2022-23 Student data checking tool:

2022-23 Transparency attainment data summary technical document and rebuild instructions

Contents

Summary

1. This document describes the algorithms applied to the Student (22056) record submitted to the designated data body (DDB) to determine a provider's Transparency attainment data.

2. The Transparency attainment data summary presents the number (headcount) of UK-domiciled undergraduate qualifiers in 2022-23 at a provider by qualification classification achieved, mode of study, ethnicity, gender, and English Index of Multiple Deprivation (EIMD) 2019. This data forms a provider's transparency information.[1]

3. The Transparency attainment data summary can be used by higher education providers to verify and correct their 2022-23 individualised student data before signing off the final data with the DDB.

4. Throughout the document, fields taken or derived from the Student record are shown in capitals.

5. The Transparency attainment data summary is formed of two files. These are:

   a. Transparency attainment data summary workbook
      'TR22_Attainment_DCT_XXXXXXXX.xlsx'

   b. Transparency attainment data summary individualised file
      'TR22_Attainment_DCT_XXXXXXXX_IND.csv'

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

6. Details of how to use individualised files are given on the Office for Students (OfS) website.[2]

7. These files can be accessed through the HESA Data Platform.[3] Once you have processed submitted data, your file can be downloaded from the 'Iris' report, under the section 'Additional Collection Reports'.

8. This document is aimed at readers with in-depth knowledge of the data. Readers are advised to refer to the 'Student record coding manual'[4] when using this document.

9. Readers should refer to the '2022-23 Student characteristics technical document and rebuild instructions'[5] for definitions of field names referenced in the algorithms in this document with the prefix 'IP'.

---

[1] For details of the transparency information, see 'Condition F1: Transparency information' in Securing student success: Regulatory framework for higher education in England.

[2] See The OfS website - working with individualised files.

[3] See HESA Data Platform.

[4] See Student record coding manual.

[5] See the '2022-23 Student data checking tool SC22 data summary technical document' in the documentation section of the OfS website - Data checking tool.

10. We will document any changes or corrections to the algorithms in this document on the 'Data checking tool' pages of the OfS website under 'Updates'.[6]

---

[6] See <u>The OfS website - Data checking tool</u>.

Changes from previous publications

11. This section describes any changes to the algorithms since the equivalent was published for the 2020-21 algorithms in '2020-21 Transparency attainment data summary: HESA Student, HESA Student Alternative and ILR technical document and rebuild instructions' available from the 'Data checking tool documentation archive' section of the OfS website.[7] This was the most recent publication of algorithms relating to the transparency attainment data summary as there was no equivalent publication relating to 2021-22 data.

12. Derivations for all derived fields have been updated to take in to account the change in data model from the previous HESA Student and HESA Student Alternative data collections to the new Student (22056) model. More information on our general approach to converting algorithms to the Student (22056) model can be found on the 'Data checking tool' pages of the OfS website under 'Data futures'.[8]

13. We have removed TRSOURCE as this document now only contains instructions for the Student Data checking tool.

---

[7] See Data checking tool documentation archive - Office for Students.

[8] See The OfS website - Data checking tool.

Transparency attainment data summary files

Workbook

14. The Transparency attainment data summary workbook is the Excel workbook 'TR22_Attainment_DCT_XXXXXXXX.xlsx' which contains the following worksheets:

Table 1: Worksheets in the Transparency attainment data summary workbook

Worksheet*	Content
Information	Workbook information
Attainment 2022-23	Table showing attainment of 2022-23 qualifiers by ethnicity, gender, and EIMD 2019, classified by mode of study and award level.

* This worksheet reference corresponds to the spreadsheet tabs.

Individualised file

15. When working through this document you must use the individualised file TR22_Attainment_DCT_XXXXXXXX_IND.csv. The individualised file contains the values of all fields used to generate the figures in the 'Attainment 2022-23' worksheet of the Transparency attainment data summary workbook.

16. Fields taken from the Student return are shown in capitals in Table 2. Fields prefixed with 'TR' are derived and are shown in capitals in Table 3.

17. Specific instructions for rebuilding the Transparency attainment data summary workbook from the individualised file are given in this document.

18. The individualised file contains one row per engagement (each unique combination of UKPRN, SID and NUMHUS).

Fields used to create the Transparency attainment data summary

19. The Student (22056) fields used to generate the Transparency attainment data summary are listed in Table 2.

Table 2: Student (22056) fields used in the Transparency attainment data summary

Entity	Name	Description
Engagement	UKPRN	UK Provider Reference Number
Engagement	SID	Student identifier
Engagement	NUMHUS	Engagement number
Student	SEXID	Sex identifier
Student	ETHNIC	Ethnicity
Entry Profile	PERMADDPOSTCODE	Postcode of the student's permanent home address on entry to the engagement.
Qualification subject	QUALSUBJECT	Subject(s) of the qualification
Student Initiative	STUINITID	Student initiative identifier
Course*	PREREQUISITE	Prerequisite qualifications for the course
Course Initiative*	COURSEINITID	Course initiative identifier
Course initiative*	COURSEINITVALIDFROM	Course initiative valid from date
Course initiative*	COURSEINITVALIDTO	Course initiative valid to date
Curriculum accreditation*	CURACCID	Curriculum accreditation identifier
Curriculum accreditation*	CURACCIDVALIDFROM	Accreditation valid from date
Curriculum accreditation*	CURACCIDVALIDTO	Accreditation valid to date
Student course session	SCSSTARTDATE	Student course session start date
Student course session	SCSENDDATE	Student course session end date
Student accreditation aim	STUACCID	Student accreditation aim accreditation identifier
Engagement derived field	Z_MODEGRP2	Categorises the student's mode of study, based on the latest student course session
Engagement derived field	Z_STATUSEND	Student status at the end of the last student course session

Entity	Name	Description
Engagement derived field	Z_PERMADDGRP4	Categorises the student's permanent address information from the latest entry profile submitted
Engagement derived field	Z_POPQUAL_CYC	Qualifications obtained population identifier
Engagement derived field	Z_QLEVEL_CYC	The student's highest qualification awarded during the cycle.
Engagement derived field	Z_QCLASS_CYC	The result of student's highest qualification awarded during the cycle.
Engagement derived field	Z_ACT_CYC	Categorises whether the student was active during the cycle
Engagement derived subject	Z_SUBJHECOS	Records the latest HECoS subject information based on the latest student course session

*The values of these fields are taken from the latest student course session associated with the student's engagement.

Description of derived fields

20. The OfS derived fields used to generate the Transparency attainment data summary are listed in Table 3.

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

Name	Description	Paragraph
TRDENT	Indicates students on a dentistry programme of study	21
TRAWARDLEVELNUM	Level of higher education record	22
TRAWARD_DETAIL	Detailed grouping of qualification awarded to student	23
TRAWARDLEVEL	Broad grouping of qualification awarded to student	24
TRAWARD	Broad grouping of qualification awarded to student	25
TRAPPRENTICE	Indicates students studying on an apprenticeship	26
TRMODE	Mode of study linked to qualification	27
TRDOM	Broad level domicile of student	28
TRETHNIC	Ethnicity of student	29
TRSEX	Sex of student	30
TREIMD	English Index of Multiple Deprivation (EIMD) 2019 quintile of student's address prior to entry	31
TRDEGCLASS	Degree classification of award	32
TRATTAINPOP	Indicator for inclusion of student in the Transparency attainment population	33

Derived fields for calculating Transparency attainment

TRDENT

21. This field indicates whether a student has at least one instance of a 'Dentistry', 'Pre-clinical dentistry' or 'Clinical dentistry' programme of study. This field has the same definition as IPDENT.

Value	Description	Definition
1	The student has at least one instance of a 'Pre-clinical dentistry' or 'Clinical dentistry' programme of study	Z_SUBJHECOS in (100266, 100268, 100275)
0	The student does not have an instance of a 'Pre-clinical dentistry' or 'Clinical dentistry' programme of study	Otherwise

TRAWARDLEVELNUM

22. This field indicates the level of study for higher education records, using the sector recognised standards as set out in condition of registration B5.[9]

---

[9] See Sector-recognised standards.

Value	Description	Definition
8	Doctoral degree	Z_QLEVEL_CYC in (D0003, D0004, E0000, E0001, E0002, E0003, E0004, L0000)
7	Masters' degree, postgraduate diplomas, postgraduate certificates	Z_QLEVEL_CYC in (L0001, L0002, , M0002, M0003, M0004, M0006, M0007, M0008, M0009, M0010, M0011, M0012, M0013, M0015, M0016, M0017, M0018, M0020, M0021, M0022, M0023,)
6	Bachelors' degrees, graduate certificates and diplomas	Z_QLEVEL_CYC in (H0003, H0004, H0005, H0006, H0007, H0008, H0009, H0010, H0012, H0013, H0014, H0015, H0016, H0018, H0019, I0001)
5	Foundation degrees, diplomas of higher education and other higher diplomas	Z_QLEVEL_CYC in (I0002, I0004, I0005, I0006, I0007, I0008, I0009, I0010, I0012, J0000, J0001, J0002, J0003, J0004, J0005, J0006, J0007, J0010, J0011, J0012)
4	Certificates of higher education	Z_QLEVEL_CYC in (C0000, C0001, C0002, C0003, C0004, C0005, C0006, C0007, C0008, C0009)
BLANK	No qualification awarded or qualification not applicable to higher education qualifications framework	Z_QLEVEL_CYC in (P0002, Z9)

TRAWARD_DETAIL

23. This field allocates the qualification awarded to the student during the reporting year to a level of qualification awarded.

Value	Description	Definition
PHD	PhD and MPhil	Z_QLEVEL_CYC in (D0003, L0000)
OTHL7_Q_R	Other Level 7 research-based qualification	Z_QLEVEL_CYC in (L0001)
MASTER	Masters'	Z_QLEVEL_CYC in (M0003, M0004, M0006, M0007)
PGCE	PGCE and other postgraduate Initial Teacher Training (ITT)	Z_QLEVEL_CYC in (H0013, M0016)
DTLLS_PG	Postgraduate diploma in teaching in the lifelong learning sector	Z_QLEVEL_CYC in (M0020)
PGCERT	Postgraduate certificate	Z_QLEVEL_CYC in (M0012)
PGDIP	Postgraduate diploma	Z_QLEVEL_CYC in (M0009)
PROCONGRAD	Professional, conversion and other graduate entry programmes	Z_QLEVEL_CYC in (H0009, H0010, H0014, I0002, I0005, I0006, I0007) or Z_QLEVEL_CYC in (H0016, I0010) and

Value	Description	Definition
		PREREQUISITE = 02
ENHANCED	Enhanced first degree (or integrated masters)	Z_QLEVEL_CYC in (H0004, M0002)
MEDVETDENT	Pre-registration first degree with honours leading towards obtaining eligibility to register to practise with the General Medical Council, General Dentistry Council (as a dentist) or the Royal College of Veterinary Surgeons	Z_QLEVEL_CYC in (H0003, H0005, I0001) and at least one value of QUALAWARDACCID in (05901, 12001, 05803) where Z_QAWARDHMRK_CYC = 1
FIRST	First degree	Z_QLEVEL_CYC in (H0003, H0005 I0001) and not above
CTLLS	Certificate in teaching in the lifelong learning sector	Z_QLEVEL_CYC = C0006
DET	Diploma in Education and Training	Z_QLEVEL_CYC = I0008
DIPHE	Diploma of Higher Education (DipHE)	Z_QLEVEL_CYC = J0002
DTLLS	Diploma in teaching in the lifelong learning sector	Z_QLEVEL_CYC in (H0015, I0009)
FOUDEG	Foundation degree	Z_QLEVEL_CYC in (J0000, J0001)
HIGHCERT	Higher certificate	Z_QLEVEL_CYC = C0000
HNC	Higher national certificate	Z_QLEVEL_CYC = C0001
HND	Higher national diploma	Z_QLEVEL_CYC = J0003
PTLLS	Preparing to teach in the lifelong learning sector	Z_QLEVEL_CYC = C0005
OTHL[X]_Q	Other Level X qualification, where X is the level indicated by TRAWARDLEVELNUM	Z_QLEVEL_CYC in (C0002, C0003, C0004, C0007, C0009, E0000, E0001, E0002, E0004, H0006, H0007, H0008, H0012, H0019, I0004, J0004, J0005, J0006, J0007, J0010, J0012, M0008, M0010, M0011, M0013, M0015, M0017, M0018, M0021, M0023) or (Z_QLEVEL_CYC in (H0016, I0010) and PREREQUISITE ≠ 02)
OTHL[X]_U	Other Level X unit, where X is the level indicated by TRAWARDLEVELNUM	Z_QLEVEL_CYC in (C0008, D0004, E0003, H0018, I0012, J0011, L0002, M0022)
NONE	No qualification	Z_QLEVEL_CYC = Z9
FE	Not higher education	Z_QLEVEL_CYC = P0002

TRAWARDLEVEL

24. This field allocates the qualification awarded to the student during the reporting year to a broad grouping.

Value	Description	Definition
PHD	PhD and MPhil	TRAWARD_DETAIL in (PHD, HIGHER)
OPGR	Other postgraduate research	TRAWARD_DETAIL = OTHL7_Q_R
PGTM	Postgraduate taught masters'	TRAWARD_DETAIL = MASTER
PGCE	PGCE	TRAWARD_DETAIL = PGCE
OPGT	Other postgraduate taught	TRAWARD_DETAIL in (DTLLS_PG, OTHL7_Q, OTHL8_Q, PGCERT, PGDIP)
PUGD	Degrees including a postgraduate component	TRAWARD_DETAIL in (ENHANCED, MEDVETDENT)
PUGO	Other qualifications with a postgraduate component	TRAWARD_DETAIL = PROCONGRAD
PGCREDIT	Credit at a postgraduate level	TRAWARD_DETAIL in (OTHL7_CC, OTHL8_CC, OTHL7_U, OTHL8_U)
DEG	First degree	TRAWARD_DETAIL = FIRST
OUG	Other undergraduate	TRAWARD_DETAIL in (CERTED, CET, CTLLS, DET, DIPHE, DTLLS, FOUDEG, HIGHCERT, HND, HNC, PTLLS, UNICERT, OTHL6_Q, OTHL5_Q, OTHL4_Q, OTHHE_Q, UGUNSPEC)
UGCREDIT	Credit at an undergraduate level	TRAWARD_DETAIL in (FDBC, OTHL4_CC, OTHL5_CC, OTHL6_CC, OTHL4_U, OTHL5_U, OTHL6_U, OTHHE_CC, OTHHE_U)
NONE	No qualification	TRAWARD_DETAIL = NONE
FE	Not higher education	TRAWARD_DETAIL = FE

TRAWARD

25. This field simplifies the broad grouping of the qualification awarded to the student during the reporting year, allocated by TRAWARDLEVEL.

Value	Description	Definition
FDEG	First degree	TRAWARDLEVEL = DEG, PUGD
OUG	Other undergraduate	TRAWARDLEVEL = OUG, PUGO
OTH	Other including no award	Otherwise

TRAPPRENTICE

26. This field indicates whether the student is studying on an apprenticeship at any level. This field has the same definition as IPAPPRENTICE.

Value	Description	Definition
1	The student is studying on an apprenticeship at any level	At least one value of student initiative, STUINITID, in (004, 020) or
		(in the latest student course session, at least one value of course initiative, COURSEINITID, in (004, 020) and (COURSEINITVALIDFROM < SCSENDDATE or SCSENDDATE = BLANK) and (COURSEINITVALIDTO >= SCSSTARTDATE or COURSEINITVALIDTO = BLANK))
0	The student is not studying on an apprenticeship	Otherwise

TRMODE

27. This field classifies the qualification mode of study. This field has the same definition as IPMODE but does not classify qualifications as 'Writing up' – these are instead classified as 'Other'.

Value	Description	Definition
APP	Apprenticeship	TRAPPRENTICE = 1 and Z_MODEGRP2 in (01, 02, 03) and Z_ACT_CYC = 1
FT	Full-time	Z_MODEGRP2 in (01, 02) and Z_ACT_CYC = 1 and not above
PT	Part-time	Z_MODEGRP2 = 03 and Z_ACT_CYC = 1 and not above
OTH	Other	Otherwise

TRDOM

28. This field classifies whether the learner is domiciled in England, Northern Ireland, Scotland, Wales, outside of the UK or their domicile is unknown. This field has the same definition as IPDOM.

Value	Description	Definition
E	England	Z_PERMADDGRP4 = 01
N	Northern Ireland	Z_PERMADDGRP4 = 02
S	Scotland	Z_PERMADDGRP4 = 03
W	Wales	Z_PERMADDGRP4 = 04
EU	European Union	Z_PERMADDGRP4 = 06
OTHER	Other international	Z_PERMADDGRP4 = 05, 07
UNKNOWN	Unknown domicile	Z_PERMADDGRP4 = Z9

TRETHNIC

29. This field classifies the ethnicity of the student. This field has the same definition as IPETHNIC.

Value	Description	Definition
A	Asian	ETHNIC in (100, 101, 102, 103, 104, 119)
B	Black	ETHNIC in (120, 121, 139)
M	Mixed	ETHNIC in (140, 141, 142, 159)
W	White	ETHNIC in (160, 161, 162, 166, 167, 169, 179)
O	Other	ETHNIC in (163, 164, 165, 168, 170, 180, 899)
U	Refused, unknown or not collected	ETHNIC in (997, 998, 999)

TRSEX

30. This field classifies the sex of the student. This field has the same definition as IPSEX.

Value	Description	Definition
F	Female	SEXID = 10
M	Male	SEXID = 11
O	Other	SEXID = 12
U	Unknown	SEXID in (96, 99)

TREIMD

31. This field shows the English Index of Multiple Deprivation (EIMD) 2019 quintile of the postcode of the student's address prior to entry.[10] To assign an EIMD 2019 quintile to a student's address prior to entry, the postcode is matched to the National Statistics Postcode Lookup (NSPL) for the United Kingdom (May 2022 release) to determine the Lower Layer Super Output Area (LSOA) in which the postcode exists. The EIMD 2019 quintile of the LSOA is then assigned.

---

[10] For information on the English Index of Multiple Deprivation 2019, see English indices of deprivation 2019.

Value	Description	Definition
1, 2, 3, 4, 5	Assigned EIMD 2019 quintile 1 (most deprived) to 5 (least deprived)	TRDOM = E and POSTCODE matched to postcode and LSOA in NSPL
U	EIMD 2019 quintile unknown	TRDOM = E and (POSTCODE = BLANK or POSTCODE not matched to postcode and LSOA in NSPL)
NA	Student domicile outside of England or unknown domicile	TRDOM ≠ E

TRDEGCLASS

32. This field indicates the degree classification awarded.

Value	Description	Definition
1	First class honours degree	TRAWARD = FDEG and Z_QCLASS_CYC = 0001
2_1	Upper second class honours degree	TRAWARD = FDEG and Z_QCLASS_CYC = 0002
2_2	Lower second class honours degree	TRAWARD = FDEG and Z_QCLASS_CYC = 0003
3	Third class honours degree	TRAWARD = FDEG and Z_QCLASS_CYC = 0004
UN	Degree awarded without classification	TRAWARD = FDEG and not above
NA	Otherwise	Otherwise

TRATTAINPOP

33. This field indicates whether the student is included in the Transparency attainment population.

Value	Description	Definition
1	Student is counted in the Transparency attainment population	TRDOM = E, N, S, W and TRAWARD = FDEG, OUG and Z_POPQUAL_CYC = 1
0	Student is not counted in the Transparency attainment population.	Otherwise

Rebuild instructions

34. The individualised file, TR22_Attainment_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 individualised files are given on the OfS website.[11]

35. Specific instructions to rebuild the Transparency attainment data summary figures in the Transparency attainment data summary workbook, TR22_Attainment_DCT_XXXXXXXX.xlsx, are provided in this section.

36. Additional fields not described in the field tables or algorithms above have been included in the individualised file for reference. These additional fields are listed in Table 4.

Table 4: Additional fields included in the Transparency attainment data summary Student individualised file

Field name	Description
COURSEID	Course identifier
COURSETITLE	Course title
QUALPROPORTION	The proportion allocated to each subject contributing to a qualification

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

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

39. The headcount present in the summary workbook is the number of records present in the individualised data file once the relevant filters have been applied.

40. To filter to the Transparency attainment population, TRATTAINPOP = 1 should be selected.

41. Filters can then be applied to TRMODE, TRDEGCCLASS, TREIMD, TRETHNIC and TRSEX to reproduce the headcount of qualifiers (number of records) in the Transparency attainment summary workbook.

42. For example, to reproduce the headcount for upper second class honours degrees awarded to female qualifiers who studied full-time you would apply the following filters:

- TRATTAINPOP = 1

- TRDEGCLASS = 2_1

- TRMODE = FT

- TRSEX = F.

The number of records in the filtered dataset represents the headcount in TR22_Attainment_DCT_XXXXXXXX.xlsx.

43. To reproduce the headcount of classified first degrees awarded for a given mode of study and characteristic, filter to the required mode and characteristic split and select TRAWARD = 1, 2_1, 2_2, 3.

44. For example, to reproduce the headcount of classified first degrees awarded to qualifiers from EIMD quintile 1 areas who studied full-time you would apply the following filters:

- TRATTAINPOP = 1

- TRDEGCLASS = 1, 2_1, 2_2, 3

- TRMODE = FT

- TREIMD = 1

The number of records in the filtered dataset represents the headcount in TR22_Attainment_DCT_XXXXXXXX.xlsx.