

Office for  
Students



# Calculating student numbers

Technical document

**Enquiries to** [studentnumbers@officeforstudents.org.uk](mailto:studentnumbers@officeforstudents.org.uk)

**Publication date** 30 November 2018

# Summary

1. This document details the algorithms applied to Higher Education Statistics Agency (HESA) student record, HESA alternative provider student record and Individualised Learner Record (ILR) data in order to determine a provider's higher education and further education student numbers. We will use this approach to assess applications for degree awarding powers and university title, calculate Office for Students registration fees, and determine whether a provider is required to participate in the Teaching Excellence and Student Outcomes Framework.
2. This will apply to degree awarding powers applications received on or after 1 April 2018, university title applications received on or after 1 April 2019, and applications for transfer from the further education to the higher education sector received on or after 1 August 2019.
3. This document is aimed at readers with in-depth knowledge of the data. Readers are advised to have a copy of the HESA student record coding manual, HESA alternative provider student record coding manual and ILR specification for the relevant year, and the Office of Qualifications and Examinations Regulation (Ofqual) Register Data Extract Specification to hand when using this document.
4. Some providers submit both a HESA data return and an ILR return. Student instances should not be returned in both collections, except in the case of higher apprenticeships. If a provider submits a HESA return it should return any records associated with higher apprenticeships to both HESA and the ILR. Therefore to avoid double counting these records, we have excluded from the population all records related to higher apprenticeships from the ILR for providers that submit HESA data.

## Note on terminology

5. The algorithms below apply to three different data collections, each of which has a different structure. In the HESA student return, the fields that we use are associated with a student instance. In the HESA alternative provider student return, the fields that we use are (mainly) associated with a student instance period; there may be many instance periods per student instance. In the ILR return, the fields that we use are associated with a course; a learner may study multiple courses.
6. Below, we use the generic term 'record' to refer to either an instance, an instance period or a learner-course combination, depending on the data source.
7. The year '20Y1-Y2' below refers to the academic year starting on 1 August 20Y1 and ending on 31 July 20Y2.

## 20Y1-Y2 fields used in the student number algorithms

**Table 1: HESA student fields used in the student number calculations**

| Entity      | Field     | Description   |
|-------------|-----------|---|
| Course      | COURSEAIM | General qualification aim of course   |
| Course      | TTCID     | Type of teacher training course, only completed for students on courses that lead to teaching qualifications <sup>1</sup> |
| Instance    | COMDATE   | Start date of instance  |
| Instance    | ENDDATE   | End date of instance  |
| Instance    | EXCHANGE  | Identifies whether the student is an exchange student   |
| Instance    | LOCSDY    | Location of study   |
| Instance    | MODE      | Mode of study   |
| Instance    | NUMHUS    | Student instance identifier   |
| Instance    | STULOAD   | Student instance full-time equivalence (FTE)  |
| Institution | UKPRN     | UK Provider Reference Number  |
| Student     | HUSID     | HESA unique student identifier  |

**Table 2: HESA alternative provider student fields used in the student number calculations**

| Entity          | Field       | Description   |
|-----------------|-------------|---|
| Course          | COURSEAIM   | General qualification aim of course   |
| Course          | TTCID       | Type of teacher training course, only completed for students on courses that lead to teaching qualifications <sup>2</sup> |
| Instance        | NUMHUS      | Student instance identifier   |
| Instance period | EXCHIND     | Identifies whether the student is an exchange student   |
| Instance period | LOCSDY      | Location of study   |
| Instance period | PERIODEND   | Instance period end date  |
| Instance period | PERIODSTART | Instance period start date  |
| Instance period | STULOAD     | Student Instance FTE  |
| Provider        | UKPRN       | UK Provider Reference Number  |
| Student         | HUSID       | HESA unique student identifier  |

<sup>1</sup> This field is only used when applying the algorithms to 2016-17 data.

<sup>2</sup> See footnote 1.

**Table 3: ILR, Learning Aim Reference Service and Ofqual fields used in the student number calculations**

| Data source | Field                         | Description  |
|-------------|-------------------------------|--|
| ILR         | LEARNACTENDDATE               | Learning actual end date   |
| ILR         | LEARNPLANENDDATE              | Learning planned end date  |
| ILR         | LEARNSTARTDATE                | Learning start date  |
| ILR         | STULOAD                       | Learner FTE  |
| ILR         | PLANLEARNHOURS                | Planned Learning Hours   |
| ILR         | PROGTYPE                      | The type of programme which the learner is undertaking   |
| ILR         | UKPRN                         | UK Provider Reference Number   |
| LARS        | ENGLANDFEHESTATUS             | Further or higher education status of a course ('England FE/HE status')  |
| LARS        | LEARNAIMREF                   | Learning aim reference   |
| LARS        | LEARNAIMREFTITLE              | Learning aim title   |
| LARS        | LEARNAIMREFTYPE               | Learning aim type  |
| LARS        | NOTIONALNVQLEVELV2            | The detailed learning aim level on the NVQ scale   |
| Ofqual      | TOTAL QUALIFICATION TIME      | The total time a learner may typically be expected to expend in achieving a qualification  |
| Ofqual      | MAXIMUM GUIDED LEARNING HOURS | The maximum number of hours a learner can be expected to be taught or instructed by their tutor or teacher to complete the qualification |
| Ofqual      | TOTAL CREDITS                 | The total credits required to obtain the qualification   |

Note: 'LARS' = 'Learning Aim Reference Service'.

## Derived fields for calculating total higher education FTE

### SNSOURCE

8. Identifies the source of the record.

| Value  | Definition   |
|--------|--|
| HESA   | Record taken from the HESA student record                      |
| HESAAP | Record taken from the HESA alternative provider student record |
| ILR    | Record taken from the ILR                                      |

## SNHESAILRFLAG

9. Identifies providers that returned both HESA and ILR data in the academic year 20Y1-Y2.

| Value | Definition   |
|-------|--|
| 1     | Provider returned both HESA and ILR data for the academic year 20Y1-Y2       |
| 0     | Provider did not return both HESA and ILR data for the academic year 20Y1-Y2 |

## SNEXCL

10. Identifies records which are excluded from the population.

| Value | Description | Definition  |
|-------|-------------|---|
| 1     | Excluded    | <p><b>SNSOURCE= HESA</b><br/>           COMDATE &gt;31/07/Y2 or<br/>           (ENDDATE &lt;01/08/Y1 and ENDDATE is not BLANK) or<br/>           MODE = 51, 63, 64, 98 or<br/>           TTCID = F<sup>3</sup> or<br/>           LOCSDY = S or<br/>           EXCHANGE = 4,G or<br/>           COURSEAIM =Z99</p> <p><b>SNSOURCE = HESAAP</b><br/>           PERIODSTART &gt; 31/07/Y2 or<br/>           (PERIODEND &lt; 01/08/Y1 and PERIODEND is not<br/>           BLANK) or<br/>           TTCID = F<sup>4</sup> or<br/>           LOCSDY = S or<br/>           EXCHIND = 1</p> <p><b>SNSOURCE= ILR</b><br/>           LEARNSTARTDATE &gt; 31/07/Y2 or<br/>           (LEARNACTENDDATE &lt; 01/08/Y1 and<br/>           LEARNACTENDDATE is not BLANK) or<br/>           LEARNAIMREF = ZPROG001 or<br/>           (SNHESAILRFLAG = 1 and PROGTYPE = 20, 21, 22,<br/>           23, 25)</p> |
| 0     | Included    | Otherwise   |

<sup>3</sup> Only used when the algorithms are applied to 2016-17 data, as TTCID = F is not a permitted value in subsequent collections.

<sup>4</sup> See footnote 3.

## SNLEVEL

11. Identifies whether the record is of higher education or further education.

| Value | Description                 | Definition   |
|-------|-----------------------------|--|
| HE    | Record of higher education  | COURSEAIM = D, E, L, M, H, I, J, C or<br>ENGLANDFEHESTATUS = H or<br>NOTIONALNVQLEVELV2 = 4, 5, 6, 7, 8, H |
| FE    | Record of further education | Otherwise  |

## SNHESTUDYLEVEL

12. Indicates the level of study for higher education records, using either the Quality Assurance Agency for Higher Education's Frameworks for higher education qualifications (for qualifications awarded using degree awarding powers) or else the Regulated qualifications framework<sup>5</sup>.

| Value | Description                 | Definition  |
|-------|-----------------------------|---|
| 6     | Records at Level 6 or above | SNLEVEL = HE and<br>(COURSEAIM = D, E, L, M, H, I00, I11, I12, I16<br>or<br>NOTIONALNVQLEVELV2 = 6, 7, 8)                                       |
| 5     | Records at Level 5          | SNLEVEL = HE and<br>(COURSEAIM=I60, I61, I70, I71, I72, I73, I74, I76, I78,<br>I79, I80, I81, I90, I91, I99, J<br>or<br>NOTIONALNVQLEVELV2 = 5) |
| 4     | Records at Level 4          | SNLEVEL=HE and not above  |

## SNTITLEHRS

13. The maximum number of hours per record as extracted from the course title on the Learning Aim Reference Service. For example, the value of SNTITLEHRS for the course listed as 'Non regulated SFA formula funded provision, Entry level, Medicine and Dentistry, 389 to 580 hrs, PW B' would be 580.

## SNTITLECREDITS

14. The maximum number of credits per record as extracted from the course title on the LARS. For example, the value of SNTITLECREDITS for the course listed as 'Innovation code, Entry level, Medicine and Dentistry, 3-5 credits' would be 5.

---

<sup>5</sup> For the Frameworks for higher education qualifications see [www.qaa.ac.uk/en/quality-code/the-existing-uk-quality-code/part-a-setting-and-maintaining-academic-standards](http://www.qaa.ac.uk/en/quality-code/the-existing-uk-quality-code/part-a-setting-and-maintaining-academic-standards).

## SNQUALHOURS

15. The total number of hours the course associated with the record is expected to take.

| Value                         | Description                                      | Definition   |
|-------------------------------|--|--|
| TOTAL QUALIFICATION TIME      | Source of hours is TOTAL QUALIFICATION TIME      | SNSOURCE = ILR and STULOAD = MISSING and TOTAL QUALIFICATION TIME ≠ MISSING                    |
| MAXIMUM GUIDED LEARNING HOURS | Source of hours is MAXIMUM GUIDED LEARNING HOURS | SNSOURCE = ILR and STULOAD = MISSING and MAXIMUM GUIDED LEARNING HOURS ≠ MISSING and not above |
| TOTAL CREDITS × 10            | Source of hours is TOTAL CREDITS                 | SNSOURCE = ILR and STULOAD = MISSING and TOTAL CREDITS ≠ MISSING and not above                 |
| SNTITLEHRS                    | Source of hours is SNTITLEHRS                    | SNSOURCE = ILR and STULOAD = MISSING and SNTITLEHRS ≠ MISSING and not above                    |
| SNTITLECREDITS × 10           | Source of hours is SNTITLECREDITS                | SNSOURCE = ILR and STULOAD = MISSING and SNTITLECREDITS ≠ MISSING and not above                |

## SNENDDATE

16. Determines the end date to be used when calculating the number of days studied (SNAYDAYSSTUDIED) in the academic year 20Y1-Y2. Used only for ILR data where SNQUALHOURS will be used to calculate the FTE of the record.

| Value           | Description                                      | Definition  |
|-----------------|--|---|
| LEARNACTENDDATE | Learner completed learning activities in 20Y1-Y2 | SNSOURCE=ILR and LEARNACTENDDATE ≠ MISSING and LEARNACTENDDATE ≤ 31/07/Y2 |
| 31/07/Y2        | Learner is still studying at the end of 20Y1-Y2  | SNSOURCE=ILR and not above  |

## SNAYDAYSSTUDIED

17. The number of days studied in the 20Y1-Y2 academic year, calculated as the numbers of days between SNENDDATE and either 1 August 20Y1 or LEARNSTARTDATE, whichever is earlier. Used only for ILR data where SNQUALHOURS will be used to calculate the FTE of the record.

## SNAVHOURSPERDAY

18. Calculates the average number of hours that would be studied each day for the duration of the course, calculated as SNQUALHOURS divided by the number of days between the LEARNSTARTDATE and LEARNPLANENDDATE. Used only for ILR data where SNQUALHOURS will be used to calculate the FTE of the record.

## SNHOURSPERAYR

19. Calculates the total number of hours studied in the academic year for each record. Used only for ILR data where SNQUALHOURS will be used to calculate the FTE of the record. We cap SNHOURSPERAYR at the value of SNQUALHOURS for the course.

- $SNHOURSPERAYR = \text{Min}(SNAVHOURSPERDAY \times SNAYDAYSSTUDIED, SNQUALHOURS)$ .

## SNHESTULOADCASE

20. Defines the fields we will use in calculating the FTE for a given higher education record.

| Value | Description   | Definition   |
|-------|---|--|
| 0     | STULOAD will be used in calculating FTE for this record     | SNLEVEL = HE and STULOAD $\neq$ MISSING  |
| 1     | STULOAD will be set to 0 for this record                    | SNLEVEL = HE and SNSOURCE = HESA, HESAAP and STULOAD = MISSING                       |
| 2     | SNQUALHOURS will be used in calculating FTE for this record | SNLEVEL = HE and SNSOURCE = ILR and STULOAD = MISSING and SNQUALHOURS $\neq$ MISSING |
| 3     | A value of 25 STULOAD will be assigned to this record       | SNLEVEL = HE and SNSOURCE = ILR and STULOAD = MISSING and not above                  |

## SNHESTULOAD

21. The STULOAD for each higher education record.

| Value                                 | Definition          |
|---------------------------------------|---------------------|
| STULOAD                               | SNHESTULOADCASE = 0 |
| 0                                     | SNHESTULOADCASE = 1 |
| $(SNHOURSPERAYR \div 540) \times 100$ | SNHESTULOADCASE = 2 |
| 25                                    | SNHESTULOADCASE = 3 |

## SNHEINSTLEVSTULOAD

22. The total STULOAD at a given SNHESTUDYLEVEL for each higher education instance<sup>6</sup>

| Value   | Definition                   |
|---|------------------------------|
| SNHESTULOAD   | SNSOURCE=HESA                |
| Sum of SNHESTULOAD across all instance periods at a given SNHESTUDYLEVEL per instance | SNSOURCE=HESAAP and SNEXCL=0 |
| Sum of SNHESTULOAD across all records for each SNHESTUDYLEVEL per student             | SNSOURCE=ILR and SNEXCL=0    |

<sup>6</sup> The ILR data does not include the concept of an 'instance'. Solely for the purpose of calculating SNHEINSTLEVSTULOAD we treat a level of study in the ILR data as analogous to an instance.



## SNHEINSTLEVSTULOADADJ

23. The capped STULOAD at a given SNHESTUDYLEVEL for each higher education instance.

- $SNHEINSTLEVSTULOADADJ = \text{Min}(SNHEINSTLEVSTULOAD, 150)$ .

## SNTOTHESTULOAD

24. SNTOTHESTULOAD is the total STULOAD per student including any capping, calculated as the sum of SNHEINSTLEVSTULOADADJ across each unique combination of instance and level per student where SNEXCL=0.

## SNTOTHEFTE

25. The total higher education FTE, calculated as  $SNTOTHESTULOAD \div 100$ .

## Additional fields required to calculate total further education FTE

### SNFEPLANLEARNSTULOAD

26. Converts PLANLEARNHOURS to STULOAD for further education records.

- $SNFEPLANLEARNSTULOAD = \text{max}(((PLANLEARNHOURS \div 540) \times 100) - SNTOTHESTULOAD, 0)$ .

### SNFESTULOADCASE

27. Determines the fields that will be used to calculate the FTE for further education records.

| Value | Description  | Definition  |
|-------|--|---|
| 0     | STULOAD will be used when calculating FTE for this record        | SNLEVEL = FE and SNSOURCE = HESA, HESAAP and STULOAD ≠ MISSING          |
| 1     | STULOAD will be set to 0 for this record                         | SNLEVEL = FE and SNSOURCE = HESA, HESAAP and STULOAD = MISSING          |
| 2     | PLANLEARNHOURS will be used when calculating FTE for this record | SNLEVEL = FE and SNSOURCE = ILR and PLANLEARNHOURS ≠ MISSING            |
| 3     | SNQUALHOURS will be used when calculating FTE for this record    | SNLEVEL = FE and SNSOURCE = ILR and SNQUALHOURS ≠ MISSING and not above |
| 4     | A value of 10 STULOAD will be assigned to this record            | SNLEVEL = FE and SNSOURCE = ILR and not above                           |

### SNNUMFERECORDS

28. The number of SNEXCL = 0, SNLEVEL = FE records per student. Used only for records from the ILR.

## SNFESTULOAD

29. The STULOAD for each further education record.

| Value  | Definition          |
|--|---------------------|
| STULOAD  | SNFESTULOADCASE = 0 |
| 0  | SNFESTULOADCASE = 1 |
| $\text{SNFEPLANLEARNSTULOAD} \div \text{SNNUMFERECORDS}$ | SNFESTULOADCASE = 2 |
| $(\text{SNHOURSPERAY} \div 540) \times 100$              | SNFESTULOADCASE = 3 |
| 10   | SNFESTULOADCASE = 4 |

## SNTOTFESTULOAD

30. The sum of SNFESTULOAD across all SNEXCL = 0 records for the student. This is capped at a maximum of 150 per student.

## SNTOTFEFTE

31. The total further education FTE, calculated as  $\text{SNTOTFESTULOAD} \div 100$ .



© The Office for Students copyright 2018

This publication is available under the Open Government Licence 3.0.

[www.nationalarchives.gov.uk/doc/open-government-licence/version/3/](http://www.nationalarchives.gov.uk/doc/open-government-licence/version/3/)