

**Office for
Students**



Calculating student numbers:

**2019-20 technical document
and rebuild instructions**

Enquiries to studentnumbers@officeforstudents.org.uk

Publication date 3 August 2020

Contents

Summary	3
Changes to the 2019-20 method	4
Changes from the 2018-19 method	4
Note on terminology	5
20Y1-Y2 fields used in the student number algorithms	5
Description of Derived Fields	8
Derived fields for calculating total higher education FTE	9
Additional fields required to calculate total further education FTE	16
Rebuild instructions	18

Summary

1. This document describes the algorithms applied to the 2019-20 Higher Education Statistics Agency (HESA) student record, HESA student alternative record and Individualised Learner Record (ILR) data, in order to determine a provider's higher education and further education student numbers.
2. 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.
3. The student numbers outputs are formed of two files. These are:
 - a. Student numbers workbook 'SN19_AAA_XXXXXXXX.xlsx'
 - b. Student numbers individualised file 'SN19_AAA_XXXXXXXX_IND.csv'

where XXXXXXXX is your provider's eight digit UK Provider Reference Number (UKPRN), 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

4. 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.
5. 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 2019-20 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.
6. 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, 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.

Changes to the 2019-20 method

7. This section describes any changes to the algorithms since it was first published on 12 August 2020.

SNQUALHOURS

This algorithm has been updated to treat 0 the same as missing when using the raw data fields TOTAL QUALIFICATION TIME, MAXIMUM GUIDED LEARNING HOURS and TOTAL CREDITS to calculate the total hours a course is expected to take (SNQUALHOURS).

SNHESTULOADCASE

SNHESTULOADCASE = 4 has been updated to include the condition SNLEVEL = HE. This change has no impact on the final FTE calculated for each student.

SNFESTULOADCASE

SNFESTULOADCASE = 5 has been updated to include the condition SNLEVEL = FE. This change has no impact on the final FTE calculated for each student.

Changes from the 2018-19 method

8. This section describes any changes to the algorithms since the equivalent was published for the 2018-19 algorithms in '2018-19 student numbers: Technical document' available from the student number data pages at www.officeforstudents.org.uk/data-and-analysis/student-number-data/

HESA student alternative

9. The name of the HESA alternative provider record has been changed to the student alternative record for the 2019-20 academic year. Algorithms and abbreviations used throughout the document have been updated accordingly.

SNEXCL16

10. The HESA student record no longer contains any apprenticeship standard 'wrappers', and the related COURSEAIM Z99 ('course aim does not apply') has consequently been removed. SNEXCL16 now excludes apprenticeship standard 'wrapper' programme aims from the ILR data only, where previously it excluded those contained in the HESA student record too.

SNHESTULOADCASE

11. An additional value of SNHESTULOADCASE (SNHESTULOADCASE = 4) has been added for excluded records (SNEXCL > 0) to make it clearer than they do not contribute to the calculation

of full time equivalence (FTE). This change has no impact on the final FTE calculated for each student.

SNHESTULOAD

12. Excluded records (SNEXCL > 0) are explicitly given a SNHESTULOAD of 0. This change has no impact on the final FTE value calculated per student.

SNFESTULOADCASE

13. An additional value of SNFESTULOADCASE (SNFESTULOADCASE = 5) has been added for excluded records (SNEXCL > 0) to make it clearer that they do not contribute to the calculation of FTE. This change has no impact on the final FTE calculated for each student.

SNFESTULOAD

14. Excluded records (SNEXCL > 0) are explicitly given a SNFESTULOAD of 0. This change has no impact on the final FTE value calculated per student.

Note on terminology

15. 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 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.

16. 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.

17. 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 2: HESA student fields used in the student number calculations

Entity	Field	Description
Course	COURSEAIM	General qualification aim of course
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 3: HESA student alternative fields used in the student number calculations

Entity	Field	Description
Course	COURSEAIM	General qualification aim of course
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
HESA derived field	XINACT01	Inactive for entire reporting year

Table 4: ILR, Learning Aim Reference Service and Ofqual fields used in the student number calculations

Data source	Field	Description
ILR	COMPSTATUS	Degree of completion of the learning activities leading to the learning aim
ILR	LEARNACTENDDATE	Learning actual end date
ILR	LEARNDELFAMCODE ¹	The value of the funding or monitoring attribute
ILR	LEARNDELFAMTYPE ¹	The type of funding or monitoring attribute
ILR	LEARNPLANENDDATE	Learning planned end date
ILR	LEARNSTARTDATE	Learning start date
ILR	STULOAD	Learner FTE
ILR	PLANLEARNHOURS	Planned learning hours
ILR	PRIORLEARNFUNDADJ	The proportion of this learning aim still to be delivered
ILR	PROGTYPE	The type of programme which the learner is undertaking
ILR	UKPRN	UK Provider Reference Number
ILR	WITHDRAWREASON	The reason why the learner has withdrawn from the learning aim
LARS	LEARNAIMREF	Learning aim reference
LARS	LEARNAIMREFTITLE	Learning aim title
LARS	LEARNAIMREFTYPE	Learning aim type

¹ The only values of LEARNDELFAMTYPE and associated LEARNDELFAMCODE field used in the calculation of student numbers are those where LEARNDELFAMTYPE = ADL.

LARS	NOTIONAL NVQ LEVEL V2	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'.

Description of Derived Fields

Table 5: Description of derived fields for recreating tables

Derived field name	Description	Paragraph
LEARNDELFAM_ADL1	Advanced Learner Loan indicator	20
SNAVHOURLSPERDAY	Average study hours per day over course	41
SNAYDAYSSTUDIED	Number of days studied in academic year	40
SNENDDATE	Last day studied in academic year	39
SNEXCL	Reason(s) for exclusion from student numbers population	31
SNEXCL1 – SNEXCL256	Student number population exclusions	22 - 30
SNFEPLANLEARNSTULOAD	Further education STULOAD calculated from PLANLEARNHOURS	49
SNFESTULOAD	STULOAD for further education record	52
SNFESTULOADCASE	Determines fields used to calculate further education FTE for a given record	50
SNHEINSTLEVSTULOAD	Total STULOAD at a given SNHESTUDYLEVEL for a given higher education instance	45
SNHEINSTLEVSTULOADADJ	Capped STULOAD at a given SNHESTUDYLEVEL for a given higher education instance	46
SNHESAILRFLAG	Providers that return both HESA and ILR data	21
SNHESTUDYLEVEL	Level of higher education record	34
SNHESTULOAD	STULOAD for higher education record	44
SNHESTULOADCASE	Determines fields used to calculate higher education FTE for a given record	43
SNHOURSPEYR	Total hours studied in academic year	42
SNLEVEL	Identifies record as higher or further education	33
SNNUMFERECORDS	Number of non-excluded further education records per student	51
SNPRIORLEARNADJ	Adjustment for prior learning	37
SNQUALHOURS	Total hours course is expected to take	38
SNSOURCE	Identifies the source of the record	18
SNTITLECREDITS	Max credits in course title	36
SNTITLEHRS	Max learning hours in course title	35
SNTOTFEFTE	Total further education further education per student	54
SNTOTFESTULOAD	Total capped further education STULOAD per student	53
SNTOTHEFTE	Total higher education FTE per student	48

SNTOTHESTULOAD	Total capped higher education STULOAD per student	47
----------------	---	----

Derived fields for calculating total higher education FTE

SNSOURCE

18. 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

19. 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_AD L1

20. 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)

21. 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

22. 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

23. 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

24. 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

25. 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 LOCSDY = S)) or (SNSOURCE = HESASAR and (EXCHIND = 1 or LOCSDY = S))
0	Included	Otherwise

SNEXCL16

26. 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

27. 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

28. 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 ² and WITHDRAWREASON = 40 and LEARNACTENDDATE = 1 August 20Y1
0	Included	Otherwise

SNEXCL128 (unavailable in DCT or AST outputs)

29. 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

SNEXCL256

30. 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 LEARNDEL FAM_AD L = 1 and NOTIONALNVQLEVELV2 = 4,5,6,
0	Included	Otherwise

² 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.

SNEXCL

31. This field indicates whether the record will be included in the calculation of student numbers. SNEXCL is the sum of all 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

32. 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

33. 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 \neq ZPROG001 and NOTIONALNVQLEVELV2 = 4, 5, 6, 7, 8, H
FE	Record of further education	LEARNAIMREF \neq ZPROG001 and not above

SNHESTUDYLEVEL

34. 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 framework³.

SNSOURCE = HESASTU, HESASAR

Value	Description	Definition
6+	Records at Level 6 or above	SNLEVEL = HE and COURSEAIM = D, E, L, M, H, I00, I11, I12, I16
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
6+	Records at Level 6 or above	SNLEVEL = HE and NOTIONALNVQLEVELV2 = 6, 7, 8
5	Records at Level 5	SNLEVEL = HE and NOTIONALNVQLEVELV2 = 5
4	Records at Level 4	SNLEVEL = HE and not above

SNTITLEHRS

35. 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

36. 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

37. 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

³ 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.

SNQUALHOURS

38. 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 = MISSING and TOTAL QUALIFICATION TIME ≠ MISSING, 0
MAXIMUM GUIDED LEARNING HOURS × SNPRIORLEARNADJ	Source of hours is MAXIMUM GUIDED LEARNING HOURS	SNSOURCE = ILR and STULOAD = MISSING and MAXIMUM GUIDED LEARNING HOURS ≠ MISSING, 0 and not above
(TOTAL CREDITS × 10) × SNPRIORLEARNADJ	Source of hours is TOTAL CREDITS	SNSOURCE = ILR and STULOAD = MISSING and TOTAL CREDITS ≠ MISSING ,0 and not above
(SNTITLEHRS) × SNPRIORLEARNADJ	Source of hours is SNTITLEHRS	SNSOURCE = ILR and STULOAD = MISSING and SNTITLEHRS ≠ MISSING and not above
(SNTITLECREDITS × 10) × SNPRIORLEARNADJ	Source of hours is SNTITLECREDITS	SNSOURCE = ILR and STULOAD = MISSING and SNTITLECREDITS ≠ MISSING and not above

SNENDDATE

39. 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 ≠ 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

40. 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.

SNVHOURSPERDAY

41. 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

42. 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.

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

SNHESTULOADCASE

43. This field 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	SNEXCL = 0 and SNLEVEL = HE and STULOAD ≠ MISSING
1	SNHESTULOAD will be set to 0 for this record	SNEXCL = 0 and SNLEVEL = HE and SNSOURCE = HESASTU, HESASAR and STULOAD = MISSING
2	SNQUALHOURS will be used in calculating FTE for this record	SNEXCL = 0 and SNLEVEL = HE and SNSOURCE = ILR and STULOAD = MISSING and SNQUALHOURS ≠ MISSING
3	A value of 25 STULOAD will be assigned to this record	SNEXCL = 0 and SNLEVEL = HE and SNSOURCE = ILR and STULOAD = MISSING and not above
4	Record is excluded. SNHESTULOAD will be set to 0 for this record	SNEXCL > 0 and SNLEVEL = HE

SNHESTULOAD

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

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

SNHEINSTLEVSTULOAD

45. This field determines the total STULOAD at a given SNHESTUDYLEVEL for each higher education instance.⁴

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

SNHEINSTLEVSTULOADADJ

46. This field calculates the capped STULOAD at a given SNHESTUDYLEVEL for each higher education instance.

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

SNTOTHESTULOAD

47. 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

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

Additional fields required to calculate total further education FTE

SNFEPLANLEARNSTULOAD

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

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

⁴ 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.

SNFESTULOADCASE

50. This field 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	SNEXCL = 0 and SNLEVEL = FE and SNSOURCE = HESASTU, HESASAR and STULOAD ≠ MISSING
1	SNFESTULOAD will be set to 0 for this record	SNEXCL = 0 and SNLEVEL = FE and SNSOURCE = HESASTU, HESASAR and STULOAD = MISSING
2	PLANLEARNHOURS will be used when calculating FTE for this record	SNEXCL = 0 and SNLEVEL = FE and SNSOURCE = ILR and PLANLEARNHOURS ≠ MISSING
3	SNQUALHOURS will be used when calculating FTE for this record	SNEXCL = 0 and SNLEVEL = FE and SNSOURCE = ILR and SNQUALHOURS ≠ MISSING and not above
4	A value of 10 STULOAD will be assigned to this record	SNEXCL = 0 and SNLEVEL = FE and SNSOURCE = ILR and not above
5	Record is excluded. SNFESTULOAD will be set to 0 for this record	SNEXCL > 0 and SNLEVEL = FE

SNNUMFERECORDS

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

SNFESTULOAD

52. 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
$(\text{SNHOURSPERAY} \div 540) \times 100$	SNFESTULOADCASE = 3
10	SNFESTULOADCASE = 4

SNTOTFESTULOAD

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

SNTOTFEFTE

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

Rebuild instructions

55. The individualised file, SN19_AAA_XXXXXXXXX_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). Specific instructions to rebuild the student number summary figures in tables 1 and 2 of the student numbers data summary workbook, SN19_AAA_XXXXXXXXX.xlsx, are provided here.
56. A number of additional fields not described in the algorithms above have been added to the individualised file. These are listed in the table below.

Table 6: 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.

57. The following instructions assume the user has the individualised file open in Microsoft Excel.
58. To apply filters to the column headings, highlight the headings. From the 'DATA' tab in the toolbar select 'Filter'.
59. 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

60. 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.
61. To rebuild the figures in Table 1, 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 SNHESTUDYLEVEL

62. To allow you to calculate the total HE FTE at each individual level (4, 5, or 6+), 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.
63. To rebuild the figures in Table 2, 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 SNHESTUDYLEVEL to 4, 5 or 6+ as required.



© The Office for Students copyright 2020

This publication is available under the Open Government Licence 3.0 except where it indicates that the copyright for images or text is owned elsewhere.

www.nationalarchives.gov.uk/doc/open-government-licence/version/3/