

TTP DEMOGRAPHIC DATA QUALITY REPORT

DATA SOURCE

DATA SUPPLIER:	Census Office
DESCRIPTION:	2011 Census (Entire enumerated population where subject is a student at term time address or subject is not a student) taken on 27 th March 2011.
COVERAGE:	Northern Ireland
NUMBER OF RECORDS:	1,720,645 ¹ http://www.nisra.gov.uk/census/2011/results.html 1. This varies from the full Census count as it does not include imputed/CUE records

DATA BACKGROUND

A census of population and housing, even in a country as small as Northern Ireland, is a major undertaking and much time, effort and resources are devoted to it. The Census in Northern Ireland is taken under the authority of the Census Act (Northern Ireland) 1969, as amended by the Census (Confidentiality)(Northern Ireland) Order 1991, which gives powers for taking a census from time to time provided that, at the commencement of the year in which a census is to be taken, at least five years have elapsed since the commencement of the year in which a census was last taken. The Act also gives powers for the census to be taken on a particular day.

The 2011 Census in Northern Ireland was conducted on 27 March 2011, in line with arrangements across the rest of the UK. The Census placed a legal obligation on every household in which someone was usually resident on Census Day, and on every person who was a usual resident of a communal establishment, to complete a Census form. This contrasts to the 1991 Census which collected information on both usual residents and visitors present on Census night.

Students and school children in full-time education studying away from home were enumerated as resident at their term-time address. Basic demographic information only (name, sex, age, marital status and relationship) was collected at their 'home' or vacation address. HM Forces permanently posted in Northern Ireland were enumerated at their actual address of residence unless they were married and unaccompanied by their spouse. In these cases they were included at the address they shared with their spouse. In contrast, HM Forces posted in Northern Ireland on short term tours of 3 – 4 months or less (Roulement troops) were not included in the Northern Ireland Census. Arrangements were made for these troops to be enumerated at their home garrisons.

It was not possible to enumerate all persons by the conventional method, and special arrangements had to be made for certain population groups, such as the Armed Forces, prisoners, students, shipping personnel, refugees and asylum seekers, the visually impaired, persons sleeping rough, travelling fairs and circuses, and others, though standard forms were used in all cases.

The Northern Ireland Statistics and Research agency (NISRA) undertook a number of important quality control measures in the interest of delivering high quality results. This involved the development of robust 'Downstream' processes to, for example:

- Identify and address any duplicate households
- Identify any missing items in the data and impute the required values in accordance with methodology agreed across the UK.
- Undertake an assessment of coverage (Census Coverage Survey) and adjust results as appropriate.

It should be noted that the number of records quoted relates to the entire enumerated population. It does not include those individuals who have been imputed which accounts for 4.8% of the total records in Census 2011. If matching to a 100% complete administrative dataset the maximum match rate that can be achieved is 95.2%.

DEMOGRAPHIC INFORMATION

Information available to the TTP for linkage purposes -

- Name information which includes Forename, Midname and Surname
- Address information at the time of the Census
- Sex
- Date of birth
- UPRN

Name, Address, Sex, Date of birth and UPRN are also available for other household members. This can be used for matching purposes and linkage quality checks.

DATA QUALITY & PRE-PROCESSING

Record linkage is highly dependent on the quality of the data being linked. Many key variables can be presented quite differently between datasets as the information has been collected and recorded using different methods. This can greatly complicate record linkage unless understood ahead of time. Pre-processing can be used to standardise these variables into a consistent format prior to linkage.

There are 2,582 records which we do not hold any demographic information for. It would be impossible to link these records. These leaves

NAME INFORMATION- Forename and Surname is available 99.99% of records. Midname is populated where applicable.

- In some instances multiple names will be recorded as the Forename. This can be edited to leave the first name that appears as the Forename and any additional names as the Midname;
- Punctuation and characters which do not appear in the English alphabet (e.g. Á , % , \$...) can also appear in the name information. A function can be used to replace these with usable characters;

ADDRESS INFORMATION- All records have address information available. Postcode details are 99.89% populated.

- Postcode information is standardised prior to linkage.

SEX- 99.99% of records contain sex.

- Sex information should be standardised prior to linkage.

DATE OF BIRTH- All records contain the Date of Birth. However, when DOB is parsed into its constituent parts (day, month and year) there are some missing values:

- DOB_Day: <0.01% of records incomplete
- DOB_Month: <0.5% of records incomplete
- DOB_Year: <0.5% of records incomplete

POINTER UPRN¹- This information 98% populated.

1. UPRN (Unique property reference number) is the NI Government's definitive address register's unique number.

MATCHING METHODS

Using the demographic information provided the following matching can be undertaken-

- Person matching
- Address matching
- Associative matching (other household members)

PREVIOUS LINKAGE MATCH RATES

The Census 2011 information provided by Census Office has previously been matched to health data as part of the Northern Ireland Longitudinal Study (NILS). This typically produces an adjusted match rate of approximately 98.9%; More information can be found on the [NILS Metadata](#).

The Census 2011 information has also been matched to death data as part of the Northern Ireland Mortality Study (NIMS) and previous ADRC-NI projects. This typically produces a raw match rate of approximately 91% (97% adjusted). More information can be found on the [NIMS Metadata](#).

As previously mentioned, the imputation rate (4.8%) should be taken into account when calculating match rates. When linked to other administrative datasets the maximum potential match rate would be 95.2%.

EVALUATION

The demographic data from the 2011 Census is of very good quality and can be used for matching without extensive pre-processing.

It should be noted that when matching Census 2011 to a 100% Northern Ireland dataset we will be unable to match a small number of records which relate to individuals who were imputed in the Census data and for which we hold no demographic information (approximately 4.8%).