



Understanding Society: Higher Education Codes

USER GUIDE

Version 2, October 2024

Contents

| | |
|---|----------|
| 1. Introduction | 2 |
| 1.1 Data correction | 2 |
| 2. Data structure | 3 |
| 2.1 Contents of the Higher Education Codes dataset | 3 |
| 2.2 Data inconsistencies- Wave 12 corrections for Higher Education Codes data | 5 |
| 2.3 Data linkage | 5 |
| 3. Data access..... | 6 |
| 4. Citation | 6 |
| Citing this User Guide | 6 |
| 5. Help and support..... | 6 |

1. Introduction

This user guide provides information on the Higher Education Codes dataset ([SN 8578](#)), its file structure and data collection process.

The dataset contains information on the higher education institutions attended by survey respondents in the main Understanding Society Study asked at Waves 5, 11, 12, 13, and 14 ([SN6614](#)). Namely respondents awarded undergraduate and postgraduate level qualifications by institutions in the United Kingdom. The dataset contains higher education institution identifiers (**instid**) and/or reference numbers (**ukprn**). The dataset can be linked to publicly available records using the **ukprn** and **instid** identifiers. It can also be combined with Understanding Society's main survey data files using the variable **pidp**.

Wave 5 of the main Study collected the name of the institutions from which the respondents obtained their degrees and qualifications, as well as the institutional identifier (**instid**) for each of them. From Wave 11 onwards, the data also contain the UK Provider Reference Number (**ukprn**) as well as the institutional identifier **instid** for each of them. Additionally, in Wave 11 you can find information on whether the participants have a UK higher education qualification and if they finished higher education in or after 1995. There is also an additional variable to indicate if each participant has provided consent for their data to be linked. Where consent is given, the Higher Education Codes data may be linked in future by the Understanding Society team using the **ukprn** and **instid** identifiers to create an administrative linked dataset.

1.1 Data correction

Due to some complications in the data collection process for Wave 11, a correction to the data occurred in Wave 12. Wave 12 collected **instids** and **ukprn** codes but only for the sample affected by the issue at Wave 11. For more information on the variables included in each wave, see section 2.1 Contents of the Higher Education Codes data file. For further details on the issues that affected the Wave 11 data, see section 2.2 Data inconsistencies.

To learn more about the Main Study, take a look at the [study overview](#) and the [user guide](#).

2. Data structure

2.1 Contents of the Higher Education Codes dataset

This dataset contains Understanding Society Special Licence data for responding adults (16+), including answers from proxy interviews. In Wave 5, adult respondents who finished studying a higher education degree in 1995 or after at a UK institution were asked about the higher education establishment that they attended. The Higher Education Codes dataset contains the institution identifiers for up to four higher education establishments attended per respondent. In Wave 11 and the correction that followed in Wave 12, the data were re-gathered on the same basis. In addition to the higher education institution identifiers, the UK Provider Reference Number (**ukprn**) of the establishment has also been included.

Wave 5 (Table 1) contains **e_heinst** to **e_heinst4**, the higher education institution where they received a degree, and **e_heiname** to **e_heiname4**, the name of the higher education institution.

Table 1: List of variables for Wave 5

| | |
|-------------------|---|
| pidp | Cross-wave unique person identifier |
| e_heinst | First higher education institution ID |
| e_heinst2 | Second higher education institution ID |
| e_heinst3 | Third higher education institution ID |
| e_heinst4 | Fourth higher education institution ID |
| e_heiname | Name of first higher education institution |
| e_heiname2 | Name of second higher education institution |
| e_heiname3 | Name of third higher education institution |
| e_heiname4 | Name of fourth higher education institution |

For Wave 11 (Table 2), **k_heinstid1** to **k_heinstid4** provide the institutional identifiers (**instid**), variables **k_heukprn1** to **k_heukprn4** include information on each institution's UK Provider Reference Number (**ukprn**), **k_heukqual** identifies if the respondent has a UK higher education qualification, **k_hesacons** provides information on whether the respondent has given consent for their data to be linked to the Higher Education Statistical Agency (HESA) records, and **k_hefinish95** indicates whether the respondent finished higher education in or after 1995. Please note that the information found in variables **e_heinst** to **e_heinst4** in Wave 5 can be found in variables **heinstid1** to **heinstid4** in subsequent waves.

Table 2: List of variables for Wave 11

| | |
|--------------------|--|
| pidp | Cross-wave unique person identifier |
| k_heinstid1 | First higher education institution ID |
| k_heinstid2 | Second higher education institution ID |

| | |
|---------------------|---|
| k_heinstid3 | Third higher education institution ID |
| k_heinstid4 | Fourth higher education institution ID |
| k_heukprn1 | First higher education institution ukprn reference number |
| k_heukprn2 | Second higher education institution ukprn reference number |
| k_heukprn3 | Third higher education institution ukprn reference number |
| k_heukprn4 | Fourth higher education institution ukprn reference number |
| k_heukqual | Has a UK higher education qualification - computed |
| k_hesacons | HESA consent |
| k_hefinish95 | Finished higher education in or after 1995 - computed |

For Wave 12 (Table 3), **l_heinstid1** to **l_heinstid4** and **l_heinstidnew** are identifiers of the institution the participants attended, and **l_heukprn1** to **l_heukprn4** and **l_heukprnnew** contain the UK Provider Reference Number (**UKRN**) for each institution. Please note that the Wave 12 correction release also introduced a set of variables marked as ‘**new**’ that record any qualifications gained since the last time the respondents were interviewed (**l_heinstidnew** and **l_heukprnnew**). More information on the Wave 12 corrections is provided in section 2.2 Data inconsistencies.

| Table 3: List of variables for Wave 12 | |
|---|--|
| pidp | Cross-wave unique person identifier |
| l_heinstid1 | First higher education institution ID |
| l_heinstid2 | Second higher education institution ID |
| l_heinstid3 | Third higher education institution ID |
| l_heinstid4 | Fourth higher education institution ID |
| l_heinstidnew | Higher education institution ID acquired since last interview |
| l_heukprn1 | First higher education institution ukprn reference number |
| l_heukprn2 | Second higher education institution ukprn reference number |
| l_heukprn3 | Third higher education institution ukprn reference number |
| l_heukprn4 | Fourth higher education institution ukprn reference number |
| l_heukprnnew | Higher education institution ukprn reference number acquired since last interview |

Subsequent waves follow the same structure as Wave 12, with the consistent release of variables **heinstid1** to **heinstid4**, **heukprn1** to **heukprn4**, and **heinstidnew** and **heukprnnew**.

2.2 Data inconsistencies- Wave 12 corrections for Higher Education Codes data

The Higher Education Codes dataset released in Wave 12 corrects complications in the data collection process for Wave 11. Changes to the questionnaire part-way through the fieldwork resulted in discrepancies in the data that contained a substantial number of missing values upon release. The participants affected by the changes were interviewed again in Wave 12. Their responses were collected on the same basis as before and are included in the Higher Education Codes dataset released in Wave 12 to supplement and correct part of the missing values previously released in Wave 11.

During fieldwork for Wave 11, the questionnaire module underwent some re-work that changed the questions and the order some questions were asked. Specifically, when fieldwork began, consent for linkage was collected from participants, those who consented, were then asked about the institution they attended. Partway through fieldwork this was reversed. All participants reporting a higher education qualification were first asked which institution they received it from, and then asked to consent to data linkage. This was to ensure that all holders of higher education qualifications were asked their institution's details. As the changes took place part-way through data collection for Wave 11, this resulted in a Higher Education Codes file that lacked data for part of the sample.

To resolve the issue in Wave 12, respondents were asked to report their institution details if they had not been collected in Wave 11. The questionnaire also included a question about whether they have received any new qualification since the last time they were interviewed in Wave 11. These are recorded in variables **heinstidnew** and **heukprnnew**. The responses were compiled and released in Wave 12 as a correction to the Wave 11 Higher Education Codes file.

2.3 Data linkage

The Higher Education Codes dataset only includes information about the participants' higher education institution attendance. Researchers can use the **ukprn** and **instid** identifiers to link to publicly available datasets on the higher education institutions. Where consent is given, the Higher Education Codes data may be linked in future by the Understanding Society team using the **ukprn** and **instid** identifiers to create an administrative linked dataset, making it possible to explore educational outcomes in more detail. Additionally, to find further information on the sample members and their families, this data file can be combined with Understanding Society's main survey data files. The variable **pidp** can be used to link to the main survey individual level files.

You can find more details on [Education Data Linkage](#) on the Understanding Society website.

3. Data access

The data from the Understanding Society: Higher Education Codes study are available from the UK Data Service. The Special Licence (SL) or safeguarded version, SN 8578, can be found here: <https://beta.ukdataservice.ac.uk/datacatalogue/studies/study?id=8578>

This dataset is categorized as Special Licence and therefore has more restrictive access conditions, details of which can be found in the “Access data” tab at the above link. In addition, access to these Special Licence datasets requires users to complete forms justifying why the data is needed. Further details on the application process can be found here: <https://ukdataservice.ac.uk/find-data/access-conditions/>

This data file is also provided as part of Understanding Society main survey data release (SN6614).

4. Citation

The bibliographic reference for this study is as follows:

University of Essex, Institute for Social and Economic Research. (2024). *Understanding Society: Waves 5, 11-14, 2013-14, 2019-2023: Special Licence Access, Higher Education Codes*. [data collection]. 5th Edition. UK Data Service. SN: 8578, DOI: <http://doi.org/10.5255/UKDA-SN-8578-5>

All works which use or refer to these materials should acknowledge these sources by means of bibliographic citation. To ensure that such source attributions are captured for bibliographic indexes, citations must appear in footnotes or in the reference section of publications.

Citing this User Guide

When citing this User Guide, you can use the citation of this particular version quoted below. Note that where an online version is available on the Understanding Society website it is always the most up to date.

Institute for Social and Economic Research. (2024), *Understanding Society: Higher Education Codes, 2013-2023, User Guide, Version 2, October, 2024*, Colchester: University of Essex.

5. Help and support

If users have further questions about the **Higher Education Codes** file, please contact User Support via email usersupport@understandingsociety.ac.uk or through our [User Forum](#).