

Office for
Students



2020-21 Individualised Learner Record (ILR) data summary

**2020-21 Learner characteristics
technical document and rebuild instructions**

Enquiries to ILRData@officeforstudents.org.uk

Publication date 11 August 2021

Contents

Summary	2
Learner characteristics workbook	3
Learner characteristics individualised file	3
ILR fields used to create the learner characteristic tables	4
Description of derived fields	6
Rebuild instructions	10
Learner characteristics	10
Taught location	14
Teaching provider	14

Summary

1. This document describes the learner characteristics (LC) data summary files generated by the 2020-21 Individualised Learner Record (ILR) data summary and the algorithms used to generate the data summary tables and rebuild instructions. Throughout the document, fields taken or derived from the ILR are shown in capitals.
2. The learner characteristics data summary can be used by higher education providers to verify and correct their 2020-21 individualised student data before submitting their data to the Education and Skills Funding Agency.
3. The learner characteristics data output is formed of two files. These are:
 - a. The learner characteristics data summary workbook 'LC20_AAA_XXXXXXXX.xlsx'.
 - b. The learner characteristics individualised file 'LC20_AAA_XXXXXXXX_IND.csv'.

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

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/.

Table 1: File name abbreviations used in learner characteristics outputs

File abbreviation	Source of outputs
DCT	Data checking tool
AST	Amendment submission tool

4. These files can be accessed from the '2020-21 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 is aimed at readers with in-depth knowledge of the data. Readers are advised to refer to the 'Specification of the Individualised Learner Record for 2020 to 2021' (available from the Education and Skills Funding Agency via <https://guidance.submitlearnerdatabeta.fasst.org.uk/previous/20-21/ilr>).
6. We will document any changes or corrections to the algorithms in this document on the OfS website: www.officeforstudents.org.uk/data-and-analysis/data-checking-tool/2020-21-ilr-data-checking-tool/ under 'Updates'.
7. 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.
8. 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/).

9. All fields with the prefix 'LC' are derived specifically for this data summary and are written in this document.

Learner characteristics workbook

The Learner characteristics workbook is the Excel workbook 'LC20_AAA_XXXXXXXX.xlsx', which contains the following worksheets:

Table 2: Worksheets in the learner characteristics workbook

Worksheet*	Title
Coversheet	Title page
Learner characteristics	Number of learners by characteristic
Taught location	Number of learners by regional location of study
Teaching provider	Number of learners by subcontractual partner

* This worksheet reference corresponds to the spreadsheet tabs.

Learner characteristics individualised file

10. When working through this document it is necessary to use the individualised file, LC20_AAA_XXXXXXXX_IND.csv. The individualised file contains the values of all the fields used to generate the data summary tables.
11. 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 3 and 4 respectively. Fields prefixed with 'IP' or 'LC' are derived, all others are taken directly from the ILR or the Learning Aim Reference Service (LARS) with its search engine Find a Learning Aim (FaLA) as indicated in Table 3.
12. Specific instructions for rebuilding the learner characteristics workbook from the individualised file are provided in this document.
13. The fields we use from the ILR return are associated with a course, identified by its learning aim reference (LEARNAIMREF); a learner may study multiple courses. We will use the generic term 'learner' in this technical document and the learner characteristics data summary workbook to refer to one row of data in the individualised file which corresponds to a learner-course combination.

ILR fields used to create the learner characteristic tables

14. The ILR fields used to generate the data summary are listed in Table 3.

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

Name	Description	Dataset
UKPRN†	UK provider reference number	ILR
LEARNREFNUMBER†	Learner reference number	ILR
AIMSEQNUMBER†	Learning aim data set sequence	ILR
ULN†	Unique learner number	ILR
NUMHUS†	Learner instance identifier	ILR
DATEOFBIRTH	Date of birth	ILR
DELLOCPOSTCODE	Learning delivery location postcode	ILR
DOMICILE	The country code of the learner's permanent or home address prior to entry to the course	ILR
ETHNICITY	The ethnicity of the learner	ILR
FUNDCOMP	The learner's completion status with respect to that year of programme of study being funded	ILR
FUNDLEV	The level of the programme of study, expressed in terms of OfS HESES definitions	ILR
HEPOSTCODE	Higher education centre location postcode	ILR
LEARNAIMREF	Learning aim reference	ILR
LEARNAIMREFTITLE	Learning aim title	ILR
LEARNFAM_DLA	To indicate if the learner is in receipt of disabled students' allowance	ILR
LEARNPLANENDDATE	Learning end date	ILR
LEARNSTARTDATE	Learning start date	ILR
LLDDCAT‡	The most significant or primary learning difficulty, disability or health problem that impacts on the learner's education	ILR
LLDDHEALTHPROB	Indicates whether the learner has a learning difficulty, disability or health problem	ILR

Name	Description	Dataset
MODESTUD	The mode of study expressed in terms of the OfS Higher Education Students Early Statistics survey (HESES) definitions.	ILR
MSTUFEE	The major source of tuition fees for the learner	ILR
PARTNERUKPRN	Subcontracted UKPRN	ILR
PCFLDCS, PCSLDCS, PCTLDCS	Proportion taught in LDCS_CO1-CO3 subject	ILR
POSTCODEPRIOR	Postcode of the address of the learner prior to enrolment.	ILR
QUALENT3	Qualification on entry	ILR
SEX	Sex of learner	ILR
SPECFEE	Records any special or non-standard fees	ILR
TYPEYR	Identifies which of the four basic types of instance year is applicable to the learner	ILR
YEARSTU	The year number that the learner is in since enrolling for a course leading to the learner's qualification aim (whether or not the intended subject or class has changed)	ILR
NOTIONALNVQLEVEL2	Level of NVQ	LARS
LEARNAIMREFTYPE	Learning aim type	LARS
LDCS_CO1, LDCS_CO2, LDCS_CO3	Learning directory classification system codes	LARS

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

‡ Where the learner 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.

Description of derived fields

15. The derived fields used to generate the data summary are listed in Table 4.

Table 4: List of derived fields used in the data summary

Derived field name	Description	Paragraph
LCHEAIMFLAG	Identifies records of higher education	16
LCHEENTITYPOP	Identifies records where the learner higher education entity has been at least partially completed	17
IPBIRTHDATE	Date of birth of the learner	8
IPSTARTAGE	Age of the learner on 31 August in the year they commence their studies	8
IPSTARTAGEBAND	Age category	8
IPSEX	Sex	8
IPDISABLE	Indicates whether a disability has been reported	8
IPDISABLETYPE	Disability type	8
IPETHNIC	The learner's ethnicity in broad groupings	8
IPETHNICDETAIL	The learner's ethnicity in detail	8
IPPOSTCODE	Postcode prior to enrolment	8
IPLOCPOSTCODE	Location of study postcode	8
OFSQAIM	Recognised higher education qualification aim	19
LCDOM	Domicile	18
LCENTQUALGRP	Highest qualification on entry	20
LCNUTS2REGION	Region of study	21
LCSBJ_CAHXX-XX*	The full person equivalent for each Common Aggregation Hierarchy (CAH2) codes per learner	22

* There is one column for each of the 35 CAH2 codes. The full person equivalent for each of these subjects is recorded for each instance.

LCHEAIMFLAG

16. This field identifies higher education level learning aims.

Value	Description	Definition
1	The learning aim is at higher education level (Level 4 or above)	LEARNAIMREF ≠ ZPROG001 and NOTIONALNVQLEVEL2 = (4, 5, 6, 7, 8, H)
0	Otherwise	Otherwise

LCHEENTITYPOP

17. This field identifies records where there is evidence that the learning delivery higher education entity has been at least partially filled in. The fields used to identify this are those where the ILR schema mandates that the field is populated. This field identifies the same population as the equivalent field in the quality control data summary, QCHEENTITYPOP. This is documented in the ILR quality control data summary document on the OfS website (www.officeforstudents.org.uk/data-and-analysis/data-checking-tool/2020-21-ilr-data-checking-tool/ under 'Documentation').

Value	Description	Definition
1	The learning delivery higher education entity has been at least partially filled in	TYPEYR ≠ BLANK or MODESTUD ≠ BLANK or FUNDEV ≠ BLANK or FUNDCOMP ≠ BLANK or YEARSTU ≠ BLANK or MSTUFEE ≠ BLANK or SPECFEE ≠ BLANK
0	Otherwise	Otherwise

LCDOM

18. Where the learner is domiciled.

Value	Description	Definition
E	England	DOMICILE = XF or (DOMICILE in (XJ, XK, GB) and IPPOSTCODE is in England) or (DOMICILE in (ZZ, BLANK) and IPPOSTCODE is in England)
S	Scotland	DOMICILE = XH or (DOMICILE in (XJ, XK, GB) and IPPOSTCODE is in Scotland) or (DOMICILE in (ZZ, BLANK) and IPPOSTCODE is in Scotland)
W	Wales	DOMICILE = XI or (DOMICILE in (XJ, XK, GB) and IPPOSTCODE is in Wales) or (DOMICILE in (ZZ, BLANK) and IPPOSTCODE is in Wales)
N	Northern Ireland	DOMICILE = XG or (DOMICILE in (XJ, XK, GB) and IPPOSTCODE is in Northern Ireland) or (DOMICILE in (ZZ, BLANK) and IPPOSTCODE is in Northern Ireland)
EU	European Union	DOMICILE in (AI, AN, AQ, AT, AW, AX, BE, BG, BL, BM, BQ, CH, CW, CY, CZ, DE, DK, EE, ES, EU, FI, FK, FO, FR, GF, GI, GL, GP, GR, GS, HR, HU, IC, IE, IO, IS, IT, KY, LI, LT, LU, LV, MF, MQ, MS, MT, NC, NL, NO, PF, PL, PM,

Value	Description	Definition
		PN, PT, RE, RO, SE, SH, SI, SK, SX, TC, TF, VG, WF, XA, XD, XE, YT)
UNKNOWN	Unknown Domicile	DOMICILE in (ZZ, BLANK) and IPPOSTCODE = BLANK or IPPOSTCODE invalid
OTHER	Not EU or UK	Otherwise

OFSQAIM

19. This field allocates qualification aims to broad categories of higher education. For a full definition of this field please refer to '2020-21 ILR data checking tool: Classifying learning aims technical document' (available under 'Documentation' at www.officeforstudents.org.uk/data-and-analysis/data-checking-tool/2020-21-ilr-data-checking-tool/).

LCENTQUALGRP

20. This field assigns the broad grouping of the learner's highest qualification on entry (QUALENT3).

Value	Description	Definition
HEPG	Higher education: Postgraduate level	QUALENT3 = DUK, DZZ, D80, M41, M44, M71, M80, M90, MUK, MZZ, H71
HEFD	Higher education: First degree level	QUALENT3 = M2X, H11, HUK, HZZ, JUK
HEOUG	Higher education: Other undergraduate level	QUALENT3 = H80, J10, J20, J30, J48, J80, C20, C30, C44, C80, C90
BACC	Baccalaureate	QUALENT3 = P62, P63
FOUND	Foundation course	QUALENT3 = J49
ACCESS	Access course	QUALENT3 = X00, X01
ALEVEL	A-levels	QUALENT3 = P50
BTECD	BTEC diploma	QUALENT3 = P41
BTECC	BTEC certificate	QUALENT3 = P42
L3UCAS	Level 3 qualifications of which are all subject to UCAS tariff	QUALENT3 = P93
OLEVEL3	Other Level 3 qualification	QUALENT3* = P (excluding P41, P42, P50, P62, P63, P93)
NONE	No formal qualifications	QUALENT3 = X02, X03, X05
OTHERS	Other qualifications (unknown level, or below Level 3)	QUALENT3* = Q, R, X04
UNKNOWN	Unknown qualifications	Otherwise

* The first character of QUALENT3 is used.

LCNUTS2REGION

21. 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 (www.ons.gov.uk/methodology/geography/geographicalproducts/postcodeproducts/).

LCSBJ-CAHXX-XX

22. These 35 fields assign the full person equivalent for each instance to a CAH2 subject or subjects where XX-XX represents the CAH2 code. For example, a learner who is studying for a first degree in biology in 2020-21 will have one full-person equivalent (FPE = 1; a headcount measure) recorded under LCSBJ-CAH03-01. A learner who is studying a joint course first degree with equal proportions of mathematics and physics in 2020-21 will have each subject assigned an FPE of 0.5 (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). This will be recorded under LCSBJ-CAH09-01 and LCSBJ-CAH07-01 for mathematics and physics, respectively. The FPE associated with each LCSBJ-CAHXX-XX code is derived using PCFLDCS, PCSLDCS and PCTLDCS. Where PCFLDCS, PCSLDCS and PCTLDCS do not sum to 1, FPE has been scaled to reflect this.

Rebuild instructions

23. The individualised file, 'LC20_AAA_XXXXXXXXX_IND.csv', can be used to interrogate the data and our calculation of derived fields by filtering to certain groups of records.
24. Specific instructions to rebuild the learner characteristics data summary workbook, 'LC20_AAA_XXXXXXXXX.xlsx', are provided here. 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).
25. The following sections of this document detail how to rebuild the populations of learners found in the sheets of the data summary workbook. Each sheet includes information which needs to be filtered from the individualised file. This is summarised in the table below.

Table 5: Learner characteristic data summary tables created using 2020-21 ILR data

Worksheet*	Title
Coversheet	Title page
Learner characteristics	Number of learners by characteristic
Taught location	Number of learners by regional location of study
Teaching provider	Number of learners by subcontractual partner

* This worksheet reference corresponds to the spreadsheet tabs.

26. To rebuild the headcounts in the workbook summary tables, apply the following filters to the individualised files:
- LCHEAIMFLAG = 1 and LCHEENTITYPOP = 1.
 - 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.
 - Once all the required filters have been applied the number of learners is returned by summing the number of records.

Learner characteristics

27. This sheet shows the number of learners categorised by various characteristics. The possible options that are displayed on this sheet are listed in the tables below.

Age

Row value	Rebuild algorithm (LCHEAIMFLAG = 1 and LCHEENTITYPOP = 1 and...)
Under 21	IPSTARTAGEBAND = U21
21 to 25	IPSTARTAGEBAND = 21_25
26 to 30	IPSTARTAGEBAND = 26_30

Row value	Rebuild algorithm (LCHEAIMFLAG = 1 and LCHEENTITYPOP = 1 and...)
31 to 40	IPSTARTAGEBAND = 31_40
41 to 50	IPSTARTAGEBAND = 41_50
51 and over	IPSTARTAGEBAND = 51+
Unknown	IPSTARTAGEBAND = U

Sex

Row value	Rebuild algorithm (LCHEAIMFLAG = 1 and LCHEENTITYPOP = 1 and...)
Male	IPSEX = 1
Female	IPSEX = 2

Disability

Row value	Rebuild algorithm (LCHEAIMFLAG = 1 and LCHEENTITYPOP = 1 and...)
Disability reported	
In receipt of DSA	IPDISABLE = Y and LEARNFAM_DLA = 1
Not in receipt of DSA	IPDISABLE = Y and LEARNFAM_DLA ≠ 1
No disability reported	
In receipt of DSA	IPDISABLE = N and LEARNFAM_DLA = 1
Not in receipt of DSA	IPDISABLE = N and LEARNFAM_DLA ≠ 1

Disability type

Row value	Rebuild algorithm (LCHEAIMFLAG = 1 and LCHEENTITYPOP = 1 and...)
None	IPDISABLETYPE = NONE
Multiple impairments	IPDISABLETYPE = MULTI
Social or communication impairment	IPDISABLETYPE = SOC
Cognitive or learning impairment	IPDISABLETYPE = COG
Mental health condition	IPDISABLETYPE = MH
Sensory, medical or physical impairments	IPDISABLETYPE = PHY

Ethnicity

Row value	Rebuild algorithm (LCHEAIMFLAG = 1 and LCHEENTITYPOP = 1 and...)
Asian or Asian British	IPETHNIC = A
Black or African or Caribbean or Black British	IPETHNIC = B
Mixed or Multiple ethnic groups	IPETHNIC = M
Other ethnic group	IPETHNIC = O

Row value	Rebuild algorithm (LCHEAIMFLAG = 1 and LCHEENTITYPOP = 1 and...)
White	IPETHNIC = W
Unknown	IPETHNIC = U

Entry qualifications

Row value	Rebuild algorithm (LCHEAIMFLAG = 1 and LCHEENTITYPOP = 1 and...)
Higher education: Postgraduate level	LCENTQUALGRP = HEPG
Higher education: First degree level	LCENTQUALGRP = HEFD
Higher education: Other undergraduate level	LCENTQUALGRP = HEOUG
Baccalaureate	LCENTQUALGRP = BACC
Foundation course	LCENTQUALGRP = FOUND
Access course	LCENTQUALGRP = ACCESS
A-levels	LCENTQUALGRP = ALEVEL
BTEC diploma	LCENTQUALGRP = BTECD
BTEC certificate	LCENTQUALGRP = BTECC
Level 3 qualifications of which are all subject to UCAS tariff	LCENTQUALGRP = L3UCAS
Other Level 3 qualifications	LCENTQUALGRP = OLEVEL3
No formal qualifications	LCENTQUALGRP = NONE
Other qualifications (unknown level, or below Level 3)	LCENTQUALGRP = OTHERS
Unknown	LCENTQUALGRP = UNKNOWN

Domicile

Row value	Rebuild algorithm (LCHEAIMFLAG = 1 and LCHEENTITYPOP = 1 and...)
England	LCDOM = E
Other UK	LCDOM in (S, W, N)
EU	LCDOM = EU
Not EU or UK	LCDOM = OTHER
Unknown	LCDOM = UNKNOWN

Subject of study (CAH2)

28. The number of learners per subject is returned by summing values (FPE) in the required LCSBJ_CAHXX-XX column.

Row value	Rebuild algorithm (LCHEAIMFLAG = 1 and LCHEENTITYPOP = 1 and sum of LCSBJ_CAHXX_XX)
Medicine and dentistry (CAH01-01)	LCSBJ_CAH01_01
Pharmacology, toxicology and pharmacy (CAH02-02)	LCSBJ_CAH02_02
Nursing and midwifery (CAH02-04)	LCSBJ_CAH02_04
Medical sciences (CAH02-05)	LCSBJ_CAH02_05
Allied health (CAH02-06)	LCSBJ_CAH02_06
Biosciences (CAH03-01)	LCSBJ_CAH03_01
Sport and exercise sciences (CAH03-02)	LCSBJ_CAH03_02
Psychology (CAH04-01)	LCSBJ_CAH04_01
Veterinary sciences (CAH05-01)	LCSBJ_CAH05_01
Agriculture, food and related studies (CAH06-01)	LCSBJ_CAH06_01
Physics and astronomy (CAH07-01)	LCSBJ_CAH07_01
Chemistry (CAH07-02)	LCSBJ_CAH07_02
General, applied and forensic sciences (CAH07-04)	LCSBJ_CAH07_04
Mathematical sciences (CAH09-01)	LCSBJ_CAH09_01
Engineering (CAH10-01)	LCSBJ_CAH10_01
Materials and technology (CAH10-03)	LCSBJ_CAH10_03
Computing (CAH11-01)	LCSBJ_CAH11_01
Geography, earth and environmental studies (CAH26-01)	LCSBJ_CAH26_01
Architecture, building and planning (CAH13-01)	LCSBJ_CAH13_01
Sociology, social policy and anthropology (CAH15-01)	LCSBJ_CAH15_01
Economics (CAH15-02)	LCSBJ_CAH15_02
Politics (CAH15-03)	LCSBJ_CAH15_03
Health and social care (CAH15-04)	LCSBJ_CAH15_04
Law (CAH16-01)	LCSBJ_CAH16_01
Business and management (CAH17-01)	LCSBJ_CAH17_01
Media, journalism and communications (CAH24-01)	LCSBJ_CAH24_01
English studies (CAH19-01)	LCSBJ_CAH19_01
Celtic studies (CAH19-02)	LCSBJ_CAH19_02
Languages and area studies (CAH19-04)	LCSBJ_CAH19_04
History and archaeology (CAH20-01)	LCSBJ_CAH20_01
Philosophy and religious studies (CAH20-02)	LCSBJ_CAH20_02

Row value	Rebuild algorithm (LCHEAIMFLAG = 1 and LCHEENTITYPOP = 1 and sum of LCSBJ_CAHXX_XX)
Creative arts and design (CAH25-01)	LCSBJ_CAH25_01
Performing arts (CAH25-02)	LCSBJ_CAH25_02
Education and teaching (CAH22-01)	LCSBJ_CAH22_01
Combined and general studies (CAH23-01)	LCSBJ_CAH23_01

Taught location

29. This sheet summarises the number of learners by NUTS Level 2 regions as defined in LCNUTS2REGION.

Teaching provider

30. Where a learner has been recorded with a subcontractual partner this sheet summarises the number of learners by filtering to a PARTNERUKPRN = UKPRN.

Teaching provider name	UKPRN	Rebuild algorithm (LCHEAIMFLAG = 1 and LCHEENTITYPOP = 1 and...)
Taught at provider	UKPRN of provider	PARTNERUKPRN = 0
Taught provider name	UKPRN	PARTNERUKPRN = UKPRN



© The Office for Students copyright 2021

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/