

Office for
Students



Supporting information about constructing student outcome and experience indicators for use in OfS regulation

Instructions for rebuilding OfS datasets

Enquiries to providermetrics@officeforstudents.org.uk

Publication date 20 January 2022

Contents

Summary	2
Guidance for using this document	3
Enquiries and feedback	4
Rebuild instructions	5
Structure of this document	5
How to use the individualised student data files to rebuild a student outcome or experience indicator	5
How to select students for a particular view of student populations	6
How to rebuild student outcome and experience measures	7
Rebuilding split indicators	22
Rebuilding data about the size and shape of higher education provision	22
Annex A: Selecting course and student characteristics for the purposes of TEF and regulation of student outcomes	26
Annex B: Rebuilding split indicators for access and participation data resources	37
List of abbreviations used in this document	40

Summary

1. The Office for Students (OfS) has issued a consultation about the construction of student outcome and experience measures to be used in our regulation of student outcomes and in the Teaching Excellence Framework (TEF). This document has been published as supporting information alongside the consultation, to aid providers and other stakeholders in understanding the technical detail of the definitions and approaches we have proposed. We anticipate that some readers of the consultation proposals, particularly those at higher education providers with in-depth knowledge of the student data collected by the Higher Education Statistics Agency (HESA), or the Education and Skills Funding Agency (ESFA), will find the information in this supporting document useful for exploring the impact of our proposals on their own student data.
2. The definitions and approaches described within this document are directly aligned to our consultation proposals, and remain subject to change upon conclusion of the consultation exercise. We expect to publish similar information to that found in this document alongside publication of the consultation outcomes later in 2022, and when we implement our final approach to constructing student outcome and experience measures.
3. The adoption of the proposed definitions and approaches may impact other OfS functions that utilise the same measures of student experiences and outcomes as the TEF and student outcomes regulation, or otherwise rely on the same underlying data and categorisations of students and their activities in higher education. Definitions that contribute to the construction of indicators used within our regulation of access and participation are included in this document, with the purpose of illustrating the wider potential impact of the consultation proposals on the access and participation data dashboard in particular.
4. As a result, this document covers definitions that contribute to all of the data indicators that the OfS constructs as numerical measures for understanding the outcomes and experiences that a provider delivers for its students at different stages of the student lifecycle in higher education:
 - a. Access to higher education study.
 - b. Continuation in, and completion of, the study of higher education qualifications.
 - c. Student views and perceptions of different aspects of their higher education experience.
 - d. Achievement and the awards made to higher education students at the end of their studies.
 - e. Progression into the labour market and other destinations after leaving higher education.
5. The indicators 4a and 4d are only used for access and participation; we are proposing to use 4c only for the TEF, 4b and 4e for both regulation of student outcomes and the TEF. Throughout this document, content that is relevant only to the regulation of student outcomes is highlighted **green**, content that is relevant only to the TEF is highlighted in **salmon** and content that is relevant only to access and participation data dashboard is highlighted **mauve**.
6. To further support providers in exploring the practical effect of implementing our proposals within their own student data, the OfS has also released data resources to providers via the

OfS Portal (including individualised student data files and workbooks showing data and indicators at provider level, in dashboard and excel formats). Providers may wish to consider one or more of these resources when reading through this document.

Guidance for using this document

7. This is one of a series of supporting technical documents that provide details of the definitions and methods that the OfS has proposed to use in constructing student outcome and experience measures.¹ Readers may want to consider this document alongside the following document(s) and resources in particular:
 - a. Supporting information about constructing student outcome and experience indicators for use in OfS regulation: Description and methodology
 - b. Supporting information about constructing student outcome and experience indicators for use in OfS regulation: Core algorithms
 - c. Individualised student data supplied to individual providers, containing data relating only to their own students
 - d. Workbooks in dashboard and Excel formats, supplied to individual providers, showing their own and sector level data and indicators.
8. This document details how individualised student data can be used to rebuild the student outcome and experience indicators. It is intended to support higher education providers to gain a detailed understanding of the OfS's use of individualised student data returns in the construction of these indicators. This document is aimed at readers with in-depth knowledge of the HESA Student, HESA Student Alternative or Individualised Learner Record (ILR) student data, and with access to individualised student data files that the OfS has made available to providers via the OfS portal. The purpose of these individualised student data files is to support understanding of how a provider's own students have been categorised according to the OfS's algorithms and definitions, and how these can be used to reproduce (or drill further into) the student outcome and experience indicators that the OfS has produced.

Important information for using individualised files

1. Individualised student data has been supplied to individual providers, containing data relating only to their own students. These data files contain sensitive personal data items and show how individual students have been categorised according to the algorithms defined in 'Supporting information about constructing student outcome and experience indicators for use in OfS regulation: Core algorithms'. For data protection reasons, this level of data cannot be made publicly available, and providers must ensure that access to and use of individualised student-level data complies with the General Data Protection Regulation.

¹ See www.officeforstudents.org.uk/publications/student-outcomes-and-teaching-excellence-consultations/outcome-and-experience-data/.

2. Individualised student data files contain data relating to students who are registered by a provider, and are intended to provide transparency about the student outcome and experience measures that the OfS calculates.
3. The individualised files are provided as separate files for each academic year. The files names begin with IND_JAN2022_CONSULT_YY_XXXXXXXX, where YY identifies the academic year (e.g. '18' corresponds to academic year 2018-19) and XXXXXXXX is the provider's UKPRN. Individualised files are provided at **Common Aggregation Hierarchy (CAH) 3** level, meaning a student will have one row of data for every different **CAH level 3 group** they are studying.
4. This means that to replicate the headcounts shown in the indicators calculated by the OfS, users must **sum the SUBWT field** contained in each of the year-specific individualised files, and **total these across the years relevant** to a given indicator (as described through this document; note that for some indicators SUBWT needs to be weighted, as described in the indicator specific instructions). Summing all the rows in a single file for a particular field will give an inflated result when students have multiple rows per person.
5. If opening the individualised files in Microsoft Excel, users may wish to represent the year's data within a pivot table, to filter and summarise the data to best effect.

Enquiries and feedback

9. For enquiries regarding the methods described in this document and feedback, please contact providermetrics@officeforstudents.org.uk.

Rebuild instructions

Structure of this document

10. The remainder of this document will explain how to rebuild a provider's own indicators from their individualised student data files.
11. Paragraphs 16 to 19 describe how to rebuild a student outcome or experience indicator.
12. Paragraphs 20 to 24 explain how to extract information for different views of a provider's student population. Specifically, it will lay out how to select students that are:
 - a. Registered by the provider.
 - b. Taught or registered by the provider.
 - c. Taught by the provider.
 - d. Registered but not taught by the provider (these students are included in the partnership population, together with students that are neither registered nor taught, but whose awards are validated by the provider).
13. Paragraphs 25 to 73 detail how to calculate the student outcome and experience indicators from the student data. In all cases, the indicators are each shown separately for full-time, part-time and apprenticeship cohorts, and for the levels of study described by this document.
14. Paragraphs 79 to 93 describe how to rebuild summaries about the overall size and shape of provision.
15. Paragraphs 74 to 78 and Annexes A and B explain how to restrict the indicators data further to replicate split indicator figures.

How to use the individualised student data files to rebuild a student outcome or experience indicator

16. All student outcome and experience measures, except for those split by year, report an '**all years aggregate**' indicator based on available data across the most recent four student cohorts that are relevant to the measure in question. For progression, only two years of data are currently available, but the time series will be extended to four years as the data becomes available.
17. The years included in the time series are labelled year 1 to year 4, where year 1 is the earliest and year 4 the most recent. They correspond to different academic years depending on the measure and mode of study. For example, year 1 for full-time continuation indicators corresponds to measuring the outcomes of 2015-16 higher education entrants, while for part-time continuation indicators this is 2014-15 entrants. For full-time degree outcomes measures year 1 is 2016-17. For details see 'Description and methodology' at www.officeforstudents.org.uk/publications/student-outcomes-and-teaching-excellence-consultations/outcome-and-experience-data/.

18. Each individualised file contains data for a single academic year, as identified by the file name and by the field IPBASEYEAR in the data.
19. To rebuild a student outcome or experience indicator in a given mode of study:
- Select the individualised files that contain data for the years specified in the measure's rebuild instructions that follow below, through the conditions placed on IPBASEYEAR (split indicators showing one or more points of the four-year time series can be rebuilt using only the file corresponding to the academic year in question).
 - If using multiple years, it is useful to stack data from different files together. To facilitate this task, all individualised files include the same fields in the same order.
 - Filter the data to the desired view of a provider's student population (see paragraphs 20 to 24), mode and level of study (see indicator specific instructions) and, if needed, the relevant attribute for constructing a split indicator (paragraphs 74 to 78 and Annex A).
 - Calculate numerators and denominators by applying the population restrictions described through this document and summing either SUBWT, or SUBWT weighted by another field, as specified by each measure's instructions, across individualised files.
 - Divide numerators by denominators and multiply by 100 to obtain the indicators (calculation of the compound indicator to measure completion outcomes requires a different step, which is detailed below).

How to select students for a particular view of student populations

20. The individualised student data files include all students that are registered by a provider, regardless of which provider they are taught by.
21. The **registered population** is obtained by placing no restriction on the IPUKPRNRC or IPUKPRNTC fields. The registered population is only used for access and participation purposes.
22. The **taught or registered (TorR) population** comprises of all students that are either taught by or registered at a provider. It includes students that are covered by the registered population and students that are only taught by the provider (IPUKPRNRC = XXXXXXXX or IPUKPRNTC = XXXXXXXX, where XXXXXXXX is the UKPRN of the provider). This population cannot be rebuilt from the individualised files, for providers that teach students that are not registered with them. For providers that only teach students they have registered, the TorR population coincides with the registered population.
23. The **taught population** comprises of all students that are taught by a provider, regardless of which provider they are registered with. For providers that only teach students they have registered, the taught population can be obtained filtering to IPUKPRNTC = XXXXXXXX, where XXXXXXXX is the UKPRN of the provider. This population cannot be rebuilt from the individualised files for providers that teach students that are not registered with them. The taught population is only used for regulating student outcomes purposes.

24. The **partnership population**, which is only used for the regulation of student outcomes, is obtained by combining students that are:
- a. Registered but not taught by the provider: to obtain these from the individualised student data files, select records with IPUKPRNRC = XXXXXXXX and IPUKPRNTC ≠ XXXXXXXX, where XXXXXXXX is the UKPRN of the provider.
 - b. Neither registered nor taught, but whose awards are validated by the provider: these students are not included in the individualised student data files, for data protection reasons. While it follows that the partnership population cannot be rebuilt from a provider's individualised files if it validates awards for students it neither teaches nor registers, construction of the indicators which include these students follows exactly the same steps as described through this document. The OfS identifies students for whom a provider is the validating body, but neither teaching nor registering provider, by selecting records with IPAWARDBOD = XXXXXXXX and IPUKPRNRC ≠ XXXXXXXX and IPUKPRNTC ≠ XXXXXXXX, where XXXXXXXX is the UKPRN of the provider.

How to rebuild student outcome and experience measures

Continuation indicators

25. For **full-time and apprenticeship** students, continuation measures are constructed and reported for each cohort of entrants to higher education courses between 2015-16 and 2018-19. Their continuation outcomes are measured as at a census point one year and 15 days after the date on which they commenced their studies.
26. For **part-time** students, continuation measures are constructed and reported for each cohort of entrants to higher education courses between 2014-15 and 2017-18. Their continuation outcomes are measured as at a census point two years and 15 days after the date on which they commenced their studies.
27. A further description of the continuation measures is provided in paragraphs 35 to 47 of the document 'Descriptions and methodology'.
28. To rebuild a continuation indicator, firstly open the individualised files corresponding to the relevant years, as detailed by the IPBASEYEAR restrictions in the table below. Then select students who have studied at the relevant mode and level of study as illustrated in the table below. Postgraduate level indicators are only used for the purposes of student outcomes regulation. If constructing a split indicator, restrict further to students with the attribute relevant to the split indicator in question, using the filters described in paragraphs 74 to 78 and in Annexes A and B.

How to select students by...	Full-time	Part-time	Apprenticeship
Year	IPBASEYEAR = 2015, 2016, 2017, 2018	IPBASEYEAR = 2014, 2015, 2016, 2017	IPBASEYEAR = 2015, 2016, 2017, 2018
Mode	IPSTARTMODE = FT	IPSTARTMODE = PT	IPSTARTMODE = APPR
Level	<ul style="list-style-type: none"> • postgraduate research, defined by IPLEVEL = PHD • postgraduate taught masters', defined by IPLEVEL = PGTM • PGCE, defined by IPLEVEL = PGCE • other postgraduate, defined by IPLEVEL = OPGR, OPGT, PUGO • first degree, defined by IPLEVEL = DEG • other undergraduate, defined by IPLEVEL = OUG • undergraduate with postgraduate components, defined by IPLEVEL = PUGD • all undergraduate, defined by IPLEVEL = DEG, OUG, PUGD 		<ul style="list-style-type: none"> • all postgraduate, defined by IPLEVEL = PHD, PGTM, PGCE, OPGR, OPGT, PUGO • all undergraduate, defined by IPLEVEL = DEG, OUG, PUGD

29. Calculate numerator and denominator by applying the following population restrictions and adding up the **SUBWT** field. To obtain the indicator or split indicator, divide numerator by denominator and multiply by 100.

Population restrictions:

Either:

- apply DFAPAPPEXCL = 0 for access and participation data resources
- make no restriction for student outcome indicators used in quality and standards regulation and in the TEF

Then:

- Denominator of the indicator:
 - full-time and apprenticeship: IPENTRANTEXCL = 0 and IPCONINDFULL_Y1 ≠ TRANSFER
 - part-time: IPENTRANTEXCL = 0 and IPCONINDFULL_Y2 ≠ TRANSFER
- Numerator of the indicator:
 - full-time and apprenticeship: IPENTRANTEXCL = 0 and IPCONINDFULL_Y1 = QUALIFIED, CONTINUING, TRANSFER_COLLAB
 - part-time: IPENTRANTEXCL = 0 and IPCONINDFULL_Y2 = QUALIFIED, CONTINUING, TRANSFER_COLLAB

Completion indicators: cohort tracking

30. **Full-time and apprenticeship** completion measures constructed using the cohort tracking method are reported for each cohort of entrants to higher education courses between 2012-13 and 2015-16. Their completion outcomes are tracked to a census point four years and 15 days after the date on which they commenced their studies.

31. **Part-time** completion measures constructed using the cohort-tracking method are reported for each cohort of entrants to higher education courses between 2010-11 and 2013-14. Their completion outcomes are tracked to a census point six years and 15 days after the date on which they commenced their studies.

32. A further description of the cohort tracking methodology for completion measures is provided in paragraphs 48 to 60 of the document 'Descriptions and methodology'.

33. To rebuild a cohort-tracking completion indicator, firstly open the individualised files corresponding to the relevant years, as detailed by the IPBASEYEAR restrictions in the table below. Then select students who have studied at the relevant mode and level of study as illustrated in the table below. Postgraduate level indicators are only used for the purposes of student outcomes regulation. If constructing a split indicator, restrict further to students with the attribute relevant to the split indicator in question, using the filters described in paragraphs 74 to 78 and in Annex A.

How to select students by...	Full-time	Part-time	Apprenticeship
Year	IPBASEYEAR = 2012, 2013, 2014, 2015	IPBASEYEAR = 2010, 2011, 2012, 2013	IPBASEYEAR = 2012, 2013, 2014, 2015
Mode	IPSTARTMODE = FT	IPSTARTMODE = PT	IPSTARTMODE = APPR
Level	<ul style="list-style-type: none"> • postgraduate research, defined by IPLEVEL = PHD • postgraduate taught masters', defined by IPLEVEL = PGTM • PGCE, defined by IPLEVEL = PGCE • other postgraduate, defined by IPLEVEL = OPGR, OPGT, PUGO • first degree, defined by IPLEVEL = DEG • other undergraduate, defined by IPLEVEL = OUG • undergraduate with postgraduate components, defined by IPLEVEL = PUGD • all undergraduate, defined by IPLEVEL= DEG, OUG, PUGD 		<ul style="list-style-type: none"> • all postgraduate, defined by IPLEVEL = PHD, PGTM, PGCE, OPGR, OPGT, PUGO • all undergraduate, defined by IPLEVEL= DEG, OUG, PUGD

34. Calculate numerator and denominator by applying the following population restrictions and adding up the **SUBWT** field. To obtain the indicator or split indicator, divide numerator by denominator and multiply by 100.

Population restrictions:

Either:

- apply DFAPAPPEXCL = 0 for access and participation data resources
- make no restriction for student outcome indicators used in quality and standards regulation and in the TEF

Then:

- Denominator of the indicator:
 - full-time and apprenticeship: IPENTRANTEXCL = 0 and IPCONINDFULL_Y4 ≠ TRANSFER
 - part-time: IPENTRANTEXCL = 0 and IPCONINDFULL_Y6 ≠ TRANSFER
- Numerator of the indicator:
 - full-time and apprenticeship: IPENTRANTEXCL = 0 and IPCONINDFULL_Y4 = QUALIFIED, CONTINUING, TRANSFER_COLLAB
 - part-time: IPENTRANTEXCL = 0 and IPCONINDFULL_Y6 = QUALIFIED, CONTINUING, TRANSFER_COLLAB

Completion indicators: compound indicator

35. The completion measures constructed using the compound indicator use withdrawal proportions in a particular academic year to inform the measure, looking at withdrawals from different stages of study. The measures are informed by students who withdrew from higher education courses between 2016-17 and 2019-20.

36. A further description of the compound indicator methodology for completion measures is provided in paragraphs 61 to 73 of the document 'Descriptions and methodology'.

37. A student is considered to have withdrawn in a given year if either of the following applies:

- a. They are recorded in the HESA or ILR student records with a date of leaving the programme of study that falls within that academic year and all of the following are true:
 - They have not qualified at any higher education level
 - They have not changed to another engagement at the same provider with a different broad level of study
 - They have not transferred to study a higher education qualification at another provider.
- b. It is the second consecutive year in which the student has been recorded as dormant.

38. A compound indicator is created **for a given year** by:

- a. Identifying all of the students who withdraw from higher education study at the provider in question in that year.
- b. Dividing this group of withdrawing students into six entry cohorts, each defined by the year in which they started their programme of study. The indicator uses six entry cohorts,

meaning it will include the cohort that started in the same year in which we have identified the withdrawals, and in the five previous years.

- c. Identifying the number of students who started a relevant higher education qualification in each of the six corresponding entry years.
- d. Calculating six cohort withdrawal proportions as the number of withdrawing students in each entry cohort, divided by the number of students who started in the corresponding entry year.
- e. Totalling the cohort withdrawal proportions across all six entry cohorts, and subtracting this total from 100 per cent to give the compound completion indicator.

39. Table 1 below shows the six entry cohort years used to calculate the compound indicators for each of the most recent four years of available data. The year the indicator refers to is the column heading, and the year of withdrawal. The entry cohort years referred to in the row labels work back from the year of withdrawal in increments of one. This creates a matrix, within which each cell represents the cohort withdrawal proportion to be calculated. For instance, for the 2016-17 compound indicator, the entry cohort year four of six refers to higher education entry in 2013-14: this cell is highlighted yellow in Table 1 to help illustrate the example that follows. The cohort withdrawal proportion for this cell is calculated as the rate at which students who started courses in 2013-14 entrants withdraw from study in 2016-17. All of the other cells in the matrix can be interpreted in the same way, to give the six cohort withdrawal proportions required for each of the four separate compound indicators.

Table 1: Entry cohort years relevant to calculation of the six cohort withdrawal proportions required to construct the compound indicator for each withdrawal year 2016-17 to 2019-20

	2016-17 withdrawal year, basis for the 2016-17 compound indicator	2017-18 withdrawal year, basis for the 2017-18 compound indicator	2018-19 withdrawal year, basis for the 2018-19 compound indicator	2019-20 withdrawal year, basis for the 2019-20 compound indicator
Entry cohort year one	2016-17	2017-18	2018-19	2019-20
Entry cohort year two	2015-16	2016-17	2017-18	2018-19
Entry cohort year three	2014-15	2015-16	2016-17	2017-18
Entry cohort year four	2013-14	2014-15	2015-16	2016-17
Entry cohort year five	2012-13	2013-14	2014-15	2015-16
Entry cohort year six	2011-12	2012-13	2013-14	2014-15
Compound indicator for the given year	100% - sum of this column	100% - sum of this column	100% - sum of this column	100% - sum of this column

40. A compound indicator is created for the **aggregate of the most recent four years** in the same way as the single year indicator, but the six cohort entry points (e.g. entry cohort year one in Table 1) each combine students who contribute to the calculations across the four single years. For each cohort entry point the withdrawal proportion is calculated by following these steps:

- a. Sum the number of students who contribute to the 2016-17, 2017-18, 2018-19, 2019-20 compound indicators as withdrawing from study, having started at the cohort entry point in question.
- b. Sum the number of students who contribute to the 2016-17, 2017-18, 2018-19, 2019-20 compound indicators as starting their study at the cohort entry point in question.
- c. Calculate the cohort withdrawal proportion for the cohort entry point in question as (a) divided by (b).

41. For example, the withdrawal proportion for the all-years aggregate for cohort entry year two is:

- a. Sum the number of students who withdrew from study on the following basis:
 - i. 2015-16 entrants withdrawing in 2016-17 (contributes to the 2016-17 compound indicator)
 - ii. 2016-17 entrants withdrawing in 2017-18 (contributes to the 2017-18 compound indicator)
 - iii. 2017-18 entrants withdrawing in 2018-19 (contributes to the 2018-19 compound indicator)
 - iv. 2018-19 entrants withdrawing in 2019-20 (contributes to the 2019-20 compound indicator).
- b. Sum the number of students who started their study in 2015-16, 2016-17, 2017-18 and 2018-19.
- c. Calculate the cohort withdrawal proportion for cohort entry year two as a) divided by b).

42. The all years aggregate compound completion indicator is calculated by summing the cohort withdrawal proportions across all six cohort entry points, and subtracting this total from 100 per cent to give the compound completion indicator. This is further illustrated in Table 2.

Table 2: Calculating cohort entry year withdrawal proportions for the purposes of constructing the compound indicator for the aggregate of the most recent four years

Withdrawing students by entry cohort year (step a)	Entry cohort year entrants (step b)	Entry cohort year withdrawal proportion (step a, divided by step b)
2016-17 entrants who withdrew in 2016-17 + 2017-18 entrants who withdrew in 2017-18 + 2018-19 entrants who withdrew in 2018-19 + 2019-20 entrants who withdrew in 2019-20 = entry cohort year one numerator	2016-17 entrants + 2017-18 entrants + 2018-19 entrants +	Entry cohort year one numerator / Entry cohort year one denominator

Withdrawing students by entry cohort year (step a)	Entry cohort year entrants (step b)	Entry cohort year withdrawal proportion (step a, divided by step b)
	2019-20 entrants = entry cohort year one denominator	
2015-16 entrants who withdrew in 2016-17 + 2016-17 entrants who withdrew in 2017-18 + 2017-18 entrants who withdrew in 2018-19 + 2018-19 entrants who withdrew in 2019-20 = entry cohort year two numerator	2015-16 entrants + 2016-17 entrants + 2017-18 entrants + 2018-19 entrants = entry cohort year two denominator	Entry cohort year two numerator / Entry cohort year two denominator
2014-15 entrants who withdrew in 2016-17 + 2015-16 entrants who withdrew in 2017-18 + 2016-17 entrants who withdrew in 2018-19 + 2017-18 entrants who withdrew in 2019-20 = entry cohort year three numerator	2014-15 entrants + 2015-16 entrants + 2016-17 entrants + 2017-18 entrants = entry cohort year three denominator	Entry cohort year three numerator / Entry cohort year three denominator
2013-14 entrants who withdrew in 2016-17 + 2014-15 entrants who withdrew in 2017-18 + 2015-16 entrants who withdrew in 2018-19 + 2016-17 entrants who withdrew in 2019-20 = entry cohort year four numerator	2013-14 entrants + 2014-15 entrants + 2015-16 entrants + 2016-17 entrants = entry cohort year four denominator	Entry cohort year four numerator / Entry cohort year four denominator
2012-13 entrants who withdrew in 2016-17 + 2013-14 entrants who withdrew in 2017-18 + 2014-15 entrants who withdrew in 2018-19 + 2015-16 entrants who withdrew in 2019-20 = entry cohort year five numerator	2012-13 entrants + 2013-14 entrants + 2014-15 entrants + 2015-16 entrants = entry cohort year five denominator	Entry cohort year five numerator / Entry cohort year five denominator
2011-12 entrants who withdrew in 2016-17 + 2012-13 entrants who withdrew in 2017-18 + 2013-14 entrants who withdrew in 2018-19 + 2014-15 entrants who withdrew in 2019-20 = entry cohort year six numerator	2011-12 entrants + 2012-13 entrants + 2013-14 entrants + 2014-15 entrants = entry cohort year six denominator	Entry cohort year six numerator / Entry cohort year six denominator
All years aggregate compound indicator		100% - sum of this column

43. To rebuild a compound completion indicator, firstly open the individualised files corresponding to the relevant years of entry and years of withdrawal. Then select students who have studied at the relevant mode and level of study as illustrated in the table below. Postgraduate level indicators are only used for the purposes of student outcomes regulation. If constructing a split indicator, restrict further to students with the attribute relevant to the split indicator in question, using the filters described in Annex A. Instructions to rebuild a single year indicator are included in the Population restrictions (yellow box) below.

How to select students by...	Full-time	Part-time	Apprenticeship
Year	Year of withdrawal = 2016, 2017, 2018, 2019. Year of entry as per Table 1 above.		
Mode	IPSTARTMODE = FT	IPSTARTMODE = PT	IPSTARTMODE = APPR
Level	<ul style="list-style-type: none"> • postgraduate research, defined by IPLEVEL = PHD • postgraduate taught masters', defined by IPLEVEL = PGTM • PGCE, defined by IPLEVEL = PGCE • other postgraduate, defined by IPLEVEL = OPGR, OPGT, PUGO • first degree, defined by IPLEVEL = DEG • other undergraduate, defined by IPLEVEL = OUG • undergraduate with postgraduate components, defined by IPLEVEL = PUGD • all undergraduates, defined by IPLEVEL = DEG, OUG, PUGD 		<ul style="list-style-type: none"> • all postgraduates, defined by IPLEVEL = PHD, PGTM, PGCE, OPGR, OPGT, PUGO • all undergraduates, defined by IPLEVEL = DEG, OUG, PUGD

Population restrictions:

Single year indicator

a. Calculate six withdrawal proportions for each of the six cohort entry years corresponding to the withdrawal year in question (as per Table 1), by applying the following population restrictions and adding up the **SUBWT** field:

- Numerator of the withdrawal proportion: IPENTRANTEXCL = 0² and IPCIRESLT = DORMANT, NO QUALIFICATION and IPBASEYEAR = cohort entry year and IPCIENDED = withdrawal year.

² A student's entry cohort can normally be established from the year of their IPCOMDATE, but the OfS actually relies on data linking to determine this: a student's entry cohort year will be the year of data in which that individual has been included in the individualised files released to providers with IPENTRANTEXCL = 0.

- Denominator of the withdrawal proportion: IPENTRANTEXCL = 0 and IPBASEYEAR = cohort entry year. Exclude students with IPCIRESLT = TRANSFER EXTERNAL and IPCIENDED ≤ (IPBASEYEAR + 5).

b. Calculate the compound completion indicator by summing the six cohort withdrawal proportions and subtracting the total from 100 per cent.

All years aggregate indicator

c. An all years aggregate compound indicator can be calculated using the method described above for the single year indicator, by summing the four entry cohort year numerators and denominators (as shown in Table 2), prior to calculating the withdrawal proportions. For instance, for the withdrawal proportion **for cohort entry year two** apply the following restrictions:

- Numerator of the withdrawal proportion: IPENTRANTEXCL = 0 and IPCIRESLT = DORMANT, NO QUALIFICATION and either (IPBASEYEAR = 2015-16 and IPCIENDED = 2016-17) or (IPBASEYEAR = 2016-17 and IPCIENDED = 2017-18) or (IPBASEYEAR = 2017-18 and IPCIENDED = 2018-19) or (IPBASEYEAR = 2018-19 and IPCIENDED = 2019-20).
- Denominator of the withdrawal proportion: IPENTRANTEXCL = 0 and IPBASEYEAR = 2015-16, 2016-17, 2017-18, 2018-19. Exclude students with IPCIRESLT = TRANSFER EXTERNAL and IPCIENDED ≤ (IPBASEYEAR + 5).

d. Calculate withdrawal proportions for the other five cohort entry years, adapting the years restrictions as described in Table 2.

e. Calculate the compound completion indicator by summing the six cohort withdrawal proportions and subtracting the total from 100 per cent.

Progression indicators

44. Measures of progression to managerial and professional employment or further study beyond graduation are constructed based on responses to the Graduate Outcomes (GO) survey.

45. Progression measures are constructed and reported for each cohort of leavers from higher education qualifications in 2017-18 and 2018-19, as currently data is only available for these two years. The time series will be extended as more data becomes available.

46. A further description of the progression methodology is provided in paragraphs 74 to 90 of the document 'Description and methodology'.

47. To rebuild a progression indicator, firstly open the individualised files corresponding to the academic years 2017-18 and 2018-19 (IPBASEYEAR = 2017, 2018), or a single file for single year indicators. Then select students who have studied at the relevant level and mode of study. Full-time students can be identified using IPSTARTMODE = FT, part-time students can be identified using IPSTARTMODE = PT and apprenticeships students can be identified using IPSTARTMODE = APPR.

48. Progression indicators are reported separately for leavers at each of the following levels (postgraduate level indicators are only used for the purposes of student outcomes regulation):
- a. postgraduate research, defined by IPLEVEL= PHD
 - b. postgraduate taught masters', defined by IPLEVEL = PGTM
 - c. PGCE, defined by IPLEVEL = PGCE
 - d. other postgraduate, defined by IPLEVEL = OPGR, OPGT, PUGO
 - e. first degree, defined by IPLEVEL = DEG
 - f. other undergraduate, defined by IPLEVEL = OUG
 - g. undergraduate with postgraduate components, defined by IPLEVEL = PUGD
 - h. all undergraduates, defined by IPLEVEL = DEG, OUG, PUGD
49. If constructing a split indicator, restrict further to students with the attribute relevant to the split indicator in question, using the filters described in paragraphs 74 to 78 and in Annexes A and B.
50. Calculate numerator by applying the following population restrictions and adding up the **SUBWT** field, **multiplied by IPEMPINDNUM**. Calculate the denominator by applying the following population restrictions and adding up the field SUBWT. To obtain the indicator, divide numerator by denominator and multiply by 100.
51. Response rates are calculated in the same way, using the restrictions below. For the progression indicators to be reportable, the response rate for the GO survey must be above the minimum threshold of 50 per cent of the target population. Response rates are calculated separately for each provider view, mode and level of study, and for each split indicator.

Population restrictions:

Either:

- apply DFAPAPPEXCL = 0 for access and participation data resources
- make no restriction for student outcome indicators used in quality and standards regulation and in the TEF

Then:

- Denominator of the **indicator**: IPEMPINDPOP = 1
- Numerator of the **indicator**: SUBWT x IPEMPINDNUM, where IPEMPINDNUM ≠ 0 and IPEMPINDPOP = 1
- Denominator of the **response rate**: IPEMPEXCL = 0
- Numerator of the **response rate**: IPEMPRRNUM = 1

Student experience indicators

52. Our related consultations have proposed that student experience measures are used for TEF purposes. These measures are constructed using data from the National Student Survey (NSS) and report the percentage of students whose survey response identified that they were satisfied with different aspects of their student experience. The measures are constructed for each of the question scales included in the NSS questionnaire:

- the teaching on my course
- learning opportunities
- assessment and feedback
- academic support
- organisation and management
- learning resources
- learning communities
- student voice
- overall satisfaction.

53. A further description of the student experience measures is provided in paragraphs 91 to 100 of the document 'Description and methodology'.

54. Student experience measures are constructed and reported for students on higher education courses surveyed between 2017-18 and 2020-21, corresponding to IPBASEYEAR = 2016 to 2019, which defines the survey's target list.

55. The individualised student data files do not include the relevant fields for rebuilding student experience indicators, due to these survey responses being strictly confidential and requiring anonymisation before they are made available to providers. While this means that these indicators cannot be rebuilt using the individualised student data files, we describe the process here for completeness.

56. To rebuild student experience indicators, data is selected for the years 2016 to 2019, or a single year for single year split indicators. Full-time students can be identified using IPSTARTMODE = FT, part-time students can be identified using IPSTARTMODE = PT and apprenticeship students can be identified using IPSTARTMODE = APPR.

57. Responses are reported separately for undergraduate students at each of the following levels:

- a. first degree, defined by IPLEVEL = DEG
- b. other undergraduate, defined by IPLEVEL = OUG
- c. undergraduate with postgraduate components, defined by IPLEVEL = PUGD

d. all undergraduates, defined by IPLEVEL= DEG, OUG, PUGD

58. If constructing a split indicator, the datasets are restricted further to students with the attribute relevant to the split indicator in question, using the filters described in paragraphs 74 to 78 and in Annex A.

59. The variables IPNSSxxxxRESPOND, where xxxx identifies the scale, represent the count of questions within that scale which had a valid response. The variables IPNSSxxxxAGREE, where xxxx identifies the scale, represent the count of questions within that scale which the student mostly or definitely agrees with. Counts of questions need to be weighted by SUBWT.

60. The numerator is calculated by applying the population restrictions below and adding up the **SUBWT field multiplied by (IPNSSxxxxAGREE/ IPNSSxxxxRESPOND)**. The denominator is calculated by applying the population restrictions described below and adding up the field SUBWT. To obtain the indicator, the numerator is divided by denominator and multiplied by 100.

Table 3: Calculation of numerators for student experience indicators

Scale	Percentage agree calculation
The teaching on my course	$(IPNSSSTEACHAGREE / IPNSSSTEACHRESPOND) \times SUBWT$
Learning opportunities	$(IPNSSLOPPAGREE / IPNSSLOPPRESPOND) \times SUBWT$
Assessment and feedback	$(IPNSSASSESSAGREE / IPNSSASSESSRESPOND) \times SUBWT$
Academic support	$(IPNSSACADAGREE / IPNSSACADRESPOND) \times SUBWT$
Organisation and management	$(IPNSSORGAGREE / IPNSSORGRESPOND) \times SUBWT$
Learning resources	$(IPNSSLRESAGREE / IPNSSLRESRESPOND) \times SUBWT$
Learning communities	$(IPNSSLCOMAGREE / IPNSSLCOMRESPOND) \times SUBWT$
Student voice	$(IPNSSSVOCAGREE / IPNSSSVOCRESPOND) \times SUBWT$
Overall satisfaction	$(IPNSSSOVSATAGREE / IPNSSSOVSATRESPOND) \times SUBWT$

61. Response rates are calculated using the restrictions below. For the student experience indicators to be reportable, the response rate for the NSS survey as a whole must be above the

minimum threshold of 50 per cent of the target population. Response rates are calculated separately for each provider view, mode and level of study, and for each split indicator.

Population restrictions:

Student experience indicators cannot be rebuilt using the individualised student data files, as they do not include the relevant fields due to data protection. Denominators and numerators for the student experience indicators are calculated as follows:

- Denominator of the **indicator**: $IPNSSRESPONSE = 1$ and $IPNSSINDEXCL = 0$ and $IPNSSxxxxRESPOND > 0$
- Numerator of the **indicator**: $(IPNSSxxxxAGREE / IPNSSxxxxRESPOND) \times SUBWT$, where $IPNSSRESPONSE = 1$ and $IPNSSINDEXCL = 0$
- Denominator of the **response rate**: $IPNSSRESRATEEXCL = 1$
- Numerator of the **response rate**: $IPNSSRESPONSE = 1$

Access to higher education indicators

62. Access to higher education measures are used in the access and participation data dashboard. They are constructed and reported for each cohort of UK-domiciled undergraduate entrants on higher education courses. The indicator expresses the number of entrants with a particular attribute as a percentage of all entrants, referenced where possible to the UK population of 18-year-olds who possess the same attribute.
63. A further description of the access indicators methodology is provided in paragraphs 101 to 111 of the document 'Description and methodology'.
64. To rebuild an access indicator, firstly open the individualised files corresponding to the academic years 2016-17 to 2019-20 ($IPBASEYEAR = 2016, 2017, 2018, 2019$). Full-time students can be identified using $IPSTARTMODE=FT$, part-time students can be identified using $IPSTARTMODE = PT$ and apprenticeship students can be identified using $IPSTARTMODE = APPR$.
65. Access indicators are reported separately for entrants at each of the following levels:
 - a. first degree, defined by $IPLEVEL = DEG$
 - b. other undergraduate, defined by $IPLEVEL = OUG$
 - c. undergraduate with postgraduate components, defined by $IPLEVEL = PUGD$
 - d. all undergraduates, defined by $IPLEVEL= DEG, OUG, PUGD$
66. If constructing a split indicator, the datasets are restricted further to students with the attribute relevant to the split indicator in question, using the filters described in paragraphs 74 to 78 and in Annex B.

67. Calculate numerator and denominator by applying the following population restrictions and adding up the **SUBWT** field. To obtain the indicator or split indicator, divide numerator by denominator and multiply by 100.

Population restrictions:

- Denominator of the indicator: DFAPAPPEXCL = 0 and IPACCEXCL = 0 and all students in scope for the split type
- Numerator of the indicator: DFAPAPPEXCL = 0 and IPACCEXCL = 0 and students with the attribute

Degree outcomes indicators

68. Degree outcomes measures are used in the access and participation data dashboard. They are constructed and reported for each cohort of undergraduate degree qualifiers who were awarded classified first degrees between 2016-17 and 2019-20. Degree outcomes are measured as the percentage of leavers that were awarded 'first' or 'upper second' (2:1) degree classifications.

69. A further description of the degree outcomes methodology is provided in paragraphs 112 to 116 of the document 'Description and methodology'.

70. To rebuild a degree outcomes indicator, firstly open the individualised files corresponding to the academic years 2016-17 to 2019-20 (IPBASEYEAR = 2016, 2017, 2018, 2019) who have studied at the relevant level and mode of study. Full-time students can be identified using IPSTARTMODE=FT, part-time students can be identified using IPSTARTMODE = PT and apprenticeship students can be identified using IPSTARTMODE = APPR.

71. Degree outcomes measures are only reported for undergraduate degree qualifiers (level 6+, identified using IPAWARDLEVEL = DEG, PUGD) who were awarded classified degrees (IPDOEGCLASS ≠ UNCLASS).

72. If constructing a split indicator, the datasets are restricted further to students with the attribute relevant to the split indicator in question, using the filters described in paragraphs 74 to 78 and in Annex B.

73. Calculate numerator and denominator by applying the following population restrictions and adding up the **SUBWT** field. To obtain the indicator or split indicator, divide numerator by denominator and multiply by 100.

Population restrictions:

- Denominator of the indicator: DFAPAPPEXCL = 0 and IPDOQUALPOP = 1
- Numerator of the indicator: DFAPAPPEXCL = 0 and IPDOQUALPOP = 1 and IPDOEGCLASS = FIRST, 2_1

Rebuilding split indicators

74. Indicators are also reported in the form of a series of split indicators. These split indicators break down the indicator populations derived by following the instructions above and report them for various study, course and student characteristics.
75. **All years aggregate** indicators include all relevant students, without any further break down. They can be rebuilt using data from the individualised files that correspond to the relevant years for each indicator, restricting only by mode and level of study, as described in paragraph **Error! Reference source not found.**
76. **Year split** indicators can be rebuilt by using a single individualised student data file, for the required academic year. Year 1 to 4 correspond to different years depending on the indicator and mode, as described above in the indicator specific instructions. The earliest year in the time series is labelled 'Year 1' and the most recent year is labelled 'Year 4'.
77. The **study and course split** indicators produced for the purposes of TEF and student outcomes regulation datasets can be rebuilt using the filters and variables described in Annex A.
78. The **student characteristic split indicators for the purposes of TEF and regulating student outcomes datasets** can be rebuilt using the filters and variables described in Annex A. The student characteristic split indicators **for the purposes of access and participation data resources** would be rebuilt using the filters and variables described in Annex B.

Rebuilding data about the size and shape of higher education provision

79. Additional data is provided on student numbers and on the overall size and shape of provision **for the purposes of TEF and regulating student outcomes datasets only:**
- a. Summary of student numbers
 - b. Partnership arrangements
 - c. Size and shape of provision.
80. In all tables, data is presented for the populations of all students, entrants and qualifiers, which can be selected using the following filters.

Population restrictions for all tables:

- All students: IPCONTEXTPOP = 1, 2, 3, 4
- Entrants: IPCONTEXTPOP = 1, 2
- Qualifiers: IPCONTEXTPOP = 1, 3

81. The data is expressed as headcount student numbers, as well as in percentage terms. Headcount numbers can be derived by summing up the variable SUBWT. Percentages express the proportion of the corresponding count data within the column.

82. All aspects of the summary data are available for each year from 2016-17 to 2019-20 individually and for an aggregate of the four years. Individual years of data can be rebuilt by using the IPBASEYEAR variable.

Summary of student numbers

83. In this table the overall, entrant and qualifier student populations are broken down between students studying mainly in the UK, students enrolled in offshore transnational education (TNE) and students studying mainly abroad.

Population restrictions:

- Student studying mainly in the UK: IPHECAT = 2, 5 and IPAYDUP = 0
- Offshore transnational education: the underlying data for this part of the summary table is not available as part of the individualised files
- Students studying mainly abroad: IPHECAT =1 and IPAYDUP=0

84. Students studying mainly in the UK are classified by mode (full-time, part-time and apprenticeship) and level of study, plus those studying for credit or modules. Table 4 below explains what variables and filters should be applied to rebuild these numbers filtering by mode and level of study.

Table 4: Variables to use to rebuild mode and level of study

Data item	Filter
Full-time	IPSTARTMODE= FT
Part-time	IPSTARTMODE = PT
Apprenticeship	IPSTARTMODE = APPR
All undergraduates	IPLEVEL = DEG, OUG, PUGD
All other undergraduate	IPLEVEL= OUG
Other undergraduate level 5+	IPLEVEL = OUG and IPLEVELNUM = 5,6
Other undergraduate level 4	IPLEVEL= OUG and IPLEVELNUM= 0,4
First degree	IPLEVEL= DEG
Undergraduate with postgraduate components	IPLEVEL= PUGD
All postgraduates	IPLEVEL = OPGT, PUGO, PGTM, PHD, OPGR, PGCE
Other postgraduate	IPLEVEL= OPGT, PUGO, OPGR
PGCE	IPLEVEL= PGCE
Postgraduate taught masters'	IPLEVEL= PGTM

Data item	Filter
Postgraduate research	IPELLEVEL= PHD
Unspecified qualification aim	IPELLEVEL= UGUNSPEC, PGUNSPEC
Credit or modules	IPELLEVEL= UGCREDIT, PGCREDIT

85. Data on **offshore TNE** is sourced from the HESA Aggregate Offshore record (AOR). Table 5 shows how the levels of study recorded on the AOR dataset have been mapped to the levels of study reported.

Table 5: Mapping between levels of study recorded in the AOR and level of study reported

Level of study reported	Level of study recorded in the AOR dataset
All undergraduates	Level = C, H, I, J
Other undergraduate	Level = J, C
First degree	Level = H, I
All postgraduates	Level = D, E, L, M
Postgraduate taught masters'	Level= M
Postgraduate taught doctorates	Level = E
Postgraduate research	Level = D, L

86. The population of **students studying mainly abroad** is not disaggregated by mode of study. Table 4 above provides details about how to rebuild the data for particular levels of study.

Partnership arrangements

87. The partnership arrangements table shows the population of students studying mainly in the UK.

Population restrictions:

Students studying mainly in the UK: IPHECAT = 2, 5 and IPAYDUP = 0

88. Data is broken down by student population (all students, entrants and qualifiers), by type of partnership, mode and broad level of study (undergraduate or postgraduate). The table below details how to select students within different types of partnership, and which summaries can be rebuilt from individualised data files. Mode and broad level of study can be rebuilt using the restrictions described in Table 4 above.

All registered or subcontracted in (taught or registered population)	IPIKPRNRC = XXXXXXXX or IPIKPRNTC = XXXXXXXX, where XXXXXXXX is the provider's 8-digit UKPRN	Can be rebuilt only for providers with no subcontracted in students
Registered and taught by this provider	IPIKPRNRC = XXXXXXXX and IPIKPRNTC = XXXXXXXX	Can be rebuilt for all providers

Only registered by this provider (subcontracted out)	IPUKPRNRC = XXXXXXXXX and IPUKPRNTC ≠ XXXXXXXXX	Can be rebuilt for all providers
Only taught by this provider (subcontracted in)	IPUKPRNRC ≠ XXXXXXXXX and IPUKPRNTC = XXXXXXXXX	Cannot be rebuilt
Only validated by this provider	IPUKPRNRC ≠ XXXXXXXXX and IPUKPRNTC ≠ XXXXXXXXX and IPAWARDBOD = XXXXXXXXX	Cannot be rebuilt

89. Percentages in this table are calculated by column, within a particular student population and type of partnership. The sum of all percentages in a year and within the same student population, type of partnership, mode and level of study equals 100.

Size and shape of provision

90. The size and shape of provision tables include only students who are mainly in the UK and actively studying for a qualification aim.

Population restrictions:

Students studying mainly in the UK and actively studying for a qualification aim: IPHECAT = 5 and IPAYDUP = 0

91. Each table reports on a single year of data. The columns group student numbers by mode and broad level of study (undergraduate and postgraduate). Table 4 above details how to select students based on these characteristics. In the rows, the data is broken down by student population (all students, entrants and qualifiers) and then by course and student characteristics as per Annex A.

92. Percentages in this table are calculated by column, within a particular student population and course/student characteristic. For each mode and level of study, percentages within a given student population and characteristic add up to 100.

93. The course and student characteristics are listed and described in Annex A.

Annex A: Selecting course and student characteristics for the purposes of TEF and regulation of student outcomes

1. The filters and variables described in Table A1 can be used to rebuild split indicators and summaries on the size and shape of provision based on **study and course characteristics**.

Table A1: Study and course characteristics and rebuild instructions for split indicators and for data on size and shape of provision

Attributes and rebuild instructions for...		
Characteristic	Split indicators	Data on size and shape of provision
Teaching arrangements	<ul style="list-style-type: none"> • IPUKPRNRC = XXXXXXXX and IPUKPRNTC = XXXXXXXX for Registered and taught by this provider, where XXXXXXXX is the provider's 8-digit UKPRN • IPUKPRNRC = XXXXXXXX and IPUKPRNTC ≠ XXXXXXXX for Only registered by this provider (subcontracted out) • IPUKPRNRC ≠ XXXXXXXX and IPUKPRNTC = XXXXXXXX for Only taught by this provider (subcontracted in) • IPAWARDBOD = XXXXXXXX and IPUKPRNRC ≠ XXXXXXXX and IPUKPRNTC ≠ XXXXXXXX for Only validated by this provider (this split is only used in regulation of student outcomes) 	<p>Same as for split indicators. In addition:</p> <ul style="list-style-type: none"> • IPUKPRNRC = XXXXXXXX or IPUKPRNTC = XXXXXXXX, where XXXXXXXX is the provider's 8-digit UKPRN for All registered or subcontracted in (taught or registered population)
Course type	<ul style="list-style-type: none"> • ILEVEL= OUG and ILEVELNUM= 5, 6 for Other undergraduate level 5+ • ILEVEL= OUG and ILEVELNUM= 0, 4 for Other undergraduate level 4 	<ul style="list-style-type: none"> • ILEVEL = DEG and IPSTARTMODE = FT and IPFOUNDEYEAR= 1 for First degree with integrated foundation year

Attributes and rebuild instructions for...		
Characteristic	Split indicators	Data on size and shape of provision
	<ul style="list-style-type: none"> • IPLEVEL = DEG and IPSTARTMODE = FT and IPFOUNDEYEAR = 1 for First degree with integrated foundation year 	
Subject	<p>(Note: for student experience indicators, if IPINTERCALATE = 1 and IPSBJ_CAH2 = CAH01-01 or CAH05-01, then IPSBJ_CAH2 should be replaced by IPINTSBJ_CAH2 in the conditions below)</p> <ul style="list-style-type: none"> • IPSBJ_CAH2 = CAH17-01 for Business and management • IPSBJ_CAH2 = CAH25-01 for Creative arts and design • IPSBJ_CAH2 = CAH25-02 for Performing arts • IPSBJ_CAH2 = CAH22-01 for Education and teaching • IPSBJ_CAH2 = CAH11-01 for Computing • IPSBJ_CAH2 = CAH10-01 for Engineering • IPSBJ_CAH2 = CAH10-03 for Materials and technology • IPSBJ_CAH2 = CAH23-01 for Combined and general studies • IPSBJ_CAH2 = CAH19-01 for English studies • IPSBJ_CAH2 = CAH20-01 for History and archaeology • IPSBJ_CAH2 = CAH19-04, CAH19-02 for Languages and area studies³ • IPSBJ_CAH2 = CAH24-01 for Media, journalism and communications • IPSBJ_CAH2 = CAH20-02 for Philosophy and religious studies 	

³ Celtic studies (CAH19-02) has been combined with Languages and area studies due to the small number of students within this category.

Attributes and rebuild instructions for...		
Characteristic	Split indicators	Data on size and shape of provision
	<ul style="list-style-type: none"> • IPSBJ_CAH2 = CAH15-02 for Economics • IPSBJ_CAH2 = CAH15-04 for Health and social care • IPSBJ_CAH2 = CAH16-01 for Law • IPSBJ_CAH2 = CAH15-03 for Politics • IPSBJ_CAH2 = CAH15-01 for Sociology, social policy and anthropology • IPSBJ_CAH2 = CAH01-01 for Medicine and dentistry • IPSBJ_CAH2 = CAH05-01 for Veterinary sciences • IPSBJ_CAH2 = CAH06-01 for Agriculture, food and related studies • IPSBJ_CAH2 = CAH13-01 for Architecture, building and planning • IPSBJ_CAH2 = CAH26-01 for Geography, earth and environmental studies • IPSBJ_CAH2 = CAH03-01 for Biosciences • IPSBJ_CAH2 = CAH07-02 for Chemistry • IPSBJ_CAH2 = CAH07-04 for General, applied and forensic sciences • IPSBJ_CAH2 = CAH09-01 for Mathematical sciences • IPSBJ_CAH2 = CAH07-01 for Physics and astronomy • IPSBJ_CAH2 = CAH03-02 for Sport and exercise sciences • IPSBJ_CAH2 = CAH02-06 for Allied health • IPSBJ_CAH2 = CAH02-05 for Medical sciences • IPSBJ_CAH2 = CAH02-04 for Nursing and midwifery 	

Attributes and rebuild instructions for...		
Characteristic	Split indicators	Data on size and shape of provision
	<ul style="list-style-type: none"> IPSBJ_CAH2 = CAH02-02 for Pharmacology, toxicology and pharmacy IPSBJ_CAH2 = CAH04-01 for Psychology 	
Course length	Not used	IPCRSELGTHGRP ≤ 1 for Less than 1 year IPCRSELGTHGRP = 1 for 1 year IPCRSELGTHGRP = 2 for 2 years IPCRSELGTHGRP = 3+ for 3 years or more
Year of study type	Not used	IPSANDWICH = 1 for Sandwich year

2. The filters and variables described in Table A2 can be used to rebuild split indicators and summaries on the size and shape of provision based on **student characteristics**.

Table A2: Student characteristics and rebuild instructions for split indicators and for data on size and shape of provision

Attributes and rebuild instructions for...		
Characteristic	Split indicators	Data on size and shape of provision
Age on entry (to higher education programme)	IPLEVEL (IPAWARDLEVEL for degree outcomes measures) = DEG, OUG, PUGD and <ul style="list-style-type: none"> IPSTARTAGEBAND = U21 for Under 21 years IPSTARTAGEBAND = 21_25, 26_30 for 21 to 30 years IPSTARTAGEBAND = 31_40, 41_50, 51+ for 31 years and over 	Same as for split indicators (including restrictions on IPLEVEL and IPSTARTAGEBAND). In addition: <ul style="list-style-type: none"> IPSTARTAGEBAND = U for Unknown

Attributes and rebuild instructions for...		
Characteristic	Split indicators	Data on size and shape of provision
	<p>IPLEVEL (IPAWARDLEVEL for degree outcomes measures) = OPGR, OPGT, PUGO, PGTM, PGCE, PHD and IPSTARTAGEBAND \neq U and</p> <ul style="list-style-type: none"> • IPSTARTAGE < 25 for Under 25 years • IPSTARTAGE \geq 25 and IPSTARTAGE < 31 for 25 to 30 years • IPSTARTAGEBAND = 31_40, 41_50, 51+ for 31 years and over 	
Disability	<ul style="list-style-type: none"> • IPDISABLE = Y for Disability reported • IPDISABLE = N for No disability reported 	Not used
Disability type	Not used	<ul style="list-style-type: none"> • IPDISABLETYPE = COG for Cognitive or learning difficulties • IPDISABLETYPE = MH for Mental health conditions • IPDISABLETYPE = MULTI for Multiple or other impairments • IPDISABLETYPE = NONE for No disability reported or unknown disability type • IPDISABLETYPE = PHY for Sensory, medical or physical impairments • IPDISABLETYPE = SOC for Social or communication impairments

Attributes and rebuild instructions for...		
Characteristic	Split indicators	Data on size and shape of provision
Ethnicity	IPDOM = E, N, S, W and <ul style="list-style-type: none"> • IPETHNIC = A for Asian • IPETHNIC = B for Black • IPETHNIC = M for Mixed • IPETHNIC = O for Other • IPETHNIC = W for White 	Same as for split indicators (including restrictions on IPDOM). In addition: <ul style="list-style-type: none"> • IPETHNIC = U or IPDOM in (EU, OTHER, UNKNOWN) for Unknown or not applicable
Sex	<ul style="list-style-type: none"> • IPSEX = 2 for Female • IPSEX = 1 for Male 	Same as for split indicators. In addition: <ul style="list-style-type: none"> • IPSEX = 9 for Other sex
Sexual orientation	Not used	This field is only available for providers that submit data to the HESA Student Record. The reported categories are: <ul style="list-style-type: none"> • IPSEXORT = 01, 02, 03 for Lesbian, gay or bisexual (LGB) • IPSEXORT = 05 for Not heterosexual or LGB • IPSEXORT = 04 for Heterosexual • IPSEXORT = 98, NA, BLANK for Unknown or not applicable

Attributes and rebuild instructions for...		
Characteristic	Split indicators	Data on size and shape of provision
Association between characteristics of students (ABCS) (continuation) quintile (Only relevant to continuation indicators)	<p>Students are divided into quintiles based on a set of characteristics and on their interactions⁴. The characteristics used for ABCS (continuation) depend on the mode of study, as reported by IPSTARTMODE. The characteristics used for apprenticeship students are the same as those used for full-time students. This attribute is only available for UK-domiciled students on undergraduate courses at higher education providers in England (IPDOM = E, N, S, W and IPLEVEL = DEG, OUG, PUGD and IPCOUNTRY = E) and only for students included in the continuation population. The categories reported for this split are:</p> <ul style="list-style-type: none"> • IPCONABCS = 1 for Quintile 1 • IPCONABCS = 2, 3 for Quintile 2 or Quintile 3 • IPCONABCS = 4, 5 for Quintile 4 or Quintile 5 	<p>Same as for split indicators (including restrictions on IPDOM, IPLEVEL and IPCOUNTRY). In addition, the following categories are reported:</p> <ul style="list-style-type: none"> • IPCONABCS = 2 for Quintile 2 • IPCONABCS = 3 for Quintile 3 • IPCONABCS = 4 for Quintile 4 • IPCONABCS = 5 for Quintile 5 • IPCONABCS = 99, BLANK for Unknown or not applicable⁵
Deprivation quintile (Index of Multiple Deprivation, IMD)	<p>IPDOM = E, N, S, W and</p> <ul style="list-style-type: none"> • IPIMDNATION = E1, E2 (providers in England), N1, N2 (providers in Northern Ireland), W1, W2 (providers in Wales), S1, S2 (providers in Scotland) for Quintile 1 or Quintile 2 • IPIMDNATION = E3, E4, E5 (providers in England), N3, N4, N5 (providers in Northern Ireland), W3, W4, W5 	<p>Same as for split indicators (including restrictions on IPDOM). In addition:</p> <ul style="list-style-type: none"> • IPIMDNATION = E1 (providers in England), N1 (providers in Northern Ireland), W1 (providers in Wales), S1 (providers in Scotland) for Quintile 1 • IPIMDNATION = E2 (providers in England), N2 (providers in Northern Ireland), W2 (providers in Wales), S2 (providers in Scotland) for Quintile 2

⁴ For more information and for a list of the characteristics, see Associations between characteristics of students (ABCS) 2021 at www.officeforstudents.org.uk/publications/associations-between-characteristics-of-students-abcs-2021/.

⁵ This category includes students who are in the entrant population but are excluded from the continuation population.

Attributes and rebuild instructions for...		
Characteristic	Split indicators	Data on size and shape of provision
	(providers in Wales), S3, S4, S5 (providers in Scotland) for Quintile 3, Quintile 4 or Quintile 5	<ul style="list-style-type: none"> • IPIMDNATION = E3 (providers in England), N3 (providers in Northern Ireland), W3 (providers in Wales), S3 (providers in Scotland) for Quintile 3 • IPIMDNATION = E4 (providers in England), N4 (providers in Northern Ireland), W4 (providers in Wales), S4 (providers in Scotland) for Quintile 4 • IPIMDNATION = E5 (providers in England), N5 (providers in Northern Ireland), W5 (providers in Wales), S5 (providers in Scotland) for Quintile 5 • IPIMDNATION = UNKNOWN, NA for Unknown or not applicable
Domicile	<ul style="list-style-type: none"> • IPDOM = E, N, S, W for UK • IPDOM = EU, OTHER for Non-UK 	<ul style="list-style-type: none"> • IPDOM = E, N, S, W for UK • IPDOM = EU for EU • IPDOM = OTHER for Other international • IPDOM = UNKNOWN for Unknown
Eligibility for free school meals (FSM)	<p>Coverage is restricted to undergraduate students under 21 years of age on entry, who were found in the NPD KS4 data in 2009-10 or later (IPFSMPOP = 1). The reported categories are:</p> <ul style="list-style-type: none"> • IPFSMSTATE = 1 for Eligible • IPFSMSTATE = 0 for Not eligible 	
Geography of employment quintile (Only relevant to progression indicators)	<p>Coverage is restricted to students who responded to the GO survey (IPEMPRESPONSE = 1). The reported categories are:</p> <ul style="list-style-type: none"> • IPGOTTWAQUINTILE = 1 for Quintile 1 	<p>Coverage is restricted to students who responded to the GO survey (IPEMPRESPONSE = 1). The reported categories are:</p> <ul style="list-style-type: none"> • IPGOTTWAQUINTILE = 1 for Quintile 1

Attributes and rebuild instructions for...		
Characteristic	Split indicators	Data on size and shape of provision
	<ul style="list-style-type: none"> • IPGOTTWAQUINTILE = 2, 3 for Quintile 2 or Quintile 3 • IPGOTTWAQUINTILE = 4, 5 for Quintile 4 or Quintile 5 	<ul style="list-style-type: none"> • IPGOTTWAQUINTILE = 2 for Quintile 2 • IPGOTTWAQUINTILE = 3 for Quintile 3 • IPGOTTWAQUINTILE = 4 for Quintile 4 • IPGOTTWAQUINTILE = 5 for Quintile 5
Entry qualifications		<ul style="list-style-type: none"> • IPENTQUALBROAD = 1 for A-levels (AAA or higher) • IPENTQUALBROAD = 2 for A-levels (ABB or higher) • IPENTQUALBROAD = 3 for A-levels (BCC or higher) or international baccalaureate • IPENTQUALBROAD = 4 for A-levels (CDD or higher) • IPENTQUALBROAD = 5 for A-levels (DDD or lower), other L3 (105 tariff points or higher) or two A-levels and one BTEC • IPENTQUALBROAD = 6 for HE-level qualifications on entry • IPENTQUALBROAD = 7 for BTECs (at least DDM), or one A-level and two BTECs • IPENTQUALBROAD = 8 for BTECs (lower than DDM) • IPENTQUALBROAD = 9 for Unspecified qualifications held by non-UK domiciled students • IPENTQUALBROAD = 10 for Access courses, foundation courses, or other L3 at 65 tariff points or higher

Attributes and rebuild instructions for...		
Characteristic	Split indicators	Data on size and shape of provision
		<ul style="list-style-type: none"> • IPENTQUALBROAD = 11 for None, unknown or other entry qualifications
Socio-economic background	Not used	<p>This field is only available for providers that submit data to the HESA Student Record. It is calculated for years 2015-16 onwards and only for UK domiciled undergraduate students (IPDOM = E, S, W, N and IPBASEYEAR ≥ 2015 and IPLEVEL = DEG, OUG, PUGD). The reported categories are:</p> <ul style="list-style-type: none"> • IPSEC = 1, 2 for Higher managerial, administrative and professional occupations • IPSEC = 3, 4 for Intermediate occupations • IPSEC = 5, 6, 7 for Routine and manual occupations • IPSEC = 8 for Never worked and long-term unemployed • IPSEC = NA, BLANK for Unknown or not applicable
Study locations	Not used	<ul style="list-style-type: none"> • IPDL= 1 for Distance learning • IPSTUDYLOCTYPE = L_01 for Local • IPSTUDYLOCTYPE = M_00, M_01 for Not local • IPDL ≠ 1 and IPSTUDYLOCTYPE ≠ L_01, M_00, M_01 for Unknown
Tracking underrepresentation by MSOA (TUNDRA MSOA)	Not used	<p>This field is only calculated for England domiciled undergraduate students (IPDOM = E and IPLEVEL = DEG, OUG, PUGD). The reported categories are:</p>

Attributes and rebuild instructions for...		
Characteristic	Split indicators	Data on size and shape of provision
		<ul style="list-style-type: none"> • IPTUNDRALOOKUP = 1 for Quintile 1 • IPTUNDRALOOKUP = 2 for Quintile 2 • IPTUNDRALOOKUP = 3 for Quintile 3 • IPTUNDRALOOKUP = 4 for Quintile 4 • IPTUNDRALOOKUP = 5 for Quintile 5 • IPIMDNATION = UNKNOWN, NA for Unknown or not applicable

Annex B: Rebuilding split indicators for access and participation data resources

1. The student characteristic split indicators for the purposes of access and participation data resources would be rebuilt using the filters and variables described in Table B1.

Table B1: Student characteristics split indicators for access and participation data resources

Split	Attributes and rebuild instructions
Age on entry SplitType = AgeOnCommencement	IPLEVEL (IPAWARDLEVEL for degree outcomes measures) = DEG, OUG, PUGD and IPSTARTAGEBAND = U21 for Young Under 21 21_25, 26_30, 31_40, 41_50, 51+ for Mature_Age21andOver 21_25 for Age21_25 26_30 for Age26_30 31_40 for Age31_40 41_50 for Age41_50 51+ for Age51andOver
Disability SplitType = Disability	IPDISABLE = Y for Disabled IPDISABLE = N for NoKnownDisability
Disability type SplitType = DisabilityType	IPDISABLETYPE = COG for CognitiveAndLearning MH for MentalHealth MULTI for MultipleImpairments PHY for SensoryMedicalAndPhysical SOC for SocialAndCommunication NONE for NoKnownDisabilityType
Eligibility for free school meals SplitType = FSMEligibility	IPFSMPOP = 1 and IPFSMSTATE = 1 for Eligible for free school meals during their schooling 0 for Not eligible for free school meals during their schooling
Deprivation quintile (IMD) SplitType = EnglishIMDQuintile_2019	IPIMDNATION ≠ UNKNOWN and IPIMDNATION = E1 for IMDQ1 E2 for IMDQ2 E3 for IMDQ3 E4 for IMDQ4 E5 for IMDQ5 E1, E2 for IMDQ1_2 E3, E4, E5 for IMDQ3_5 E1, E3, E4, E5 for IMDQ1345

Split	Attributes and rebuild instructions
	<p>E1, E2, E4, E5 for IMDQ1245 E1, E2, E3, E5 for IMDQ1235 E1, E2, E3, E4 for IMDQ1234 IPIMDNATION \neq UNKNOWN and IPSTARTAGE = 18 and IPIMDNATION =</p> <p>E1 for IMDQ1_Age18 E2 for IMDQ2_Age18 E3 for IMDQ3_Age18 E4 for IMDQ4_Age18 E5 for IMDQ5_Age18</p>
Participation of local areas (POLAR4) SplitType = POLAR4Quintile	IPPOLAR4 \neq U and IPSTARTAGEBAND = U21 and IPPOLAR4 = <p>1 for POLAR4Q1 2 for POLAR4Q2 3 for POLAR4Q3 4 for POLAR4Q4 5 for POLAR4Q5 1, 2 for POLAR4Q1_2 3, 4, 5 for POLAR4Q3_5 2, 3, 4, 5 for POLAR4Q2345 1, 3, 4, 5 for POLAR4Q1345 1, 2, 4, 5 for POLAR4Q1245 1, 2, 3, 5 for POLAR4Q1235 1, 2, 3, 4 for POLAR4Q1234</p> IPPOLAR4 \neq U and IPSTARTAGE = 18 and IPPOLAR4 = <p>1 for POLAR4Q1_Age18 2 for POLAR4Q2_Age18 3 for POLAR4Q3_Age18 4 for POLAR4Q4_Age18 5 for POLAR4Q5_Age18</p>
Sex SplitType = Sex	IPSEX \neq 9 and IPSEX = <p>2 for Female 1 for Male</p> IPSEX \neq 9 and IPSTARTAGE = 18 and IPSEX = <p>2 for Female_Age18 1 for Male_Age18</p>
Intersection of deprivation quintile and ethnicity SplitType = Int_IMDEthnicity	IPETHNIC \neq U and IPIMDNATION \neq NA, UNKNOWN and IPIMDNATION = E1, E2 and IPETHNIC = A, B, M, O for IMDQ12_ABMO

Split	Attributes and rebuild instructions
	IPIMDNATION = E3, E4, E5 and IPETHNIC = A, B, M, O for IMDQ345_ABMO IPIMDNATION = E1, E2 and IPETHNIC = W for IMDQ12_White IPIMDNATION = E3, E4, E5 and IPETHNIC = W for IMDQ345_White
Intersection of POLAR4 quintile and ethnicity SplitType = Int_POLAREthnicity	IPSTARTAGEBAND = U21 and IPETHNIC \neq U and IPPOLAR4 \neq U and IPPOLAR4 = 1, 2 and IPETHNIC = A, B, M, O for POLARQ12_ABMO IPPOLAR4 = 3, 4, 5 and IPETHNIC = A, B, M, O for POLARQ345_ABMO IPPOLAR4 = 1, 2 and IPETHNIC = W for POLARQ12_White IPPOLAR4 = 3, 4, 5 and IPETHNIC = W for POLARQ345_White
Intersection of deprivation quintile and sex SplitType = Int_IMDsex	IPSEX \neq 9 and IPIMDNATION \neq NA, UNKNOWN and IPIMDNATION = E1, E2 and IPSEX = 1 for IMDQ12_Male IPIMDNATION = E3, E4, E5 and IPSEX = 1 for IMDQ345_Male IPIMDNATION = E1, E2 and IPSEX = 2 for IMDQ12_Female IPIMDNATION = E3, E4, E5 and IPSEX = 2 for IMDQ345_Female
Intersection of POLAR4 quintile and sex SplitType = Int_POLARSex	IPSTARTAGEBAND = U21 and IPSEX \neq 9 and IPPOLAR4 \neq U and IPPOLAR4 = 1, 2 and IPSEX = 1 for POLARQ12_Male IPPOLAR4 = 3, 4, 5 and IPSEX = 1 for POLARQ345_Male IPPOLAR4 = 1, 2 and IPSEX = 2 for POLARQ12_Female IPPOLAR4 = 3, 4, 5 and IPSEX = 2 for POLARQ345_Female

List of abbreviations used in this document

Abbreviation	Meaning
ABCS	Associations between characteristics of students (analyses)
AOR	Aggregate Offshore Record
CAH	Common Aggregation Hierarchy
ESFA	Education and Skills Funding Agency
FSM	Free school meals
GO	Graduate Outcomes (survey)
HESA	Higher Education Statistics Agency
ILR	Individualised Learner Record
IMD	Indices of Multiple Deprivations 2019
KS4	Key stage 4
MSOA	Middle layer super output area
NPD	National Pupil Database
NSS	National Student Survey
OfS	Office for Students
ONS	Office for National Statistics
PGCE	Postgraduate Certificate in Education
POLAR	Participation of local areas (classification)
TEF	Teaching Excellence Framework
TNE	Transnational education
TorR	Taught or registered
TUNDRA	Tracking underrepresentation of areas



© The Office for Students copyright 2022

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/