

# Calculating student numbers:

2020-21 technical document  
and rebuild instructions

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

**Publication date** 4 August 2021

**Last updated** 16 February 2022

# Contents

<b>Summary</b>	<b>3</b>
<b>Changes</b>	<b>5</b>
Changes in the 2020-21 method	5
Note on terminology	6
Student numbers workbook	6
Student numbers individualised file	7
20Y1-Y2 fields used to create the student number outputs	7
<b>Description of derived fields</b>	<b>10</b>
Derived fields for calculating total higher education FTE	12
Additional fields required to calculate total further education FTE	20
Additional fields required to calculate total undergraduate higher education FTE relevant to TEF	21
<b>Rebuild instructions</b>	<b>24</b>

# Summary

1. This document describes the algorithms applied to the 2020-21 Higher Education Statistics Agency (HESA) Student record, HESA Student alternative record and Individualised Learner Record (ILR) data. These are used to determine a provider's higher education and further education student numbers. Throughout the document, fields taken or derived from HESA or ILR records are shown in capitals.
2. The student numbers outputs are formed of two files. These are:
  - a. Student numbers workbook 'SN20\_AAA\_XXXXXXXX.xlsx'
  - b. Student numbers individualised file 'SN20\_AAA\_XXXXXXXX\_IND.csv'

where XXXXXXXX is the UK Provider Reference Number (UKPRN) for the provider, and AAA identifies the output type as listed in Table 1.

**Table 1: File name abbreviations used in student number outputs**

File abbreviation	Source of outputs
DCT	Data checking tool
AST	Amendments submission tool
PCO	Post-collection outputs

3. The data in the student numbers data checking tool and amendments submission tool outputs are provided as a data quality tool, to aid providers in getting their data correct and must not be used for any other purpose.
4. The final HESA Student, HESA Student alternative and ILR R14 data will be used to generate the final student numbers, which will be released in the 2020-21 post-collection outputs. These will be the numbers used by the Office for Students (OfS) for regulatory purposes, such as assessing applications for degree awarding powers and university title, and calculating OfS registration fees. The post-collection outputs will be based on the data originally submitted by the provider and signed off as accurate by the accountable officer. We use amended data for these purposes only when this use has been formally approved by the OfS.
5. Some providers submit both a HESA data return and an ILR return. Student instances meeting either of the criteria listed below should be reported to both HESA and the Education and Skills Funding Agency (ESFA):
  - a. Students on higher and degree apprenticeships for whom you receive direct funding from the ESFA (PROGTYPE 20, 21, 22, 23 or 25).
  - b. Students on higher education instances (Level 4 or above) who are financed by Advanced Learner Loans.

We would expect all other student instances to be returned to either HESA or the ESFA, but not both. For those providers that return both HESA and ILR data, the student instances listed above will be excluded from their ILR data to avoid double counting.

6. 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 Student alternative 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.
7. We will document any changes or corrections to the algorithms in this document on the OfS website ([www.officeforstudents.org.uk/data-and-analysis/student-number-data/](http://www.officeforstudents.org.uk/data-and-analysis/student-number-data/)).

# Changes

## Changes in the 2020-21 method

8. This section describes any changes to the algorithms since it was first published on 4 August 2021.

### SNLEVELNUM

9. The field SNLEVEL has been updated to remove the courseaim I13 which was introduced for Level 5, as the courseaim has now been updated to J13 in the 2020-21 HESA student and HESA Student alternative record.

10. The field SNHESTUDYLEVEL has been renamed to SNLEVELNUM and has been expanded to include SNLEVELNUM = 7 and 8.

SNSOURCE = HESASTU, HESASAR

Value	Description	Definition
8	Records at Level 8	SNLEVEL = HE and COURSEAIM = D00, D01, D90, E00, E13, E40, E43, E90, L00
7	Records at Level 7	SNLEVEL = HE and COURSEAIM = L80, L90, L91, L99, M00, M01, M02, M10, M11, M13, M16, M22, M26, M28, M40, M41, M42, M43, M44, M45, M50, M70, M71, M72, M73, M76, M78, M79, M80, M86, M88, M90, M91, M99
6	Records at Level 6	SNLEVEL = HE and COURSEAIM = I00, I11, I12, I16, H00, H11, H12, H13, H16, H18, H22, H23, H24, H41, H42, H43, H50, H60, H61, H62, H70, H71, H72, H76, H78, H79, H80, H81, H88, H90, H91, H99
5	Records at Level 5	SNLEVEL = HE and COURSEAIM = I60, I61, I70, I71, I72, I73, I74, I76, I78, I79, I80, I81, I90, I91, I99, J
4	Records at Level 4	SNLEVEL = HE and not above

SNSOURCE = ILR

Value	Description	Definition
8	Records at Level 8	SNLEVEL = HE and NOTIONALNVQLEVELV2 = 8
7	Records at Level 7	SNLEVEL = HE and NOTIONALNVQLEVELV2 = 7
6	Records at Level 6	SNLEVEL = HE and NOTIONALNVQLEVELV2 = 6
5	Records at Level 5	SNLEVEL = HE and NOTIONALNVQLEVELV2 = 5
4	Records at Level 4	SNLEVEL = HE and not above

### Additional fields for undergraduate students relevant to TEF

The additional fields SNREL\_TEF, SNHEINSTLEVSTULOAD\_TEF, SNHEINSTLEVSTULOADADJ\_TEF, SNTOTHESTULOAD\_TEF, SNTOTHEFTE\_TEF have been added to calculate the total undergraduate HE FTE relevant to the Teaching Excellence Framework (TEF).

### Note on terminology

- 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 Student alternative return, the fields that we are use are (mainly) associated with a student instance period; there may be multiple 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.
- 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.
- The year '20Y1-Y2' below refers to the academic year starting on 1 August 20Y1 and ending on 31 July 20Y2.

### Student numbers workbook

- The student numbers workbook is the Excel workbook SN20\_AAA\_XXXXXXXXX.xlsx which contains the following worksheets:

**Table 2: Worksheets in the student numbers workbook**

Worksheet*	Title
Information	Information sheet
Student Numbers	Number of student FTEs by broad level and HE study level

\* This worksheet reference corresponds to the spreadsheet tabs.

## Student numbers individualised file

15. When working through this document it is necessary to use the individualised file, SN20\_AAA\_XXXXXXXXX\_IND.csv. The individualised file contains the values of all fields used to generate student numbers.
16. Fields taken from the individualised student return or derived as part of the student numbers outputs are shown in capitals using the names given in Tables 3, 4, 5 and 6.
17. Specific instructions for rebuilding the student numbers workbook from the individualised file are provided in this document.
18. The individualised file contains one record per instance in the latest academic year. For HESA Student and HESA Student alternative data there is one row per student instance (each unique combination of UKPRN, HUSID and NUMHUS), and for ILR data there is one row per learning aim for a learner (each unique combination of UKPRN, LEARNREFNUMBER, LEARNAIMREF and AIMSEQNUMBER).

## 20Y1-Y2 fields used to create the student number outputs

19. The fields used to generate the student numbers outputs are listed in Tables 3, 4 and 5.

**Table 3: HESA Student fields used in the student number calculations**

Name	Description	Column in individualised file
COURSEAIM	General qualification aim of course	AV
COMDATE	Start date of instance	AX
ENDDATE	End date of instance	AY
EXCHANGE	Identifies whether the student is an exchange student	AZ
LOCSDY	Location of study	BA
MODE	Mode of study	BB
NUMHUS	Student instance identifier	G
STULOAD	Student instance full-time equivalence (FTE)	BC
UKPRN	UK Provider Reference Number	A
HUSID	HESA unique student identifier	H

**Table 4: HESA Student alternative fields used in the student number calculations**

Name	Description	Column in individualised file
COURSEAIM	General qualification aim of course	AV
NUMHUS	Student instance identifier	G
EXCHIND	Identifies whether the student is an exchange student	BD

Name	Description	Column in individualised file
LOCSDY	Location of study	BA
PERIODEND	Instance period end date	BE
PERIODSTART	Instance period start date	BF
STULOAD	Student instance FTE	BC
UKPRN	UK Provider Reference Number	A
HUSID	HESA unique student identifier	H
XINACT01	Inactive for entire reporting year	BG

**Table 5: ILR, Learning Aim Reference Service and Ofqual fields used in the student number calculations**

Field	Description	Column in individualised file
COMPSTATUS	Degree of completion of the learning activities leading to the learning aim	BH
LEARNACTENDDATE	Learning actual end date	BL
LEARNDELFAMCODE	The value of the funding or monitoring attribute	BM
LEARNDELFAMTYPE	The type of funding or monitoring attribute	BM
LEARNPLANENDDATE	Learning planned end date	BN
LEARNSTARTDATE	Learning start date	BO
STULOAD	Learner FTE	BC
PLANLEARNHOURS	Planned learning hours	BR
PRIORLEARNFUNDADJ	The proportion of this learning aim still to be delivered	BS
PROGTYPE	The type of programme which the learner is undertaking	BT
UKPRN	UK Provider Reference Number	A
WITHDRAWREASON	The reason why the learner has withdrawn from the learning aim	BU
LEARNAIMREF	Learning aim reference	BI
QUAL_TIT	Learning aim title	BJ
QUAL_TYP	Learning aim type	BK
NVQ_LV	Notional NVQ level	BP
NOTIONALNVQLEVELV2	The detailed learning aim level on the NVQ scale	BQ
TOTAL QUALIFICATION TIME	The total time a learner may typically be expected to expend in achieving a qualification	BV



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	BW
TOTAL CREDITS	The total credits required to obtain the qualification	BX

## Description of derived fields

20. The derived fields used to generate the student numbers outputs are listed in Table 5.

**Table 6: Description of derived fields for recreating tables**

Derived field name	Description	Paragraph	Column in individualised file (HESA/ILR output)
LEARNDELFAM_ADL1	Advanced Learner Loan indicator	23	BM
SNAVHOURSPERDAY	Average study hours per day over course	44	Z
SNAYDAYSSTUDIED	Number of days studied in academic year	43	Y
SNENDDATE	Last day studied in academic year	42	X
SNEXCL	Reason(s) for exclusion from student numbers population	34	K
SNEXCL1 – SNEXCL256	Student number population exclusions	25 - 33	AM-AU
SNFEPLANLEARNSTULOAD	Further education STULOAD calculated from PLANLEARNHOURS	52	AH
SNFESTULOAD	STULOAD for further education record	55	AK
SNFESTULOADCASE	Determines fields used to calculate further education FTE for a given record	53	AI
SNHEINSTLEVSTULOAD	Total STULOAD at a given SNLEVELNUM for a given higher education instance	48	AD
SNHEINSTLEVSTULOADADJ	Capped STULOAD at a given SNLEVELNUM for a given higher education instance	49	P
SNHEINSTLEVSTULOAD_TEF	Total STULOAD at a given SNLEVELNUM for a given undergraduate higher education instance relevant to TEF	60	AF
SNHEINSTLEVSTULOADADJ_TEF	Capped STULOAD at a given SNLEVELNUM for a given undergraduate higher education instance relevant to TEF	61	S
SNHESAILRFLAG	Providers that return both HESA and ILR data	24	C

SNLEVELNUM	Level of higher education record	37	M
SNHESTULOAD	STULOAD for higher education record	47	AC
SNHESTULOADCASE	Determines fields used to calculate higher education FTE for a given record	46	AB
SNHOURSPERAYR	Total hours studied in academic year	45	AA
SNLEVEL	Identifies record as higher or further education	36	L
SNNUMFERECORDS	Number of non-excluded further education records per student	54	AJ
SNPRIORLEARNADJ	Adjustment for prior learning	40	V
SNQUALHOURS	Total hours course is expected to take	41	W
SNREL_TEF	Undergraduate higher education provision relevant to TEF	58	Q
SNSOURCE	Identifies the source of the record	21	D
SNTITLECREDITS	Max credits in course title	39	U
SNTITLEHRS	Max learning hours in course title	38	T
SNTOTFEFTE	Total further education FTE per student	57	N
SNTOTFESTULOAD	Total capped further education STULOAD per student	56	AL
SNTOTHEFTE	Total higher education FTE per student	51	O
SNTOTHESTULOAD	Total capped higher education STULOAD per student	50	AE
SNTOTHEFTE_TEF	Total higher education undergraduate FTE per student relevant to TEF	63	R
SNTOTHESTULOAD_TEF	Total capped higher education undergraduate STULOAD per student relevant to TEF	62	AG

## Derived fields for calculating total higher education FTE

### SNSOURCE

21. This field identifies the source of the record.

Value	Definition
HESASTU	Record taken from the HESA Student record
HESASAR	Record taken from the HESA Student alternative record
ILR	Record taken from the ILR

22. Where an algorithm cannot be applied in the same way to each SNSOURCE, this will be indicated in the description of the algorithm.

### LEARNDEL FAM\_ADL1

23. This field contains the value of LEARNDEL FAMCODE where LEARNDEL FAMTYPE = ADL.

### SNHESAILRFLAG (Only available in ILR DCT or AST. Available in all post-collection outputs)

24. This field 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

### SNEXCL1

25. This field identifies records that start after the academic year of interest.

Value	Description	Definition
1	Excluded	(SNSOURCE = HESASTU and COMDATE > 31/07/Y2) or (SNSOURCE = HESASAR and PERIODSTART > 31/07/Y2) or (SNSOURCE = ILR and LEARNSTARTDATE > 31/07/Y2)
0	Included	Otherwise

## SNEXCL2

26. This field identifies records that finish before the academic year of interest.

Value	Description	Definition
1	Excluded	(SNSOURCE = HESASTU and ENDDATE < 01/08/Y1 and ENDDATE is not BLANK) or (SNSOURCE = HESASAR and PERIODEND < 01/08/Y1 and PERIODEND is not BLANK) or (SNSOURCE = ILR and LEARNACTENDDATE < 01/08/Y1 and LEARNACTENDDATE is not BLANK)
0	Included	Otherwise

## SNEXCL4

27. This field identifies records in the HESA Student data that are dormant or on sabbatical, or records in the HESA Student alternative data that have been inactive for the reporting year.

Value	Description	Definition
1	Excluded	(SNSOURCE = HESASTU and MODE = 51, 63, 64, 98) or (SNSOURCE = HESASAR and XINACT01 = 1)
0	Included	Otherwise

## SNEXCL8

28. This field identifies incoming exchange students in the HESA Student and HESA Student alternative data.

Value	Description	Definition
1	Excluded	((SNSOURCE = HESASTU and (EXCHANGE = 4,G or LOCSYD = S)) or (SNSOURCE = HESASAR and (EXCHIND = 1 or LOCSYD = S))
0	Included	Otherwise

## SNEXCL16

29. This field identifies apprenticeship standard 'wrapper' programme aims in the ILR data.

Value	Description	Definition
1	Excluded	SNSOURCE = ILR and LEARNAIMREF = ZPROG001
0	Included	Otherwise

## SNEXCL32

30. This field identifies higher apprentices in the ILR that have been returned by providers that also submit HESA Student data.

Value	Description	Definition
1	Excluded	SNSOURCE = ILR and SNHESAILRFLAG = 1 and PROGTYPE = 20, 21, 22, 23, 25
0	Included	Otherwise

## SNEXCL64

31. This field identifies an ILR record which has been closed to correct an incorrect LEARNPLANENDDATE.

Value	Description	Definition
1	Excluded	SNSOURCE = ILR and COMPSTATUS = 4 <sup>1</sup> and WITHDRAWREASON = 40 and LEARNACTENDDATE = 1 August 20Y1
0	Included	Otherwise

## SNEXCL128 (unavailable in DCT or AST outputs)

32. This field identifies records in the ILR which we believe relate to teaching provision rather than registered provision. This will be identified if two records returned by two providers for the same student are found to have the same LEARNAIMREF, the courses start within 31 days of each other and the UKPRN for one record matches the PARTNERUKPRN of the other record. The record with missing PARTNERUKPRN will then be flagged.

Value	Description	Definition
1	Excluded	SNSOURCE = ILR and the student appears to exist in another provider's ILR and this record will be excluded from the student numbers calculation
0	Included	Otherwise

---

<sup>1</sup> ILR records returned to the ESFA with COMPSTATUS=3 and WITHDRAWREASON=40 will have the value of COMPSTATUS changed to 4 by the ESFA. This is replicated in the ILR data checking and amendment submission tools.

## SNEXCL256

33. This field identifies students in the ILR on higher education instances financed by an Advanced Learner Loan that have been returned by providers that also submit HESA Student data.

Value	Description	Definition
1	Excluded	SNSOURCE = ILR and SNHESAILRFLAG = 1 and LEARNDELFAM_ADL = 1 and NOTIONALNVQLEVELV2 = 4,5,6,
0	Included	Otherwise

## SNEXCL

34. This field indicates whether the record will be included in the calculation of student numbers SNEXCL captures exclusions using applicable values from the table below.

Value	Description	Definition
1	Record starts after academic year of interest	SNEXCL1 = 1
2	Record ends before academic year of interest	SNEXCL2 = 1
4	Student is dormant, on sabbatical or inactive for the entire reporting year	SNEXCL4 = 1
8	Incoming exchange student	SNEXCL8 = 1
16	Apprenticeship standard 'wrapper' programme aim	SNEXCL16 = 1
32	Higher apprentices returned in ILR by provider that also submits HESA data	SNEXCL32 = 1
64	ILR records closed to correct LEARNPLANENDDATE	SNEXCL64 = 1
128	ILR record returned by the teaching provider also returned by the registering provider	SNEXCL128 = 1
256	HE student financed by Advanced Learner Loan returned in ILR by provider that also submits HESA data	SNEXCL256 = 1
0	Otherwise	None of the above

35. This field contains the exclusion reason(s) for the instance. It is calculated as  $(1 \times \text{SNEXCL1}) + (2 \times \text{SNEXCL2}) + \dots + (256 \times \text{SNEXCL256})$ .

## SNLEVEL

36. This field identifies whether the record is of higher education or further education.

SNSOURCE = HESASTU, HESASAR

Value	Description	Definition
HE	Record of higher education	COURSEAIM = D, E, L, M, H, I, J, C
FE	Record of further education	COURSEAIM = P, Q, R, S, X

SNSOURCE = ILR

Value	Description	Definition
HE	Record of higher education	LEARNAIMREF ≠ ZPROG001 and NOTIONALNVQLEVELV2 = 4, 5, 6, 7, 8, H
FE	Record of further education	LEARNAIMREF ≠ ZPROG001 and not above

## SNLEVELNUM

37. This field 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>2</sup>.

SNSOURCE = HESASTU, HESASAR

Value	Description	Definition
8	Records at Level 8	SNLEVEL = HE and COURSEAIM = D00, D01, D90, E00, E13, E40, E43, E90, L00
7	Records at Level 7	SNLEVEL = HE and COURSEAIM = L80, L90, L91, L99, M00, M01, M02, M10, M11, M13, M16, M22, M26, M28, M40, M41, M42, M43, M44, M45, M50, M70, M71, M72, M73, M76, M78, M79, M80, M86, M88, M90, M91, M99
6	Records at Level 6	SNLEVEL = HE and COURSEAIM = I00, I11, I12, I16, H00, H11, H12, H13, H16, H18, H22, H23, H24, H41, H42, H43, H50, H60, H61, H62, H70, H71, H72, H76, H78, H79, H80, H81, H88, H90, H91, H99
5	Records at Level 5	SNLEVEL = HE and COURSEAIM = I60, I61, I70, I71, I72, I73, I74, I76, I78, I79, I80, I81, I90, I91, I99, J
4	Records at Level 4	SNLEVEL = HE and not above

SNSOURCE = ILR

Value	Description	Definition
8	Records at Level 8	SNLEVEL = HE and NOTIONALNVQLEVELV2 = 8
7	Records at Level 7	SNLEVEL = HE and NOTIONALNVQLEVELV2 = 7
6	Records at Level 6	SNLEVEL = HE and NOTIONALNVQLEVELV2 = 6

<sup>2</sup> For the Frameworks for higher education qualifications see <https://www.qaa.ac.uk/en/quality-code/qualifications-frameworks>.



Value	Description	Definition
5	Records at Level 5	SNLEVEL = HE and NOTIONALNVQLEVELV2 = 5
4	Records at Level 4	SNLEVEL = HE and not above

### SNTITLEHRS

38. For ILR records only. This field identifies the maximum number of hours per record as extracted from the course title on the Learning Aim Reference Service (LARS). 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

39. For ILR records only. This field identifies 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.

### SNPRIORLEARNADJ

40. For ILR records only. This field identifies the funding adjustment for prior learning expressed as a decimal value.

Value	Definition
PRIORLEARNFUNDADJ/100	SNSOURCE = ILR and PRIORLEARNFUNDADJ > 0
0	SNSOURCE = ILR and PRIORLEARNFUNDADJ = 0

### SNQUALHOURS

41. For ILR records only. This field determines the total number of hours the course associated with the record is expected to take.

Value	Description	Definition
TOTAL QUALIFICATION TIME × SNPRIORLEARNADJ	Source of hours is TOTAL QUALIFICATION TIME	SNSOURCE = ILR and STULOAD = BLANK and TOTAL QUALIFICATION TIME ≠ BLANK, 0
MAXIMUM GUIDED LEARNING HOURS × SNPRIORLEARNADJ	Source of hours is MAXIMUM GUIDED LEARNING HOURS	SNSOURCE = ILR and STULOAD = BLANK and MAXIMUM GUIDED LEARNING HOURS ≠ BLANK, 0 and not above
(TOTAL CREDITS × 10) × SNPRIORLEARNADJ	Source of hours is TOTAL CREDITS	SNSOURCE = ILR and STULOAD = BLANK and TOTAL CREDITS ≠ BLANK, 0 and not above
(SNTITLEHRS) × SNPRIORLEARNADJ	Source of hours is SNTITLEHRS	SNSOURCE = ILR and STULOAD = BLANK and

Value	Description	Definition
		SNTITLEHRS ≠ BLANK and not above
(SNTITLECREDITS × 10) × SNPRIORLEARNADJ	Source of hours is SNTITLECREDITS	SNSOURCE = ILR and STULOAD = BLANK and SNTITLECREDITS ≠ BLANK and not above

## SNENDDATE

42. For ILR records only. This field determines the end date to be used when calculating the number of days studied (SNAYDAYSSTUDIED) in the academic year 20Y1-Y2. Used only 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 ≠ BLANK and LEARNACTENDDATE ≤ 31/07/Y2
31/07/Y2	Learner is still studying at the end of 20Y1-Y2	SNSOURCE = ILR and not above

## SNAYDAYSSTUDIED

43. For ILR records only. This field determines 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 later. Used only where SNQUALHOURS will be used to calculate the FTE of the record.

## SNAVHOURSPERDAY

44. For ILR records only. This field 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 where SNQUALHOURS will be used to calculate the FTE of the record.

## SNHOURSPERAYR

45. For ILR records only. This field calculates the total number of hours studied in the academic year for each record. Used only where SNQUALHOURS will be used to calculate the FTE of the record. We cap SNHOURSPERAYR at the value of SNQUALHOURS for the course.

- $SNHOURSPERAYR = (SNAVHOURSPERDAY \times SNAYDAYSSTUDIED)$  or SNQUALHOURS, whichever is lowest.

## SNHESTULOADCASE

46. This field defines the fields we will use in calculating the FTE for a given higher education record.

Value	Description	Definition
0	Students with STULOAD returned	SNEXCL = 0 and SNLEVEL = HE and STULOAD ≠ BLANK
1	Students in HESA records with no STULOAD returned	SNEXCL = 0 and SNLEVEL = HE and SNSOURCE = HESASTU, HESASAR and STULOAD = BLANK
2	Students in ILR with no STULOAD returned and with SNQUALHOURS returned	SNEXCL = 0 and SNLEVEL = HE and SNSOURCE = ILR and STULOAD = BLANK and SNQUALHOURS ≠ BLANK
3	Students in ILR with no STULOAD returned and no SNQUALHOURS returned	SNEXCL = 0 and SNLEVEL = HE and SNSOURCE = ILR and STULOAD = BLANK and not above
4	Students where their record is excluded	SNEXCL > 0 and SNLEVEL = HE

## SNHESTULOAD

47. This field determines the STULOAD for each higher education record.

Value	Definition
STULOAD	SNHESTULOADCASE = 0
0	SNHESTULOADCASE = 1 or 4
$(\text{SNHOURSPERAYR} \div 540) \times 100$	SNHESTULOADCASE = 2
25	SNHESTULOADCASE = 3

## SNHEINSTLEVSTULOAD

48. This field determines the total STULOAD at a given SNLEVELNUM for each higher education instance.<sup>3</sup>

Value	Definition
SNHESTULOAD	SNSOURCE = HESASTU
Sum of SNHESTULOAD across all instance periods at a given SNLEVELNUM per instance	SNSOURCE = HESASAR
Sum of SNHESTULOAD across all records for each SNLEVELNUM per student	SNSOURCE = ILR

<sup>3</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

49. This field calculates the capped STULOAD at a given SNLEVELNUM for each higher education instance.

- $SNHEINSTLEVSTULOADADJ = SNHEINSTLEVSTULOAD$  or 150, whichever is lowest.

## SNTOTHESTULOAD

50. 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.

## SNTOTHEFTE

51. This field calculates the total higher education FTE, calculated as  $SNTOTHESTULOAD \div 100$ .

## Additional fields required to calculate total further education FTE

### SNFEPLANLEARNSTULOAD

52. For ILR records only. This field converts PLANLEARNHOURS to STULOAD for calculation of further education FTE.

- $SNFEPLANLEARNSTULOAD = (((PLANLEARNHOURS \div 540) \times 100) - SNTOTHESTULOAD)$  or 0, whichever is highest.

### SNFESTULOADCASE

53. This field determines the fields that will be used to calculate the FTE for further education records.

Value	Description	Definition
0	Students with STULOAD returned	$SNEXCL = 0$ and $SNLEVEL = FE$ and $SNSOURCE = HESASTU, HESASAR$ and $STULOAD \neq BLANK$
1	Students in HESA records with no STULOAD returned	$SNEXCL = 0$ and $SNLEVEL = FE$ and $SNSOURCE = HESASTU, HESASAR$ and $STULOAD = BLANK$
2	Students in ILR with PLANLEARNHOURS returned	$SNEXCL = 0$ and $SNLEVEL = FE$ and $SNSOURCE = ILR$ and $PLANLEARNHOURS \neq BLANK$
3	Students in ILR with SNQUALHOURS returned and no PLANLEARNHOURS returned	$SNEXCL = 0$ and $SNLEVEL = FE$ and $SNSOURCE = ILR$ and $SNQUALHOURS \neq BLANK$ and not above
4	Students in ILR fitting none of the above	$SNEXCL = 0$ and $SNLEVEL = FE$ and $SNSOURCE = ILR$ and not above
5	Students where their record is excluded	$SNEXCL > 0$ and $SNLEVEL = FE$

## SNNUMFERECORDS

54. For ILR data only. This field identifies the number of records per student with SNLEVEL = FE and SNEXCL = 0.

## SNFESTULOAD

55. This field determines the STULOAD for each further education record.

Value	Definition
STULOAD	SNFESTULOADCASE = 0
0	SNFESTULOADCASE = 1 or 5
$\text{SNFEPLANLEARNSTULOAD} \div \text{SNNUMFERECORDS}$	SNFESTULOADCASE = 2
$(\text{SNHOURLSPERAY} \div 540) \times 100$	SNFESTULOADCASE = 3
10	SNFESTULOADCASE = 4

## SNTOTFESTULOAD

56. This field calculates the sum of SNFESTULOAD across all records for the student. This is capped at a maximum of 150 per student.

## SNTOTFEFTE

57. This field determines the total further education FTE, calculated as  $\text{SNTOTFESTULOAD} \div 100$ .

## Additional fields required to calculate total undergraduate higher education FTE relevant to TEF

### SNREL\_TEF

58. Identifies undergraduate provision relevant to TEF.

59. SNSOURCE = HESASTU, HESASAR

Value	Description	Definition
1	Provision in-scope for TEF	SNEXCL = 0 and COURSEAIM in (C13, C20, C30, C41, C42, C43, C77, C78, C80, C90, C99, M22, M26, M28, H00, H11, H12, H13, H16, H18, H22, H23, H24, H41, H42, H43, H70, H76, H79, H80, H90, H91, H99, I00, I11, I12, I16, I70, I74, I76, I78, I79, I80, I90, I91, I99, J10, J13, J16, J30, J20, J26, J41, J42, J43, J45, J76, J80, J90, J99)
0	Provision not in-scope for TEF	Otherwise

SNSOURCE = ILR

Value	Description	Definition
1	Provision in-scope for TEF	SNEXCL = 0 and ( SNLEVELNUM in (4,5,6) or QUAL_TYP in (6003, 1458, 1459) or  ( LEARNAIMREF in (60102548, 60104624, 60104636, 60105185, 6010580X, 60112281, 60114629,60116225, 60123837, 60124453, 60125032, 6012717X, 60132644, 60153507, 60161991, 60179752, 60181229, 60305757) and (NOTIONALNVQLEVELV2 in (4, 5, 6, 7, 8, H) or NVQ_LV in (4, 5, H)) ) or  ( QUAL_TYP = 1449 and (NOTIONALNVQLEVELV2 in (4, 5, 6, 7, 8, H) or NVQ_LV in (4, 5, H)) ) )
0	Provision not in-scope for TEF	Otherwise

### SNHEINSTLEVSTULOAD\_TEF

60. The total STULOAD at a given SNLEVELNUM for each undergraduate student higher education instance<sup>4</sup> relevant to TEF

Value	Definition
SNHESTULOAD	SNSOURCE=HESASTU SNREL_TEF = 1
Sum of SNHESTULOAD across all undergraduate students relevant to TEF instance periods at a given SNLEVELNUM per instance	SNSOURCE=HESAAP and SNREL_TEF = 1
Sum of SNHESTULOAD across all undergraduate students relevant to TEF records for each SNLEVELNUM per student	SNSOURCE=ILR and SNREL_TEF = 1

### SNHEINSTLEVSTULOADADJ\_TEF

61. The capped STULOAD at a given SNLEVELNUM for each undergraduate student higher education instance relevant to TE.

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

<sup>4</sup> The ILR data does not include the concept of an 'instance'. Solely for the purpose of calculating SNHEINSTLEVSTULOAD\_TEF we treat a level of study in the ILR data as analogous to an instance.

### **SNTOTHESTULOAD\_TEF**

62. SNTOTHESTULOAD\_TEF is the total STULOAD per student including any capping, calculated as the sum of SNHEINSTLEVSTULOADADJ\_TEF across each unique combination of instance and level per student.

### **SNTOTHEFTE\_TEF**

63. The total undergraduate students higher education FTE relevant to TEF, calculated as  $SNTOTHESTULOAD\_TEF \div 100$ .

## Rebuild instructions

64. The individualised file, SN20\_AAA\_XXXXXXXX\_IND.csv, can be used to interrogate the data and our calculation of derived fields by filtering to certain groups of records. Full details of how to access and use an individualised file are given on the OfS website ([www.officeforstudents.org.uk/data-and-analysis/supplying-data/working-with-individualised-files](http://www.officeforstudents.org.uk/data-and-analysis/supplying-data/working-with-individualised-files)).
65. Specific instructions to rebuild the student number summary figures in tables 1 and 2 of the student numbers data summary workbook, SN20\_AAA\_XXXXXXXX.xlsx, are provided here.
66. A number of additional fields not described in the algorithms above have been added to the individualised file. These are listed in Table 7.

**Table 7: Additional fields in the individualised file**

Field name	Description
STUDENTIDTYPE	The type of student identifier field for the record. This will be HUSID for records from the HESA Student record or HESA Student alternative record, and LEARNREFNUMBER for records from the ILR.
STUDENTID	The HUSID or LEARNREFNUMBER for the record.
REBUILDHEFEFLAG	A column provided to enable the figures in Table 1 to be rebuilt.
REBUILDL456FLAG	A column provided to enable the figures in Table 2 to be rebuilt.

67. The following instructions assume the user has the individualised file open in Microsoft Excel.
68. To apply filters to the column headings, highlight the headings. From the 'DATA' tab in the toolbar select 'Filter'.
69. You will need to use the SUBTOTAL function with option 9 to sum the values in the relevant columns once the filters have been applied. Using the SUBTOTAL function instead of the SUM function will ensure that you only include those values that have been filtered to.

**Table 1 Split of further and higher education provision**

70. The fields SNTOTFEFTE and SNTOTHEFTE are calculated per student. Where a student instance extends over more than one row of data, SNTOTFEFTE and SNTOTHEFTE will be the sum of the further education and higher education FTE respectively across all rows of data. The REBUILDHEFEFLAG is provided to allow you to easily de-duplicate the data, and only count the total further education and higher education FTE once per student. Please note that REBUILDHEFEFLAG should not be used to determine which individual rows of data have been included or excluded from the calculation of student numbers. Instead, the SNEXCL field should be used to determine this.
71. To rebuild the figures in table 1 of the workbook, filter to REBUILDHEFEFLAG = 1. The subtotals for SNTOTFEFTE and SNTOTHEFTE will give the totals for the further education and higher education FTE respectively. You do not need to filter by SNLEVEL as this has already been accounted for in generating REBUILDHEFEFLAG.



## **Table 2 Split of higher education by SNLEVELNUM**

72. To allow you to calculate the total HE FTE at each individual level (4, 5, 6, 7 or 8), the REBUILDL456FLAG is provided. This field can be used to select one row of data per student per level of higher education study, to prevent double counting the FTE where an individual has multiple aims at the same level of study.
73. To rebuild the figures in table 2 of the workbook, first clear the filter from REBUILDHEFEFLAG. Filter to REBUILDL456FLAG = 1. The subtotal for SNHEINSTLEVSTULOADADJ, once divided by 100 to produce an FTE, will give the total FTE in Table 2. To rebuild the figures for each level of study, filter SNLEVELNUM to 4, 5, 6, 7 or 8 as required.

## **Table 3 Higher education provision relating to undergraduate levels of study relevant to TEF (SNREL\_TEF)**

74. To rebuild the figures in table 3 of the workbook, first clear any filters. Filter to REBUILDL456FLAG = 1 and SNREL\_TEF = 1. The subtotal for SNHEINSTLEVSTULOADADJ\_TEF, once divided by 100 to produce an FTE, will give the total FTE in Table 3.



© The Office for Students copyright 2022

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/)