NILS Database Metadata

Document by:

NILS-Core

NISRA

This document should be used in conjunction with the data dictionary and the completed working papers. The combined set of information provides full metadata for the NILS database.

The NILS Data Dictionary will provide the value labels for all the variables. This document includes a description of variables and some value labels that are essential for the understanding of the NILS structure and NILS sample membership.

**Relevant working papers:**

1. The Northern Ireland Longitudinal Study – An Introduction
2. NILS Data Matching Methodology
3. Linking the Northern Ireland Longitudinal Study to the 2011 Census

(Please click [here](http://www.qub.ac.uk/research-centres/NILSResearchSupportUnit/UsefulDocuments/) to find out more about our working papers)

The owner of this document is NILS-Core and any changes to the document should be suggested to NILS-Core who will make amendments if required.

|  |  |  |
| --- | --- | --- |
|  |  |  |

**Table of Contents:**

[Current database structure 5](#_Toc367098690)

[X-Files/Variables 5](#_Toc367098691)

[Date of Registration vs Date of Occurrence for Events 6](#_Toc367098692)

[Date-stamped database 7](#_Toc367098693)

[NILS DataBase Diagram 9](#_Toc367098694)

[Metadata for CORENILSDATA 10](#_Toc367098695)

[Source 12](#_Toc367098696)

[CORENILSDATA: Table 1[q3] – Source Distributions 12](#_Toc367098697)

[StatusHistory\_Full 13](#_Toc367098698)

[StatusHist 13](#_Toc367098700)

[CORENILSDATA: Table 2 – StatusHist Descriptions 14](#_Toc367098701)

[CORENILSDATA: Table 3 [q7] – StatusHist Distributions 15](#_Toc367098702)

[CORENILSDATA: Table 4 [q8] – StatusHist Distributions by Download 16](#_Toc367098703)

[Current\_Address\_SOA 16](#_Toc367098704)

[CORENILSDATA: Table 5 [q9] – Current\_Address level of SOA Assignment 16](#_Toc367098705)

[CORENILSDATA: Table 6 [q10] – Current\_Address\_SOA level of Assignment by Source 17](#_Toc367098706)

[Current\_Address\_XUPRN 18](#_Toc367098707)

[CORENILSDATA: Table 7 – Current\_Address level of XUPRN Assignment 18](#_Toc367098708)

[CORENILSDATA: Table 8 [Q12] – Current\_Address\_XUPRN level of Assignment 18](#_Toc367098709)

[Date of Death Month and Year 19](#_Toc367098710)

[CORENILSDATA: Table 9 [q17] – Year of Death Distribution 19](#_Toc367098711)

[CORENILSDATA: Table 10 [q18] – Year and month of Death Distribution 20](#_Toc367098712)

[Metadata for XAGES 21](#_Toc367098713)

[XAGES: Table 11 [q22] – Average Age April 2001 and April 2011 21](#_Toc367098714)

[Metadata for EVENTS 21](#_Toc367098715)

[Event\_Type 23](#_Toc367098716)

[EVENTS: Table 12 [q24] – Number of Events by Event\_Type 23](#_Toc367098717)

[EVENTS: Table 13 [q25] – Unique Birth and Death Events 23](#_Toc367098718)

[EVENTS: Table 14 [q26] – Births – Combination of Baby, Mum and Dads 23](#_Toc367098719)

[Metadata for BIRTHSTATS 24](#_Toc367098720)

[Metadata for DEATHSSTATS 26](#_Toc367098721)

[Metadata for ADDRESS\_HISTORY 28](#_Toc367098722)

[What is an address change? 30](#_Toc367098723)

[Quality Assurance 30](#_Toc367098724)

[Data Issues: 30](#_Toc367098725)

[ADDRESS\_HISTORY: Table 15 [q31] – Number of Address Changes 31](#_Toc367098726)

[SOURCE 31](#_Toc367098727)

[ADDRESS\_HISTORY: Table 16 – [q32] Distribution of Source 31](#_Toc367098728)

[Type 32](#_Toc367098729)

[ADDRESS\_HISTORY: Table 17 [q33] – Distribution of Type 32](#_Toc367098730)

[Current\_Flag 32](#_Toc367098731)

[ADDRESS\_HISTORY: Table 18 [q34] – Distribution of Current Flag 32](#_Toc367098732)

[ADDRESS\_HISTORY: Table 19 [q35] – Distribution of Current Flag for all ‘Live records at April 2012’ 32](#_Toc367098733)

[Note 541800 the number of live records at the October 2008 download [q35a]. 32](#_Toc367098734)

[SOA2001 33](#_Toc367098735)

[ADDRESS\_HISTORY: Table 20 [q37]– Coverage of SOA2001 33](#_Toc367098736)

[XUPRN 33](#_Toc367098737)

[ADDRESS\_HISTORY: Table 21 [q38]– Coverage of XUPRN 33](#_Toc367098738)

[Metadata for MIGRATION\_EVENTS 34](#_Toc367098739)

[Type 36](#_Toc367098740)

[MIGRATION\_EVENTS: Table 22 [q41] – Type of Event 36](#_Toc367098741)

[MIGRATION\_EVENTS: Table 23 [q42] – Type of Event by DateMoved 36](#_Toc367098742)

[Out of\_SOA2001 37](#_Toc367098743)

[MIGRATION\_EVENTS: Table 24 [q43] – Coverage of Outof\_SOA2001 37](#_Toc367098744)

[MIGRATION\_EVENTS: Table 25 [q44] – Coverage of Outof\_SOA2001 by ChangeType 37](#_Toc367098745)

[MIGRATION\_EVENTS: Table 26 [q45] – Coverage of Outof\_SOA2001 by type 38](#_Toc367098746)

[INTO\_SOA2001 38](#_Toc367098747)

[MIGRATION\_EVENTS: Table 27 [q46] – Coverage of Into\_SOA2001 38](#_Toc367098748)

[MIGRATION\_EVENTS: Table 28 [q47] Coverage of Into\_SOA2001 by type 3](#_Toc367098749)9

[MIGRATION\_EVENTS: Table 29 [q48] – Coverage of Into\_SOA2001 and Out of SOA by ChangeType 39](#_Toc367098750)

[Outof\_XUPRN](#_Toc367098751) 40

[MIGRATION\_EVENTS: Table 30 [q50] – Coverage of Outof\_XUPRN by ChangeType 40](#_Toc367098752)

[MIGRATION\_EVENTS: Table 31 [q51] – Coverage of Outof\_XUPRN by DateMoved 40](#_Toc367098753)

[INTO\_UPRN 41](#_Toc367098754)

[MIGRATION\_EVENTS: Table 32 [q52] – Coverage of Into\_XUPRN by ChangeType 41](#_Toc367098755)

[MIGRATION\_EVENTS: Table 33 [q53] – Coverage of Into\_XUPRN by DateMoved 41](#_Toc367098756)

[Metadata for CENSUSP\_2001 42](#_Toc367098757)

[Metadata for CENSUSHH\_2001 44](#_Toc367098758)

[Metadata for CENSUSP\_2011 4](#_Toc367098759)6

[Metadata for CENSUSHH\_2011 48](#_Toc367098760)

[Metadata for CENSUSP\_1991 50](#_Toc367098761)

[Metadata for CENSUSHH\_1991 52](#_Toc367098762)

[Metadata for Properties 53](#_Toc367098763)

[Metadata for 2001 Household Imputation\_Flags 55](#_Toc367098764)

[Person Imputation Flags: Table 34 – Distributions 56](#_Toc367098765)

[Person Imputation Flags: Chart 1 – Distributions 57](#_Toc367098766)

[Metadata for 2001 Household Imputation Flags 58](#_Toc367098767)

[Household Imputation Flags: Table 35 – Distributions 58](#_Toc367098768)

[More detail on Item Imputation 59](#_Toc367098769)

[More detail on Record swapping 60](#_Toc367098770)

**Current database structure**

The NILS Database has been redesigned (in November 2011) and the set of tables and the associated fields will form the NILS database. This has been further updated with the 2011 Census tables.

A database diagram is included on Page 9 showing how the different tables relate to each other.

Any changes to the structure of the NILS, addition of new variables etc. must be identified by NILS-Core and NILS-RSU as soon as possible and a decision made by NILS-Core about whether they should be added to the database.

Further work will be completed by Population Statistics section of DMB on the BSO other household members to determine the quality and usefulness for researchers.

# X-Files/Variables

There are some tables and variables that are required for processing of derived variables for researchers but are restricted or confidential and so should not be released to the NILS researchers. Some of these variables will be added to the database but flagged with an X – the NILS-RSU need to ensure these are not released to the researchers.

To date many research projects have required the addition of some information not on the current dataset and therefore reliance on NILS-Core. Examples of previous X-File/Variables work include:

Project 002 used Day, Month and Year of Date of Death to determine the temperature at time of death, 7 day, 14 day, 21 day and 28 day lags. For this NILS-Core had to use Day of Death and COA that is not available to NILS researchers.

Projects 001 & 007 used individual property IDs to determine migration into and out of Nursing and Residential Homes. NILS-Core had to use unique property IDs and address to extract this information.

Project 019 on alcohol related deaths required access to all causes of death to extract information on underlying and associated causes of deaths relating to alcohol deaths.

The following table shows the common variables required for the X-Files/Variables and how NILS-Core plan to integrate them into the NILS database.

|  |  |
| --- | --- |
| NILS | Solution |
| Property indicator | Include XUPRN on all datasets where appropriate : corenilsdata, address\_history, migration\_events, censushousehold |
| Capital Value | Create an encrypted property Number called XUPRN. Properties table will have XUPRN as key identifier with all property information such as Capital Value |
| Settlement Bands | Attach to the census household table and the properties table |
| Age at specific points in time | An additional table called XAges has been created with the age of each person at each download from the BSO. For the December 2013 database this will include 23 different ages. |
| Detailed Cause of Death | An additional table XDeath\_Details has been created containing detailed ICD09 and ICD10 codes for each NILS death. |

Any variable that is preceded by an X must not be made available to researchers unless specific justification has been included in the RAG application form and approved by the RAG.

As each project is submitted for RAG approval it should be clear to NILS-RSU what variables/tables are not available in the current NIMS or NILS databases. It is up to the NILS-RSU to quickly identify these and work with the NILS-Core to get the derived variable onto the NIMS/NILS database (if suitable) so that others in the future can use these variables (if approved).

# Date of Registration vs Date of Occurrence for Events

Vital events (including births and deaths) have 2 different dates associated with them. One is the date of the event taking place (e.g. date of birth and date of death) called the occurrence date and the other is the date of registration i.e. the date the event was registered with the Registrar in the local offices.

Statistics on the number of events by Date of Registration do not change. Statistics on the number of events by Date of Occurrence will change because of late registrations of events.

The Northern Ireland the Registrar General’s annual report and other Vital Events publications use the Date of Registration to produce the finalised fixed number of events. The Vital Events statistical coded data are not finalised until after the publication of the Registrar General’s annual report. This is because an intense QA process takes place verifying any anomalies in the data.

The following table, using births data, highlights the delay in getting finalised information on Births Occurrence. To date NILS researchers have wanted data based on Date of Occurrence. SLS and LS release data to researchers based on Date of Registration. NILS-Core have met with Vital Events colleagues and are currently working on increasing the frequency of data downloads to improve timeliness.

|  |  |  |  |
| --- | --- | --- | --- |
| **Occurrence\*** | **Registration** | **Publication\*\*** | **Months to NILS-Core for matching** |
| Jan-13 | Feb-13 | Dec-13 | 23 |
| Feb-13 | Mar-13 | Dec-13 | 22 |
| Mar-13 | Apr-13 | Dec-13 | 21 |
| Apr-13 | May-13 | Dec-13 | 20 |
| May-13 | Jun-13 | Dec-13 | 19 |
| Jun-13 | Jul-13 | Dec-13 | 18 |
| Jul-13 | Aug-13 | Dec-13 | 17 |
| Aug-13 | Sep-13 | Dec-13 | 16 |
| Sep-13 | Oct-13 | Dec-13 | 15 |
| Oct-13 | Nov-13 | Dec-13 | 14 |
| Nov-13 | Dec-13 | Dec-13 | 13 |
| Dec-13 | Dec-13 | Dec-13 | 12 |
| Dec-13 | Jan-13 | Dec-14 | 24 |

# Date-stamped database

The plan has been to release a NILS database twice a year. The new NILS database structure has been created and date stamped ‘NILS\_RSU\_ DEC2013’. The databases will be called NILS\_RSU\_MMMYYYY.

Each database will have a table called DATA\_RELEASED which indicates the time period which the data in each table covers. The NILS\_RSU\_ DEC2013 database DATA\_RELEASED table contents are shown below:

Any additional changes to the database such as the definition/format of a variable, improved coverage of will be recorded in the DATA\_RELEASED table.

|  |  |
| --- | --- |
| Data [q1a] | Date reference |
| CORENILSDATA | APRIL 2001 to APRIL 2012 (23 downloads) |
| XAGES | APRIL 2001 to APRIL 2012 (23 downloads) |
| ADDRESS\_HISTORY | APRIL 2001 to APRIL 2012 (23 downloads) |
| MIGRATION\_EVENTS | APRIL 2001 to APRIL 2012 (23 downloads) |
| EVENTS: BIRTHS OF NILS MEMBERS | JANUARY 1997 to DECEMBER 2011 (OCCURRENCES) |
| EVENTS: BIRTHS TO NILS MOTHERS | JANUARY 1997 to DECEMBER 2011 (OCCURRENCES) |
| EVENTS: BIRTHS TO NILS FATHERS | JANUARY 1997 to DECEMBER 2011 (OCCURRENCES) |
| EVENTS: DEATHS OF NILS MEMBERS | JULY 2001 to DECEMBER 2011 (OCCURRENCES) |
| BIRTHSSTATS | JANUARY 1997 to DECEMBER 2011 |
| DEATHSSTATS | JULY 2001 to DECEMBER 2011 |
| CENSUSP\_1991 | 22ND APRIL 1991 NO CHANGES |
| CENSUSHH\_1991 | 23rd APRIL 1991 NO CHANGES |
| CENSUSP\_2001 | 29TH APRIL 2001 NO CHANGES |
| CENSUSHH\_2001 | 29TH APRIL 2001 NO CHANGES |
| CENSUS01\_RELATIONSMATRIX | 29TH APRIL 2001 NO CHANGES |
| CENSUSP\_2011 | 27TH MARCH 2011 NO CHANGES |

The following is an early draft of the NILS Database Diagram. This needs altered when additional tables such as stillbirths, marriages, widowerhoods, infant deaths to NILS members etc. are added.

# NILS DataBase Diagram

# Metadata for CORENILSDATA

|  |  |
| --- | --- |
| Database Name | NILS |
| Table Name | CORENILSDATA |
| Table Description | This is the basic record for all NILS members. A person is chosen as a NILS member if their day and month of the BSO date of birth falls on one of the 104 NILS dates.  All NILS members in the study since 2001 are included in this table. |
| Table size (number of rows) | 646,134[q1] records in NILS\_RSU\_DEC2013 (always growing) |
| Unique Identifiers | NILSID |
| Tables linked to | via NILSID : CENSUSP\_2001, MIGRATION\_EVENTS,  EVENTS, ADDRESS\_HISTORY  via CURRENT\_ADDRESS\_XUPRN : PROPERTIES |
| Frequency of the Data | Latest information included for April 2012[q2]. Updated and released every 6 months |
| Variables | |  |  | | --- | --- | | Variable Name | Variable Description | | NILSID | Unique identifier | | SOURCE | This gives the source of the first time the patients record joined the NILS (and the BSO). Many records have a source of 200104 (April download in 2001)  Format YYYY04 (April download in YYYY) and YYYY10 (October download in year YYYY) | | GENDER | Gender as recorded by the BSO | | STATUSHISTORY\_FULL | Full status history of the person. This field is variable length and it has one status flag (0, E, L, D) for each download – see notes below | | STATUSHIST | Summarised version of STATUSHISTORY\_FULL – see notes below | | CURRENT\_ADDRESS\_SOA, | Super Output Area of current address | | CURRENT\_ADDRESS\_XUPRN | Anonymised property ID of current address | | DODMMM | Month of death if status = ‘D’ | | DODYYYY | Year of death if status = ‘D’ | |
| Variable Values | Gender – M= Male; F = Female  STATUSHISTORY\_FULL, STATUSHIST  0 – patient not on Health register  L – Live patient on Health register  E – patient flagged as gone away on the Health Register  D – patient flagged as deceased on the Health Register  CURRENT\_ADDRESS\_SOA  890 valid SOA codes  XXXXXX is a missing value (normally invalid postcodes)  000000 is no code available (valid postcodes)  CURRENT\_ADDRESS\_XUPRN is an anonymised property ID that can be used to link to the Propertiestable |

## 

## Source

This identifies the number of records that were added to the NILS in each 6 month period. The majority of NILS records were added in April 2001. The first four digits indicate the year and the last two digits indicates which download (04 = 1st download of the year in April, 10 = 2nd download of the year in October).

The downloads are referred to as April and October but the exact date is determined by the date the BSO took the quarterly extract of their data for GP payment purposes.

### CORENILSDATA: Table 1[q3] – Source Distributions

|  |  |
| --- | --- |
| **Source** | **Records** |
| 200104 | 514527 |
| 200110 | 4818 |
| 200204 | 4910 |
| 200210 | 5276 |
| 200304 | 4679 |
| 200310 | 5042 |
| 200404 | 5394 |
| 200410 | 4859 |
| 200504 | 6272 |
| 200510 | 6185 |
| 200604 | 6631 |
| 200610 | 6410 |
| 200704 | 7612 |
| 200710 | 7208 |
| 200804 | 6860 |
| 200810 | 7063 |
| 200904 | 6104 |
| 200910 | 6144 |
| 201004 | 5797 |
| 201010 | 6342 |
| 201104 | 6185 |
| 201110 | 6097 |
| 201204 | 5719 |

**The source is the indication of when the person first appeared on the BSO downloads and is not an indication of when the person became live in the NILS. The StatusHistory\_Full should be used for that.**

**In each download the BSO provides information on all live people, including people who used to live in NI and are flagged as emigrated.**

## StatusHistory\_Full

The NILS\_RSU\_DEC2013 database includes information on all downloads and so the length of the StatusHistory\_Full field is same as the number of downloads (including the 2012 download = 23 characters). This number will change in future date-stamped databases when more downloads are added.

This variable gives the detail of the status of each NILS member at each download and it is essential that this is understood. The following table gives the most common values for StatusHistory\_Full and some have a brief description to illustrate what this variable means.

## StatusHist

The StatusHist is a summary version of the StatusHistory\_Full. It is created by removing multiple versions of the same letter. This allows a summary of the status of a person but does not allow you to identify the timing of any status change.

### CORENILSDATA: Table 2 – StatusHist Descriptions

| **StatusHistory\_Full** | **StatusHist** | **Description** |
| --- | --- | --- |
| LLLLLLLLLLLLLLLLLLLLLLL | L | Live throughout the study |
| 000000000000000000000LL | 0L | Not in the study at the beginning but joined and is currently live |
| 000000000000LLLLLLLLLLL | 0L |
| 000000000000000LLLLLLLL | 0L |
| 0000000000000000000LLLL | 0L |
| 00000000000000000000LLL | 0L |
| 0000000000000000000000L | 0L |
| 00000000000000LLLLLLLLL | 0L |
| 00000000000000000LLLLLL | 0L |
| 0000000000000000LLLLLLL | 0L |
| 000000000000000000LLLLL | 0L |
| 000000LLLLLLLLLLLLLLLLL | 0L |
| 000LLLLLLLLLLLLLLLLLLLL | 0L |
| 0000000LLLLLLLLLLLLLLLL | 0L |
| 00000LLLLLLLLLLLLLLLLLL | 0L |
| LLLLLLLLLEEEEEEEEEEEEEE | LE | Live in the study at the beginning but has since been flagged as emigrated |
| LLLLLLLLEEEEEEEEEEEEEEE | LE |
| LLLLLLLLLLLLDDDDDDDDDDD | LE |
| LLLEEEEEEEEEEEEEEEEEEEE | LE |
| LLLLLLLLLLLLDDDDDDDDDDD | LD | Live in the study at the beginning but has since been flagged as deceased |
| LLDDDDDDDDDDDDDDDDDDDDD | LD |
| LLLLLLLLLLLLLLLLDDDDDDD | LD |

The following table shows the different combinations for StatusHist.

### CORENILSDATA: Table 3 [q7] – StatusHist Distributions

|  |  |
| --- | --- |
| **StatusHist** | **Records** |
| L | 411518 |
| 0L | 107158 |
| LD | 44069 |
| LE | 33000 |
| 0LE | 18670 |
| LEL | 10172 |
| EL | 6597 |
| 0LEL | 2867 |
| L0 | 2155 |
| ELE | 1472 |
| LELE | 1366 |
| E0L | 1054 |
| 0LD | 912 |
| L0L | 848 |
| 0LELE | 725 |
| 0L0 | 564 |
| LELEL | 422 |
| ELEL | 401 |
| LE0 | 248 |

A person could be flagged as a ‘0’ starting position if they were born after April 2001 or immigrated to NI. In addition, if someone had a date of birth change on the BSO from a non-NILS date to a NILS date they would initially appear as a ‘0’.

Note: A person who was live in the study could drop out of the study – i.e. end in a ‘0’. This could happen if they were removed from the BSO records totally, had a date of birth changed to a non-NILS date of birth or, were among a small number of records with a duplicate CHI or NHAIS number (about 500 per 1.6 million)

There may be some inconsistent looking records: for example: DL DE, E0; D0. – These are likely to be administrative errors and are very small in number.

The following table shows the number of records in each of the statuses. As expected the number of ‘0’ is smallest in the latest download – i.e. all babies and immigrants now have full records.

### CORENILSDATA: Table 4 [q8] – StatusHist Distributions by Download

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Sum of Count** | **Column Labels** |  |  |  |  |
| **Row Labels** | **0** | **D** | **E** | **L** | **Total** |
| 01st Download | 131607 |  | 10061 | 504466 | 646134 |
| 02nd Download | 126858 | 1802 | 11807 | 505667 | 646134 |
| 03rd Download | 122009 | 4017 | 13829 | 506279 | 646134 |
| 04th Download | 116790 | 6045 | 16241 | 507058 | 646134 |
| 05th Download | 112155 | 8164 | 18107 | 507708 | 646134 |
| 06th Download | 107169 | 10165 | 19822 | 508978 | 646134 |
| 07th Download | 101817 | 12244 | 22568 | 509505 | 646134 |
| 08th Download | 97060 | 14174 | 23845 | 511055 | 646134 |
| 09th Download | 92594 | 16316 | 25436 | 511788 | 646134 |
| 10th Download | 86337 | 18221 | 29464 | 512112 | 646134 |
| 11th Download | 79640 | 20419 | 32690 | 513385 | 646134 |
| 12th Download | 73302 | 22437 | 34002 | 516393 | 646134 |
| 13th Download | 65651 | 24869 | 36332 | 519282 | 646134 |
| 14th Download | 58492 | 26784 | 37893 | 522965 | 646134 |
| 15th Download | 51734 | 29008 | 41228 | 524164 | 646134 |
| 16th Download | 44718 | 30984 | 43189 | 527243 | 646134 |
| 17th Download | 38698 | 33258 | 45080 | 529098 | 646134 |
| 18th Download | 32598 | 35124 | 45884 | 532528 | 646134 |
| 19th Download | 26867 | 37350 | 48310 | 533607 | 646134 |
| 20th Download | 20605 | 39283 | 49331 | 536915 | 646134 |
| 21st Download | 14504 | 41444 | 51559 | 538627 | 646134 |
| 22nd Download | 8600 | 43351 | 53310 | 540873 | 646134 |
| 23rd Download | 3065 | 45506 | 55763 | 541800 | 646134 |
|  |  |  |  |  |  |

## Current\_Address\_SOA

Current address is the last known address for the person. It can also be extracted from the address\_history table where current\_flag = ‘C’. The following table shows the level of current\_address\_SOA coverage.

### CORENILSDATA: Table 5 [q9] – Current\_Address level of SOA Assignment

|  |  |  |  |
| --- | --- | --- | --- |
| **SOA Code** | **Description** | **Number of Records** | **% Distribution** |
| XXXXXXXX or null | missing (invalid/missing postcodes) | 5866 | 1 |
| Valid SOA | Valid SOA Code | 640268 | 99 |
| All Records |  | 646134 | 100 |

The following table shows that the percentage of assigned SOAs is high no matter when the record was added to the NILS.

### CORENILDATA: Table 6 [q10] – Current\_Address\_SOA level of Assignment by Source

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Source** | **No SOA Assigned** | **Assigned Valid SOA** | **Total** | **% Assigned** |
| 200104 | 198 | 509356 | 514527 | 99 |
| 200110 | 1 | 4797 | 4818 | 100 |
| 200204 | 3 | 4879 | 4910 | 99 |
| 200210 | 1 | 5241 | 5276 | 99 |
| 200304 | 1 | 4642 | 4679 | 99 |
| 200310 | 2 | 4998 | 5042 | 99 |
| 200404 | 1 | 5366 | 5394 | 99 |
| 200410 | 4 | 4828 | 4859 | 99 |
| 200504 | 2 | 6236 | 6272 | 99 |
| 200510 | 5 | 6149 | 6185 | 99 |
| 200604 | 1 | 6597 | 6631 | 99 |
| 200610 | 3 | 6378 | 6410 | 100 |
| 200704 | 4 | 7578 | 7612 | 100 |
| 200710 | 11 | 7170 | 7208 | 99 |
| 200804 | 12 | 6825 | 6860 | 99 |
| 200810 | 11 | 7028 | 7063 | 100 |
| 200904 | 6 | 6082 | 6104 | 100 |
| 200910 | 6 | 6112 | 6144 | 99 |
| 201004 | 9 | 5774 | 5797 | 100 |
| 201010 | 14 | 6305 | 6342 | 99 |
| 201104 | 10 | 6156 | 6185 | 100 |
| 201110 | 8 | 6075 | 6097 | 100 |
| 201204 | 9 | 5696 | 5719 | 100 |

## Current\_Address\_XUPRN

The coverage of XUPRN is lower than SOA but is still high at 94%. 39769 records do not have a valid property identifier in NILS\_RSU\_NOV2011 current addresses.

### CORENILSDATA: Table 7 – Current\_Address level of XUPRN Assignment

|  |  |  |  |
| --- | --- | --- | --- |
| **SOA Code** | **Description** | **Number of Records** | **% Distribution** |
| Missing | Missing | 39769 | 6 |
| Valid UPRN | Valid unique property ID | 606365 | 94 |
| All Records |  | 646134 | 100 |

The following table shows that there is a slight change in the level of assignment of a property id since the introduction of the NHAIS system. Work is ongoing to improve this coverage.

### CORENILSDATA: Table 8 [q12] – Current\_Address\_XUPRN level of Assignment

|  |  |  |  |
| --- | --- | --- | --- |
| **Source** | **Total** | **Valid UPRN** | **% of Valid UPRN** |
| 200104 | 514527 | 483052 | 94 |
| 200110 | 4818 | 4524 | 94 |
| 200204 | 4910 | 4609 | 94 |
| 200210 | 5276 | 4961 | 94 |
| 200304 | 4679 | 4375 | 94 |
| 200310 | 5042 | 4706 | 93 |
| 200404 | 5394 | 5060 | 94 |
| 200410 | 4859 | 4580 | 94 |
| 200504 | 6272 | 5857 | 93 |
| 200510 | 6185 | 5793 | 94 |
| 200604 | 6631 | 6188 | 93 |
| 200610 | 6410 | 6056 | 94 |
| 200704 | 7612 | 7139 | 94 |
| 200710 | 7208 | 6792 | 94 |
| 200804 | 6860 | 6379 | 93 |
| 200810 | 7063 | 6627 | 94 |
| 200904 | 6104 | 5677 | 93 |
| 200910 | 6144 | 5753 | 94 |
| 201004 | 5797 | 5419 | 93 |
| 201010 | 6342 | 5908 | 93 |
| 201104 | 6185 | 5784 | 94 |
| 201110 | 6097 | 5759 | 94 |
| 201204 | 5719 | 5367 | 94 |

## Date of Death Month and Year

The total number of people have been flagged as deceased in the NILS\_RSU\_DEC2013 download via StatusHistory\_Full is 45506 [q13]. All records should have a date of death [q14] but this is not always the case.

The date of death recorded is the date that the BSO staff entered onto the CHI or NHAIS when they were notified of the death. This may have come from the GRO, a GP, a family member or a through a data cleansing exercise.

The following table shows the distribution of year of death. Some records will have a year of death pre-2001 but were included as a NILS member because the BSO did not know they were deceased until after the April 2001 download. This represents 1% of the records which have a date of death attached.

### CORENILSDATA: Table 9 [q17] – Year of Death Distribution

|  |  |
| --- | --- |
| **Year of Death** | **Count of Records** |
| 2001 | 3171 |
| 2002 | 4108 |
| 2003 | 4251 |
| 2004 | 4181 |
| 2005 | 4097 |
| 2006 | 4355 |
| 2007 | 4232 |
| 2008 | 4373 |
| 2009 | 4268 |
| 2010 | 4308 |
| 2011 | 3355 |
| 2012 | 63 |

There is a smaller number of records in the latest year because only deaths notified to the BSO by April/October have been included. Therefore this does not represent a full year’s data. These deceased records provide the basis for death events although not all of them will have a GRO link.

### CORENILSDATA: Table 10 [q18] – Year and month of Death Distribution

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **YEAR** | **JAN** | **FEB** | **MAR** | **APR** | **MAY** | **JUN** | **JUL** | **AUG** | **SEP** | **OCT** | **NOV** | **DEC** | **TOTAL** |
| 2001 | 37 | 39 | 49 | 280 | 347 | 319 | 313 | 352 | 338 | 351 | 335 | 411 | 3171 |
| 2002 | 388 | 346 | 343 | 321 | 315 | 343 | 309 | 351 | 331 | 330 | 343 | 388 | 4108 |
| 2003 | 397 | 361 | 358 | 334 | 348 | 334 | 324 | 330 | 351 | 367 | 380 | 367 | 4251 |
| 2004 | 404 | 342 | 354 | 373 | 356 | 305 | 317 | 319 | 327 | 357 | 332 | 395 | 4181 |
| 2005 | 358 | 363 | 393 | 348 | 338 | 349 | 317 | 320 | 294 | 308 | 354 | 355 | 4097 |
| 2006 | 370 | 347 | 418 | 380 | 337 | 358 | 375 | 328 | 343 | 367 | 364 | 368 | 4355 |
| 2007 | 418 | 395 | 390 | 361 | 313 | 309 | 316 | 331 | 307 | 366 | 336 | 390 | 4232 |
| 2008 | 424 | 385 | 406 | 387 | 318 | 340 | 338 | 314 | 328 | 321 | 389 | 423 | 4373 |
| 2009 | 479 | 374 | 380 | 369 | 302 | 308 | 316 | 310 | 329 | 355 | 361 | 385 | 4268 |
| 2010 | 433 | 334 | 373 | 360 | 384 | 354 | 306 | 323 | 322 | 344 | 381 | 394 | 4308 |
| 2011 | 423 | 354 | 365 | 383 | 357 | 362 | 319 | 330 | 321 | 75 | 37 | 29 | 3355 |
| 2012 | 30 | 23 | 10 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 63 |

# Metadata for XAGES

|  |  |
| --- | --- |
| Database Name | NILS |
| Table Name | XAGES |
| Table Description | This is an x-file, i.e. it should not be freely given to researchers. It is only for the use by the NILS-RSU to generate extracts.  All NILS members in the study since 2001 are included and age is only shown if the person is a live member.  Ages are calculated using BSO date of birth. This is the most consistent form of age and is the only one available for all NILS members. |
| Table size (number of rows) | 646134 [q19] records in NILS\_RSU\_DEC2013 (always growing) |
| Unique Identifiers | NILSID |
| Tables linked to | via NILSID : NILSCOREDATA |
| Frequency of the Data | At each new download an additional age variable is created. In NILS\_RSU\_DEC2013 there are a total of 23 age points.  Updated and released every 6 months |
| Variables | |  |  | | --- | --- | | Variable Name | Variable Description | | NILSID | Unique identifier | | AgeatApr01 | Age calculated at April 2001 | | AgeatOct01 |  | |  |  | |  |  | | .. |  | | Ageatoct11 |  | | AgeatApr12 | Age calculated at April 2012 | |
| Variable Values | Missing value is where the NILS member is not flagged as Live – where the StatusHistory\_Full value is ‘0’, ‘D’ or ‘E’  Age ranges from 0 to approx 110 |

### XAGES: Table 11 – Average Age April 2001 and April 2012 [q22]

|  |  |  |
| --- | --- | --- |
|  | **April 2001** | **April 2012** |
| Average Age | 35.8 | 37.7 |
| Number of Records with Age | 504466 | 541800 |
| Total Records | 646134 | 646134 |

# 

# Metadata for EVENTS

|  |  |
| --- | --- |
| Database Name | NILS |
| Table Name | EVENTS |
| Table Description | This table gives the linking id for all the vital events to a NILS member. It currently includes 4 types of events relating to births and deaths but will be expanded over time to include widower hoods, stillbirths, infant deaths, marriages etc.  Each NILS member can have many events.  Only those with a valid link are included. Those that were expected to be matched but where a match could not be found are excluded. |
| Table size (number of rows) | 327589 [q23] records in NILS\_RSU\_DEC2013 (always growing) |
| Unique Identifiers | NILSID |
| Tables linked to | via LINKID  Birth of NILS member – GROBID to BIRTHSSTATS  Birth of baby to NILS member – GROBID to BIRTHSSTATS  Death of NILS member – GRODID to DEATHSTATS |
| Frequency of the Data | At each new release of the data. See metadata for table called DATARELEASED |
| Variables | |  |  | | --- | --- | | Variable Name | Variable Description | | NILSID | Unique identifier | | EVENT\_TYPE\_NAME  EVENT\_TYPE\_CODE | Textual description of the event type  Coded description of event | | LINKID | ID for the relevant table. For a birth event this will be a GROBID and will begin with a ‘B’. For a death event this will be a GRODID and will begin with a D. | |
| Variable Values | EVENT\_TYPE\_CODE  And EVENT\_TYPE\_NAME currently has the following values:  BF - BIRTH TO NILS DAD  BM - BIRTH TO NILS MUM  BB - BIRTH  DL - DEATH |

**Event\_Type**

The following table gives the distribution of the event type.

### EVENTS: Table 12 [q24] – Number of Events by Event\_Type

|  |  |
| --- | --- |
| EVENT\_TYPE | Number of Events |
| DEATH | 42291 |
| BIRTH | 97933 |
| BIRTH TO NILS DAD | 87848 |
| BIRTH TO NILS MUM | 99517 |

A birth can be represented 1, 2 or 3 times depending on whether the baby, mother and father are NILS members. Death registrations can only be linked to one NILS member and therefore will only be in the Events table once. This may change when infant deaths to NILS members are included.

### EVENTS: Table 13 [q25] – Unique Birth and Death Events

|  |  |
| --- | --- |
| Number of Birth event Records | 285298 |
| Number of Unique birth registrations | 207684 |
| Number of Unique Death registrations | 42291 |

The following table shows the number of records for each baby, mum and dad combination.

### EVENTS: Table 14 [q26] – Births – Combination of Baby, Mum and Dads

|  |  |
| --- | --- |
| **Combination** | **Number of Records** |
| Mum Only | 53731 |
| Baby Only | 52603 |
| Dad Only | 45021 |
| Baby & Mum (no dad) | 20619 |
| Mum & Dad (no baby) | 18116 |
| Baby & Dad (no mum) | 17660 |
| Baby, mum & dad | 7051 |

The probability of the birth having a NILS mum is 0.28, the probability of the birth having a NILS dad is 0.28 and the probability of the baby being a NILS member is 0.28.

The probability of each birth having a mum, dad or baby is 0.627. Therefore approximately 62.7% of published births to NI residents should be included in the NILS.

# Meta Data for BIRTHSTATS

|  |  |
| --- | --- |
| Database Name | NILS |
| Table Name | BIRTHSSTATS |
| Table Description | This table gives the statistical coded information for the birth registrations. This information is provided by the informant at the time of birth registration and coded/validated by the teams in the GRO and in DMB. |
| Table size (number of rows) | 214801 [q27] records in NILS\_RSU\_DEC2013 (always growing) |
| Unique Identifiers | GROBID |
| Tables linked to | via LINKID TO EVENTS |
| Frequency of the Data | At each new release of the data more births will be included. See metadata for table called DATARELEASED |
| Variables | |  |  |  |  | | --- | --- | --- | --- | | GROBID | REGYR | REGMONTH | REGCOUN | | OCCYR | OCCMONTH | HMEADDCO | POB | | POB05 | OUTSIDENI | SEX | FAGE | | MAGE | TOTALPREV | TSB | TLB | | MLB | MB | MARSTATOFPARENTS |  | | DUROFMARR | PREVMARR | SOCIALCL | SOCIALCL01 | | FEMPSTAT | MEmpStat | EMPSTAT | EMPSTATNEW | | MEMPSTATNEW | FEMPTSTATNEW | MOCCCDE | FOCCCDE | | MSOCCLASS | FSOCLASS | SOA2001 |  | |
| Variable Values | See Data Dictionary for details of Variables included. |
|  | Legal requirements: Babies born in Northern Ireland must be registered within 42 days of the birth.  Who can register a birth? For married couples either parent can register the birth on their own. However, in the case of a child born to an unmarried couple, the name of the father may only be recorded in the entry of birth if both parents attend and sign the registration together or a declaration of paternity is produced.  Declaration of paternity: An unmarried father who registers the birth of his child jointly with the child's natural mother, and has his name recorded on the birth registration form, will for children born on or after 15 April 2002, acquire parental responsibilities.  The following people may also register the birth:   * grandmother * grandfather * uncle or aunt of the baby who has knowledge of the birth * any person present at the birth * any person having charge of the child * the occupier of the premises where the baby was born or * District Registrars in Northern Ireland.   The following information is required to register a birth:   * a birth registration form filled in by the person registering the birth (usually the mother) * full name of the baby - you can register your child's name in any language providing you use any unicode character * sex and date of birth of the baby * district and place of birth of the baby * full names and dates of birth of parents * full addresses and occupations of the parents. |

# Meta Data for DEATHSSTATS

|  |  |
| --- | --- |
| Database Name | NILS |
| Table Name | DEATHSSTATS |
| Table Description | This table gives the statistical coded information for the death registrations. This information is provided by the informant at the time of death registration and coded/validated by the teams in the GRO and in DMB.  Some deaths can be very late registrations due to coroners cases etc. |
| Table size (number of rows) | 42291 [q29] records in NILS\_RSU\_DEC2013 (always growing) |
| Unique Identifiers | GRODID |
| Tables linked to | via GRODID to LINKID in EVENTS |
| Frequency of the Data | At each new release of the data more deaths will be included. See metadata for table called DATARELEASED |
| Variables | GRODID REGDIST REGMNTH  RegYr BRTHMNTH BRTHYR  DTHMNTH DTHYR Age  COUNTRYOFBIRTH COUNTRYOFUSUALRESIDENCE  EmpStat EmpStat05 HMEDIST  Hypertension ICD10CHAP MAINCAUSE  MARITAL OccCde OutsideNI PLACEOFDEATHCODED Sex SOA\_USRES SocialClass SocialClass01 TYPEOFDT TYPEOFDT05 |
| Variable Values | See Data Dictionary for details of Variables included. |
|  | Registering a death: deaths are normally registered within five days.  When and where to register a death: In Northern Ireland a death should be registered within five days, to allow funeral arrangements to be made. This is with the exception of deaths which have been referred to the coroner.  A death can be registered with the registrar in the district in which the person died, or in the district in which the person normally lived, if within Northern Ireland.  A death which occurs in Northern Ireland, can be registered by:   * any relative of the deceased who has knowledge of the details required to be registered (this includes a relative by marriage) * a person present at the death * a person taking care of the funeral arrangements * the executor or administrator of the deceased's estate * the governor, matron or chief officer of a public building where the death occurred * a person living in and responsible for a house, lodgings or apartments where the death occurred * a person finding, or a person taking charge, of the body   Most deaths are registered by a relative of the deceased. The registrar would normally only allow one of the other people listed above to do so if there are no relatives available or they cannot be traced.  Information required to register a death:   * full name and surname of the deceased * date and place of death and usual address * marital status (single, married, widowed or divorced) * date and place of birth * occupation of the deceased (if the deceased was a wife or widow, the full name and occupation of her husband or deceased husband) will be required * if the deceased was a child, the full names and occupation of the father will be required, or where the parents are not married, the full name and occupation of the mother will be required * maiden surname (if the deceased was a woman who had married) * the name and address of the deceased's GP * details of any pension apart from a state pension that the deceased may have held |

# Metadata for ADDRESS\_HISTORY

|  |  |
| --- | --- |
| Database Name | NILS |
| Table Name | ADDRESS\_HISTORY |
| Table Description | Information on all addresses that each NILS member has lived at. One record is added for each person for each download.  This data forms the basis of all migration data for the NILS |
| Table size (number of rows) | 13050085 [q30] in NILS\_RSU\_DEC2013 (always growing) |
| Unique Identifiers | NILSID |
| Tables linked to | Via NILSID : CORENILSDATA  Via XUPRN : PROPERTIES |
| Frequency of the Data | Latest information included for April 2012 in the NILS\_RSU\_ DEC2013. Updated and released every 6 months |
| Variables | |  |  | | --- | --- | | Variable Name | Variable Description | | NILSID | Unique identifier | | SOURCE | This gives the download for the address record | |  |  | | CURRENT\_FLAG | This indicates which address is the current address (the current address is included in the NILSCOREDATA table) | | TYPE | This indicates whether the address changed, and if so why. See below for descriptions | | SOA2001 | The Super Output Area that the address is in | | XUPRN | The anonymised properties id | | PREVADD | This gives the download for the previous address record | | PREV\_SOA2001 | The Super Output Area of the previous address | | PREV\_XUPRN | The anonymised properties id of the previous address | |
| Variable Values | SOURCE  Format YYYY04 (April download in YYYY and YYYY10 (October download in year YYYY)  CURRENT\_FLAG  C – Current Address  P – Previous Address   |  |  | | --- | --- | | TYPE |  | | NC | No Address Change | | AC | Address Change | | EM | Emigration | | NR | Original record from the first NILS download in April 2001 | | RE | Reentrant | |  |  |   SOA2001  890 valid SOA codes  XXXXXX is a missing value (normally invalid postcodes) |
| What is an address change? Address changes have been notified to the BSO by the patient, the GP or through a data cleansing exercise. Any new address is entered using QuickAddress software and the system automatically updates the date of address change field. The new postcode and/or unique property reference numbers are stored. An address change on the NILS is determined by a change in the postcode or UPRN. Quality Assurance As each new download is loaded onto the NILS the migration events are created. The number of address changes is sent to the BSO for quality assurance. Data Issues: Some data cleansing of the BSO data may show up as an address change  This data does not pick up moves that occur within a six-month period. For example:   * if someone moves several times between 2 BSO downloads this would only be picked up as 1 move with the address set to the latest BSO download address * if someone moved out of an address and back into the same address between 2 downloads this would not be recorded on NILS.   The BSO changed IT systems from the Central Health Index (CHI) to the NHAIS system in 2005. There was a slight increase in the number of address changes (and hence migration events) at this time since the BSO worked closely with GPs to remove people from their lists who had emigrated but the BSO had not been notified.  If someone returns to Northern Ireland the Health Card Registration System is searched and if the original record is found it is reactivated for that person. In a small number of cases this does not happen and a new Health and Care number is created. Subsequent matching exercises may link this person back to the old NHS\_number or CHI number and therefore duplicate records can be identified. Approximately 15-50 of these occur in every 6 monthly download of the NILS.  Paul Barr has analysed the BSO migration data and compared against the 2000-2001 census migration data to look at the timing of the migration events. In summary his research has shown that the address change (migration) events did happen but that there was a delay in notifying the BSO. This differs based on socio-economic characteristics. Therefore caution should be used when analysing the download of the move. | |

### 

### ADDRESS\_HISTORY: Table 15 [q31] – Number of Address Changes

|  |  |
| --- | --- |
| **Number of Address Changes** | **Records** |
| 1 | 164274 |
| 2 | 67621 |
| 3 | 30387 |
| 4 | 12408 |
| 5 | 5277 |
| 6 | 2378 |
| 7 | 1100 |
| 8 | 524 |
| 9 | 258 |
| 10+ | 242 |

## SOURCE

The following table gives the number of records for each Source variable i.e. it gives an indication of the number of records with addresses.

### ADDRESS\_HISTORY: Table 16 – [q32] Distribution of Source

|  |  |
| --- | --- |
| **Source** | **Records** |
| 200104 | 514555 |
| 200110 | 519304 |
| 200204 | 524153 |
| 200210 | 529372 |
| 200304 | 534007 |
| 200310 | 538993 |
| 200404 | 544345 |
| 200410 | 549102 |
| 200504 | 540358 |
| 200510 | 546616 |
| 200604 | 553313 |
| 200610 | 559651 |
| 200704 | 567304 |
| 200710 | 574462 |
| 200804 | 581220 |
| 200810 | 588236 |
| 200904 | 594256 |
| 200910 | 600356 |
| 201004 | 606087 |
| 201010 | 612349 |
| 201104 | 618450 |
| 201110 | 624350 |
| 201204 | 629881 |

## 

## Type

The following table gives the detail on the type of address change. The large numbers of records relate to the first address (NR) and the address changes (AC) that are notified.

### ADDRESS\_HISTORY: Table 17 [q33] – Distribution of Type

|  |  |
| --- | --- |
| **Type** | **Records** |
| AC | 497835 |
| EM | 71048 |
| NC | 11825422 |
| NR | 646162 |
| RE | 10253 |

## Current\_Flag

The following table shows the number of current addresses (this should be the same number as the number of records in CORENILSDATA) and previous addresses.

### ADDRESS\_HISTORY: Table 18 [q34] – Distribution of Current Flag

|  |  |
| --- | --- |
| **Current\_Flag** | **Number of Records** |
| C | 646162 |
| P | 12404558 |

### ADDRESS\_HISTORY: Table 19 [q35] – Distribution of Current Flag for all ‘Live records at April 2012’

|  |  |
| --- | --- |
| **Current\_Flag** | **Number of Records** |
| C | 541800 |
| P | 10536459 |

## Note 541800 the number of live records at the December 2013 download [q35a].

## SOA2001

### ADDRESS\_HISTORY: Table 20 [q37] – Coverage of SOA2001

|  |  |  |  |
| --- | --- | --- | --- |
| **SOA Code** | **Description** | **Number of Records** | **% Distribution** |
| XXXXXXXX or null | missing (invalid/missing postcodes) | 124862 | 1 |
| Valid SOA | Valid SOA Code | 12925858 | 99 |
| Total Records |  | 13050720 | 100 |

## XUPRN

### ADDRESS\_HISTORY: Table 21 [q38] – Coverage of XUPRN

|  |  |  |  |
| --- | --- | --- | --- |
| **XUPRN** | **Description** | **Number of Records** | **% Distribution** |
| Null | No property ID available | 749859 | 6 |
| Valid XUPRN | Valid property IDs | 12300861 | 94 |
| Total Records |  | 13050720 | 100 |

# Meta Data for MIGRATION\_EVENTS

|  |  |
| --- | --- |
| Database Name | NILS |
| Table Name | Migration\_Events |
| Table Description | Information on all migration events as derived from the Health Card registration system and in particular the Address\_History database. This includes internal migration, immigrants, emigration and information on people who have reentered. |
| Table size (number of rows) | 634100 [q39] in NILS\_RSU\_ DEC2013 (always growing) |
| Unique Identifiers | NILSID |
| Tables linked to | CORENILSDATA |
| Frequency of the Data | Latest information included for April 2013. Updated and released every 6 months |
| Variables | |  |  | | --- | --- | | Variable Name | Variable Description | | NILSID | Unique identifier | | DATEMOVED | This is only a proxy for date moved - determined by the date the BSO is notified and NILS updated in the six-monthly downloads | | CHANGETYPE | Variable that identifies the different migration types | | OUTOF\_SOA2001 | The SOA code of the SOA that the person moved out of | | INTO\_SOA2001 | The code of the SOA that the person moved into | | OUTOF\_XUPRN | The anonymised code of the property that the person moved out of | | INTO\_XUPRN | The anonymised code of the property that the person moved into | | ORDER | The order of the migration event since April 2001. 1 is the first , 2 is the second …. | |
| Variable Values | CHANGETYPE has the following values. Percentages are given as approximate distribution [q40].  IM – Immigrant – 10%  RE – Re-Entrant – 2%  EM – Emigration – 11%  AC – Address Change – 77%  INTO\_SOA2001  890 valid SOA codes  NULL is not required  XXXXXX is a missing value (normally invalid postcodes)  OUT OF\_SOA2001  890 valid SOA codes  NULL is not required  XXXXXX is a missing value (normally invalid postcodes)  XUPRN is an anonymised property ID that can be used to link to the Property\_Data table |
| *How migration events were created?*  A download of demographic data is taken every six months from the BSO. The address information of the latest download is compared with the previous download using Unique Property Reference Numbers, postcodes and date of address changes from the BSO system. If there has been a change in address then this is recorded as a migration event. There are 4 different types of migration event and all are included in the table below. They are distinguished by the ChangeType variable.  *Address Change (Internal\_Migration)*  If the person is flagged on the current download as live and was on the previous download as live then the event is classed as an internal migration event and the ChangeType is set to ‘AC’ (short for address change). The ‘into’ SOA and XUPRN are set to the current address and the ‘outof’ SOA and XUPRN are set to the previous address.  *Emigration*  If the person was previously Live and has now been flagged as moved out of NI then an emigration record is created. The ChangeType is set to ‘EM’. The ‘out of’ SOA and XUPRN are set to the previous address and the ‘into’ SOA and XUPRN are set to null.  *Immigration*  If the person has not been on the NILS (i.e. since 2001) system before, and for those born on/after 1997 and we have not created a birth link for them, they are treated as an immigration event and the ChangeType is set to ‘IM’. The ‘out of’ SOA and XUPRN are set to null and the ‘into’ SOA and XUPRN are set to the current address.  *Re-entrants*  If the NILS member has been on the NILS system before but is currently flagged as ‘away’ their return is treated as a re-entrant event and the ChangeType is set to ‘RE’. The ‘out of’ SOA and XUPRN are set to the address the NILS member had prior to emigration and the ‘into’ SOA and XUPRN are set to the current address.  *Order*  The order is set at each extract and is the rank order of the event. If a NILS member has 5 events they are sorted in date order with the earliest event getting an order of 1 and the latest event getting an order of 5. | |

## 

## Type

The following table shows the number of each type of migration event from 2001 - 2012.

### MIGRATION\_EVENTS: Table 22 [q41] – Type of event

|  |  |  |
| --- | --- | --- |
| ***ChangeType*** | ***Description of Change Type*** | **Number of Records** |
| *AC* | *Internal Migration* | 491347 |
| *EM* | *Emigrants* | 71045 |
| *IM* | *Immigrants* | 61456 |
| *RE* | *Re-entrants* | 10252 |
| *All Types* | *All Migration Events* | 634100 |

### MIGRATION\_EVENTS: Table 23 [q42] – Type of event by DateMoved

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **DATEMOVED** | **AC** | **EM** | **IM** | **RE** | **TOTAL** |
| 200110 | 19967 | 2865 | 2891 | 484 | 26207 |
| 200204 | 20317 | 3131 | 1992 | 419 | 25859 |
| 200210 | 20272 | 3533 | 2197 | 464 | 26466 |
| 200304 | 19215 | 2896 | 1933 | 400 | 24444 |
| 200310 | 20228 | 2761 | 2117 | 415 | 25521 |
| 200404 | 19428 | 3867 | 2422 | 373 | 26090 |
| 200410 | 24908 | 2385 | 1989 | 460 | 29742 |
| 200504 | 12723 | 4077 | 3106 | 370 | 20276 |
| 200510 | 22367 | 4905 | 2901 | 439 | 30612 |
| 200604 | 19687 | 4116 | 3557 | 362 | 27722 |
| 200610 | 24331 | 2469 | 3106 | 536 | 30442 |
| 200704 | 23954 | 3557 | 4037 | 497 | 32045 |
| 200710 | 22629 | 2911 | 3589 | 619 | 29748 |
| 200804 | 20613 | 4378 | 3425 | 469 | 28885 |
| 200810 | 20814 | 3109 | 3348 | 488 | 27759 |
| 200904 | 20539 | 2925 | 2840 | 420 | 26724 |
| 200910 | 21097 | 2048 | 2475 | 602 | 26222 |
| 201004 | 20273 | 3405 | 2444 | 426 | 26548 |
| 201010 | 20037 | 2194 | 2466 | 563 | 25260 |
| 201104 | 19809 | 3138 | 2574 | 405 | 25926 |
| 201110 | 34180 | 2846 | 2416 | 585 | 40027 |
| 201204 | 43959 | 3529 | 3631 | 456 | 51575 |

\*\* the dates used are date of the download.

## Out of\_SOA2001

The following table shows the proportion of records that had an SOA allocated. ‘XXXXXX’ are those that could not be allocated (most likely because the postcode was missing). ‘0000000’ are those where the postcode was valid but the CPD has not allocated an SOA.

### MIGRATION\_EVENTS: Table 24 [q43] – Coverage of Outof\_SOA2001

| **OUTOF\_ SOA2001** | **Description** | **Records** | **% of all migration records** |
| --- | --- | --- | --- |
| NULL | Not required or not linked to the CPD | 69407 | 10.9 |
| Valid Records |  | 564693 | 89.1 |
| All Records |  | 634100 | 100.0 |

### MIGRATION\_EVENTS: Table 25 [q44] – Coverage of Outof\_SOA by ChangeType

The following table shows the number and proportion of records that had an Outof\_SOA allocated (excluding missing, ‘000000’ or ‘XXXXXX’ values) for each type of migration event.

|  |  |  |  |
| --- | --- | --- | --- |
| **ChangeType** | **OUTOF\_SOA Assigned Records** | **All Records** | **% Distribution** |
| AC | 484388 | 491347 | 99 |
| EM | 70088 | 71045 | 99 |
| IM | 0 | 63530 | 0 |
| RE | 10217 | 10252 | 100 |

### MIGRATION\_EVENTS: Table 26 [q45] – Coverage of Outof\_SOA by DateMoved

|  |  |  |
| --- | --- | --- |
| **DateMoved** | **OUTOF\_SOA2001 Records** | **% OF ALL MIGRATION** |
| 201204 | 47253 | 99 |
| 201110 | 36765 | 98 |
| 201104 | 23097 | 99 |
| 201010 | 22581 | 99 |
| 201004 | 23880 | 99 |
| 200910 | 23520 | 99 |
| 200904 | 23632 | 99 |
| 200810 | 24172 | 99 |
| 200804 | 25200 | 99 |
| 200710 | 25897 | 99 |
| 200704 | 27727 | 99 |
| 200610 | 26925 | 98 |
| 200604 | 23924 | 99 |
| 200510 | 27265 | 98 |
| 200504 | 16891 | 98 |
| 200410 | 27033 | 97 |
| 200404 | 23326 | 99 |
| 200310 | 23146 | 99 |
| 200304 | 22204 | 99 |
| 200210 | 23875 | 98 |
| 200204 | 23474 | 98 |
| 200110 | 22906 | 98 |

\*\* Immigration events excluded

## INTO\_SOA2001

### MIGRATION\_EVENTS: Table 27 [q46] – Coverage of Into\_SOA2001

| **INTO\_SOA2001** | **Description** | **Records** | **% of all migration records** |
| --- | --- | --- | --- |
| NULL | Not required or not linked to CPD | 76293 | 12 |
| Valid Records |  | 557807 | 88 |
| All Records |  | 634100 | 100 |

The following tables show the number and proportion of records that had an Into\_SOA allocated (excluding missing, ‘000000’ or ‘XXXXXX’ values) for each type of migration event.

### MIGRATION\_EVENTS: Table 28 [q47] Coverage of Into\_SOA2001 by ChangeType

| **ChangeType** | **INTO\_SOA Assigned Records** | **All Records** | **% Distribution** |
| --- | --- | --- | --- |
| AC | 486547 | 491347 | 99 |
| IM | 63100 | 63530 | 99 |
| RE | 10216 | 10252 | 100 |
| EM | 0 | 71045 | 0 |

### MIGRATION\_EVENTS: Table 29 [q48] – Coverage of Into\_SOA2001 and Outof\_SOA by DateMoved

|  |  |  |
| --- | --- | --- |
| **DateMoved** | **INTO\_SOA Records** | **% DISTRIBUTION** |
| 201204 | 47195 | 98 |
| 201110 | 36867 | 99 |
| 201104 | 22559 | 99 |
| 201010 | 22875 | 99 |
| 201004 | 22937 | 99 |
| 200910 | 23962 | 99 |
| 200904 | 23617 | 99 |
| 200810 | 24411 | 99 |
| 200804 | 24123 | 98 |
| 200710 | 26446 | 99 |
| 200704 | 28198 | 99 |
| 200610 | 27788 | 99 |
| 200604 | 23450 | 99 |
| 200510 | 25560 | 99 |
| 200504 | 16049 | 99 |
| 200410 | 27176 | 99 |
| 200404 | 22023 | 99 |
| 200310 | 22597 | 99 |
| 200304 | 21369 | 99 |
| 200210 | 22799 | 99 |
| 200204 | 22600 | 99 |
| 200110 | 23206 | 99 |

\*\* Emigration events excluded

## 

## Outof\_XUPRN

The following table shows the proportion of records that had a unique property ID allocated (excluding missing or postcode values) for each type of migration event and the time-period of the migration event.

A total of 106568 [q49] records did not have a valid XUPRN– these were not required (for Emigration events) and not assigned.

### MIGRATION\_EVENTS: Table 30 [q50] – Coverage of Outof\_XUPRN by ChangeType

|  |  |  |  |
| --- | --- | --- | --- |
| **ChangeType** | **OUTOF\_XUPRN** | **TOTAL RECORDS** | **% ASSIGNED** |
| AC | 452588 | 491347 | 92 |
| EM | 65030 | 71045 | 92 |
| IM | 0 | 61456 | 0 |
| RE | 9914 | 10252 | 97 |

### MIGRATION\_EVENTS: Table 31 [q51] – Coverage of Outof\_XUPRN by DateMoved

|  |  |  |
| --- | --- | --- |
| **DateMoved** | **OUTOF\_XUPRN** | **% ASSIGNED** |
| 200110 | 21351 | 92 |
| 200204 | 21951 | 92 |
| 200210 | 22395 | 92 |
| 200304 | 20862 | 93 |
| 200310 | 21841 | 93 |
| 200404 | 21998 | 93 |
| 200410 | 24839 | 90 |
| 200504 | 15793 | 92 |
| 200510 | 25523 | 92 |
| 200604 | 22677 | 94 |
| 200610 | 25100 | 92 |
| 200704 | 26197 | 94 |
| 200710 | 24426 | 93 |
| 200804 | 23621 | 93 |
| 200810 | 22736 | 93 |
| 200904 | 22232 | 93 |
| 200910 | 22147 | 93 |
| 201004 | 22526 | 93 |
| 201010 | 21262 | 93 |
| 201104 | 21754 | 93 |
| 201110 | 31471 | 84 |
| 201204 | 44830 | 94 |

\*\* Immigration events excluded

## INTO\_UPRN

### MIGRATION\_EVENTS: Table 32 [q52] – Coverage of Into\_XUPRN by ChangeType

|  |  |  |  |
| --- | --- | --- | --- |
| **ChangeType** | **INTO\_XUPRN** | **Total Records** | **% Assigned** |
| AC | 457039 | 491347 | 93 |
| EM | 0 | 71045 | 0 |
| IM | 56613 | 61456 | 92 |
| RE | 9914 | 10252 | 97 |

### MIGRATION\_EVENTS: Table 33 [q53] – Coverage of Into\_XUPRN by DateMoved

|  |  |  |
| --- | --- | --- |
| **DateMoved** | **INTO\_XUPRN** | **% Assigned** |
| 200110 | 22059 | 95 |
| 200204 | 21442 | 94 |
| 200210 | 21878 | 95 |
| 200304 | 20389 | 95 |
| 200310 | 21104 | 93 |
| 200404 | 20844 | 94 |
| 200410 | 25518 | 93 |
| 200504 | 14252 | 88 |
| 200510 | 23996 | 93 |
| 200604 | 21910 | 93 |
| 200610 | 25840 | 92 |
| 200704 | 26366 | 93 |
| 200710 | 24722 | 92 |
| 200804 | 22289 | 91 |
| 200810 | 22795 | 92 |
| 200904 | 22048 | 93 |
| 200910 | 22404 | 93 |
| 201004 | 21358 | 92 |
| 201010 | 21315 | 92 |
| 201104 | 20852 | 92 |
| 201110 | 35442 | 95 |
| 201204 | 44743 | 93 |

# Meta Data for CENSUSP\_2001

|  |  |
| --- | --- |
| Database Name | NILS |
| Table Name | CENSUSP\_2001 |
| Table Description | This table gives the statistical coded information from the 2001 Census for all NILS members with a census link and all people living in a household with a NILS member  This table only includes enumerated population  Also includes some students identified as being away at term time. To filter these out users should use to the STU\_INDP0 and TTADDP0 fields |
| Table size (number of rows) [q54] | 1060172 records ( 452604 NILS\_Members, 607568 Other Household Occupants) records in NILS\_RSU\_DEC2013 (static) |
| Unique Identifiers | NILSID |
| Tables linked to | via CENSUSHID to CENSUSHH\_2001 |
| Frequency of the Data | A small number of records may be added at each download because of additional matching exercises |
| Variables | NILSID (Only included for NILS members, Blank for non-NILS members)  NILSID NILS\_MEMBERP0 CENSUSHID0 CENSUSPID0 PRSN\_TYPP0 POSCOMMP0  SEXP0 AGEP0 WRKPEN\_INDP0  MARITALP0 STU\_INDP0 TTADDP0  STUACCP0 KNOWIRISHP0 CMMNTY\_BCKGRNDP0  RLGNP0 COBP0 ETH\_GRPP0  ETHNICITYP0 GHEALTHP0 LLTIP0  UNPAIDCAREP0 LIVARRP0 HHRPP0  FRPP0 DEPCHLDP0 DEPPERSP0  FAMSTATP0 GENINFAMP0 MIGSTATP0  COUNADD1YRP0 MIG\_OAP0 MIG\_PARLP0  MIG\_HBP0 MIG\_ELBP0 MIG\_NUTSP0  MGRPP0 EDQUAL\_HIGHP0 ACTLWKP0  ECACTP0 OCCUPP0 INDUSTRYP0  NSSECP0 NOHOURSP0 EMPSTATP0  YEARLSTWRKDP0 COMPSIZEP0 TRAVWRKP0  TRVWDISTP0 TRVWDISTGROUPP0 WP\_LOCP0  WP\_COUNTRYP0 TRV\_OAP0 TRV\_PARLP0  READIRISHP0 SPEAKIRISHP0 UNDERIRISHP0  WRITEIRISHP0 RELPRACP0 RELUPBRP0  EDLEV01P0 EDLEV02P0 EDLEV03P0  EDLEV04P0 EDLEV05P0 EDLEV06P0  EDLEV07P0 EDLEV08P0 EDLEV09P0  EDLEV10P0 EDLEV11P0 EDLEV12P0  EDLEV13P0 SOAADD1YRP0 SOAENUMP0  IMPUTEDPERSON\_NOTAWAYSTUDENTP0 SETTLEMENTBANDP0  (see metadata section on imputation variables for the following: )  LLTI\_IMPP0 GHEALTH\_IMPP0 RELG\_IMPP0  RELUPBR\_IMPP0 EDU\_IMPP0 GENDER\_IMPP0  AGE\_IMPP0 ETHNIC\_IMPP0 MARITAL\_IMPP0  ACTLW\_IMPP0 UNPAIDCARE\_IMPP0 |
| Variable Values | See Data Dictionary for details of Variables included. |

# Meta Data for CENSUSHH\_2001

|  |  |
| --- | --- |
| Database Name | NILS |
| Table Name | CENSUSHH\_2001 |
| Table Description | This table gives the statistical coded information from the 2001 census for households with a NILS member with a census link  This table only includes households with the enumerated population |
| Table size (number of rows) | 318316 records in NILS\_RSU\_DEC2013 (static) |
| Unique Identifiers | CENSUSHID |
| Tables linked to | via CENSUSHID to CENSUSP\_2001 to get NILS members  or other household members |
| Frequency of the Data | A small number of records may be added at each download because of additional matching exercises |
| Variables | CENSUSHID0 SOAENUM CETYP\_H0  CEMANTYPEH0 CECOMBTYPEH0 CEREGSTATH0  HHOCCSTATH0 HHSIZEH0 ACCTYPEH0  TENUREH0 SELDCONTH0 HHROOMS\_COUNTH0  HHROOMSREQH0 PERSPERROOMH0 OCCRATH0  BATHSHOWH0 CENTHEATH0 LWST\_FLR\_LVLH0  HHFLRS\_COUNTH0 HHCARS\_COUNTH0 HH17PLS\_COUNTH0  HHWRKG\_COUNTH0 HHPEN\_COUNTH0 HHADULST\_COUNTH0  HHDEPCHLD\_COUNTH0 HHSTUHOME\_COUNTH0  HHADEMP\_COUNTH0 HHCOMPH0 HHFAMTYPEH0  HHPENSH0 HHADCHLDSTRH0 HHDEPCHLDH0  HHCARERSH0 HHLLTIH0 STUAWAYH0  HHCOMBACKH0 HHETHRELH0 HHETHSTRH0  HHMIG\_INDH0 HHMIG\_OAH0 HHMIG\_PARLH0  HHMIG\_HBH0 HHMIG\_ELBH0 HHMIG\_NUTSH0  HHWKRSTRANSH0 HRP\_SEXH0 HRP\_AGEH0  HRP\_MARSTATH0 HRP\_COMMBACKH0 HRP\_RELH0  HRP\_COBH0 HRP\_ETHGRPH0 HRP\_EDHLQH0  HRP\_ECACTH0 HRP\_OCCH0 HRP\_INDH0  HRP\_NSSH0 HRP\_SOCGRDH0 HHFAM\_COUNTH0  HHCARERS\_COUNTH0 HHLLTI\_COUNTH0 HHPEOPLE\_COUNTH0  CEAGE\_ELDH0 CEAGE\_ADH0 CEAGE\_CLHDH0  CETYP\_PHYH0 CETYP\_LDH0 CETYP\_MHH0  CETYP\_CNVH0 CETYP\_DRUGH0 CETYP\_TLH0  CETYP\_CIH0 CETYP\_AIH0 CETYP\_ELDH0  CETYP\_STH0 CETYP\_PRSH0 CETYP\_NURSH0  CETYP\_AFH0 CETYP\_HMLSSH0 CETYP\_OTHH0  CETYP\_NOURH0 OWNERSHIPH0 LNDLRDH0  NS\_DEP\_EMPH0 NS\_DEP\_EDUH0 NS\_DEP\_HEAH0  NS\_DEP\_HOUSH0 NS\_DEP\_TENH0 XUPRN  (see metadata section on imputation variables for the following: )  CENTHEAT\_IMPH0  CARS\_IMPH0  TENURE\_IMPH0 |
| Variable Values | See Data Dictionary for details of Variables included. |

# Meta Data for CENSUSP\_2011

|  |  |
| --- | --- |
| Database Name | NILS |
| Table Name | CENSUSP\_2011 |
| Table Description | This table gives the statistical coded information from the 2011 census for all NILS members with a census link and all people living in a household with a NILS member  This table only includes enumerated population  Also includes some students identified as being away at term time. To filter these out users should use to the STUDENTP1 and TERMINDP1 fields. People who are not identified as permanent residents of NI can be excluded using the INTENTIONP1 field. |
| Table size (number of rows) | 1075590 records (77737 NILS Members, 597853 Other Household Occupants) records in NILS\_RSU\_DEC2013 (static) |
| Unique Identifiers | NILSID |
| Tables linked to | via CENSUSHID1 to CENSUSHH\_2011  NILSID NILS\_MEMBERP1 CENSUSPID1  CENSUSHID1 ACTLWP1 ADULTLSP1  AGGECOP1 AGGMAINLANGP1 AGGMAINLANGPRFP1  AMAINLANGPRFP1 DCHP1 ECOCATP1  ECOP1 ELARP1 FMSP1  FRPP1 HLQP1 HRPP1  INDEP1 INDP1 LARP1  LRESP1 MAINLANGPRFP1 NATIDBP1  NATIDEP1 NATIDSP1 NATIDUSP1  NATIDWP1 NSSECP1 OCCP1P1  OCCP2P1 OCCP3P1 PENEXACTP1  PENP01P1 PENP1 PSSPUKP1  PSSPIP1 PSSPOP1 PSSPNP1  PTRANSP1 QUALS01P1 QUALS02P1  QUALS03P1 QUALS04P1 QUALS05P1  QUALS06P1 QUALS07P1 QUALS08P1  QUALS09P1 QUALS10P1 QUALS11P1  QUALS12P1 QUALS13P1 QUALS14P1  RLARP1 STAINDP1 STAP1  UNEMPHISTP1 WF65PLP1 WRKAGEP1  EMPLOYP1 TRANSPORTP1 HEACONDBP1  MIGORIGP1 HEACONBP1 USLANRP1  INDUSTRY\_CODEP1 HEACONCIP1 USLANSP1  AGGLASTYRWRKP1 DISABILITYP1 OCCP1  ILANUP1 LASTYRWRKP1 HEACONP1  COBP1 GENINFAMP1 NATIDNIP1  PSPTELP1 AVAILWORKP1 HEALTHP1  EVERWORKP1 HEACONLTP1 IRISH1P1  ILANRP1 YRARR\_YEARP1 RELBTP1  MARSTATP1 EMPLYGRPP1 LOOKWORKP1  ULSTER2P1 AGEARRP1 WAITWORKP1  YRADINTP1 CPRP1 RELBUP1  PSSPRTP1 YRARRP1 HEACONMDP1  VOLWORKP1 RELBTBUIP1 HEACONMLP1  NATIDOP1 MAINLANGP1 STUDENTP1  ILANWP1 HOURSP1 IRISH2P1  LANGPRFP1 ETHP1 HEACONDP1  CARERP1 USLANWP1 IDENINTP1  WKPLINDP1 NATIDIP1 HEACONOCP1  INTENTIONP1 TERMINDP1 SCHOOLAGEP1  HEACONLDP1 ULSTER1P1 POSITIONP1  RESIDENCE\_TYPEP1 SEXP1 USLANUP1  EMPSTATP1 IDENUKP1 ILANSP1  WKPLINTP1 HEACONCDP1 AGEP1  HEACONMHCP1 ETHFULLP1 A1YR\_ELB\_2011P1  A1YR\_HSCT\_2011P1 A1YR\_HSSB\_2011P1  A1YR\_LA\_CODE\_2011P1 A1YR\_NUTS3\_2011P1  A1YR\_SOA\_CODE\_2011P1 WPS\_ELB\_2011P1  WPS\_HSCT\_2011P1 WPS\_HSSB\_2011P1 WPS\_LA\_CODE\_2011P1  WPS\_NUTS3\_2011P1 WPS\_SOA\_CODE\_2011P1 |
| Frequency of the Data | A small number of records may be added at each download because of additional matching exercises |
| Variables | NILSID (Only included for NILS members, Blank for non-NILS members)  List of variable need updated because of trailing 1  (see metadata section on imputation variables for the following: )  AGE\_IMPP1 AVAILWORK\_IMPP1 CARER\_IMPP1  COB\_IMPP1 DISABILITY\_IMPP1 EMPLYGRP\_IMPP1  EMPSTAT\_IMPP1 EVERWORK\_IMPP1 HEALTH\_IMPP1  HOURS\_IMPP1 IDENINT\_IMPP1 IDENUK\_IMPP1  INDUSTRY\_CODE\_IMPP1 INTENTION\_IMPP1  LANGPRF\_IMPP1 LASTYRWRK\_IMPP1 LOOKWORK\_IMPP1  MAINLANG\_IMPP1 MARSTAT\_IMPP1 PSPTEL\_IMPP1  PSSPRT\_IMPP1 RESIDENCE\_TYPE\_IMPP1 SEX\_IMPP1  STUDENT\_IMPP1 TERMIND\_IMPP1 TRANSPORT\_IMPP1  WAITWORK\_IMPP1 WKPLIND\_IMPP1 WKPLINT\_IMPP1  YRADINT\_IMPP1 YRARR\_YEAR\_IMPP1 |
| Variable Values | See Data Dictionary for details of Variables included. |

# Meta Data for CENSUSHH\_2011

|  |  |
| --- | --- |
| Database Name | NILS |
| Table Name | CENSUSHH\_2011 |
| Table Description | This table gives the statistical coded information from the 2011 census for households with a NILS member with a census link.  This table only includes households with the enumerated population |
| Table size (number of rows) | 349650 records in NILS\_RSU\_DEC2013 (static) |
| Unique Identifiers | CENSUSHID1 |
| Tables linked to | via CENSUSHID1 to CENSUSP\_2011 to get NILS members  or other household members |
| Frequency of the Data | A small number of records may be added at each download because of additional matching exercises |
| Variables | CENSUSHID1 ADAPT1H1 ADAPT2H1  ADAPT3H1 ADAPTHH1 ADAPTOH1  ADAPTPMH1 ADAPTVH1 ADAPTWH1  ADEMH1 ADTH1 AGEGRP17H1  AGEGRP24H1 AGEGRP64H1 AGEGRP65H1  AGGETHH1 AHCH1 AHTH1  CARSNOH1 CEAH1 CECTMCEWSH1  CENHEATH1 CGHH1 CLIENTS01H1  CLIENTS02H1 CLIENTS03H1 CLIENTS04H1  CLIENTS05H1 CLIENTS06H1 CLIENTS07H1  CLIENTS08H1 CLIENTS09H1 CLIENTS10H1  CLIENTS11H1 CLIENTS12H1 CLIENTS13H1  CLIENTS14H1 CLIENTS15H1 CLIENTS16H1  CLIENTS17H1 CLIENTS18H1 CLIENTS19H1  CLIENTS20H1 CLIENTS21H1 HRPCOBH1  CRSH1 DEPEDH1 DEPEMH1  DEPHDH1 DEPHSH1 DEPRIVEDH1  DEPTNH1 DPCH1 EILAH1  ESTNATUREH1 ETHH1 FAMH1  HH\_STRUCTUREH1 HHCH1 HHLDLANGH1  HHLSH1 IHC1H1 IHC2H1  ILAH1 ILLADULTH1 ILLH1  ILLLITTLEH1 ILLLOTH1 LANDLORDH1  MEIGH1 MNAGEMNTH1 HEACONAH1  HEACONH1 NSSH1 NSTAH1  P17H1 PENH1 PPROOMH1  ROOMREQH1 ROOMSH1 SELFCONH1  SIZH1 STAH1 TEMPNATUREH1  TENH1 TENUREH1 TYPACCOMH1  USHC1H1 USHC2H1 NORDH1  TENDH1 HHSDH1 EA\_ELB\_2011H1  EA\_HSCT\_2011H1 EA\_HSSB\_2011H1 EA\_LA\_CODE\_2011H1  EA\_NUTS3\_2011H1 EA\_SOA\_CODE\_2011H1  XUPRN  (see metadata section on imputation variables for the following: )  TYPACCOM\_IMPH1 TENURE\_IMPH1 SELFCON\_IMPH1  ROOMS\_IMPH1 LANDLORD\_IMPH1 CENHEAT\_IMPH1  CARSNO\_IMPH1 |
| Variable Values | See Data Dictionary for details of Variables included. |

# Meta Data for CENSUSP\_1991

|  |  |
| --- | --- |
| Database Name | NILS |
| Table Name | CENSUSP\_1991 |
| Table Description | This table gives the statistical coded information from the 1991 census for all NILS members with a census link and all people living in a household with a NILS member. This is a partial link only that will be completed in September 2014  This table only includes enumerated population  What about not usually resident (i.e. students away at term-time) |
| Table size (number of rows) | 321576 records in NILS\_RSU\_DEC2013 (static) |
| Unique Identifiers | NILSID |
| Tables linked to | via CENSUSHID9 to CENSUSHH\_1991 |
| Frequency of the Data | A small number of records may be added at each download because of additional matching exercises |
| Variables | NILSID CENSUSPID9 CENSUSHID9  RECTYPP9 DISTRICTP9 WARDP9  PERSNOP9 ELBP9 HSSBP9  NUTS3P9 PARLCONP9 USADDCP9  GENDER1P9 GENDERP9 AGEP9  AGE1P9 AGE2P9 MARITALP9  MARITAL1P9 MARFEMP9 COBP9  COB1P9 RELIGIONP9 RELIGN1P9  IRISHP9 LTILLP9 RELATIONP9  WHEREP9 TOTCHILDP9 LYCHILDP9  TERMCODEP9 UALYCODEP9 MIGTYPE1P9  PRIMARYP9 SECONDP9 ECONACTP9  EMPSTATP9 EMPSTAT1P9 ESTSIZEP9  CURRENTP9 OLDJOBP9 HOURSP9  PARTFULLP9 ECACMARWP9 MWPTEMPP9  MAJOCCP9 SOCCLASSP9 SOCCLAS2P9  TRANSPORP9 WKADDCDP9 FIRQUALP9  SECQUALP9 THIRQUALP9 FOURQUALP9  FIFQUALP9 SIXQUALP9 VOCSUBP9  VOCLEVP9 HEADHOUSP9 WMHOHP9  CHECSUPPP9 TEN60P9 AMEN60P9  NOPER2P9 NOPERUR2P9 ESTTYPEP9  DWELTYPEP9 ACCTYPEP9 ACCDWELP9  TENUREP9 SHARACCP9 NOROOM1P9  DENSITYP9 NOCARP9 FAMTYPEP9  PENSHOUSP9 AGECOMBP9 NOEARN1P9  NODECH1P9 PENSCOMBP9 MIGDETP9  MIGDISTP9 |
| Variable Values | See Data Dictionary for details of Variables included. |

# Metadata for CENSUSHH\_1991

|  |  |
| --- | --- |
| Database Name | NILS |
| Table Name | CENSUSHH\_1991 |
| Table Description | This table gives the statistical coded information from the 1991 census for households with a NILS member with a census link.  This table only includes households with the enumerated population |
| Table size (number of rows) | 233028 records in NILS\_RSU\_DEC2013 (static) |
| Unique Identifiers | CENSUSHID9 |
| Tables linked to | via CENSUSHID9 to CENSUSP\_1991 to get NILS members  or other household members |
| Frequency of the Data | A small number of records may be added at each download because of additional matching exercises |
| Variables | CENSUSHID9 RECTYPEH9 DISTRICTH9  WARDH9 ELBH9 HSSBH9  NUTS3H9 PARLCONH9 NOMALEH9  NOFEMH9 NOPERSH9 NOPER2H9  NOMALEURH9 NOFEMLURH9 NOPERURH9  NOPERUR2H9 ESTTYPEH9 DWELTYPEH9  ACCTYPEH9 ACCDWELH9 ACCDWEL2H9  TENUREH9 SHARACCH9 NOROOMH9  NOROOM1H9 DENSITYH9 BATHSHOWH9  TOILETH9 HEATINGH9 WATERH9  SEWAGEH9 NOCARH9 FAMTYPEH9  PENSHOUSH9 AGECOMBH9 NOEARNH9  NOEARN1H9 NODECHH9 NODECH1H9  PENSCOMBH9 MIGDETH9 MIGDISTH9  SEGCESH9 USADDHHH9 GENDERHHH9  AGEHHH9 MARHHH9 COBHHH9  RELHHH9 MIGD1H9 |
| Variable Values | See Data Dictionary for details of Variables included. |

# Metadata for Properties

|  |  |
| --- | --- |
| Database Name | NILS |
| Table Name | PROPERTIES |
| Table Description | This table gives the coded information from the Land and Property Services POINTER and Valuation lists.  **Includes all XUPRNs used throughout the NILS whether they have property characteristics or not.** |
| Table size (number of rows) | 512082 records in NILS\_RSU\_DEC2013 (this will grow as each of the migration and vital events tables are updated) |
| Unique Identifiers | XUPRN  XUPRN will take the form of letters – for example:  ARELSEAEI  WPWWWEAEI  SDALNAAEI  DSEIPNAEI |
| Tables linked to | Via XUPRN to:  CORENILSDATA  ADDRESS\_HISTORY  MIGRATION\_EVENTS  CENSUSHH\_2001 |
| Frequency of the Data | Latest download from February 2010. |
| Variables | |  | | --- | | XUPRN | | settlementband | | CV\_Non\_Ex | | CV\_Ex | | Primary\_Class | | Sub\_Class | | Type | | List\_Descpn | | Hab\_Space | | Anc\_Space | | Hab\_Rooms | | Total\_Bed | | Baths | | Half\_Baths | | Heating | | Year\_Built | | Storeys | | Floor | | Parking | | Glazing | |
| Variable Values | See Data Dictionary for details of Variables included.  CV\_NON\_EX is the domestic capital value based in 2005. There are currently no plans to change this to a more up to date capital value. |
| Issues | About 95% of records have a XUPRN attached. About 90-95% of property records have detailed property information from the valuation list. This gives about 85% of records with detailed property information.  LPS is currently working on improving the link between the XUPRN and the valuation list.  <http://www.lpsni.gov.uk/index/property_rating.htm> |

# Metadata for 2001 Household Imputation\_Flags

|  |  |
| --- | --- |
| Database Name | NILS |
| Table Name | Imputation\_Flags included on 2001 Census Person table |
| Table Description | Question level imputation flags for all person-level variables |
| Table size (number of rows) | 1,627,931 |
| Unique Identifiers | censuspid |
| Tables linked to | Census\_p2001 |
| Frequency of the Data | Generated for 2001. Not updateable |
| Variables | |  |  | | --- | --- | | Variable Name | Imputation Flag for: | | CensusPID |  | | LLTI\_IMP | LLTI | | GHEALTH\_IMP | Ghealth | | RELG\_IMP | Religion | | RELUPBR\_IMP | Religion upbringing | | EDU\_IMP | Education | | GENDER\_IMP | Gender | | AGE\_IMP | Age | | ETHNIC\_IMP | Ethnic | | MARITAL\_IMP | Marital | | ACTLW\_IMP | Actlw | | UNPAIDCARE\_IMP | Unpaidcare | |
| Variable Values | All variables have same values/descriptions  0 No change  1 Non response (edited/missing) |
| How they were created?  Imputation flags have been created for the NILS. These compare the original scanned response with the basic edit checks completed to the final census database that was used for published outputs and for the NILS variables. Not all variables are included in the original scanned database and therefore imputation flags can only be created for a subset of variables.  The processes in the 2001 census which may have affected/changed the response to person or household questions are:   1. Edit and Imputation System (EDIS - see below for further detail) 2. Record Swapping – a small (undisclosed) percentage of the households had unique IDs swapped | |
| For example:  Genpuk in original census\_postedit\_person file was compared with data3/sex in the final census database. Where the codes in the 2 files were the same the imputation flag (gender\_imp) was set to zero. Any change between the 2 files was set to 1. This includes those where a valid value was changed and those where a missing value in the original file was changed to a valid value in the final file. | |

The following table shows the distribution of each of the imputation flags.

## Person Imputation Flags: Table 34 – Distributions

|  | **Variable Value** | **Number of Records** | **% of records** |
| --- | --- | --- | --- |
| LLTI\_IMP |  |  |  |
| No Change | 0 | 1,542,941 | 95 |
| Non response (edited/missing) | 1 | 84,990 | 5 |
|  |  |  |  |
| GHEALTH\_IMP |  |  |  |
| No Change | 0 | 1,592,546 | 98 |
| Non response (edited/missing) | 1 | 35,385 | 2 |
|  |  |  |  |
| RELG\_IMP |  |  |  |
| No Change | 0 | 1,393,942 | 86 |
| Non response (edited/missing) | 1 | 233,989 | 14 |
|  |  |  |  |
| RELUPBR\_IMP |  |  |  |
| No Change | 0 | 1,553,322 | 95 |
| Non response (edited/missing) | 1 | 74,609 | 5 |
|  |  |  |  |
| EDU\_IMP |  |  |  |
| No Change | 0 | 1,521,629 | 94 |
| Non response (edited/missing) | 1 | 106,302 | 7 |
|  |  |  |  |
| GENDER\_IMP |  |  |  |
| No Change | 0 | 1,613,478 | 99 |
| Non response (edited/missing) | 1 | 14,453 | 1 |
|  |  |  |  |
| AGE\_IMP |  |  |  |
| No Change | 0 | 1,583,559 | 97 |
| Non response (edited/missing) | 1 | 44,372 | 3 |
|  |  |  |  |
| ETHNIC\_IMP |  |  |  |
| No Change | 0 | 1,560,520 | 96 |
| Non response (edited/missing) | 1 | 67,411 | 4 |
|  |  |  |  |
| MARITAL\_IMP |  |  |  |
| No Change | 0 | 1,600,715 | 98 |
| Non response (edited/missing) | 1 | 27,216 | 2 |
|  |  |  |  |
| ACTLW\_IMP |  |  |  |
| No Change | 0 | 1,088,983 | 67 |
| Non response (edited/missing) | 1 | 538,948 | 33 |
|  |  |  |  |
| UNPAIDCARE\_IMP |  |  |  |
| No Change | 0 | 33,919 | 2 |
| Non response (edited/missing) | 1 | 1,594,012 | 98 |

## Person Imputation Flags: Chart 1 – Distributions

Metadata for 2001 Household Imputation Flags

|  |  |
| --- | --- |
| Database Name | NILS |
| Table Name | Imputation\_Flags included on Census Household tables |
| Table Description | Question level imputation flags for selected household-level variables |
| Table size (number of rows) | 592,802 |
| Unique Identifiers | CensusHID |
| Tables linked to | Census\_HH2001 |
| Frequency of the Data | Generated for 2001. Not updateable |
| Variables | |  |  | | --- | --- | | Variable Name | Imputation Flag for: | | CensushID |  | | CARS\_IMP | CARS | | TENURE\_IMP | TENURE | | CENTHEAT\_IMP | CENTHEAT | |
| Variable Values | All variables have same values/descriptions  0 No change  1 Non response (edited/missing) |
| How they were created?  see as above for person imputation flags | |
| For example:  see as above for person imputation flags | |

The following table shows the distribution of each of the imputation flags.

## Household Imputation Flags: Table 35 – Distributions

|  |  |  |  |
| --- | --- | --- | --- |
| CARS\_IMP |  |  |  |
| No Change | 0 | 565,968 | 96% |
| Non response (edited/missing) | 1 | 26,834 | 5% |
|  |  |  |  |
| TENURE\_IMP |  |  |  |
| No Change | 0 | 557,771 | 94% |
| Non response (edited/missing) | 1 | 35,031 | 6% |
|  |  |  |  |
| CENTHEAT\_IMP |  |  |  |
| No Change | 0 | 573,562 | 97% |
| Non response (edited/missing) | 1 | 19,240 | 3% |

## More detail on Item Imputation

Although full completion of the Census form is a statutory requirement, it is recognised that some Census forms will not be fully completed for all questions, and it is impractical to return them to each household to insist upon full completion. Further, the completion rate for most questions exceeded 90 per cent.

The Census White Paper (Cm 4253, 1999) announced that a system would be developed to impute responses to omitted questions in otherwise complete Census forms. This was developed and applied to all questions, with the exception of religion (although it was applied to the derived variable of community background).

The purpose of the Census is to produce a statistical portrait of the population, and the sole purpose of item imputation is to ensure that the portrait of the population is as complete and accurate as possible. One major benefit of the use of item imputation is that there is no longer a residual category in Census outputs labeled ‘Not stated’.

Users generally appear to find this beneficial and seem to be content with the application of item imputation. Accordingly, the use of item imputation was incorporated into the 2001 methodology.

*Methodology*

The adjustment of Census results for respondents who either failed to answer a question, answered inconsistently or answered incorrectly was made possible using an Edit and Donor Imputation System (EDIS) that was devised for the 2001 Census. The system was created to fill in a number of gaps in the records for enumerated people and households. At a later stage in processing the database was adjusted using the One Number Census process.

EDIS contained four initial components, these were:

* Multi-tick rules when more than one box was ticked but only one option was allowed;
* Range checks to prevent answers being outside an acceptable range;
* Filter rules to resolve some inconsistencies and to decide which fields should be set to 'No Code Required' where questions were answered but should not have been; and
* Edit rules to deal with missing items or responses which appeared to be in error or inconsistent when compared with other data.

After the application of these components the Imputation component was applied. The basis for the Imputation component is to search for a single “donor” person to supply all the missing variables for a recipient person. The method searched for a donor person who was similar using a number of other Census variables. A series of criteria were drawn up to determine what was meant by ‘similar’. A suitable selection of variables known as Primary Matching Variables was defined to match on for each missing item. Values were copied from the donor person to fill the missing values on the record of the recipient person.

If more than one suitable donor person was found a donor was selected from a similar household. This was based on the age, sex, marital status and relationship between the people in the household. For the Community Background, Ethnicity, Language, Address one year ago and Country of birth variables, the system also considered the responses given by the rest of the household. If there was still more than one suitable donor the person in the geographically closest household was picked.

A similar method was applied for household variables (e.g. tenure) and people living in communal establishments. If several people in a household had missing responses or some of the responses to the household questions were missing the system tried to select all the donors from the same household in order to preserve household structure.

An initial paper which details the EDIS methodology more fully can be found at

<http://www.statistics.gov.uk/census2001/pdfs/ag0013.pdf>

It should be noted that this paper details the methodology as proposed in August 2000 and some small changes in application occurred since. This issue will also be described in full in the forthcoming 2001 Census Quality Report.

The application of the EDIS system means that missing responses have been catered for in all Census topics (except a person’s current religion). The system was designed to remove bias that would otherwise have been created in the final statistics by missing responses. [Extracted from 2001 Census Methodology Paper]

## More detail on Record swapping

This procedure adds uncertainty to data by swapping the geographical location of a small sample of households with that of another household in the same District Council (or group of District Councils). The procedure was designed such that the integrity of swapped data was not substantially different among key variables from that of unswapped data. The percentage of records swapped and the basis on which they are swapped must remain confidential.

# Metadata for 2011 Imputation\_Flags

It is anticipated that further work will be carried out to produce further item imputation fields for 2011 Census data. On completion of this work, NILS Core will update the Metadata document with detailed information on imputation akin to what has been produced for the 2001 Census.

This is planned to be completed by early Spring 2014.