

Population and migration statistics transformation in England and Wales, technical topic guide: 2023

Supplementary information on research using administrative data for the future of population and migration statistics.

Contact:
Vicky Collison, Paulina
Galezewska, Elizabeth Pereira,
Emilie Woodhall
2023Consultation@ons.gov.uk
+44 1329 444972

Release date:
26 June 2023

Next release:
To be announced

Table of contents

1. [Overview](#)
2. [Population stock](#)
3. [International migration](#)
4. [Internal migration](#)
5. [Alternative population bases](#)
6. [Labour market status](#)
7. [Income](#)
8. [Ethnicity](#)
9. [Income by ethnicity](#)
10. [Occupied address](#)
11. [Communal establishments](#)
12. [Housing stock](#)
13. [Housing by ethnicity](#)
14. [Education](#)
15. [Morbidity and mortality](#)
16. [Veterans](#)

17. [Estimation of travel to work matrices](#)
18. [Outcomes over time \(Refugee Integration Outcomes\)](#)
19. [Glossary](#)
20. [Cite this methodology](#)

1 . Overview

This topic guide provides further information on topics that are part of the evidence being published to support the public consultation on the National Statistician's forthcoming recommendation on the future of population statistics. This includes definitions, methods and limitations or caveats related to research using administrative data and can be used to gain high level understanding of topics and their methods.

This technical topic guide accompanies the [Local authority case studies](#) and [Research Overview](#) being published as part of the consultation.

The statistics and research referred to in the topic guide are in the research and development stage and while some are [Experimental Statistics](#), they currently have limited use for policy or decision-making. If you use these data, please include the Experimental Statistics label in your output.

Across many of the topics detailed in this guide, there are key differences between our administrative-based methods and those currently used in the production of official statistics. These differences, generally in the methods and/or definitions used, mean that making direct comparisons can be challenging. More information on how specific topics may be impacted is available in the following relevant sections.

2 . Population stock

What is being measured

Usually resident population at the mid-year point (June). Outputs are produced by age, sex and local authority. We have also produced Lower layer Super Output Area estimates for five local authorities.

How is it being measured

The dynamic population model (DPM) uses statistical modelling techniques to combine a range of data sources, the [Statistical Population Datasets](#), [international migration estimates](#), [estimates of internal moves](#) and demographic insights to estimate the population and population change. This provides a coherent statistical framework for more timely population statistics known as the [admin-based population estimates \(ABPEs\)](#).

Things you need to know

To provide the best possible population estimates from the DPM we incorporate Census 2021-based mid-year estimates (MYE) as our best picture of the population in June 2021. ABPEs for June 2021 are very similar to Census 2021-based MYE.

The [Statistical Population Dataset \(SPD\)](#) is a key input and approximates the usually resident population using administrative data before any modelling is applied.

Related links

[Admin-based population estimates: provisional estimates for local authorities in England and Wales, 2011 to 2022](#)

- Admin-based population estimates for all local authorities in England and Wales from the dynamic population model.

[Dynamic population model, improvements to data sources and methodology for local authorities in England and Wales: 2011 to 2022](#) - Developments of methods and data used in the dynamic population model.

[Developing Statistical Population Datasets, England and Wales: 2021](#) - Aggregate comparisons between the Statistical Population Dataset version 4.0 (v4.0) and Census 2021.

[Transforming population statistics, comparing 2021 population estimates in England and Wales](#) - Evaluating progress towards a transformed population statistics system, using comparisons between census-based and admin-based population estimates and Census 2021.

[Understanding quality of Statistical Population Dataset in England and Wales using the 2021 Census - Demographic Index linkage](#) - Analysis of Statistical Population Dataset version 4.0 2021 using a linkage between Census 2021 and the Demographic Index.

3 . International migration

What is being measured

We produce experimental and provisional estimates of long-term international migration flows using administrative data known as the admin-based migration estimates. Our latest estimates use different data sources and methods for each nationality grouping. We currently publish estimates on immigration, emigration and net migration for non-European Union (EU) nationals, EU nationals and British nationals. We continue to use the United Nations' (UN) definition of a long-term migrant: a person who moves to a country other than that of their usual residence for at least a year.

How is it being measured

For migration estimates of non-EU nationals, we use [Home Office Borders and Immigration data](#), which combines visa and travel information. Our latest methodology to estimate the migration of EU nationals is based on our [methods for measuring international migration using Registration and Population Interaction Database \(RAPID\) administrative data](#). RAPID currently provides the best insight into the migration of EU nationals. While we continue our research into which administrative data source is best to measure migration patterns of British nationals, we continue to measure this group using the [International Passenger Survey \(IPS\)](#).

Things you need to know

In the future, the revision of long-term international migration statistics will be an important part of the production of these estimates. Provisional estimates are released with the expectation they may be revised as more complete data become available. In addition, our methods are still evolving, and we will therefore revise the estimates as our methods mature. Further information on our approach to revisions can be found in our [Population and International Migration Statistics Revisions Policy](#).

Related links

[Long-term international migration, provisional: year ending December 2022](#) - Experimental and provisional estimates of UK international migration, 2018 to 2022.

[Methods to produce provisional long-term international migration estimates](#) - An explanation of the methods used to produce the latest provisional experimental statistics on migration flows into and out of the UK.

[International migration research, progress update: May 2023](#) - An update on international migration methods and research.

4 . Internal migration

What is being measured

Estimates of internal moves within England and Wales at local authority level, as well as cross-border flows to and from Scotland and Northern Ireland, between one mid-year point and the next, by age and sex.

How is it being measured

Estimates of internal moves and cross-border flows are produced using information on address changes in administrative data, which is considered a proxy for residential moves. Different methods are currently being explored to account for limitations of administrative data, such as lagging in updates to address information.

Things you need to know

Official estimates of internal migration and cross-border moves are produced using a [different methodology](#). These estimates are lagged: for example, mid-2022 estimates are published in 2023. As part of the work to provide more timely population estimates, provisional internal migration estimates are calculated using data sources that are available shortly after the reference date. The [dynamic population model \(DPM\)](#) scales the provisional estimates according to how they relate to the official estimates in previous years.

Related links

[Population estimates for the UK, mid-2020: methods guide - Office for National Statistics \(ons.gov.uk\)](#) - official population estimates method guide, section 5 and appendix 12 describes the internal migration method.

[Dynamic population model, improvements to data sources and methodology for local authorities, England and Wales: 2011 to 2022 - Office for National Statistics \(ons.gov.uk\)](#) - latest update on the dynamic population model, input data and methods used, including a section on internal migration and cross-border moves methodology.

[Dynamic population model for local authority case studies in England and Wales - Office for National Statistics \(ons.gov.uk\)](#) - section on internal migration and cross-border moves methods.

5 . Alternative population bases

What is being measured

We are looking at alternative population bases in response to user demand and are exploring what is possible. There are many potential ways of defining populations and migrations flows, such as the [experimental daytime population estimates](#) we have produced for 14 local authorities, which estimate population size by time of day.

How is it being measured

Daytime population estimates use survey, administrative, commercial and open data from a range of sources.

Things you need to know

Official population statistics estimate the usually resident population at their place of residence. The population moves over the course of the day as people engage in different activities like work or study, thus the daytime accuracy of these traditional estimates changes in line with population mobility. Our experimental statistics aim to estimate the number of people present in a particular place at a given date and time.

Related links

[Population and migration estimates - exploring alternative definitions: May 2023](#) - Considering new ways of estimating the population to enhance our existing statistics.

6 . Labour market status

What is being measured

Labour market status (economic activity) of the usually resident population of England and Wales, aged 16 years and over, by tax year.

How is it being measured

The admin-based labour market status (ABLMS) statistics are produced using a combination of income and education administrative data sources, using the [Statistical Population Dataset \(SPD\)](#) to provide a base population. Rules are used to assign a labour market status to each individual and to make decisions where multiple labour market statuses are possible.

Things you need to know

These statistics are the result of feasibility research using a methodology different to that currently used to produce [labour market statistics](#). There are key differences in our methodology along with the limitations of administrative data, which impact their use in practice. They have limited use for policy or decision-making.

Related links

[Feasibility research into admin-based labour market status for England and Wales: tax year ending 2016](#) - Feasibility research into producing a measure of admin-based labour market status (ABLMS) using administrative data for the tax year ending 2016.

[Comparison of labour market data sources](#) - The strengths and weaknesses of the main data sources we use to produce the labour market figures, including the advantages of new administrative data sources and limitations of some of our published figures.

7 . Income

What is being measured

Experimental gross and net, individual and occupied address income statistics for small areas in England and Wales by tax year.

How is it being measured

The admin-based income statistics (ABIS) are produced using a variety of administrative data sources from the Department for Work and Pensions and HM Revenue and Customs to determine a total income value per individual per tax year, using the [Statistical Population Dataset \(SPD\)](#) to provide a base population. Development of the ABIS is guided by the [Canberra Group Handbook on Household Income Statistics](#), which provides guidance on defining and measuring components of income.

Things you need to know

The ABIS are [Experimental Statistics](#). They currently have limited use for policy or decision-making, as both the income measure and coverage are in development. The [income estimates for small areas](#) are the recommended estimates to use for household income for small areas.

Related links

[Admin-based income statistics, England and Wales: tax year ending 2018](#) - Experimental gross and net, individual and occupied address income statistics for small areas using administrative data from Pay As You Earn, Self Assessment and benefit systems.

[Admin-based income statistics QMI](#) - Quality and Methodology Information (QMI) for admin-based income statistics in England and Wales, detailing the strengths and limitations of the data, methods used, and data uses and users.

[Income and earnings statistics guide](#) - Explains the relationship between income and earnings data and outlines the statistics produced by the ONS, Department for Work and Pensions and HM Revenue and Customs.

8 . Ethnicity

What is being measured

Ethnic group of the usually resident population of England and Wales at 5- and 18-category level produced at the national, regional, local authority and Lower layer Super Output Area (LSOA) levels. The admin-based ethnicity statistics (ABES) produces high-level ethnicity by age band and sex produced at the national, regional and local authority levels.

How is it being measured

The [Statistical Population Dataset \(SPD\)](#) is used to provide a base population to which ethnicity information is added from eight [administrative sources](#). A [number of rules](#) are used to select one ethnicity per individual where multiple are recorded. This method produces the admin-based ethnicity dataset (ABED), which the ABES are generated from.

Things you need to know

Our [official ethnicity statistics](#) currently use the census. This transformed approach uses a different methodology and aims to produce more frequent and timely updates, as well as producing outputs at lower levels of geography. Some differences can be observed in the admin-based versus census-based outputs, for example, the proportion of individuals in the “White” category is [higher in the ABES than in census](#).

Additionally, the 18-category breakdown for “White” and “Other” is not accurate because of data issues resulting from some data sources not following the harmonised ethnicity standard.

At more granular levels of geography especially LSOA level, the statistics produced are impacted by the statistical disclosure control rules applied to the outputs meaning we are unable to publish ethnicity for every LSOA.

Related links

[Developing admin-based ethnicity statistics for England and Wales: 2020](#) - Research update on producing population statistics by ethnic group for England and Wales from administrative data, with comparisons with Census 2021 estimates.

[Producing admin-based ethnicity statistics for England: methods, data and quality](#) - An overview of methods, data sources and data quality for the feasibility research on producing statistics on the population by ethnic group.

9 . Income by ethnicity

What is being measured

Subnational multivariate median net individual income by ethnicity statistics for the usually resident population of England and Wales using administrative data sources.

How is it being measured

To produce the admin-based income by ethnicity statistics (ABIES), this research combines the [admin-based income statistics \(ABIS\) dataset](#) with the [admin-based ethnicity dataset \(ABED\) V3.0](#). These datasets are derived from multiple administrative data sources, which have been linked to produce statistics about income by ethnic group for England and Wales. Both datasets use the [Statistical Population Dataset \(SPD\)](#) as a population base.

Things you need to know

Given the ABIES uses the [SPD](#), [ABIS](#) and the [ABED](#) as their base, any quality or coverage considerations in their methods or outputs also apply to the ABIES. The income by ethnic group data are based on the proportion of the population that have income information identified and a stated ethnic group for in the ABIES dataset. Those for whom we are able to identify income information may not be representative of the population at large. The smaller sizes of some of the ethnic groups make them more susceptible to outlier income amounts, which could explain some of the differences seen. The ABIS are still under development and do not capture some components of income, such as investment income and income from rental and royalties. These gaps have the potential to affect the accuracy of the income figures by ethnic group.

Related links

[Developing subnational multivariate income by ethnicity statistics from administrative data, England and Wales: tax year ending 2018](#) - Update on feasibility research producing income by ethnic group statistics for England and Wales from administrative data.

[Developing subnational multivariate income by ethnicity statistics from administrative data, England: tax year ending 2016](#) - Feasibility research on an initial case study producing income by ethnic group statistics for England from administrative data.

[Methods for producing multivariate population statistics using administrative and survey sources \(PDF\)](#) - This paper provides an outline for the programme of methodological work to produce multivariate population outputs, which are primarily based on administrative data but use survey and other data sources to provide robust outputs that account for missingness and other data problems.

10 . Occupied address

What is being measured

The number and resident count of occupied addresses in England and Wales, comparing administrative data against the census. Outputs produced by national (England and Wales) and local authority level, using 2021 data.

How is it being measured

The [Statistical Population Dataset \(SPD\)](#) is used to provide a base population. We then use [Unique Property Reference Number \(UPRN\)](#) information from the Personal Demographic Service and English School Census to assign people to an address, removing those without a UPRN. Those found in communal establishments and other non-household addresses are removed and used in the admin-based communal establishment dataset (ABCED). This creates the admin-based occupied address dataset (ABOAD), referred to in previous publications as the admin-based household estimate (ABHE).

Things you need to know

Our methods for producing household statistics differ between the census and the ABOAD. It is not currently possible to produce [census-defined household](#) information using administrative data. Instead, the ABOAD looks at occupied addresses, which are addresses with at least one usual resident. This difference in definitions means that the data may not be directly comparable. We found that less than 1% of UPRNs in the census contained more than one household. Work is ongoing to understand the areas and the populations potentially impacted by this definitional difference.

In addition, not all records in the SPD are assigned a UPRN and in some instances, we do not know if the UPRN is a household or communal establishment address. Further work is being conducted to understand the impact of the differences in definitions, to explore those records without a UPRN and those we cannot confidently assign to the household population to understand what bias might be introduced during our processes.

Related links

[Population and household estimates, England and Wales: Census 2021](#) - Census 2021 unrounded population and household estimates for England and Wales, by sex and single year of age.

[Research Outputs: An update on developing household statistics for an Administrative Data Census](#) - This article presents an update on our research into producing household statistics from administrative data sources. These Research Outputs are not official statistics.

11 . Communal establishments

What is being measured

The number of people in England and Wales living in a communal establishment (CE). Produced by age and sex at a local authority (LA) level and CE type at LA.

How is it being measured

Using data from the Ministry of Justice (MoJ) and Higher Education Statistics Agency (HESA), alongside those flagged as living in communal establishments in the admin-based occupied address dataset (ABOAD) method, the admin-based communal establishment dataset (ABCED) is produced.

Things you need to know

The CE address information in the ABCED has a different reference period compared with the occupied address data used in the ABOAD, meaning the data are not directly comparable. In future releases this issue will be rectified.

Census CE statistics require a person to usually reside within a CE to be counted as part of this population. We are not able to establish the duration of an individual's residence in a CE using our current administrative data sources to determine whether the duration meets the minimum six months residence.

These data should not be used for decision making purposes. These figures are not directly comparable with official HESA halls of residence figures or MoJ data on prisoners.

Further investigation of coverage of administrative data is required for a small number of living arrangements because of addressing challenges (for example, transient population groups, temporary accommodation sites), access restrictions (for example, royal households, embassies) and/or because their residents may not interact with administrative sources or surveys in the same way as other populations (for example, those who are homeless). These are known as Special Population Groups.

Related links

[Communal establishment residents, England and Wales: Census 2021 Communal establishment residents, England and Wales: Census 2021](#) - Information about people who live in communal establishments, including age, sex and type of establishment, Census 2021 data.

[Understanding quality of Statistical Population Dataset in England and Wales using the 2021 Census - Demographic Index linkage](#) - Analysis of Statistical Population Dataset version 4.0 2021 using a linkage between Census 2021 and the Demographic Index.

[Understanding quality of linked administrative data sources in England and Wales, using the 2021 Census – Demographic Index linkage](#) - Analysis of linkage between the Demographic Index and linked census and Census Coverage Survey to understand the quality of administrative data sources in England and Wales.

[Design of Address Frame, Collection and Coverage Assessment and Adjustment of Communal Establishments in 2021 Census](#) – Insight into the collection of CEs and SPG in the Census 2021.

12 . Housing stock

What is being measured

Number of all and occupied household addresses by accommodation type, number of bedrooms, number of rooms, number of bathrooms, and property build period, at local authority level in England and Wales.

How is it being measured

All and occupied addresses and their characteristics are identified using the admin-based housing stock (ABHS) dataset. The ABHS was produced by linking a residential Address Frame to Valuation Office Agency (VOA) data to obtain property characteristic information, and to the [admin-based occupied address dataset \(ABOAD\)](#) to identify occupied addresses.

Things you need to know

The residential [Address Frame](#) was built to suit Census 2021 operational purposes, rather than for use with administrative data. The ABHS therefore has more residential addresses than the final Census 2021 estimates.

Methodological challenges in the ABOAD dataset means that it identifies a lower proportion of occupied addresses (especially flats) than Census 2021. This is therefore reflected in the ABHS.

The VOA accommodation type (Census 2021 definition) variable is derived from VOA property type and VOA dwelling code variables to resemble the eight category Census 2021 accommodation type variable as closely as possible, while adding an additional ninth category for annexes. Our research identified explainable difference between VOA, data that is collected by trained surveyors, and self-reported census data. For example, census respondents prefer to define their property as “semi-detached” instead of “end-terrace”.

The VOA number of bedrooms variable is defined differently to the census number of bedrooms question. Most notably, the VOA exclude rooms smaller than four square metres and include rooms built as bedrooms even if not used as such. The VOA also state that studio or bedsit accommodation with a combined living room and bedroom should be recorded as having one room and one bedroom.

Related links

[Admin-based housing stock profile for England and Wales: 2020](#) - Summary statistics demonstrating the feasibility of producing census-like housing stock statistics using the admin-based housing stock version 1.0 (ABHS V1.0) dataset for 2020.

[Developing admin-based housing stock statistics for England and Wales: 2020](#) - An overview of methods, data sources and data quality for the feasibility research on producing census-like housing stock statistics using administrative data.

[Admin-based accommodation type statistics for England and Wales, feasibility research: 2011](#) - Further research demonstrating the improved potential of Valuation Office Agency (VOA) data to provide detailed information on accommodation type, and examining how VOA data compare with the 2011 Census data, and a request for feedback on the usefulness of these statistics.

13 . Housing by ethnicity

What is being measured

Subnational multivariate housing by ethnicity statistics for the usually resident population of England and Wales using administrative data sources.

How is it being measured

To create the [admin-based housing by ethnicity dataset \(ABHED\)](#), we linked the admin-based ethnicity dataset version 4.0 (ABED V4.0) for 2021, the [admin-based occupied address dataset \(ABOAD\)](#) for 2021, and the [admin-based housing stock version 1.0](#) (ABHS V1.0) dataset for 2021. The [Statistical Population Dataset \(SPD\)](#) version 4.0 (SPD V4.0) provides the population base for both the ABES and ABOAD.

Things you need to know

The ABHS V1.0 provides information about residential addresses, so communal establishments and special population groups are removed from our dataset. All analysis is at the individual level.

We have used the ethnicity variable from ABED V4.0, which makes use of 2011 Census data in addition to multiple administrative data sources to [derive an individual's ethnic group](#).

Given that the ABHED uses the ABHS, ABED and the ABOAD as their base, any quality and coverage considerations for these datasets needs to be accounted for when using the ABHED as well as the quality and coverage issues for the SPDs.

Related links

[Developing subnational multivariate housing by ethnicity statistics from administrative data, England and Wales: 2020](#) - Feasibility research on producing housing by ethnicity statistics for England and Wales from administrative data.

[Occupied address-level ethnicity measures for multivariate statistics, England and Wales: 2020](#) - Approaches used to identify occupied address-level ethnicity for feasibility research on producing multivariate statistics for England and Wales from administrative data.

14 . Education

What is being measured

Highest level of qualification in 2011 for individuals aged 16 to 25 years, who studied in government-funded education in England.

How is it being measured

Initial research used the feasibility All Education Dataset for England (AEDE). The AEDE was created and supplied by the Department for Education (DfE) and brings together attainment data from the National Pupil Database (NPD) for the academic year ending 2002 to the academic year ending 2015, with data from the Individualised Learner Record (ILR) and the Higher Education Statistics Agency (HESA) for the academic year ending 2003 to the academic year ending 2015.

Things you need to know

The feasibility AEDE was a one-off supply from DfE, and only includes attainment data for individuals aged between 14 and 29 years on 31 August 2015, and who were in government-funded education in England. This offers an insight into a large proportion of first-time entrants to the labour market and consequently an understanding of whether this group is equipped with the skills to meet market demands.

Differences between the feasibility AEDE and other sources like the Annual Population Survey (APS) or census, can be explained by the different data collection methods, the different time periods to which the data relate and differences in defining full attainment of qualification levels.

We are exploring the feasibility of linking NPD, ILR and HESA as well as Welsh equivalent of the NPD and Lifelong Learning Wales Record (LLWR). This dataset would improve the population coverage by including data for more recent academic years and covering learners in England and Wales.

Related links

[Admin-based qualification statistics, feasibility research: England](#) - Early research to demonstrate the potential of administrative data to provide information on educational qualifications, which has been collected by the census since 1961.

[Feasibility All Education Dataset for England \(AEDE\)](#) - Overview giving a high-level view of new data sources included in the Administrative Data Research Outputs.

15 . Morbidity and mortality

What is being measured

We have compared responses to Census 2021 with morbidity and mortality information derived from electronic health records to assess coherence. This research aims to address a knowledge gap regarding the comparison and coherence of these very different measures of health.

How is it being measured

Census 2021 included questions related to self-reported health status and disability. These responses were compared with measures of morbidity derived from electronic health records (EHRs), such as those routinely collected in primary care and hospital settings through the course of diagnosing, treating and managing illness. We analysed the coherence of these two sets of measures and validated these against benchmark administrative measures of health (hospitalisation, all-cause death) and employment characteristics.

Things you need to know

The [General Practice Extraction Service \(GPES\) Data for Pandemic Planning and Research \(GDPPR\)](#) extract used in this analysis includes approximately 40,000 medical codes out of approximately 1 million available for use by general practitioners. While many prevalent chronic health conditions are included within the scope of the extract, we were not able to identify people with some common conditions, such as generalised anxiety disorder. In addition, we only used the EHR to derive flags for a range of conditions by using any evidence of a diagnosis in a 10-year window. EHRs can be used to measure the severity of certain conditions like asthma by using the detailed information on the type and dosage of medications that are prescribed to patients.

Although we had access to Census 2021 data for individuals in both England and Wales, we only had access to EHRs for individuals in England, hence the study population does not include people in Wales.

Related links

[Measuring morbidity: comparing self-reported responses to electronic health records, England: 2021](#)

16 . Veterans

What is being measured

We have conducted feasibility research on producing statistics on the population of England and Wales who have previously served in the UK armed forces as of 30 June 2021 using Census 2021 and administrative data. Produced by age, sex, region and local authority.

How is it being measured

The [Statistical Population Dataset \(SPD\)](#) has been used as the population base. We have joined on the Service Leavers Database (SLD), which covers all veterans that have left the UK armed forces since 1975, and Census 2021, which provides data for veterans in England and Wales as at Census Day, 21 March 2021 to identify those who left the armed forces prior to 1975.

Things you need to know

Official statistics of the veteran population are from [Census 2021](#). This feasibility research is our initial exploration of the potential to produce veteran statistics using administrative data. It is not the finalised method and further work is needed before we can produce robust estimates.

Related links

[UK armed forces veterans, England and Wales: Census 2021](#) - UK armed forces veterans population who have either previously served in the regular forces, reserve forces or both, Census 2021 data.

[Feasibility research on producing UK armed forces veteran statistics for England and Wales: 2021](#) - Feasibility research on producing statistics on the population who have previously served in the UK armed forces in England and Wales using the Service Leavers Database and Census 2021.

17 . Estimation of travel to work matrices

What is being measured

We have produced modelled estimates of travel to work matrices of the usual residents of England and Wales aged 16 years and over in employment with a fixed workplace at Middle layer Super Output Areas (MSOA) annually between 2012 and 2021. Travel to work matrices show movement of people from their home (origin) to their place of work (destination).

How is it being measured

We have developed a gravity model, calibrated using the Census 2011 travel to work data, and using several input datasets to estimate future year (2012 to 2021) travel to work matrices.

Things you need to know

Travel to work data are collected from the [census](#), which allows for the generation of travel to work every 10 years with no updates for years in-between. The Census 2021 collected responses during the COVID-19 pandemic. The national lockdowns, associated guidance and furlough measures will [have affected the travel to work topic](#).

These modelled estimates are the first release in a planned work programme of incremental improvements to the model and outputs. The current model can produce more timely outputs than the census (that is, annually between 2012 and 2021 and 2022 onwards as new data become available). Planned improvements include breakdowns by transport modes, industry, occupancy and socioeconomic characteristics. This segmentation will allow us to make better assumptions about travel behaviour trends by different types of commuters and allow the estimation of travel to work matrices by these segmentation groups.

Because of the lack of a representative survey of travel to work for 2021 it is not currently possible to validate the estimates. We are addressing this issue in our planned work programme, using alternative data sources. These data are experimental and should not be used for decision making purposes.

Related links

[Travel to work quality information for Census 2021](#) - Known quality information affecting travel to work data from Census 2021 in England and Wales to help users correctly interpret the statistics.

[Estimation of Travel to Work Matrices \(Data Science Campus blog post\)](#) - an overview of the project and outputs.

[Estimation of Travel to Work Matrices \(technical report\)](#)

18 . Outcomes over time (Refugee Integration Outcomes)

What is being measured

Integration outcomes for cohorts of resettled or asylum refugees in England and Wales including access to health services, housing, education, labour market, fertility, mortality and migration.

How is it being measured

Linkage of administrative data (NHS Personal Demographic Service and Home Office border systems data) and Census 2021 data to cohorts of refugees resettled in England and Wales between 2015 and 2020 under the Vulnerable Persons Resettlement Scheme and the Vulnerable Children's Resettlement Scheme (VPRS and VCRS). Equivalent cohorts for those granted asylum in England and Wales between 2015 and 2020 are also linked.

Things you need to know

Analysis is based on linked data and excludes those which did not link. Linkage rates for VPRS and VCRS refugees are very high and of good quality. We achieved lower linkage rates for asylum refugees and are investigating how to improve linkage before publishing information on asylum refugee outcomes.

There are several [reasons why some refugee records did not link](#) to the administrative or census data. For example, missed matches in the linkage (records that did not link but should have); individuals that were not recorded in the administrative data or census data because they have moved from England and Wales; individuals who have not yet registered for health services; individuals who may have left the UK, or left and returned, but were not recorded in Home Office border systems data (these data exclude entries and exits within the Common Travel Area); or individuals who may have died since arriving.

Related links

[Refugee Integration Outcomes \(RIO\) data linkage pilot](#) - Methods used to link refugee data to various administrative data in a pilot study led by the Office for National Statistics (ONS) and the Home Office.

[Early integration outcomes for refugees resettled in England and Wales: 2015 to 2021](#)

[Refugee Integration Outcomes: Census 2021 Linkage Methodology Update](#)

19 . Glossary

Administrative data

Administrative data refers to information collected primarily for administrative reasons (not research). This type of data is collected by government departments and other organisations for registration, transactions, and record-keeping, usually when delivering a service.

Communal establishment

A communal establishment is an establishment with full-time or part-time supervision providing residential accommodation, such as student halls of residence, boarding schools, armed forces bases, hospitals, care homes and prisons.

Local authority

The general term for a body administering local government services. In England, local government is administered by either single-tier or two-tier local authorities. The single-tier authorities comprise unitary authorities, metropolitan districts, and London boroughs, though some services such as transport planning are carried out by the Greater London Authority. The two-tier authorities elsewhere comprise counties and non-metropolitan districts. In Wales, there are single-tier unitary authorities.

Lower layer Super Output Area (LSOA)

Lower layer Super Output Areas (LSOAs) are made up of groups of Output Areas (OAs), usually four or five. They comprise between 400 and 1,200 households and have a usually resident population between 1,000 and 3,000 persons.

Occupied address

For this research, an occupied address is a unique property reference number (UPRN) on the Address Frame, which has been successfully linked to at least one individual in the [Statistical Population Dataset version 3.0](#) (SPD V3.0). It is different to the concept of a household, which uses a definition based on shared facilities. More information on the differences between a traditional “household” and an “occupied address” is available in our [Occupied address \(household\) estimates from Administrative Data: 2011 and 2015 release](#).

Overcoverage

Overcoverage occurs when a record is counted more than once at the same location, more than once at a different location, counted in the wrong location, or is incorrectly included.

Statistical modelling

Statistical modelling involves making a set of assumptions about underlying processes that generate data in order to make inferences or to create estimates or predictions. Often, a model is fitted to a set of observed data to establish the values of parameters that describe the relationships between variables.

Undercoverage

Undercoverage occurs when a record is incorrectly excluded from data.

Usual resident

A usual resident of the UK is anyone who, on 21 March 2021, is in the UK and has stayed, or intends to stay, in the UK for 12 months or more, or has a permanent UK address and is outside the UK and intends to be outside the UK for less than 12 months.

20 . Cite this methodology

Office for National Statistics (ONS), published 26 June 2023, ONS website, methodology, [Population and migration statistics transformation in England and Wales, technical topic guide: 2023](#)