 16–75.
- We adjust the lower age threshold by incorporating population data for 15-year-olds, all of whom are assigned to the ISCED 0–2 category.
- We adjust the upper age threshold by averaging educational attainment across the two oldest five-year age groups and assuming a linear trend across educational categories for the excluded population aged 75+. Educational shares are then estimated using municipality-level population counts for this group and added to the aggregate ISCED totals.
- Detailed age breakdowns are available.

Core Geography

- LAU, 2018 boundaries (290 units in 2018).
- No major boundary changes occurred during the period; units are matched consistently over time.

Migration Data Sources

- Migration data are publicly available from Statistics Sweden.
- Data are available for individuals with a foreign nationality (non-citizens).

[6] Because the ISCED97-3B category confers a full upper secondary qualification that permits further study, we code it as upper secondary rather than lower secondary. This choice aligns with the conventions used by Barro and Lee and by the OECD; see the validation section in the dataset paper.

Electoral Data Sources

- Election results are provided by Statistics Sweden.
- Legislative elections (Riksdag): 19 September 1982; 15 September 1985; 18 September 1988; 15 September 1991; 18 September 1994; 20 September 1998; 15 September 2002; 17 September 2006; 19 September 2010; 14 September 2014; 9 September 2018; 11 September 2022.

Electoral Geography

- See the core geography section above.

Switzerland

Human Capital and Population Data Sources

- National censuses: 1970, 1980, 1990, and 2000.
- Structural Survey 2010–2014 and 2015–2019.
- Data were provided to us in Excel format in April 2020 by the Swiss Federal Statistical Office.

Educational Qualifications

- A common set of education categories is available for all years.
- **ISCED 4–8**: *Terti'rstufe*.
- **ISCED 6–8 (corrected)**: An adjustment is applied to estimate the share of individuals with ISCED 4–8 qualifications attaining ISCED 6–8, using IPUMS-I microdata at the NUTS2 level combined with national cohort data (ten-year birth cohorts).
- **ISCED 4–5 (corrected)**: An adjustment is applied to estimate the share of individuals with ISCED 4–8 qualifications attaining ISCED 4–5, using IPUMS-I microdata at the NUTS2 level combined with national cohort data (ten-year birth cohorts).
- **ISCED 3**: *Sekundarstufe II*.
- **ISCED 2**: *Sekundarstufe I*.

Age Ranges and Adjustments to Education Estimates

- Educational attainment is measured for the population aged 15 and over.
- Detailed age breakdowns are available.

Core Geography

- LAU (municipalities), using 2019 boundaries (V1), 2025 boundaries (V2)
- Municipal data for 1970–2000 are harmonized to 2010 boundaries.
- Data for 2010–2014 use 2014 boundaries.
- Numerous municipal mergers occurred during the period; these are tracked using official municipal match files, supplemented with limited manual matching.
- Most boundary changes involve mergers. The Swiss Federal Statistical Office cautions against over-interpreting trends in municipalities affected by mergers.
- A small number of municipal splits are addressed by allocating population shares based on pre-split population distributions.

Migration Data Sources

- Migration data are publicly available from the Swiss Federal Statistical Office.
- Data are available for individuals with a foreign nationality (non-citizens).

Electoral Data Sources

- Election results are obtained from official federal government sources.
- Legislative elections (Conseil national / Nationalratswahlen): 23 October 1983; 18 October 1987; 20 October 1991; 22 October 1995; 24 October 1999; 19 October 2003; 21 October 2007; 23 October 2011; 18 October 2015; 20 October 2019.
- In Switzerland, disaggregated election results up to 2015 are reported using multiple counted ballots. We therefore convert party vote shares into notional "single votes" by multiplying party vote shares by the total number of ballots cast. As a result, total vote counts at the local level exceed the number of voters. Party vote shares are accurate, but absolute vote totals should be interpreted with caution.

Electoral Geography

- In small municipalities, disaggregated results are sometimes suppressed, with votes for major parties aggregated into an "other" category. This can generate large apparent swings over time (e.g. in Glarus).
- Postal votes and votes cast by Swiss citizens abroad are not included, leading to discrepancies between aggregate national results and the geocoded election data used here.

UK – Great Britain

Human Capital and Population Data Sources

- UK Census data for 1981, 1991, 2001, and 2011 for England, Wales, and Scotland; 2021 Census data for England and Wales only; 2022 Census data for Scotland.
- Data for Northern Ireland are not included.
- All data for England and Wales are obtained from Nomisweb. Data for Scotland in 1981, 1991, 2001, and 2022 are obtained from Casweb, while data for Scotland in 2011 are obtained from National Records of Scotland.
- 1981 data are drawn from Table SAS81-48, *Qualified Manpower* (10% sample).
- 1991 data are drawn from Table SAS84, *Qualified Manpower* (10% sample).
- 2001 data for England and Wales are drawn from Table ST105 (*Sex and age by highest level of qualification*) and Table ST115 (*Count of qualifications by sex*).
- 2001 data for Scotland are drawn from Table ST204 (*Gender and age by highest level of qualification*).
- 2011 data for England and Wales are drawn from Table DC5107EWla (*Highest level of qualification by sex by age*) and Table QS502EW (*Qualifications gained*).
- 2011 data for Scotland are drawn from Table DC5102SC (*Highest level of qualification by sex by age*).
- 2021 data for England and Wales are drawn from RM055 (*Highest level of qualification by sex*, LSOA).

Educational Qualifications

- Qualifications reported in census tables require substantial harmonization to ensure comparability across years and between England/Wales and Scotland. Coding schemes change fundamentally between the 1991 and 2001 censuses.
- Census questions in 1981 and 1991 only record post-school qualifications. As a result, it is only possible to consistently identify qualifications at ISCED 4+ over the full 1981–2011 period. We therefore report education outcomes only for ISCED 6–8.
- In 2001 and 2011, qualifications are coded into five levels.[7]
- Only in England and Wales in 1991 is there a direct measure of ISCED 6–8. In all other cases, ISCED 6–8 attainment is estimated.
- **ISCED 5–8 (2021, England and Wales)**: Level 4 qualifications or above, including degrees, higher degrees, NVQ Levels 4–5, HNC/HND, and professional qualifications. These include some ISCED 5 qualifications.

[7]England and Wales Census 2001 classification: No qualifications; Level 1; Level 2; Level 3; Level 4/5; Other/unknown. The specific qualifications included in each level differ between England/Wales and Scotland. See UK Data Service documentation.

- **ISCED 6–8 (corrected, 2021)**: To isolate ISCED 6–8, we estimate NUTS1-level gender-specific degree shares using the Labour Force Survey (LFS). These shares are applied to census counts to divide raw totals into ISCED 6–8 and ISCED 3–5.
- **ISCED 4–8 (2001, 2011)**: England and Wales Level 4/5 qualifications; Scotland Level 3 and Level 4 qualifications. For 1981 and 1991, we use Level A (higher degrees), Level B (first degrees), and Level C (post-secondary but sub-degree qualifications).[8]
- **ISCED 6–8 (corrected, 2011 England and Wales)**: We estimate ISCED 6–8 attainment using LAU-level tables that report the total number of qualifications obtained (double-counting individuals with multiple qualifications). Counts of degree and higher-degree qualifications are used, with gender shares estimated using 2001 patterns adjusted by observed changes in Scotland between 2001 and 2011.
- **ISCED 6–8 (corrected, 2001 England and Wales)**: Estimated using LAU-level counts of degree and higher-degree qualifications from census tables that report total qualifications obtained.
- **ISCED 6–8 (corrected, 2001 and 2011 Scotland)**: The highest education category includes some post-secondary vocational qualifications (mainly nursing and teaching). We use Labour Force Survey microdata to estimate the share of degree holders within this group, distinguishing between Strathclyde and the rest of Scotland.
- **ISCED 6–8 (corrected, 1991 England and Wales)**: Level A (higher degrees) and Level B (first degrees and equivalent qualifications).
- **ISCED 6–8 (corrected, 1981 UK and 1991 Scotland)**: Estimated using 5% census microdata (Sample of Anonymised Records) from the UK Data Service. These data provide substantial geographic detail, allowing estimation of the share of ISCED 6–8 attainment among those with ISCED 4–8 qualifications at the LAU level.

Age Ranges and Adjustments to Education Estimates

- Population aged 18 and over in 1981 and 1991.
- Population aged 16–74 in 2001.
- Population aged 16 and over in 2011.
- For 2001, attainment among those aged 75+ is estimated using the educational distribution of individuals in the same area from comparable birth cohorts in the 1991 census (men aged 65+ and women aged 60+).

Core Geography

- LAU, 2018 boundaries. In England and Wales, LAUs correspond to Local Authority Districts. In Scotland, LAUs generally correspond to Council Areas; in some rural areas they correspond to Local Enterprise Company (LEC) areas, which provide finer geographic detail.
- All data for England and Wales are available aggregated to LAU boundaries via the NOMIS interface.

[8] See UK Data Service documentation: UK Census 1991 Definitions.

- Data for Scotland are manually aggregated to LAUs from lower-level units: Data Zones (2011), Standard Table Postcode Sectors (2001), Wards (1991), and Postcode Sectors (1981). For 1991–2011, aggregation uses spatial overlap estimates; in 1981, area-weighted spatial interpolation is applied.

Migration Data Sources

- Migration data are publicly available from the Office for National Statistics and the Scottish Census.
- Data are available for foreign-born individuals.

Electoral Data Sources

- Raw constituency-level election results are available from the House of Commons Library.
- UK election results are reported at the parliamentary constituency level; polling-booth-level data are not available. We therefore use a spatial allocation procedure to link constituencies to local areas.
- Legislative elections (House of Commons): 9 June 1983; 11 June 1987; 9 April 1992; 1 May 1997; 7 June 2001; 5 May 2005; 6 May 2010; 7 May 2015; 8 June 2017; 12 December 2019; 4 July 2024.
- The SDP and Liberal Democrats are coded as a single party category (Liberals).

Electoral Geography

- Because election results are not reported below the constituency level, results are imputed to local authorities.
- We match 2016 census wards to parliamentary constituencies using shapefiles from the UK Data Service Census Boundary Dataset and the `sf` package in R.
- Wards are assigned shares of constituency vote totals based on spatial overlap and then aggregated to local authority units.
- For example, if a ward constitutes 0.007 of a constituency‘‘s area, it is assigned 0.7% of that constituency‘‘s vote for each party before aggregation.
- Matching is not perfectly precise, resulting in a small margin of error (below 1%) in aggregate vote totals. For example, in 2019 total votes are underestimated by approximately 100,000 out of 37 million, reflecting rounding in area shares rather than missing data.

USA

Human Capital and Population Data Sources

- Decennial population censuses for 1980, 1990, and 2000.[9]
- American Community Survey (ACS) five-year files: 2005–2009; 2008–2012; 2014–2018; and 2019–2023.[10]
- All data are obtained from IPUMS NHGIS.

Educational Qualifications

- **ISCED 6–8**: College education of four, five, or six or more years (1980); bachelor‘‘s degree, graduate degree, or professional degree (1990 Census; 2000 Census; 2005–2009, 2008–2012, 2014–2018 ACS).
- **ISCED 5**: College education of two or three years (1980); some college with no degree or associate‘‘s degree (1990 Census; 2000 Census; 2008–2012 ACS); some college with no degree or associate‘‘s degree (2005–2009 and 2014–2018 ACS).
- **ISCED 3**: High school education of three or four years; college education of one year (1980); high school graduate, GED, or alternative credential (1990 Census; 2000 Census; 2005–2009, 2008–2012, 2014–2018 ACS).
- **ISCED 2**: High school education of one or two years (1980); 9th to 12th grade with no diploma (1990 Census; 2000 Census; 2005–2009, 2008–2012, 2014–2018 ACS).
- **ISCED 0–1**: Never attended school, nursery school, or kindergarten (1980); elementary school (1–4 years, 5–7 years, or 8 years), or less than 9th grade (1990 Census; 2000 Census; 2005–2009, 2008–2012, 2014–2018 ACS).

Age Ranges and Adjustments to Education Estimates

- Educational attainment is reported for individuals aged 18 and over in all years.
- Individuals aged 15–17 are assumed to hold ISCED 0–2 qualifications and are added using age-by-gender population breakdowns.
- Detailed age breakdowns are available from 1990 onward. The 1980 census allows educational attainment to be distinguished for ages 15–24 and 25+.

Core Geography

- LAU units correspond to counties.
- County name changes and mergers are manually corrected using U.S. Census Bureau county change files.

[9] 1980: Table NTPB46 (*Sex by Age by Years of School Completed*); 1990: Table NPB45 (*Sex by Age by Educational Attainment*); 2000: Table NPCT025B (*Population 18 Years and Over by Sex by Age Groups by Educational Attainment*).

[10] Table B15001: *Sex by Age by Educational Attainment for the Population 18 Years and Over.*

- A small number of county splits or complex boundary changes are manually harmonized backward.
- Complex boundary changes in rural areas of Alaska cannot be fully reconciled over time.
- Approximately 99.5% of geographic units are successfully matched over time.

Migration Data Sources

- Migration data are publicly available from IPUMS NHGIS.
- Data are available for foreign-born individuals.

Electoral Data Sources

- Presidential vote shares by party are obtained from ICPSR and the MIT Election Data and Science Lab.
- Voter turnout is not reported in the primary source files but can be merged from auxiliary datasets.

Electoral Geography

- Due to boundary changes in rural Alaska, electoral data for Alaska are not available.