2023-24 Student numbers data checking tool

2023-24 Student numbers technical document and rebuild instructions

Contents

Summary

1. This document describes the algorithms applied to the Student (23056) record submitted to the designated data body (DDB) and the 2023-24 Individualised Learner Record (ILR) submitted to the Education and Skills Funding Agency (ESFA) to determine a provider's higher education and further education student numbers. Throughout the document, fields taken or derived from the Student (23056) or ILR records are shown in capitals.

2. We intend to use these student numbers for various regulatory purposes. These include setting registration fees, assessing applications for degree awarding powers, assessing applications for university title and university college title, and for determining mandatory participation in the Teaching Excellence Framework (TEF).

3. The data in the student numbers data checking tool outputs can be used by higher education providers to verify and correct their 2023-24 individualised student data. The outputs are also an opportunity for providers to give feedback on any potential issues with Office for Students (OfS) algorithms and must not be used for any other purpose.

4. The student numbers outputs are formed of two files. These are:

   a. Student numbers workbook 'SN23_DCT_ XXXXXXXX.xlsx'

   b. Student numbers individualised file 'SN23_DCT_ XXXXXXXX_IND.csv'

   where XXXXXXXX is the UK Provider Reference Number (UKPRN) for the provider.

5. Details of how to use individualised files are given on the Office for Students (OfS) website at: www.officeforstudents.org.uk/data-and-analysis/supplying-data/working-with-individualised-files/.

6. For the Student (23056) record, these files can be accessed through the HESA Data Platform.[1] Once you have processed submitted data, your file can be downloaded from the 'Iris' report, under the section 'Additional Collection Reports'.

7. For ILR, these files can be accessed from the '2023-24 Data checking tool' area of the OfS portal. Access details will be sent to the appropriate contacts at providers when the data checking tool is available.

8. Some providers submit both a Student (23056) and an ILR return. Student engagements meeting either of the criteria listed below should be reported to both the DDB and the 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 engagements (Level 4 or above) who are financed by Advanced Learner Loans.

---

[1] See HESA Data Platform.

We would expect all other student engagements to be returned to either the DDB or the ESFA, but not both. For those providers that return both Student (23056) and ILR data, the student engagements listed above will be excluded from the ILR data to avoid double counting.

9. This document is aimed at readers with in-depth knowledge of the data. Readers are advised to refer to a copy of the Student record – 23056 coding manual[2], 2023-24 ILR specification[3], and the Office of Qualifications and Examinations Regulation (Ofqual) Register Data Extract Specification[4] to hand when using this document.

10. We will document any changes or corrections to the algorithms in this document in the data checking tool section of the OfS website.[5]

Terminology

11. The algorithms below apply to the Student (23056) and ILR data collections.

12. In the Student (23056) return, the fields that we use are associated with a student engagement. In the ILR return the fields that we use are associated with a learning aim. A learner may study multiple learning aims within the same qualification.

13. Below, we use the generic term 'record' to refer to either a student engagement or a learning aim, depending on the data source.

14. The year '2023-24' below refers to the academic year starting on 1 August 2023 and ending on 31 July 2024.

---

[2] See Student record – 23056 coding manual.

[3] See GOV.UK - 2023-24 ILR specification.

[4] See Ofqual register data extract specification.

[5] See The OfS website – Data checking tool.

Changes from previous publications

15. This is the first publication of these algorithms relating to 2023-24 data. Any changes after initial publication will be listed here.

16. This section also describes any changes to the algorithms since the equivalent document was published for the 2022-23 algorithms in '2022-23 Calculating student numbers DCT technical document and rebuild instructions'[6] available from the data checking tool documentation archive section of the OfS website.[7]

17. Any deletions are shown with a ~~strikethrough~~. Any other changes are shown as **bold and underlined**.

Changes from the Student (22056) data checking tool

SNLEVELNUM

18. The following changes have been made to SNLEVELNUM:

    a. Replacing QUALCAT with Z_LEVEL. For these purposes QUALCAT and Z_LEVEL are equivalent and so this has no impact on the output.

    b. Removal of values 'D0005' and 'E0005' as these are no longer valid values for 2023-24.

    c. Addition of explicit values of Z_LEVEL for SNLEVELNUM=4.

    d. Addition of a '<4' category. We do not expect any qualification aims to be less than level 4, so data in this category is either incorrect or out of scope of the Student (23056) collection.

    e. Addition of an 'Unknown' category which captures records with no student course session or missing values of QUALCAT.

Value	Description	Definition
8	Records at Level 8	~~QUALCAT~~ **Z_LEVEL** in (D0003, D0004, E0000, E0001, E0002, E0003, E0004, L0000)
7	Records at Level 7	~~QUALCAT~~ **Z_LEVEL** in (L0001, L0002, L0003, ~~D0005~~, M0002, M0003, M0004, M0006, M0007, M0008, M0009, M0010, M0011, M0012, M0013, M0015, M0016, M0017, M0018, M0020, M0021, M0022, M0023, M0024, ~~E0005~~)
6	Records at Level 6	~~QUALCAT~~ **Z_LEVEL** in (H0003, H0004, H0005, H0006, H0007, H0008, H0009, H0010, H0012, H0013, H0014, H0015, H0016, H0018, H0019, H0020, I0001)

---

[6] See https://www.officeforstudents.org.uk/media/8178/2022-23-calculating-student-numbers-dct-technical-document-and-rebuild-instructions.pdf [PDF].

[7] See The OfS website - Data checking tool documentation archive.

Value	Description	Definition
5	Records at Level 5	~~QUALCAT~~ **Z_LEVEL** in (I0002, I0004, I0005, I0006, I0007, I0008, I0009, I0010, I0012, I0013, J0000, J0001, J0002, J0003, J0004, J0005, J0006, J0007, J0010, J0011, J0012)
4	Records at Level 4	~~Otherwise~~ **Z_LEVEL in (C0000, C0001, C0002, C0003, C0004, C0005, C0006, C0007, C0008, C0009)**
<4	**Records below level 4**	**Z_LEVEL = P0002**
Unknown	**Records with no student course session or missing QUALCAT**	**Z_LEVEL = Z9**

SNREL_TEF

19. The following changes have been made to SNREL_TEF:

    a. Replacing QUALCAT with Z_LEVEL. For these purposes QUALCAT and Z_LEVEL are equivalent and so this has no impact on the output.

    b. Adding the requirement that PREREQUISITE ≠ 02 in order for records with Z_LEVEL in (H0016, I0010) to be in scope for TEF. This is a correction to the algorithm used in the Student (22056) data checking tool.

Value	Description	Definition
1	Provision in-scope for TEF	SNEXCL = 0 and **(Z_LEVEL** ~~QUALCAT~~ in (C0000, C0001, C0002, C0003, C0004, C0005, C0006, C0007, C0008, C0009, H0003, H0004, H0005, H0006, H0007, H0008, H0012, H0015, ~~H0016,~~ H0018, H0019, H0020, I0001, I0004, I0008, I0009, ~~I0010~~ I0012, I0013, J0000, J0001, J0002, J0003, J0004, J0005, J0006, J0007, J0010, J0011, J0012, M0002) **or** **(Z_LEVEL in (H0016, I0010) and PREREQUISITE ≠ 02))**
0	Provision not in-scope for TEF	Otherwise

PRINONUK
20. PRINONUK has been renamed to ENGPRINONUK and moved from the Study Location entity to the Engagement entity.

Additional fields calculated
21. SNLEVEL, SNHESTULOAD, SNHEENGLEVSTULOAD, SNHEENGLEVSTULOADADJ, SNTOTHESTULOAD, SNHEENGLEVSTULOAD_TEF, SNHEENGLEVSTULOADADJ_TEF and SNTOTHESTULOAD_TEF were not previously explicitly calculated for DDB records. These fields are now calculated for DDB records.

SNTOTHEFTE

22. The algorithm for SNTOTHEFTE has now been updated to use SNTOTHESTULOAD. This means SNTOTHEFTE is now the FTE per student rather than the FTE per engagement.[8]

Changes from the 2022-23 ILR data checking tool

SNLEVELNUM

23. Values of '0' and '*BLANK*' have been added.

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~~ **NOTIONALNVQLEVELV2 = 4**
0	**Unknown HE level aim**	**SNLEVEL = HE and NOTIONALNVQLEVELV2 = H**
BLANK	**Not applicable to higher education qualifications framework**	**Otherwise**

SNLEVEL

24. A value *'BLANK'* has been added to explicitly show the value where LEARNAIMREF = ZPROG001.

Value	Description	Definition
HE	Record of higher education	SN (SNSOURCE = ILR and LEARNAIMREF ≠ ZPROG001 and NOTIONALNVQLEVELV2 = 4, 5, 6, 7, 8, H)
FE	Record of further education	LEARNAIMREF ≠ ZPROG001 and not above
BLANK	**Apprenticeship standard 'wrapper' programme aim**	**LEARNAIMREF = ZPROG001**

SNQUALHOURS

25. SNQUALHOURS now includes a new value to account for records with unknown qualification hours.

---

[8] In '2022-23 Calculating student numbers DCT technical document and rebuild instructions' SNTOTHEFTE is described as 'the total higher education FTE per *student* including any capping'. However, for DDB records, it was calculating the total higher education FTE per *engagement* including any capping. This had no impact on the total figures as this was accounted for in the rebuild instructions.

SNHESTULOAD

26. SNHESTULOAD now includes a new value to account for blank records.

SNHEINSTLEVSTULOAD

27. SNHEINSTLEVSTULOAD has been renamed to SNHEENGLEVSTULOAD.

SNHEINSTLEVSTULOADADJ

28. SNHEINSTLEVSTULOADADJ has been renamed to SNHEENGLEVSTULOADADJ.

SNHEINSTLEVSTULOAD_TEF

29. SNHEINSTLEVSTULOAD_TEF has been renamed to SNHEENGLEVSTULOAD_TEF.

SNHEINSTLEVSTULOADADJ_TEF

30. SNHEINSTLEVSTULOADADJ_TEF has been renamed to SNHEENGLEVSTULOADADJ_TEF.

SNFESTULOAD

31. SNFESTULOAD now includes a new value to account for blank records.

SNREL_TEF

32. SNREL_TEF is now derived using OFSQAIM.

Value	Description	Definition
1	Provision in-scope for TEF	SNEXCL = 0 and **OFSQAIM in (FIRST, ENHANCED, FDBC, OTHL6_Q, OTHL6_CC, OTHL6_U, HND, DET, FOUDEG, DIPHE, DTLLS, OTHL5_Q, OTHL5_CC, OTHL5_U, HNC, CERTED, UNICERT, HIGHCERT, CTLLS, PTLLS, CET, OTHL4_Q, OTHL4_CC, OTHL4_U, OTHHE_Q, OTHHE_CC, OTHHE_U)** ~~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

Student numbers output files

Workbook

33. The student numbers workbook is the Excel workbook SN23_DCT_ XXXXXXXX.xlsx which contains the following worksheets:

Table 1: Worksheets in the student numbers workbook

Worksheet*	Title
Coversheet	Student/ILR data checking tool 2023-24 student numbers data summary

* This worksheet reference corresponds to the spreadsheet tabs.

Individualised file

34. When working through this document it is necessary to use the individualised file, SN23_DCT_ XXXXXXXX_IND.csv. The individualised file contains the values of all fields used to generate student numbers.

35. Fields taken from Student (23056), the ILR or derived as part of the student numbers outputs are shown in capitals using the names given in Tables 2, 3 and 4.

36. Specific instructions for rebuilding the student numbers workbook from the individualised file are provided in this document.

37. For Student (23056) data the individualised file contains one row per engagement (each unique combination of UKPRN, SID and NUMHUS). For ILR data the individualised file contains one row per learning aim for a learner (each unique combination of UKPRN, LEARNREFNUMBER, LEARNAIMREF and AIMSEQNUMBER).

2023-24 fields used to create the student numbers outputs

38. The fields used to generate the student numbers outputs are listed in Tables 2 and 3.

Table 2: Student (23056) fields used in the student numbers calculations

Entity	Name	Description
Engagement	UKPRN	UK Provider Reference Number
Engagement	SID	Student identifier
Engagement	NUMHUS	Engagement number
Engagement	ENGSTARTDATE	Engagement start date
Engagement	ENGENDDATE	Engagement end date
Engagement	ENGPRINONUK	Identifies if the student is studying primarily outside the UK
Engagement	INCOMINGEXCHANGE	Identifies if the student is an incoming or visiting exchange student
Course	PREREQUISITE	Identifies PGR courses for funding purposes.
Engagement derived field	Z_ACT_CYC	Derived field records the student activity during the cycle
Engagement derived field	Z_STULOAD_CYC	Derived field represents the total student load to date within the cycle
Engagement derived field	Z_LEVEL	Derived field records student's course level for the engagement

Table 3: ILR, Learning Aim Reference Service (LARS) and Ofqual fields used in the student numbers calculations

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	The value of the funding or monitoring attribute
ILR	LEARNDELFAMTYPE	The type of funding or monitoring attribute
ILR	LEARNPLANENDDATE	Learning planned end date
ILR	LEARNSTARTDATE	Learning start date
ILR	STULOAD	Learner Full Time Equivalence (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

Source	Field	Description
ILR	WITHDRAWREASON	The reason why the learner has withdrawn from the learning aim
ILR	LEARNAIMREF	Learning aim reference
ILR	AIMSEQNUMBER	Learning aim data set sequence
LARS	UNITTYPE	Learning aim's unit type
LARS	LEARNAIMREFTITLE	Learning aim title
LARS	LEARNAIMREFTYPE	Learning aim type
LARS	NVQ_LV	Notional NVQ level
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

Description of derived fields

39. The derived fields used to generate the student numbers outputs are listed in Table 4.

Table 4: Description of derived fields for recreating tables

Derived field name	Description	Paragraph
SNSOURCE	Identifies the source of the record	40
LEARNDELFAM_ADL1	Advanced Learner Loan indicator	41
SNDDBILRFLAG	Identifies providers that return both Student and ILR data	42
SNEXCL1 – SNEXCL256	Student numbers population exclusions	43 – 50
SNEXCL	Reason(s) for exclusion from student numbers population	51
OFSQAIM[9]	Broad category of the learning aim	-
SNLEVEL	Identifies record as higher or further education	52
SNLEVELNUM	Level of higher education record	53
SNTITLEHRS	Max learning hours in course title	54
SNTITLECREDITS	Max credits in course title	55
SNPRIORLEARNADJ	Adjustment for prior learning	56
SNQUALHOURS	Total hours course is expected to take	57

---

[9] OFSQAIM is derived using the algorithms documented in 'HESES23 Course table information', available on the OfS website at HESES - Office for Students. The algorithm for OFSQAIM will be updated and republished the week commencing 26 August 2024.

Derived field name	Description	Paragraph
SNENDDATE	Last day studied in academic year	58
SNAYDAYSSTUDIED	Number of days studied in academic year	59 – 60
SNAVHOURSPERDAY	Average study hours per day over course	61 – 62
SNHOURSPERAYR	Total hours studied in academic year	63
SNHESTULOAD	Student load for higher education record	64
SNHEENGLEVSTULOAD	Student load per engagement (for DDB data) or level (for ILR data)	65
SNHEENGLEVSTULOADADJ	Capped student load per engagement (for DDB data) or level (for ILR data)	66
SNTOTHESTULOAD	Higher education student load per student	67
SNTOTHEFTE	Higher education FTE per student	68
SNFEPLANLEARNSTULOAD	Further education student load calculated from PLANLEARNHOURS	69
SNNUMFERECORDS	Number of non-excluded further education records per student	70
SNFESTULOAD	Student load for further education record	71
SNTOTFESTULOAD	Further education student load per student	72
SNTOTFEFTE	Further education FTE per student	73
SNREL_TEF	Identifies undergraduate higher education provision relevant to TEF	74
SNHEENGLEVSTULOAD_TEF	Student load per engagement (for DDB data) or level (for ILR data) relevant to the Teaching Excellence Framework (TEF)	75
SNHEENGLEVSTULOADADJ_TEF	Capped student load per engagement (for DDB data) or level (for ILR data) relevant to TEF	76
SNTOTHESTULOAD_TEF	Higher education student load relevant to TEF per student	77
SNTOTHEFTE_TEF	Higher education FTE relevant to TEF per student	78

Derived fields for calculating total higher education FTE

SNSOURCE

40. This field identifies the source of the record.

Value	Definition
DDB	Record taken from the DDB (Designated Data Body)
ILR	Record taken from the ILR

LEARNDELFAM_ADL1

41. For ILR records only. This field contains the value of LEARNDELFAMCODE where LEARNDELFAMTYPE = ADL.

SNDDBILRFLAG

42. For ILR records only. This field identifies providers that returned both DDB and ILR data in the academic year 2023-24.

Value	Definition
1	Provider returned both DDB and ILR data for the academic year 2023-24
0	Provider did not return both DDB and ILR data for the academic year 2023-24

SNEXCL1

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

Value	Description	Definition
1	Excluded	(SNSOURCE = DDB and ENGSTARTDATE > 31 July 2024) or (SNSOURCE = ILR and LEARNSTARTDATE > 31 July 2024)
0	Included	Otherwise

SNEXCL2

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

Value	Description	Definition
1	Excluded	(SNSOURCE = DDB and ENGENDDATE < 01 August 2023 and ENGENDDATE ≠ BLANK) or (SNSOURCE = ILR and LEARNACTENDDATE < 01 August 2023 and LEARNACTENDDATE ≠ BLANK)
0	Included	Otherwise

SNEXCL4

45. For DDB records only. This field identifies records that have been inactive for the reporting year.

Value	Description	Definition
1	Excluded	Z_ACT_CYC = 0
0	Included	Otherwise

SNEXCL8

46. For DDB records only. This field identifies incoming exchange students and students who are primarily outside the UK for their engagement.

Value	Description	Definition
1	Excluded	INCOMINGEXCHANGE ≠ BLANK or ENGPRINONUK = 01
0	Included	Otherwise

SNEXCL16

47. For ILR records only. This field identifies apprenticeship standard 'wrapper' programme aims.

Value	Description	Definition
1	Excluded	LEARNAIMREF = ZPROG001
0	Included	Otherwise

SNEXCL32

48. For ILR records only. This field identifies higher apprentices in the ILR that have been returned by providers that also submit DDB data.

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

SNEXCL64

49. For ILR records only. This field identifies a record which has been closed to correct an incorrect LEARNPLANENDDATE.

Value	Description	Definition
1	Excluded	COMPSTATUS = 4[10] and WITHDRAWREASON = 40 and LEARNACTENDDATE = 01 August 2023
0	Included	Otherwise

SNEXCL256

50. For ILR records only. This field identifies students on higher education engagements financed by an Advanced Learner Loan that have been returned by providers that also submit DDB data.

Value	Description	Definition
1	Excluded	SNDDBILRFLAG = 1 and LEARNDELFAM_ADL1 = 1 and NOTIONALNVQLEVELV2 = 4, 5, 6
0	Included	Otherwise

---

[10] 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

51. This field indicates whether the record will be included in the calculation of student numbers. The field contains the exclusion reason(s) for the engagement. It is calculated as (1 × SNEXCL1) + (2 × SNEXCL2) + …+ (256 × SNEXCL256).

SNLEVEL

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

Value	Description	Definition
HE	Record of higher education	SNSOURCE = DDB or (SNSOURCE = ILR and LEARNAIMREF ≠ ZPROG001 and NOTIONALNVQLEVELV2 = 4, 5, 6, 7, 8, H)
FE	Record of further education	LEARNAIMREF ≠ ZPROG001 and not above
BLANK	Apprenticeship standard 'wrapper' programme aim	LEARNAIMREF = ZPROG001

SNLEVELNUM

53. This field indicates the level of study for higher education records, using the sector recognised standards as set out in condition of registration B5.[11]

SNSOURCE = DDB

Value	Description	Definition
8	Records at Level 8	Z_LEVEL in (D0003, D0004, E0000, E0001, E0002, E0003, E0004, L0000)
7	Records at Level 7	Z_LEVEL in (L0001, L0002, L0003, M0002, M0003, M0004, M0006, M0007, M0008, M0009, M0010, M0011, M0012, M0013, M0015, M0016, M0017, M0018, M0020, M0021, M0022, M0023, M0024)
6	Records at Level 6	Z_LEVEL in (H0003, H0004, H0005, H0006, H0007, H0008, H0009, H0010, H0012, H0013, H0014, H0015, H0016, H0018, H0019, H0020, I0001)
5	Records at Level 5	Z_LEVEL in (I0002, I0004, I0005, I0006, I0007, I0008, I0009, I0010, I0012, I0013, J0000, J0001, J0002, J0003, J0004, J0005, J0006, J0007, J0010, J0011, J0012)
4	Records at Level 4	Z_LEVEL in (C0000, C0001, C0002, C0003, C0004, C0005, C0006, C0007, C0008, C0009)
<4	Records below level 4	Z_LEVEL = P0002
UNKNOWN	Records with no student course session or missing QUALCAT	Z_LEVEL = Z9

---

[11] See Sector-recognised standards [PDF].

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 NOTIONALNVQLEVELV2=4
0	Unknown HE level aim	SNLEVEL=HE and NOTIONALNVQLEVELV2=H
BLANK	Not applicable to higher education qualifications framework	Otherwise

SNTITLEHRS

54. 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

55. 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

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

Value	Definition
PRIORLEARNFUNDADJ ÷ 100	PRIORLEARNFUNDADJ > 0
0	PRIORLEARNFUNDADJ = 0

SNQUALHOURS

57. 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	STULOAD = BLANK and TOTAL QUALIFICATION TIME ≠ BLANK, 0
MAXIMUM GUIDED LEARNING HOURS × SNPRIORLEARNADJ	Source of hours is MAXIMUM GUIDED LEARNING HOURS	STULOAD = BLANK and MAXIMUM GUIDED LEARNING HOURS ≠ BLANK, 0 and not above
(TOTAL CREDITS × 10) × SNPRIORLEARNADJ	Source of hours is TOTAL CREDITS	STULOAD = BLANK and TOTAL CREDITS ≠ BLANK, 0 and not above
SNTITLEHRS × SNPRIORLEARNADJ	Source of hours is SNTITLEHRS	STULOAD = BLANK and SNTITLEHRS ≠ BLANK and not above
(SNTITLECREDITS × 10) × SNPRIORLEARNADJ	Source of hours is SNTITLECREDITS	STULOAD = BLANK and SNTITLECREDITS ≠ BLANK and not above
BLANK	Unknown qualification hours	Otherwise

SNENDDATE

58. 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 2023-24. Used only where SNQUALHOURS will be used to calculate the FTE of the record.

Value	Description	Definition
LEARNACTENDDATE	Learner completed learning activities in 2023-24	LEARNACTENDDATE ≠ BLANK and LEARNACTENDDATE <= 31 July 2024
31 July 2024	Learner is still studying at the end of 2023-24	Otherwise

SNAYDAYSSTUDIED

59. For ILR records only. This field determines the number of days studied in the 2023-24 academic year.

60. Calculated as the numbers of days between SNENDDATE and either 01 August 2023 or LEARNSTARTDATE, whichever is later. Used only where SNQUALHOURS will be used to calculate the FTE of the record.

SNAVHOURSPERDAY

61. For ILR records only. This field calculates the average number of hours that would be studied each day for the duration of the course.

62. 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

63. 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 = min((SNAVHOURSPERDAY × SNAYDAYSSTUDIED), SNQUALHOURS).

SNHESTULOAD

64. This field determines the higher education student load that falls within the student numbers population for each record. When using DDB data each record relates to an engagement. When using ILR data, a record refers to a learning aim.

Value	Definition
Z_STULOAD_CYC	SNSOURCE = DDB and SNEXCL = 0
STULOAD	SNSOURCE = ILR and SNEXCL = 0 and SNLEVEL = HE and STULOAD ≠ BLANK
(SNHOURSPERAYR ÷ 540) × 100	SNSOURCE = ILR and SNEXCL = 0 and SNLEVEL = HE and STULOAD = BLANK and SNQUALHOURS ≠ BLANK
25	SNSOURCE = ILR and SNEXCL = 0 and SNLEVEL = HE and STULOAD = BLANK and SNQUALHOURS= BLANK
0	SNEXCL > 0 and SNLEVEL = HE
BLANK	Otherwise

SNHEENGLEVSTULOAD

65. This field determines the total student load for each higher education engagement.[12]

Value	Definition
SNHESTULOAD	SNSOURCE=DDB
Sum of SNHESTULOAD across all records for each unique combination of student (LEARNREFNUMBER) and level (SNLEVELNUM)	SNSOURCE=ILR

SNHEENGLEVSTULOADADJ

66. This field calculates the capped student load for each higher education engagement.[13]

$$SNHEENGLEVSTULOADADJ = min(SNHEENGLEVSTULOAD, 150).$$

---

[12] The ILR data does not include the concept of an 'engagement'. For the purpose of calculating SNHEENGLEVSTULOAD we treat a level of study in the ILR data as analogous to an engagement.

[13] The ILR data does not include the concept of an 'engagement'. For the purpose of calculating SNHEENGLEVSTULOADADJ we treat a level of study in the ILR data as analogous to an engagement.

SNTOTHESTULOAD

67. The total student load per student including any capping, calculated as the sum of SNHEENGLEVSTULOADADJ per student.

SNTOTHEFTE

68. This field calculates the total higher education FTE per student including any capping, calculated as SNTOTHESTULOAD ÷ 100.

Additional fields required to calculate total further education FTE (for ILR records only)

SNFEPLANLEARNSTULOAD

69. For ILR records only. This field converts PLANLEARNHOURS to student load for calculation of further education FTE.

   - SNFEPLANLEARNSTULOAD = max(((PLANLEARNHOURS ÷ 540) × 100) - SNTOTHESTULOAD, 0).

SNNUMFERECORDS

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

SNFESTULOAD

71. For ILR records only. This field determines the student load that falls within the student numbers population for each further education record.

Value	Definition
SNFEPLANLEARNSTULOAD ÷ SNNUMFERECORDS	SNEXCL = 0 and SNLEVEL = FE and PLANLEARNHOURS ≠ BLANK
(SNHOURSPERAYR ÷ 540) × 100	SNEXCL = 0 and SNLEVEL = FE and SNQUALHOURS ≠ BLANK and not above
10	SNEXCL = 0 and SNLEVEL = FE and not above
0	SNEXCL > 0 and SNLEVEL = FE
BLANK	Otherwise

SNTOTFESTULOAD

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

SNTOTFEFTE

73. For ILR records only. This field determines the total further education FTE, calculated as SNTOTFESTULOAD ÷ 100.

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

SNREL_TEF

74. Identifies undergraduate provision relevant to TEF.

SNSOURCE = DDB

Value	Description	Definition
1	Provision in-scope for TEF	SNEXCL = 0 and (Z_LEVEL in (C0000, C0001, C0002, C0003, C0004, C0005, C0006, C0007, C0008, C0009, H0003, H0004, H0005, H0006, H0007, H0008, H0012, H0015, H0018, H0019, H0020, I0001, I0004, I0008, I0009, I0012, I0013, J0000, J0001, J0002, J0003, J0004, J0005, J0006, J0007, J0010, J0011, J0012, M0002) or (Z_LEVEL in (H0016, I0010) and PREREQUISITE ≠ 02))
0	Provision not in-scope for TEF	Otherwise

SNSOURCE = ILR

Value	Description	Definition
1	Provision in-scope for TEF	SNEXCL = 0 and OFSQAIM in (FIRST, ENHANCED, FDBC, OTHL6_Q, OTHL6_CC, OTHL6_U, HND, DET, FOUDEG, DIPHE, DTLLS, OTHL5_Q, OTHL5_CC, OTHL5_U, HNC, CERTED, UNICERT, HIGHCERT, CTLLS, PTLLS, CET, OTHL4_Q, OTHL4_CC, OTHL4_U, OTHHE_Q, OTHHE_CC, OTHHE_U)
0	Provision not in-scope for TEF	Otherwise

SNHEENGLEVSTULOAD_TEF

75. The total student load relevant to TEF for each engagement.[14]

---

[14] The ILR data does not include the concept of an 'engagement'. For the purpose of calculating SNHEENGLEVSTULOAD_TEF we treat a level of study in the ILR data as analogous to an engagement. This means that for each row of data in the individualised file relating to ILR data, SNHEENGLEVLSTULOAD_TEF shows the student load relevant to TEF for that student at that level of study, regardless of whether the given row of data contains student load relevant to TEF.

Value	Definition
SNHESTULOAD	SNSOURCE=DDB and SNREL_TEF=1
0	SNSOURCE=DDB and SNREL_TEF≠1
Sum of SNHESTULOAD across all provision relevant to TEF records (SNREL_TEF=1) for each unique combination of student (LEARNREFNUMBER) and level (SNLEVELNUM)	SNSOURCE=ILR

SNHEENGLEVSTULOADADJ_TEF

76. The capped student load relevant to TEF for each engagement.[15]

$$\text{SNHEENGLEVSTULOADADJ\_TEF} = \text{Min}(\text{SNHEENGLEVSTULOAD\_TEF}, 150).$$

SNTOTHESTULOAD_TEF

77. The total student load relevant to TEF per student, including any capping.

Value	Definition
Sum of SNHEENGLEVSTULOADADJ_TEF per student.	SNSOURCE=DDB
Sum of SNHEENGLEVSTULOADADJ_TEF across each unique combination of student and level.	SNSOURCE=ILR

SNTOTHEFTE_TEF

78. The total FTE relevant to TEF per student including any capping, calculated as SNTOTHESTULOAD_TEF÷100.

---

[15] The ILR data does not include the concept of an 'engagement'. For the purpose of calculating SNHEENGLEVSTULOADADJ_TEF we treat a level of study in the ILR data as analogous to an engagement.

Rebuild instructions

79. The individualised file, SN23_DCT _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.[16]

80. Specific instructions to rebuild figures in the student numbers data summary workbook, SN23_DCT_ XXXXXXXX.xlsx, are provided below.

81. A number of additional fields not described in the algorithms above have been added to the individualised file for ILR records. These are listed in Table 5.

Table 5: Additional fields in the individualised file

Field name	Description
STUDENTIDTYPE	The type of student identifier field for the record. This will be SID for records from the Student record and LEARNREFNUMBER for records from the ILR.
STUDENTID	The SID or LEARNREFNUMBER for the record.
REBUILDLEVFLAG	For ILR records only. A column provided to enable the figures in Table 1 and 2 to be rebuilt.
REBUILDHEFEFLAG	For ILR records only. A column provided to enable the figures in Table 3 to be rebuilt.

82. REBUILDLEVFLAG and REBUILDHEFEFLAG are provided to allow you to easily de-duplicate the data in the individualised file. REBUILDHEFEFLAG is provided to select one row of data per student. REBUILDLEVFLAG is provided to select one row of data per student per level of higher education study.

83. The rebuild flags 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.

Rebuilding the data in the summary file

84. The following instructions assume the user has the individualised file open in Microsoft Excel.

85. To apply filters to the column headings, highlight the headings. From the 'Data' tab in the toolbar select 'Filter'.

86. 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.

---

[16] See The OfS website – Working with individualised files.

Table 1 – Split of higher education by SNLEVELNUM

87. To rebuild the figures in Table 1 of the workbook, filter to REBUILDLEVFLAG = 1 if you are using ILR data, otherwise skip this step. Next, sum SNHEENGLEVSTULOADADJ and divide by 100 to produce an FTE. This will give the total FTE in Table 1. To rebuild the figures for each level of study, filter SNLEVELNUM as required.

Table 2 – Higher education provision relating to undergraduate levels of study relevant to TEF (SNREL_TEF)

88. To rebuild the figures in Table 2 of the workbook, first clear any filters. If you are using ILR data, filter to REBUILDLEVFLAG = 1. The subtotal for SNHEENGLEVSTULOADADJ_TEF, once divided by 100 to produce an FTE, will give the total FTE in Table 2.

Table 3 – Split of further and higher education provision (SNLEVEL)

89. For ILR records only. The fields SNTOTFEFTE and SNTOTHEFTE are calculated per student. Where a student has more than one learning aim, SNTOTFEFTE and SNTOTHEFTE will be the sum of the further education and higher education FTE respectively across all rows of data.

90. To rebuild the figures in table 3 of the workbook clear any previous filters and 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.