Individualised Learner Record (ILR) quality control data summary

2020-21 technical document and rebuild instructions

Contents

Summary

1. This document describes the algorithms applied to the 2020-21 Individualised Learner Record (ILR) data to build the tables in the ILR quality control data summary. These data summary tables are provided to help providers identify potential data errors and reduce the numbers of key fields within ILR data that have unknown values, issues with continuity or issues with credibility, all of which will affect our onward uses of the data. Throughout the document, fields taken or derived from the ILR are shown in capitals.

2. This document is aimed at readers with in-depth knowledge of the data. Readers are advised to have a copy of the ILR specification for 2020-21 (https://guidance.submit-learner-data.service.gov.uk/ilr) to hand when using this document.

3. The ILR quality control data summary is formed of two files. These are:

   a. ILR quality control data summary workbook 'QC20_AAA_XXXXXXXX.xlsx'

   b. ILR quality control data summary individualised file 'QC20_AAA_XXXXXXXX_IND.csv'

   Where XXXXXXXX is your provider's UK Provider Reference Number (UKPRN), and AAA identifies the output type as listed in Table 1.

Table 1: Filename abbreviations used in the ILR quality control outputs

File abbreviation	Source of outputs
DCT	Data checking tool
AST	Amendments submission tool
PCO	Post-collection outputs

ILR quality control data summary workbook

4. The ILR quality control data summary workbook is the Excel workbook QC20_AAA_XXXXXXXX.xlsx which contains the following worksheets:

Table 2: Worksheets in the quality control data summary workbook

Worksheet*	Title
Coversheet	
HE populations	Higher education populations
Modes and qualifications	Modes of study and qualifications
Unknowns	Fields returned as unknown or missing
Credibility	Credibility checks
HE learning aims	Summary of higher education learning aims
Continuity	Continuity checks

*This worksheet reference corresponds to the spreadsheet tabs.

ILR quality control data summary individualised file

5. When working through this document it is necessary to use the individualised file, QC20_AAA_XXXXXXXX_IND.csv. The individualised file contains the values of all fields used to generate the data summary tables.

6. 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 'record' in this technical document and the quality control data summary workbook to refer to one row of data in the individualised file which corresponds to a learner-course combination.

Changes from the 2019-20 ILR quality control data summary outputs

7. In 2019-20, we produced the ILR quality control data summary for the first time. After a review, we have redesigned it to focus more on data quality issues in the following areas:

    a. Unknowns: Key fields which are returned as 'unknown', 'not provided' or have missing values

    b. Credibility: Combinations of fields or values that appear less credible

    c. Continuity: Examining the consistency of learner reporting between years.

8. We have also undertaken work to ensure any questions raised on data quality concerns are more clearly communicated to providers.

9. The key differences from last year's output are as follows:

    a. We have removed all algorithms and checks looking at Discover Uni data.

    b. Checks on student characteristics, demographics and location of study have been removed from the ILR quality control data summary. The new learner characteristics data summary output has been designed to look at data in this area.

    c. Tables in the ILR quality control data summary workbook use counts of records rather than counts of full-person equivalence (FPE).

    As a result of the above changes, the algorithms in this year's technical document are substantially different from last year.

Fields used in the quality control algorithms

10. The ILR and Learning Aims Reference Service (LARS) fields used to generate the quality control data summary are listed in Table 3.

Table 3: ILR and LARS fields used in the ILR quality control data summary

Data source	Field	Description
ILR	AIMSEQNUMBER[1]	Learning aim data set sequence
ILR	COMPSTATUS	Completion status
ILR	DOMICILE	Country of domicile
ILR	ETHNICITY	Ethnicity
ILR	FUNDCOMP	Completion of year of programme of study
ILR	FUNDLEV	Level applicable to HESES
ILR	FUNDMODEL	Funding model
ILR	GROSSFEE	Gross tuition fee
ILR	HEPOSTCODE	HE centre location postcode
ILR	LDCS_CO1, LDCS_CO2, LDCS_CO3	Learning directory classification system codes
ILR	LEARNACTENDDATE	Learning actual end date
ILR	LEARNAIMREF	Learning aim reference
ILR	LEARNDELFAMCODE[2]	The value of the funding or monitoring attribute
ILR	LEARNDELFAMTYPE[2]	The type of funding or monitoring attribute
ILR	LEARNREFNUMBER	Learner reference number
ILR	LEARNSTARTDATE	Learning start date
ILR	LLDDHEALTHPROB	The nature of the learner's disability, learning difficulty and/or health problem
ILR	MODESTUD	Mode applicable to HESES
ILR	MSTUFEE	Major source of tuition fees
ILR	NUMHUS**Error! Bookmark not defined.**	Student instance identifier
ILR	OUTCOME	Indicates whether the learner achieved the learning aim
ILR	OUTGRADE	The examination grade awarded to the learner for the learning aim
ILR	PARTNERUKPRN	Subcontracted or partnership UKPRN

---

[1] These fields are not used in the algorithms but are included in the individualised file to allow easy identification of students.

[2] The only values of LEARNDELFAMTYPE and LEARNDELFAMCODE field used in the data quality algorithms are those where LEARNDELFAMTYPE = ADL or SOF.

Data source	Field	Description
ILR	PCFLDCS, PCSLDCS, PCTLDCS	Proportion taught in LDCS_CO1 – LDCS_CO3 subject
ILR	PCOLAB	Percentage not taught by this institution
ILR	POSTCODEPRIOR	Postcode prior to enrolment
ILR	PRIMARYLLDD	The most significant or primary learning difficulty, disability or health problem that impacts on the learner's education
ILR	PROGTYPE	The type of programme which the learner is undertaking
ILR	QUALENT3	Qualification on entry
ILR	SPECFEE	Special fee indicator
ILR	SSN	Student Support Number
ILR	STULOAD	Student instance full-time equivalence (FTE)
ILR	TYPEYR	Type of instance year
ILR	UKPRN	UK provider reference number
ILR	WITHDRAWREASON	Withdrawal reason
ILR	YEARSTU	Year of student on this instance
LARS	ENGLPRSCID	Identifies prescribed higher education learning aims, as defined by the Education (Prescribed Courses of Higher Education) (Wales) (Amendment) Regulations 1998
LARS	LEARNAIMREFTITLE	Learning aim reference title
LARS	NOTIONALNVQLEVELV2	The detailed learning aim level on the NVQ scale

Description of derived fields

11. The derived fields used to generate the quality control data summary are listed in Table 4.

12. The derived fields ANNIV, HESEXCL and HESTYPE are derived using the algorithms documented in the HESES20 comparison technical document available on the OfS website (insert link).

13. OFSQAIM is derived using the algorithms documents in '2020-21 ILR data checking tool: Classifying learning aims technical document', available on the OfS website (www.officeforstudents.org.uk/data-and-analysis/data-checking-tool/2020-21-ilr-data-checking-tool/).

14. A number of the derived fields described below will be calculated for both the 2019-20 and the 2020-21 ILR data. The suffix 'YY' will be used in the name of these derived fields, where 'YY' signifies that the field is based on the data from the academic year 20YY-YY+1.

Table 4: Description of derived fields for recreating tables

Derived field name	Description	Paragraph
AIMSEQNUMBER19	AIMSEQNUMBER in 2019-20 ILR	37
ANNIV	Anniversary of start date	NA
COMPSTATUS19	COMPSTATUS in 2019-20 ILR	38
HESEXCL	Reason(s) for exclusions from the HESES population	NA
HESTYPE	HESES fundability status	NA
LEARNACTENDDATE19	LEARNACTENDDATE in 2019-20 ILR	39
LEARNSTARTDATE19	LEARNSTARTDATE in 2019-20 ILR	40
LEARNDELFAM_ADL1	Advanced Learner Loan indicator	24
LEARNDELFAM_SOF1	Source of funding	25
NOTIONALNVQLEVELV2_19	NOTIONALNVQLEVELV2 in 2019-20 ILR	41
OFSQAIM	Broad category of higher education	NA
OUTCOME19	OUTCOME in 2019-20 ILR	42
QCAYDAYSSTUDIED	Number of days studied within the academic year	32
QCCAH3_1, QCCAH3_2, QCCAH3_3	CAH3 code(s) mapped to fields LDCS_CO1 to 3	34
QCDEDUP19	Derived field for deduplicating 2019-20 ILR data in the individualised file	36
QCDUPOUTCOME	Duplicate outcomes for a single programme of study indicator	47
QCENDDATE	Last day of activity in 2020-21	31
QCEXCL1-16_YY	Individual QC population exclusions for the 20YY-YY+1 data	19-21

Derived field name	Description	Paragraph
QCEXCLYY	Reason(s) for exclusion from the 20YY-YY+1 population	22
QCHEENTITYFLAG	Identifies learners that meet the definition to be returned in the learner HE entity	27
QCHEENTITYPOP	Identifies records where the learner HE entity has been at least partially completed	28
QCJACS1, QCJACS2, QCJACS3	JACS code(s) mapped to fields LDCS_CO1 to 3	33
QCLINKCAT	Linking category for the record	37
QCLINKED	Identifies records successfully linked between 2019-20 and 2020-21 ILR data	35
QCMODE	Mode of study for workbook Table 3	30
QCOMITTED	Learners that did not end their study in 2019-20, but are not returned 2020-21	45
QCPOP20	Population from 2020-21 ILR data on which most quality control checks will be focussed	
QCSOURCEYY	Identifies records that contain data from the 20YY-YY+1 ILR	18
QCYEARSTUERR	Identifies where year of study has not been incremented between years	46
QCAYDAYSSTUDIED	Number of days studied within the academic year	32
UKPRN19	UKPRN in 2019-20 ILR	43
WITHDRAWREASON19	WITHDRAWREASON in 2019-20 ILR	45
YEARSTU19	YEARSTU in 2019-20 ILR	44

Linking between years

15. We link the 2020-21 ILR data back to the 2019-20 ILR data to check the consistency of data reporting between years. The individualised file will therefore contain data from both the 2020-21 and the 2019-20 ILR returns. For brevity, only higher education records from 2019-20 will be included in the linking process and the individualised file (NOTIONALNVQLEVELV2 = ('4','5','6','7','8','H')).

16. Before linking the 2020-21 and 2019-20 ILR data, the exclusion fields QCEXCL1_YY to QCEXCL16_YY will be calculated on each year's data. Only non-excluded data from each year (QCEXCLYY = 0) will be included in the linking process.

17. A programme of study is linked between years by matching on LEARNREFNUMBER, LEARNAIMREF and LEARNSTARTDATE. Where there are multiple aims within the same year which have the same combination of LEARNREFNUMBER, LEARNAIMREF and LEARNSTARTDATE, we will link as follows:

    a. Where a single record in 2019-20 data has multiple matches in the 2020-21 data, it will be linked to all 2020-21 records.

b. If a single record in the 2020-21 data has multiple potential matches in the 2019-20 data then a single match will be picked after sorting the 2019-20 records. The potential 2019-20 matches will be sorted by LEARNACTENDDATE, from most recent to least recent with blank LEARNACTENDDATEs first, then by descending YEARSTU. The topmost 2019-20 record will be selected as the match, and the remaining 2019-20 records will not be matched.

c. If there are multiple records with the same set of LEARNREFNUMBER, LEARNAIMREF and LEARNSTARTDATE in both years of data, then both sets will be sorted by LEARNACTENDDATE, from most recent to least recent with blank LEARNACTENDDATEs first, then by descending YEARSTU. Records will then be matched by their position in the sort order, with the topmost record in 2020-21 being matched to the topmost record in 2019-20 and so on, until we run out of matches.

QCSOURCEYY

18. This field identifies those records that contain data from the 20YY-YY+1 ILR.

Value	Description
1	Record contains data from the 20YY-YY+1 ILR
0	Otherwise

QCEXCL1_YY

19. This field identifies records in the 20YY-YY+1 ILR data that start after the 20YY-YY+1 academic year.

Value	Description	Definition
1	Excluded	LEARNSTARTDATE ≥ 1 August 20YY+1
0	Included	Otherwise

QCEXCL2_YY

20. This field identifies records in the 20YY-YY+1 ILR data that finish before the 20YY-YY+1 academic year.

Value	Description	Definition
1	Excluded	LEARNACTENDDATE < 1 August 20YY and LEARNACTENDDATE ≠ BLANK
0	Included	Otherwise

QCEXCL4_YY

21. This field identifies records in the 20YY-YY+1 data that have been closed to correct an incorrect LEARNPLANENDDATE.

Value	Description	Definition
1	Excluded	COMPSTATUS = 4[3] and WITHDRAWREASON = 40 and LEARNACTENDDATE = 1 August 20YY
0	Included	Otherwise

QCEXCL8_YY

22. This field identifies the higher education records in the 20YY-YY+1 data.

Value	Description	Definition
0	Included – learning aim is at higher education level	LEARNAIMREF ≠ ZPROG001 and NOTIONALNVQLEVELV2 = (4, 5, 6, 7, 8, H)
1	Excluded	Otherwise

QCEXCL16_19

23. This field identifies records in the 2019-20 data where the learner has withdrawn from the learning aim. Applied to the 2019-20 data only.

Value	Description	Definition
1	Excluded	WITHDRAWREASON ≠ 0
0	Included	Otherwise

QCEXCLYY

24. This field indicates whether the record will be included in the population for building the summary tables in the quality control workbook. QCEXCLYY is the sum of all applicable values from the table below.

Value	Description	Definition
1	Record from 20YY-YY+1 that starts after academic year 20YY-YY+1	QCEXCL1_YY = 1
2	Record from 20YY-YY+1 that ends before academic year 20YY-YY+1	QCEXCL2_YY = 1
4	Record closed to correct LEARNPLANENDDATE	QCEXCL4_YY = 1
8	Record is not of higher education	QCEXCL8_YY = 1
16	Learner has withdrawn from learning. For 2019-20 data only	QCEXCL16_19 = 1
0	Otherwise	None of the above

25. This field contains the exclusion reason(s) for the record. It is calculated as (1 × QCEXCL1_YY) + (2 × QCEXCL2_YY) +(4 × QCEXCL4_YY) +...

---

[3] 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 change is replicated by the OfS.

LEARNDELFAM_ADL1

26. This field contains the value of LEARNDELFAMCODE where LEARNDELFAMTYPE = ADL.

LEARNDELFAM_SOF1

27. This field contains the value of LEARNDELFAMCODE where LEARNDELFAMTYPE = SOF.

QCHEENTITYFLAG

28. This field identifies those records in the 2020-21 ILR data for which the Learning Delivery HE entity should be returned. The criteria are provided in the 'Specification of the Individualised Learner Record for 2020 to 2021' (https://guidance.submitlearnerdatabeta.fasst.org.uk/previous/20-21/ilr/entity/learningdeliveryhe).

Value	Description	Definition
1	**For all providers:** Learning aims that are OfS funded, as indicated in the Learning Delivery Funding and Monitoring fields using code SOF1	QCSOURCE20 = 1 and LEARNDELFAM_SOF1 = 1
1	**For grant funded provision (FE colleges and other grant funded providers):** Learning aims that are Level 4 or above in the LARS database and are funded by the 16-19 (excluding Apprenticeships) funding (FundModel 25); Adult skills funding (FundModel 35) - non apprenticeships only; or are not funded by the ESFA (FundModel 99)	QCSOURCE20 = 1 and NOTIONALNVQLEVELV2 = (4, 5, 6, 7, 8, H) and ((FUNDMODEL = (25, 35) and PROGTYPE ≠ (2, 3, 20, 21, 22, 23, 25)) or FUNDMODEL = 99)
1	**For grant funded provision (FE colleges and other grant funded providers):** Learning aims that are identified as prescribed HE in the LARS database (using the EnglPrscID field) and are part of a higher apprenticeship programme (ProgType 20-23, 25)*	QCSOURCE20 = 1 and ENGLPRSCID = 1 and PROGTYPE = (20, 21, 22, 23, 25)
1	**For contract funded provision (private training providers and all apprenticeships started from 1 May 2017):** Learning aims that are identified as prescribed HE in the LARS database (using the EnglPrscID field)	QCSOURCE20 = 1 and PROGTYPE = (2, 3, 20, 21, 22, 23, 25) and LEARNSTARTDATE ≥ 1 May 2017 and ENGLPRSCID = 1
1	**For Advanced Learner Loans provision (all grant funded, contract funded and loans only providers):** Learning aims that are Level 4 or above in the LARS database and are funded through an Advanced Learner Loan	QCSOURCE20 = 1 and NOTIONALNVQLEVELV2 = (4, 5, 6, 7, 8, H) and FUNDMODEL = 99 and LEARNDELFAM_ADL1 = 1

Value	Description	Definition
	(FundModel 99 and LearnDelFAM = ADL1)	
0	2020-21 records for which the Learning Delivery HE entity does not need to be returned	QCSOURCE20 = 1 and not above

QCHEENTITYPOP

29. This field identifies records in the 2020-21 ILR data where there is evidence that the Learning Delivery HE entity has been at least partially completed, based on the completion of one or more of the mandatory Learning Delivery HE entity fields.

Value	Description	Definition
1	The Learning Delivery HE entity has been at least partially completed	QCSOURCE20 = 1 and (TYPEYR ≠ BLANK or MODESTUD ≠ BLANK or FUNDLEV ≠ BLANK or FUNDCOMP ≠ BLANK or YEARSTU ≠ BLANK or MSTUFEE ≠ BLANK or SPECFEE ≠ BLANK)
0	Otherwise	Otherwise

QCPOP20

30. This field identifies the population from the 2020-21 ILR data on which most of the quality control checks will be focused.

Value	Description	Definition
1	Population from the 2020-21 ILR data for most quality control checks	QCSOURCE20 = 1 and QCEXCL20 = 0 and QCHEENTITYPOP = 1
0	Otherwise	Otherwise

QCMODE

31. This field identifies the mode of study for Table 3 of the ILR quality control data summary workbook.

Value	Description	Definition
APPR	Apprenticeship	QCSOURCE20 = 1 and PROGTYPE = (2, 3, 20, 21, 22, 23, 25)
FT	Full-time (including sandwich year out), excluding apprenticeships	QCSOURCE20 = 1 and MODESTUD = (1, 2) and not above
PT	Part-time, excluding apprenticeships	QCSOURCE20 = 1 and not above

QCENDDATE

32. This field holds the last day of study in the 2020-21 academic year, to be used in calculating the number of days studied in the 2020-21 academic year (QCAYDAYSSTUDIED) and for checking that the year of study field in the 2020-21 data has been correctly incremented (QCYEARSTUERR).

Value	Description	Definition
LEARNACTENDDATE	LEARNACTENDDATE field completed	QCSOURCE20 = 1 and LEARNACTENDDATE ≠ BLANK
31 JULY 2021	LEARNACTENDDATE field has not been completed	QCSOURCE20 = 1 and not above

QCAYDAYSSTUDIED

33. This field determines the number of days studied in the 2020-21 academic year, calculated as the number of days between:

   a. LEARNSTARTDATE or 1 August 2020, whichever is later, and

   b. QCENDDATE

34. This field is only calculated for the 2020-21 ILR data (QCSOURCE20 = 1).

QCJACS1-3

35. This field shows the full four-digit Joint Academic Coding System (JACS) code that has been mapped to fields LDCS_CO1 to 3 respectively. Mapping is only applied to non-blank LDCS_COX fields. For details of this mapping, see the 'LDCS to JACS lookup' document on the OfS website (www.officeforstudents.org.uk/data-and-analysis/institutional-performance-measures/technical-documentation/). Only calculated for the 2020-21 data (QCSOURCE20 = 1).

QCCAH3_1-3

36. This field shows which of the Common Aggregation Hierarchy level 3 (CAH3) codes has been mapped to fields QCJACS1 to 3 respectively, using the current version of the Common Aggregation Hierarchy. Mapping is only applied to non-blank QCJACS fields. Where we cannot map to a subject, we set QCCAH3_X to CAH23-01-01. The mapping of JACS codes to the CAH3 codes can be found on the HESA website (https://www.hesa.ac.uk/support/documentation/hecos).

Derived fields used in continuity checks

QCLINKED

37. This field identifies those records that have been successfully linked between the 2019-20 and 2020-21 ILR data, following the method described in 'Linking between years.'

Value	Description
1	The record has been linked between the 2019-20 and 2020-21 ILR data
0	Otherwise

QCLINKCAT

38. This field categorises records in terms of how many non-excluded records with a matching set of LEARNREFNUMBER, LEARNAIMREF and LEARNSTARTDATE are found in each year of data. It can be used to help identify the linking approach taken for that particular record.

Value	Description
A	No non-excluded records with a matching set of LEARNREFNUMBER, LEARNAIMREF and LEARNSTARTDATE exist in the other year of data
B	One record in 2020-21 with one match in 2019-20
C	One record in 2020-21 with multiple matches in 2019-20
D	One record in 2019-20 with multiple matches in 2020-21
E	Multiple matching records in both years of data
BLANK	Record is excluded so will not be linked

QCDEDUP19

39. This field is used to de-duplicate the 2019-20 ILR data in the case that a single record in 2019-20 has been linked to multiple matching rows in the 2020-21 ILR data. Filtering to QCDEDUP19 = '1' will ensure that only one instance of each 2019-20 ILR record is selected.

Value	Description
1	The 2019-20 record is unique, or this is the first instance of a non-unique record
0	This record is a duplicate copy

AIMSEQNUMBER19

40. This field contains the aim sequence number (AIMSEQNUMBER) for 2019-20. This is obtained using the linking method described in 'Linking between years.'

COMPSTATUS19

41. This field contains the completion status (COMPSTATUS) for 2019-20. This is obtained using the linking method described in 'Linking between years.'

LEARNACTENDDATE19

42. This field contains the learning actual end date (LEARNACTENDDATE) for 2019-20. This is obtained using the linking method described in 'Linking between years.'

LEARNSTARTDATE19

43. This field contains the learning start date (LEARNSTARTDATE) for 2019-20. This is obtained using the linking method described in 'Linking between years.'

NOTIONALNVQLEVELV2_19

44. This field contains the Notional NVQ level version 2 (NOTIONALNVQLEVELV2) for 2019-20. This is obtained using the linking method described 'Linking between years.'

OUTCOME19

45. This field contains the learning outcome for the learning aim (OUTCOME) for 2019-20. This is obtained using the linking method described in 'Linking between years.'

UKPRN19

46. This field contains the UKPRN for 2019-20. This is obtained using the linking method described in 'Linking between years.' UKPRN and UKPRN19 will differ where providers have been involved in a merger or de-merger.

WITHDRAWREASON19

47. This field contains the withdrawal reason (WITHDRAWREASON) FOR 2019-20. This is obtained using the linking method described in 'Linking between years.'

YEARSTU19

48. This field contains the year of student on this instance (YEARSTU) for 2019-20. This is obtained using the linking method described in 'Linking between years.'

QCOMITTED

49. This field identifies 2019-20 records where the learner was recorded as continuing or intending to continue the learning aim, but no non-excluded record of the same learning aim has been found in the 2020-21 data for that learner.

Value	Description	Definition
1	Learners recorded as continuing or intending to continue in 2019-20, but the learner has not returned on the same aim in 2020-21	QCSOURCE19 = 1 and QCEXCL19 = 0 and COMPSTATUS19 = 1 and QCLINKED = 0 and QCLINKCAT ≠ (C,E)
0	Otherwise	Otherwise

QCYEARSTUERR

50. This field identifies 2020-21 records where the year of study (YEARSTU) field may be incorrect. It highlights records where the learner has continued on active study between 2019-20 and 2020-21, but the year of study has not been incremented as expected.

Value	Description	Definition
1	The record is linked between 2019-20 and 2020-21. The year of study has not been incremented as expected.	QCPOP20 = 1 and QCLINKED = 1 and COMPSTATUS19 = 1 and YEARSTU = YEARSTU19 and QCENDDDATE > (ANNIV + 14)

Value	Description	Definition
0	Otherwise	Otherwise

QCDUPOUTCOME

51. This field identifies learners that ended their study in 2019-20 and achieved their learning aim, but are returned as achieving the same learning aim in 2020-21 with a different end date.

Value	Description	Definition
1	Learner ended their study in 2019-20 and achieved their learning aim, but is returned as achieving the same aim in 2020-21 with a different end date	QCPOP20 = 1 and QCLINKED = 1 and OUTCOME19 = 1 and OUTCOME = 1 and LEARNACTENDDATE19 ≠ LEARNACTENDDATE
0	Otherwise	Otherwise

Rebuild instructions

52. The individualised file, QC20_AAA_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 (www.officeforstudents.org.uk/data-and-analysis/supplying-data/working-with-individualised-files/).

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

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

55. The figures in Tables 1 to 6 and Table 8 of the ILR quality control summary workbook, QC20_AAA_XXXXXXXX.xlsx, can be rebuilt by following the rebuild algorithm tables provided in the workbook. These specify which filters need to be applied to the individualised data to find the relevant records. For example, to rebuild the number of full-time records in Table 3: HE records by mode of study, the column QCPOP20 should be filtered to 1 and the column QCMODE to 'FT.'

56. Table 7, the summary of higher education learning aims, is focused on course subject information. Only higher education courses with non-excluded learners in the 2020-21 ILR data will be included in the table. The courses included in the table can be found by filtering to QCPOP20 = 1. Each course will have up to three rows in Table 7, depending on the number of LDCS codes it has.