2019-20 ILR data checking tool: Higher education level apprenticeship data summary

Technical algorithms and rebuild instructions

Contents

Summary

1. This document describes the higher education (HE) level apprenticeship data summary files generated by the 2019-20 Individualised Learner Record (ILR) data checking tool and the algorithms used to generate the data summary tables. Throughout the document, fields taken or derived from the ILR are shown in capitals.

2. The Office for Students (OfS) recognises that the reporting requirements for apprenticeship data differ from those of other higher education programmes. The HE-level apprenticeship data summary is intended to promote a shared understanding of how the OfS will interpret such data for the purpose of identifying issues that may affect the OfS's uses of the data. The summary uses fields taken from the ILR to categorise students that started an HE-level apprenticeship in 2019-20 by student characteristics, subcontractual partner, study region, subject and student outcome. It provides separate information on student apprenticeship programme aim and component aim starts.

3. The HE-level apprenticeship data summary is formed of two files. These are:

   a. The HE-level apprenticeship data summary workbook 'APPR19_DCT_XXXXXXXX.xlsx' (where XXXXXXXX is the UK Provider Reference Number (UKPRN) for the provider).

   b. The HE-level apprenticeships individualised file 'APPR19_DCT_XXXXXXXX_IND.csv'.

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

4. These files can be accessed from the '2019-20 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 made available.

5. This document lists the fields used to build the data summary tables, which are either taken directly from the ILR return or derived from them. The majority of the derived fields used within this output are also used in the construction of OfS institutional performance measures[1] and are prefixed with 'IP'. The algorithms used to derive these fields are documented in the institutional performance measures core algorithms document (available on the OfS website via www.officeforstudents.org.uk/data-and-analysis/institutional-performance-measures/technical-documentation/). The remaining derived fields that do not feature within the core algorithms document are described here.

6. This document also describes how the data summary tables can be rebuilt from the individualised file provided. This document is aimed at readers with in-depth knowledge of the data. Readers are advised to have a copy of 'Specification of the Individualised Learner Record for 2019 to 2020' (available from the Education and Skills Funding Agency via https://www.gov.uk/government/publications/ilr-specification-validation-rules-and-appendices-2019-to-2020) to hand when using this document.

HE-level apprenticeship data summary workbook

7. The HE-level apprenticeship data summary workbook is the Excel workbook APPR19_DCT_XXXXXXXX.xlsx, where XXXXXXXX is the UK Provider Reference Number (UKPRN) for the provider. The workbook contains the following worksheets:

Table 1: HE-level apprenticeship data summary tables created using 2019-20 ILR data

Worksheet*	Title
Coversheet	Title page and headcount of withdrawn students with withdrawal reason 'other'
Student characteristics	Student headcount by student characteristic at component level and programme level
Provision characteristics	Student headcount by subcontractual partner
Study regions	Student headcount by regional location of study
Subjects	Student headcount by subject at component level and programme level
Outcomes	Student headcount by outcome and withdrawal reason
Employers	Number of employers by student headcount

* This worksheet reference corresponds to the spreadsheet tabs.

Using the individualised file

8. When working through this document it is necessary to use the individualised file, APPR19_DCT_XXXXXXXX_IND.csv, where XXXXXXXX is the UKPRN for the provider. The individualised file contains the values of all fields used to generate the data summary tables. Fields taken from the ILR return or derived as part of the data summary tables are shown in capitals using the names given in Tables 2 and 3 respectively. Fields prefixed with 'IP', 'OFS', or 'APPR' are derived, all others are taken directly from the ILR or the Learning Aim Reference Service (LARS) as indicated in Table 2.

9. Specific instructions for rebuilding the HE-level apprenticeship data summary workbook from the individualised file are provided in this document.

10. The individualised file contains one record per subject area per learning aim per learner in the latest academic year. For example, a student who is studying for a first degree in biology in 2019-20 will have one record for that instance and will count as one full-person equivalent (FPE = 100; a headcount measure). A student who is studying a joint course first degree with equal proportions of mathematics and physics in 2019-20 will have two records for that instance; one for each subject area, with each assigned an FPE of 50 (their single headcount for the instance being apportioned across the subject areas according to the proportion of the course that relates to each subject they are studying).

ILR fields used to create the HE-level apprenticeship tables

11. The ILR fields used to generate the data summary are listed in Table 2.

Table 2: List of ILR Fields used in the data summary

Name	Description	Dataset	Column in individualised file
UKPRN†	UK provider reference number	ILR	A
LEARNREFNUMBER†	Learner reference number	ILR	B
AIMSEQNUMBER†	Learning aim data set sequence	ILR	C
ULN†	Unique learner number	ILR	D
NUMHUS†	Student instance identifier	ILR	E
COMPSTATUS	Completion status	ILR	I
DATEOFBIRTH	Date of birth	ILR	J
DELLOCPOSTCODE	Learning delivery location postcode	ILR	K
EMPID	Employer identifier	ILR	L
ETHNICITY	Ethnicity	ILR	N
FWORKCODE	Framework code	ILR/LARS	O
HEPOSTCODE	Higher education centre location postcode	ILR	P
LEARNAIMREF	Learning aim reference	ILR	Q
LEARNPLANENDDATE	Learning end date	ILR	R
LEARNSTARTDATE	Learning start date	ILR	S
LLDDCAT‡	Category of student's learning difficulty, disability, or health problem	ILR	T
LLDDHEALTHPROB	Indicates whether the student has a learning difficulty, disability, or health problem	ILR	U
PARTNERUKPRN	Subcontracted UKPRN	ILR	V
PROGTYPE	Programme type	ILR	W
QUAL_TIT	Learning aim title	ILR	X
SEX	Sex of student	ILR	Y
STDCODE	Apprenticeship standard code	ILR/LARS	Z
WITHDRAWREASON	Withdrawal reason	ILR	AA

Name	Description	Dataset	Column in individualised file
NOTIONALENDLEVEL	Level of NVQ	LARS	AB
SECTORSUBJECTAREA TIER1DESC	Tier 1 description of learning aim	LARS	AC
SECTORSUBJECTAREA TIER2DESC	Tier 2 description of learning aim	LARS	AD
STANDARDNAME	Apprenticeship standard name	LARS	AE

† These fields are not used in the calculations but are included in the individualised file to allow easy identification of students.

‡ Where the student has multiple types of learning difficulty, disability or health problem, the value of LLDDCAT with an associated value of PRIMARYLLDD = 1 is given in the individualised file.

Derived fields used to create the HE-level apprenticeship tables

12. Here we give details of the derived fields in the individualised file. These fields are used to build the key dimensions of the HE-level apprenticeship data summary tables.

13. All fields with the prefix 'IP' are derived using the algorithms documented in the institutional performance measures core algorithms document available on the OfS website (www.officeforstudents.org.uk/data-and-analysis/institutional-performance-measures/technical-documentation/).

Table 3: List of derived fields used in the data summary

Derived field name	Description	Column in individualised file
APPRAIM	Identifies programme aims from component aims	H
APPRDISABLE	Disability	AO
APPRNVQ_LEVEL	Level of NVQ	AP
APPROUTCOME	Completion status	AQ
APPRPOP	Indicates whether the student is included in the apprenticeship population	F
APPRREGION	Regional location of study	AR
IPAGEBAND	Age category	AF
IPAPPRENTICE	Apprenticeship	AG
IPDISABLETYPE	Disability type	AH

Derived field name	Description	Column in individualised file
IPENTRANT	Indicates whether the student is an entrant	G
IPETHNIC	Ethnicity	AI
IPETHNICDETAIL	Detailed ethnicity	AJ
IPFPE	Full-person equivalent	AK
IPLOCPOSTCODE	Location of study postcode	AL
IPYEAR_ST	Academic start year	AM
OFSQAIM	Recognised higher education qualification aim	AN

APPRAIM

14. This field has been provided to enable filtering of the data to programme or component aims.

Value	Description	Definition
PROGRAMME	Programme aim	LEARNAIMREF = ZPROG001
COMPONENT	Component aim	Otherwise

APPRDISABLE

15. This field indicates whether the student has a disability.

Value	Description	Definition
Y	The student is disabled	LLDDHEALTHPROB = 1 and IPDISABLETYPE ≠ NONE
N	The student is not disabled	LLDDHEALTHPROB = 2 and IPDISABLETYPE = NONE
UNKNOWN	Unknown	Otherwise

Note: This field has a similar definition to the field IPDISABLE described in the institutional performance measures core algorithms document. Where IPDISABLE can take only the values Y or N, APPRDISABLE can take an additional value, UNKNOWN. If IPDISABLE would show Y but IPDISABLETYPE = NONE, APPRDISABLE is set to UNKNOWN. Similarly, if IPDISABLE would show N but IPDISABLETYPE ≠ NONE or LLDDHEALTHPROB ≠ 2, APPRDISABLE is set to UNKNOWN.

APPRNVQ_LEVEL

16. This field indicates the notional NVQ level.

Value	Description	Definition
4	Level 4	(PROGTYPE = 25 and NOTIONALENDLEVEL = 4) or PROGTYPE = 20
5	Level 5	(PROGTYPE = 25 and NOTIONALENDLEVEL = 5) or PROGTYPE = 21
6	Level 6	(PROGTYPE = 25 and NOTIONALENDLEVEL = 6) or PROGTYPE = 22
7	Level 7	(PROGTYPE = 25 and NOTIONALENDLEVEL = 7) or PROGTYPE = 23
BLANK	Otherwise	Otherwise

APPROUTCOME

17. This field indicates the continuation outcome for the learning aim.

Value	Description	Definition
Completed	Student has achieved the learning aim	COMPSTATUS = 2
Temporarily withdrawn	Student has temporarily withdrawn from the learning aim	COMPSTATUS = 6
Withdrawn	Student has withdrawn from the learning aim	COMPSTATUS in (3, 4)
Continuing	Student is continuing the learning aim	Otherwise

APPRPOP

18. This field has been provided to enable filtering of the data to the apprenticeships population.

Value	Description	Definition
1	Student is counted in the apprenticeships population	APPRNVQ_LEVEL in (4, 5, 6, 7) and IPAPPRENTICE = 1 and (APPRAIM = PROGRAMME or (APPRAIM = COMPONENT and OFSQAIM ≠ FE))
0	Otherwise	Otherwise

APPRREGION

19. This field assigns the location of study postcode (IPLOCPOSTCODE) to broader regions of study defined by Nomenclature of Territorial Units for Statistics level 2 (NUTS level 2). This mapping is based on National Statistics Postcode Lookup (NSPL) data from May 2020. The NSPL lists all postcodes in the UK and links them to higher geographies. For more information see www.ons.gov.uk/methodology/geography/geographicalproducts/postcodeproducts/.

OFSQAIM

20. This field allocates qualification aims to broad categories of higher education. For a full definition of this field please refer to '2019-20 ILR data checking tool: Classifying learning aims technical document' (available via: www.officeforstudents.org.uk/data-and-analysis/data-checking-tools/2019-20-ilr-data-checking-tool/).

HE-level apprenticeship data summary: rebuild instructions

21. The individualised file, APPR19_DCT_XXXXXXXX_IND.csv, can be used to interrogate the data and our calculation of derived fields by filtering to certain groups of records. Specific instructions to rebuild the HE-level apprenticeship data summary workbook, APPR19_DCT_XXXXXXX.xlsx, are provided here.

22. The following sections of this document detail how to rebuild the populations of students found in the sheets of the data summary workbook. Each sheet either includes information which needs to be filtered from the individualised file by programme aim and/or component aim. This is summarised in the table below.

Table 4: HE-level apprenticeship data summary tables created using 2019-20 ILR data

Worksheet	Description	Programme and/or component aims
Coversheet	Title page and headcount of withdrawn students with withdrawal reason 'other'	Component
Student characteristics	Student headcount by student characteristic at component level and programme level	Programme and component
Provision characteristics	Student headcount by subcontractual partner	Component
Study regions	Student headcount by regional location of study	Component
Subjects	Student headcount by subject at component level and programme level	Programme and component
Outcomes	Student headcount by outcome and withdrawal reason	Component
Employers	Number of employers by student headcount	Component

23. To rebuild the headcounts in the workbook summary tables, apply the following filters to the individualised files:

   a. Restrict the data to the HE-level apprenticeship population by filtering to APPRPOP = 1.

   b. Restrict the data to only the starts population by applying the filter: IPENTRANT = 1.

   c. Restrict the data to programme or component level as required by filtering on APPRAIM to 'PROGRAMME' or 'COMPONENT', respectively.

   d. The additional filters which should be applied to the individualised file to create the population for each specific row and column of the summary tables are explained on the following pages.

   e. Once all the required filters have been applied the headcount is returned by summing the values in the IPFPE column and dividing by 100.

Coversheet

24. The coversheet includes a summary of the number of component aim starts for which the student has withdrawn from the learning aim for reason 'other'. The counts can be recalculated using the instructions in paragraph 23 and filtering the data according to the table below.

Column value	Filter
Student has withdrawn	APPROUTCOME = Withdrawn
Student has withdrawn for reason 'other'	APPROUTCOME = Withdrawn and WITHDRAWREASON = 97

Student characteristics

25. This sheet shows the number of apprenticeship starts categorised by various student characteristics per notional NVQ level. The possible options that are displayed on this sheet are listed in the tables below. The number of apprenticeship programme aim or component aim starts can be recalculated using the instructions in paragraph 23.

Notional NVQ level

Column value	Filter
Level 4	APPRNVQ_LEVEL = 4
Level 5	APPRNVQ_LEVEL = 5
Level 6	APPRNVQ_LEVEL = 6
Level 7	APPRNVQ_LEVEL = 7

Sex

Row value	Filter
Male	SEX = M
Female	SEX = F

Age

Row value	Filter
Under 21	IPAGEBAND = U21
21 to 25	IPAGEBAND = 21_25
26 to 30	IPAGEBAND = 26_30
31 to 40	IPAGEBAND = 31_40
41 to 50	IPAGEBAND = 41_50
51 and over	IPAGEBAND = 51+

Ethnicity

Row value	Filter
Asian	IPETHNIC = A
Black	IPETHNIC = B
Mixed	IPETHNIC = M
Other	IPETHNIC = O
White	IPETHNIC = W
Unknown	IPETHNIC = UNKNOWN

Disability

Row value	Filter
Disabled	APPRDISABLE = Y
Not disabled	APPRDISABLE = N
Unknown	APPRDISABLE = UNKNOWN

Provision characteristics

26. Where a HE-level apprenticeship student has been recorded with a subcontractual partner (PARTNERUKPRN), this sheet summarises the number of component aim starts per subcontractual partner and can be recalculated by filtering to a PARTNERUKPRN then using the instructions in paragraph 23.

Study regions

27. This sheet summarises the number of component aim starts by regional location of study (APPRREGION) and can be recalculated by filtering to an APPRREGION then using the instructions in paragraph 23.

Subjects

28. Where an HE-level apprenticeship student has been recorded on either an apprenticeship framework (via FWORKCODE) or an apprenticeship standard (via STDCODE) on the ILR, this has been linked to the LARS to retrieve information on the standard name, notional NVQ level, subject area tier 1 and subject area tier 2, which are populated on this sheet. These fields are all available within the individualised file and are labelled as shown in the table below. Note that apprenticeship frameworks do not have a standard name and are populated with 'Framework'.

Column value	Variable in the individualised file
Standard code	STDCODE (for apprenticeship standards) or FWORKCODE (for apprenticeship frameworks)
Standard name	STANDARDNAME
Notional NVQ level	APPRNVQ_LEVEL
Subject area tier 1	SECTORSUBJECTAREATIER1DESC
Subject area tier 2	SECTORSUBJECTAREATIER2DESC

29. The number of programme aim starts on the apprenticeship framework or apprenticeship standard are shown and can be recalculated using the instructions in paragraph 23.

30. Students returned with either an apprenticeship framework or an apprenticeship standard are linked to their component aims present in the 2019-20 ILR data and the qualification title (variable QUAL_TIT in the individualised file). There may be multiple qualification titles per framework or standard.

31. The number of component aim starts on the apprenticeship framework or apprenticeship standard are shown and can be recalculated using the instructions in paragraph 23.

Outcomes

32. This sheet summarises the completion status, as recorded on ILR via COMPSTATUS and WITHDRAWREASON, for component aim starts per notional NVQ level. The first table on this sheet summarises the number of component aim starts by student outcome (APPROUTCOME) and can be recalculated by filtering to an APPROUTCOME then using the instructions in paragraph 23.

33. The second table on this sheet shows the number of component aim starts for which the student has withdrawn categorised by the withdrawal reason (WITHDRAWREASON). The reason for withdrawal is populated using the definition of WITHDRAWREASON in the document 'Specification of the Individualised Learner Record for 2019 to 2020' (available from the Education and Skills Funding Agency via https://www.gov.uk/government/publications/ilr-specification-validation-rules-and-appendices-2019-to-2020). The number of component aims starts can be recalculated by filtering to APPROUTCOME = 'Withdrawn' and a WITHDRAWREASON using the instructions in paragraph 23.

Notional NVQ level

Column value	Filter
Level 4	APPRNVQ_LEVEL = 4
Level 5	APPRNVQ_LEVEL = 5
Level 6	APPRNVQ_LEVEL = 6
Level 7	APPRNVQ_LEVEL = 7

Student outcome

Outcome	Filter
Completed	APPROUTCOME = Completed
Continuing	APPROUTCOME = Continuing
Temporarily withdrawn	APPROUTCOME = Temporarily withdrawn
Withdrawn	APPROUTCOME = Withdrawn

Employers

34. Where a HE-level apprenticeship student has been recorded with an employer identifier (EMPID), this sheet summarises the number of employers by:

    a. The overall number of employers.

    b. Employers with 10 or more component aim starts.

    c. Employers with between two and nine (inclusive) component aim starts.

    d. Employers with one component aim starts.

OGL