

**Office for  
Students**



# **2021-22 Student post-collection outputs**

**2022-23 Degree apprenticeship allocation  
data**

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

**Publication date** 9 March 2023

# Contents

<b>Summary</b>	<b>2</b>
Degree apprenticeship funding data summary workbook	2
Degree apprenticeship funding data summary individualised files	3
Student fields	3
ILR and LARS fields	4
<b>Description of derived fields</b>	<b>6</b>
Linking to other data sources	7
Derived fields	8
<b>Rebuild instructions</b>	<b>15</b>

# Summary

1. This document describes the degree apprenticeship funding data summary supplied as part of the 2021-22 Student post-collection outputs, and the algorithms applied to the 2021-22 Student record (C21051) and 2021-22 Individualised Learner Record (ILR) R14 data that are used to calculate the 2022-23 degree apprenticeship allocation. Throughout the document, fields taken or derived from the student record or ILR are shown in capitals.
2. The degree apprenticeship funding data summary should be used by staff responsible for individualised data returns at higher education providers. This document is intended to help them understand in detail how the data has been used to calculate their provider's 2022-23 degree apprenticeship allocation.
3. The degree apprenticeship funding data summary output is formed of three files. These are:
  - a. Summary workbook 'DA21\_XXXXXXXX.xlsx', where XXXXXXXX denotes the UK Provider Reference Number, UKPRN.
  - b. Individualised file (showing both 2021-22 Student record and 2021-22 ILR R14 data) 'DA21\_XXXXXXXX\_IND.csv'.
  - c. Individualised file (showing only 2021-22 ILR R14 data) 'DA21\_ILR\_XXXXXXXX\_IND.csv'.

Details of how to use individualised files are given on the Office for Students (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/)).

4. These files can be accessed from the '2021-22 Post-collection outputs' area of the OfS portal. Details will be emailed to the named student data contacts at providers when outputs become available.
5. This document is aimed at readers with in-depth knowledge of the data. Readers are advised to refer to the Student Record Coding Manual 2021-22 (<https://www.hesa.ac.uk/collection/c21051>) and the 'Specification of the Individualised Learner Record for 2021 to 2022' (available from the Education and Skills Funding Agency (ESFA) at <https://guidance.submit-learner-data.service.gov.uk/21-22/ilr/overview>) when using this document.
6. If you identify errors in your 2021-22 Student data, 2021-22 ILR R14 data or Higher Education Students Early Statistics 2022-23 (HESES22) return that you believe have a material effect on your 2022-23 degree apprenticeship allocation, an error summary may be submitted by following the guidance in [www.officeforstudents.org.uk/data-and-analysis/amendments-to-data/](http://www.officeforstudents.org.uk/data-and-analysis/amendments-to-data/).

## Degree apprenticeship funding data summary workbook

7. The degree apprenticeship funding data summary workbook is the Excel workbook 'DA21\_XXXXXXXX.xlsx' which contains the worksheets shown in Table 1.

**Table 1: Worksheets in the degree apprenticeship funding data summary workbook**

Worksheet	Title
Cover_sheet	This displays the provider's legal name, the provider's UKPRN, and the date on which the workbook was created.
Funding_data_summary	This shows how the allocation is calculated using headcounts and funding rates.

## Degree apprenticeship funding data summary individualised files

8. When working through this technical document it is necessary to use the individualised files, 'DA21\_XXXXXXXXX\_IND.csv' and 'DA21\_ILR\_XXXXXXXXX\_IND.csv'. The individualised files contain the values of all fields used to generate the degree apprenticeship funding data summary output.
9. Fields taken from the 2021-22 Student record, 2021-22 ILR R14 data, and those derived as part of the degree apprenticeship funding data summary are shown in capitals using the names given in Tables 2, 3 and 4 respectively.
10. Specific instructions for rebuilding the degree apprenticeship funding data summary workbook from the individualised files are provided in the 'Rebuild instructions' in this document.
11. The individualised file 'DA21\_XXXXXXXXX\_IND.csv' contains one record per year of instance (each unique combination of UKPRN, HUSID, NUMHUS and STUBID) in the 2021-22 academic year. Where no Student data exists but a link has been made to the ILR, the ILR data will be shown and Student record fields will be blank. The individualised file 'DA21\_ILR\_XXXXXXXXX\_IND.csv' contains all 2021-22 ILR R14 records.

## Student fields

12. The Student fields used to generate the degree apprenticeship funding data summary are listed in Table 2.

**Table 2: List of Student fields used**

Entity	Name	Description	Column in DA21_XXXXXXXXX_IND.csv
Course	COURSEAIM	General qualification aim of course	U
Course	COURSEID	Course identifier	V
Course	CTITLE	Course title	W
Instance	COMDATE	Start date of instance	T

Entity	Name	Description	Column in DA21_XXXXXXXXX_IND.csv
Instance	ENDDATE	End date of instance	X
Instance	FESTUMK	Further education student marker	Y
Instance	INITIATIVES1-3	Initiatives	Z, AA, AB
Instance	NUMHUS	Student instance identifier	C
Provider	UKPRN	UK provider reference number	A
Student	HUSID	HESA unique student identifier	B

## ILR and LARS fields

13. The ILR and Learning Aim Reference Service (LARS) fields used to generate the degree apprenticeship funding data summary are listed in Table 3.

**Table 3: List of ILR and LARS fields used**

Name	Description	Column in DA21_XXXXXXXXX_IND.csv	Column in DA21_ILR_XXXXX_XXX_IND.csv
AIMSEQNUMBER	Aim sequence number	AD (_PROG) AS (_COMP)	C
AIMTYPE	Aim type	N/A	J
FWORKCODE	Framework code	AG (_PROG) AW (_COMP)	K
LEARNACTENDDATE	Learning actual end date	AH (_PROG) AX (_COMP)	L
LEARNAIMREF	Learning aim reference number	AI (_PROG) AY (_COMP)	M
LEARNREFNUMBER	Learner reference number	AC (_PROG) AR (_COMP)	B
LEARNSTARTDATE	Learning start date	AK (_PROG) BA (_COMP)	O

Name	Description	Column in DA21_XXXXXXXX_I ND.csv	Column in DA21_ILR_XXXXX XXX_IND.csv
NOTIONALENDLEVEL	Level of apprenticeship standard	AL (_PROG) BB (_COMP)	P
ORIGLEARNSTARTDATE	Original learning start date	AM (_PROG) BC (_COMP)	Q
PROGTYPE	Programme type	AN (_PROG) BD (_COMP)	R
PWAYCODE	Pathway	AO (_PROG) BE (_COMP)	S
QUAL_TIT	Qualification title	AP (_PROG) BF (_COMP)	T
STDCODE	Apprenticeship standard code	AQ (_PROG) BG (_COMP)	U
UKPRN	UK provider reference number	A	A

## Description of derived fields

14. The derived fields used to generate the degree apprenticeship funding data summary are listed in Table 4.

**Table 4: List of derived fields used**

Name	Description	Paragraph	Column in DA21_XXXX XXXX_IND.csv	Column in DA21_ILR_XXXXXXXXX_IND.csv
ANNIV_COMP	Anniversary of start date where AIMTYPE = 3	26	AT	D
HESAPPASSUMEDL6	Apprenticeship is assumed to be at Level 6	31	E	N/A
HESAPPCOURSEAIM20	2020-21 value of COURSEAIM	25	F	N/A
HESAPPFESTUMK3	Assumed to be in the base population using Student data only	29	G	N/A
HESAPPHESABASEPOP	In the base population using Student data only	28	H	N/A
HESAPPHESAUGPOP	In Student record with aim of undergraduate degree	27	I	N/A
HESAPPILRDEG	Component aim is for an undergraduate degree	33	J	E
HESAPPILRDEGPOP	ILR component record is for an undergraduate degree aim on a Level 6 apprenticeship	33	K	F
HESAPPILRPROGPOP	Level 6 apprenticeship	30	L	G
HESAPPPPOP	Degree apprenticeship allocation population	34	M	N/A
HESAPPSTARTER	Starter in the degree apprenticeship population	35	N	N/A

Name	Description	Paragraph	Column in DA21_XXXX XXXX_IND.csv	Column in DA21_ILR_XXXXXXX _IND.csv
HESAPPUSEREC	ILR record has been kept for use in determining if the record in the Student return is part of the degree apprenticeship allocation population	23	N/A	H
HESCOMP <sup>†</sup>	Completion status	N/A	O	N/A
HESEXCL <sup>†</sup>	Reason(s) for exclusion from the HESES population	N/A	P	N/A
HESLEVEL <sup>†</sup>	Level of study	N/A	Q	N/A
HESMODE <sup>†</sup>	Mode of study	N/A	R	N/A
HESTYPE <sup>†</sup>	Fundability status	N/A	S	N/A
OFSQAIM <sup>‡</sup>	Broad category of higher education	N/A	AE (_PROG) I AU (_COMP)	
STUBID <sup>†</sup>	Unique countable year of instance identifier	N/A	D	N/A

<sup>†</sup> The algorithms for deriving these fields are given in the 'HESES21 comparison technical document' ([www.officeforstudents.org.uk/data-and-analysis/post-collection-outputs/](http://www.officeforstudents.org.uk/data-and-analysis/post-collection-outputs/))

<sup>‡</sup> For a full definition of this field, please refer to '2021-22 ILR data checking tool: Classifying learning aims technical document' ([www.officeforstudents.org.uk/data-and-analysis/data-checking-tools/2021-22-ilr-data-checking-tool/](http://www.officeforstudents.org.uk/data-and-analysis/data-checking-tools/2021-22-ilr-data-checking-tool/)).

## Linking to other data sources

- As the level of a particular apprenticeship standard or framework cannot be determined from the Student record, we have linked to 2021-22 ILR R14 data for your provider to obtain this information. We link by combinations of first name(s), surname, date of birth, gender and (where available) home postcode and prior educational establishment. Spelling errors and other typographical errors (e.g. in dates) are taken into account.
- The ILR data we link to is filtered to programme aims for apprentices only (AIMTYPE = 1 and PROGTYPE = 2, 3, 20, 21, 22, 23, 25) and is further filtered to remove records where:
  - LEARNACTENDDATE is before 1 August 2021, or



- LEARNSTARTDATE is on or after 1 August 2022.

Where more than one record for a person exists in the ILR data after we apply these filters, we keep the record where the activity for the aim in the Student record overlaps with the activity for the programme aim in the ILR (i.e. COMDATE in the Student record is not after LEARNACTENDDATE in the ILR, and ENDDATE in the Student record is not before the LEARNSTARTDATE in the ILR). If more than one ILR programme record overlaps with the Student record, then we keep the record with the most recent start date for the programme.

17. The level (NOTIONALENDLEVEL) of a particular apprenticeship standard (STDCODE) is determined from the LARS database. The level of a particular apprenticeship framework is determined by the PROGTYPE recorded in the ILR.
18. In some cases (for example, where an ILR record is found, but there is no corresponding Student record) we need to instead determine if there is a record in the ILR for an apprentice who is aiming for a degree (AIMTYPE = 3 and PROGTYPE = 2, 3, 20, 21, 22, 23, 25 and OFSQAIM\_COMP = FIRST). The further filters and de-duplication process described in paragraph 16 are also carried out on this data.
19. Where a Student record is found, but no corresponding ILR records are found, we look at the ILR records for other records in the Student record with the same value of UKPRN x HUSID and use this to make an assumption on whether that record is for a Level 6 apprenticeship.

## Derived fields

### ILR fields

20. Where AIMTYPE = 1, an ILR field will be shown in the individualised file 'DA21\_XXXXXXXXX\_IND.csv' with the suffix '\_PROG'. Where AIMTYPE = 3, an ILR field will be shown with the suffix '\_COMP'. For example, LEARNSTARTDATE for an aim with AIMTYPE = 1 will be shown as LEARNSTARTDATE\_PROG, whilst LEARNSTARTDATE for an aim with AIMTYPE = 3 will be shown as LEARNSTARTDATE\_COMP.
21. Where multiple records with AIMTYPE = 1 exist, only one record will be displayed in the individualised file 'DA21\_XXXXXXXXX\_IND.csv', and the one displayed will be that found via the method described in paragraph 16. Similarly, where multiple records exist with AIMTYPE = 3, only one record will be displayed in the individualised file, and the one displayed will be that found via the method described in paragraphs 16 and 18.
22. The individualised file 'DA21\_ILR\_XXXXXXXXX\_IND.csv', ILR fields will be shown without the '\_PROG' and '\_COMP' suffixes. HESAPPUSEREC (paragraph 23) will indicate if the aim will be displayed in the individualised file 'DA21\_XXXXXXXXX\_IND.csv'.

### HESAPPUSEREC

23. This field shows if the ILR record has been kept (following the linking method described in paragraphs 15 to 19) for use in determining if the record in the Student return is part of the degree apprenticeship allocation population. This field is only displayed in the individualised file 'DA21\_ILR\_XXXXXXXXX\_IND.csv'.

Value	Description	Definition
PROG	This programme aim will be used in determining if the record in the Student return is part of the degree apprenticeship allocation population	AIMTYPE = 1 and record is kept following the linking method described in paragraphs 15 to 19
COMP	This component aim will be used in determining if the record in the Student return is part of the degree apprenticeship allocation population	AIMTYPE = 3 and record is kept following the linking method described in paragraphs 15 to 19
NO	This aim will not be used in determining if the record in the Student return is part of the degree apprenticeship allocation population	Otherwise

### **OFSQAIM\_COMP**

24. This field shows the value of OFSQAIM for the record with AIMTYPE = 3.

### **HESAPPCOURSEAIM20**

25. We have linked the 2021-22 Student data to 2020-21 Student data using UKPRN, HUSID and NUMHUS. This field shows the value of COURSEAIM as found in 2020-21 using this linking method.

### **ANNIV\_COMP**

26. This field contains the anniversary of the start date for the aim with AIMTYPE = 3 (LEARNSTARTDATE\_COMP) during the current academic year. Where ORIGLEARNSTARTDATE\_COMP exists, this is used instead of LEARNSTARTDATE\_COMP.

### **HESAPPHESAUGPOP**

27. This field indicates if the Student record is for a home fee undergraduate student with a full-time or part-time mode in the HESSES population with an aim of an undergraduate degree.

Value	Description	Definition
1	In the undergraduate population (using Student data only)	HESEXCL = 0 and HESMODE ≠ SWOUT and HESTYPE ≠ ISOV and COURSEAIM = H00, H11, H12, H16, H18, H22, H23, I00, I11, I12, I16 and HESLEVEL = UG
0	Otherwise	Otherwise

### HESAPPHESABASEPOP

28. This field indicates if the Student record is in the undergraduate population and is part of a higher apprenticeship (i.e. is part of the 'base population').

Value	Description	Definition
1	In the base population (using Student data only)	HESAPPHESAUGPOP = 1 and (INITIATIVES1 = K or INITIATIVES2 = K or INITIATIVES3 = K)
0	Otherwise	Otherwise

### HESAPPFESTUMK3

29. This field indicates where INITIATIVES has not been completed to indicate the student is an Apprentice but we could assume HESAPPHESABASEPOP = 1 because the further education student marker records the student as an ESFA-funded learner.

Value	Description	Definition
1	Assumed to be in the base population (using Student data only)	HESAPPHESAUGPOP = 1 and HESAPPHESABASEPOP = 0 and FESTUMK = 3
0	Otherwise	Otherwise

### HESAPPILRPROGPOP

30. This field indicates if the ILR data indicates the apprenticeship is at Level 6.

Value	Description	Definition
1	The apprenticeship is at Level 6	NOTIONALENDLEVEL_PROG = 6 or PROGTYPE_PROG = 22
0	Otherwise	Otherwise

### HESAPPASSUMEDL6

31. This field indicates if no record exists in the ILR for this student and that we have assumed the apprenticeship is at Level 6 based on the ILR records for other students with the same UKPRN and COURSEID. Where no records exist in the ILR data for this combination of UKPRN and COURSEID, the apprenticeship is assumed to be at Level 6.

Value	Description	Definition
1	Apprenticeship is assumed to be at Level 6	LEARNREFNUMBER_PROG = BLANK and (at least 50 per cent of records where LEARNREFNUMBER_PROG ≠ BLANK with the same UKPRN x COURSEID have HESAPPILRPROGPOP = 1 or no records are found in the ILR for this combination of UKPRN x COURSEID)
0	Otherwise	Otherwise

### HESAPPILRDEG

32. This field indicates if the component aim shown in the individualised file is for an undergraduate degree.

Value	Description	Definition
1	Component aim is for an undergraduate degree	OFSQAIM_COMP = FIRST
0	Otherwise	Otherwise

### HESAPPILRDEGPOP

33. This field indicates if the ILR component record is for an undergraduate degree aim on a Level 6 apprenticeship that is assumed to be in the HESES population.

Value	Description	Definition
0	Not an undergraduate degree aim as part of a Level 6 apprenticeship, or not assumed to be part of the HESES population	HESAPPILRDEG = 0 or HESAPPILRPROGPOP = 0 or (LEARNACTENDDATE_COMP ≠ BLANK and (LEARNACTENDDATE_COMP < ANNIV_COMP + 14 or LEARNACTENDDATE_COMP < 2 November 2021 or LEARNACTENDDATE_COMP – LEARNSTARTDATE_COMP < 14))
1	Otherwise	Otherwise

### HESAPPPOP

34. This field indicates if the student is in the degree apprenticeship allocation population. A student is in the population if one of the following is true:

- The student is in the base population using Student data only (HESAPPHESABASEPOP = 1), or would be in this base population if INITIATIVES = K (HESAPPFESTUMK3 = 1), and:
  - There is a Level 6 apprenticeship programme aim in LARS (HESAPPILRPROGPOP = 1), or
  - We have assumed there is a Level 6 apprenticeship programme aim where no link is found to the ILR (HESAPPASSUMEDL6 = 1).
- The student is not in the base population using Student data only (HESAPPHESABASEPOP = 0, BLANK), but a degree aim is found on the ILR that is assumed to be part of the HESES population (HESAPPILRDEGPOP = 1), and is part of a Level 6 apprenticeship (HESAPPILRPROGPOP = 1) and:
  - No record exists in the Student record (HUSID = BLANK), or
  - The COURSEAIM in the Student record is not an undergraduate degree and we have not already found another Student record that would count for this HUSID (the HUSID has no other record where HESAPPHESABASEPOP = 1 or HESFESTUMK3 = 1).
- The student is not in the base population using Student data only (HESAPPHESABASEPOP = 0), but is in the undergraduate population (HESAPPHESAUGPOP = 1) and:
  - A degree aim is found on the ILR that is assumed to be part of the HESES population (HESAPPILRDEGPOP = 1), and is part of a Level 6 apprenticeship (HESAPPILRPROGPOP = 1).

Value	Description	Definition
1	In the degree apprenticeship population	<p>((HESAPPHESABASEPOP = 1 or  HESAPPFESTUMK3 = 1) and  (HESAPPILRPROGPOP = 1 or  HESAPPASSUMEDL6 = 1)) or</p> <p>(HESAPPHESABASEPOP = 0, BLANK and  (HUSID = BLANK or  (COURSEAIM ≠ H00, H11, H12, H16, H18, H22, H23, I00, I11, I12,  I16 and there is not another record for this HUSID with  HESAPPHESABASEPOP = 1 or HESAPPFESTUMK3 = 1)) and  HESAPPILRDEGPOP = 1 and  HESAPPILRPROGPOP = 1) or</p> <p>(HESAPPHESABASEPOP = 0 and  HESAPPHESAUGPOP = 1 and  HESAPPILRDEGPOP = 1 and  HESAPPILRPROGPOP = 1)</p>
0	Otherwise	Otherwise

## HESAPPSTARTER

35. This field indicates if the student is part of the degree apprenticeship allocation population and classed as a starter.

Value	Description	Definition
1	Starter in the degree apprenticeship population	<p>HESAPPPOP = 1 and</p> <p>((HUSID ≠ BLANK and ((COMDATE ≥ 1 August 2021 and COMDATE &lt; 1 August 2022) or (COMDATE &lt; 1 August 2021 and HESAPPCOURSEAIM20 ≠ H00, H11, H12, H16, H18, H22, H23, H24, I00, I11, I12, I16, BLANK))))</p> <p>or</p> <p>(HUSID = BLANK and LEARNSTARTDATE_COMP ≥ 1 August 2021 and LEARNSTARTDATE_COMP &lt; 1 August 2022 and (ORIGLEARNSTARTDATE_COMP = BLANK or ORIGLEARNSTARTDATE_COMP ≥ 1 August 2021)))</p>
0	Otherwise	Otherwise

## Rebuild instructions

36. The individualised files, DA21\_XXXXXXXX\_IND.csv and DA21\_ILR\_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 individualised files 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/)).
37. Specific instructions to rebuild the OfS-fundable starters on degree apprenticeships (2021-22 individualised data) figure in the workbook, DA21\_XXXXXXXX.xlsx, are provided here. These instructions apply to the DA21\_XXXXXXXX\_IND.csv individualised file only – this figure cannot be rebuilt from the DA21\_ILR\_XXXXXXXX\_IND.csv individualised file.

Item in workbook	Derived field selection in individualised file
OfS-fundable starters on degree apprenticeships (2021-22 individualised data)	HESAPPSTARTER = 1 and ((HUSID ≠ BLANK and HESATYPE = HOMEF and HESCOMP = 4) or HUSID = BLANK)





© The Office for Students copyright 2023

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