

Equality and diversity data: Methodology

Overview

1. This document provides information on the methodology and the groups of students used in the Office for Students' (OfS) equality and diversity webpages.¹
2. The charts and tables are available for six different student populations:
 - i. all students – all domiciles
 - ii. all students – UK-domiciled
 - iii. entrants – all domiciles
 - iv. entrants – UK-domiciled
 - v. qualifiers – all domiciles
 - vi. qualifiers – UK-domiciled.

See the 'Definitions' section of this document for more detail on these populations.

3. The data is further broken down by the following levels of study:
 - i. All levels of higher education
 - 1) All undergraduate
 - First degrees
 - Undergraduate with postgraduate components
 - Other undergraduate
 - 2) All postgraduate
 - Postgraduate research
 - Postgraduate taught masters
 - PGCE
 - Other postgraduate.
4. There is the option to display the data for all English higher education providers or split by mode of study (i.e. full-time/part-time) or subject of study (both broad and detailed). These options are described as study characteristics.
5. Finally, the statistics are shown by one the following student characteristics:
 - Age at entry
 - Disability (broad)
 - Disability (detailed)
 - Educational disadvantage (POLAR4)
 - Ethnicity (2 groups)
 - Ethnicity (5 groups)
 - Ethnicity (17 groups)

¹ www.officeforstudents.org.uk/data-and-analysis/equality-and-diversity/

- Sex.

Detailed definitions of these student characteristics can be found in Table 3.

Methodology

6. Data was combined from the Higher Education Statistics Agency (HESA) student record², HESA student alternative record³ and the Education and Skills Funding Agency (ESFA) Individualised Learner Record (ILR)⁴ for the academic years 2010-11 to 2018-19. Duplicate student instances within each academic year were then removed from the data.
7. The data in this release has been limited to English higher education providers.
8. Over the time series there have been changes to the number of providers that return student data to each record. The number of providers returning data in each year to each data source and included in each population of students is shown in Table 1. The 'Total' is not equal to the sum of the number of providers in the ILR and HESA records for all populations in 2015-16, because one provider appeared in multiple records in that year.
9. Throughout this release, counts (of student numbers) have been rounded to the nearest five and proportions have been rounded to the nearest 0.1. Additionally, in circumstances where the combination of study and student characteristics applies to fewer than 23 individuals, then neither the count nor proportion is presented.
10. Where the student characteristic is unknown it is included in the data tables as 'No response'. This data is not included in the charts or the calculated proportions. With this data we cannot know whether it is unknown due to active decisions by the students not to report data or whether it is missing for other reasons.

² <https://www.hesa.ac.uk/collection/c18051>

³ <https://www.hesa.ac.uk/collection/c18054>

⁴ <https://www.gov.uk/government/publications/ilr-specification-validation-rules-and-appendices-2018-to-2019>

Table 1: Number of providers included in the release returning data to each data source

	2010-11	2011-12	2012-13	2013-14	2014-15	2015-16	2016-17	2017-18	2018-19
All students									
HESA student	131	131	134	136	133	134	135	136	137
HESA student alternative					62	97	95	98	97
ILR	190	200	222	308	311	340	377	402	392
Total	321	331	356	444	506	570	607	636	626
Entrants									
HESA student	131	131	134	136	133	134	135	135	137
HESA student alternative					61	97	92	95	93
ILR	189	193	217	280	291	327	362	369	364
Total	320	324	351	416	485	557	589	599	594
Qualifiers									
HESA student	131	131	132	136	133	134	135	136	137
HESA student alternative					58	95	95	93	93
ILR	186	192	212	288	290	288	328	337	340
Total	317	323	344	424	481	516	558	566	570

Definitions

11. The majority of the fields used in this release have been defined using algorithms found in the OfS's 'Technical algorithms for institutional performance measures: 2020 core algorithms' document.⁵ Any field discussed that begins with 'IP' (formerly 'B3MON') is defined in detail in that document.
12. The student populations are defined as follows:
 - i. All students – the students that are in the standard higher education registration population (IPXPSR) for those whose record is included in the learner higher education entity (IPHEENTITYFLAG), i.e. those that are enrolled in higher education during the academic year specified.
 - ii. Entrants – of the enrolled students, those who began their studies in the academic year specified (IPENTRANT).
 - iii. Qualifiers – students who were awarded a higher education-level qualification during the academic year specified (IPQUALIFIER) who were. This includes some students that qualify but are not in the standard higher education registration population as a result of being 'dormant' or 'writing up' during the year in which they complete their qualification. The qualifiers population includes those who qualified at a lower level than the level at which they were studying (e.g. a student studying for a postgraduate qualification who qualified with an undergraduate qualification). These statistics differ from the access and participation dashboard⁶ in that respect.

These populations are further split by student domicile:

- i. All domiciles – all students attending English higher education providers
 - ii. UK-domiciled – only those students domiciled in England, Scotland, Wales or Northern Ireland (IPDOM).
13. Level of study is provided at different levels of detail and is based on IPLEVEL for entrants and enrolled students; for qualifiers, IPAWARDLEVEL is used. At a detailed level, the following levels are represented:
 - i. First degrees (IPLEVEL: DEG)
 - ii. Undergraduate with postgraduate components (IPLEVEL: PUGD, PUGO)
 - iii. Other undergraduate (IPLEVEL: OUG)
 - iv. Postgraduate research (IPLEVEL: PHD)
 - v. Postgraduate taught masters (IPLEVEL: PGTM)
 - vi. PGCE (IPLEVEL: PGCE)
 - vii. Other postgraduate (IPLEVEL: OPGR, OPGT).

For undergraduate students, these levels of study match those used as part of the OfS's access and participation data.

14. The detailed levels of study have been grouped into the following broad groups:
 - i. All levels of higher education – all

⁵ See '2020 core algorithms', available from www.officeforstudents.org.uk/data-and-analysis/institutional-performance-measures/technical-documentation/

⁶ www.officeforstudents.org.uk/data-and-analysis/access-and-participation-data-dashboard/

- ii. All undergraduate – first degrees, undergraduate with postgraduate components, other undergraduate
 - iii. All postgraduate – postgraduate research, postgraduate taught masters, PGCE and other postgraduate.
15. The study and student characteristics are defined in Table 3. Subject data is presented at a detailed and broad level. Detailed subject groups are based on the Common Aggregation Hierarchy 1 (CAH1)⁷ Version 1.3 codes. Broad subjects are based on a new experimental grouping detailed below.
16. We have developed a new, experimental grouping of subjects to allow us to more easily report and visualise sector-level trends by subject. This new grouping of subjects has been produced by combining the 21 CAH1 groups produced by HESA as part of their move to introduce the Higher Education Classification of Subjects (HECoS). The CAH1 groups have been further aggregated into ten groups using the hierarchical structure of the common aggregation hierarchy. As such, when producing these new subject groups, the CAH1 groups were combined but never split.
17. The CAH1 groups were combined using a number of criteria, applied in the order of priority that follows:
- i. We have grouped the subjects that we judged to have similar aims, structures and/or entry requirements, for example medicine and dentistry and veterinary sciences.
 - ii. We also considered the student demographic profiles (age, ethnicity and sex) of the subject groups when combining them.
 - iii. Groups of disparate size were avoided. For example, based on the two previous criteria, business and management; law; and social sciences could all be combined. However, because business and management is large in terms of student numbers, we decided to keep it as a single group. Small groups (such as medicine, dentistry and veterinary sciences, or education and teaching) were allowed only where the grouping was deemed to be appropriate or where the student profile appears to be unique.
18. The new subject groups are based CAH version 1.3 and can be found in Table 2.

⁷ <https://www.hesa.ac.uk/innovation/hecos>

Table 2: New subject grouping

Subject group	CAH1 groups (Version 1.3)
Medicine, dentistry and veterinary sciences	Medicine and dentistry (CAH01) Veterinary sciences (CAH05)
Nursing, allied health and psychology	Subjects allied to medicine (CAH02) Psychology (CAH04)
Natural and mathematical sciences	Biological and sport sciences (CAH03) Physical sciences (CAH07) Mathematical sciences (CAH09)
Engineering, technology and computing	Engineering and technology (CAH10) Computing (CAH11)
Natural and built environment	Agriculture, food and related studies (CAH06) Architecture, building and planning (CAH13) Geography, earth and environmental studies (CAH26)
Law and social sciences	Social sciences (CAH15) Law (CAH16)
Business and management	Business and management (CAH17)
Humanities and languages	Language and area studies (CAH19) Historical, philosophical and religious studies (CAH20) Combined and general studies (CAH23) Media, journalism and communications (CAH24)
Education and teaching	Education and teaching (CAH22)
Design, and creative and performing arts	Design, and creative and performing arts (CAH25)

19. This is a new way of grouping subjects for the purposes of reporting on sector-level trends, and we are actively seeking feedback. Any comments or feedback can be sent to official.statistics@officeforstudents.org.uk.

Changes

20. This release has a number of differences to our 2019 equality and diversity data publication. These changes reflect changes to the OfS core algorithms, detailed elsewhere.⁸ The major differences that affect this publication are summarised below:
- i. Level of study (IPLEVEL) has changed resulting in a small number of students being attributed a different level of study as before.

⁸ See '2019 to 2020 changes to core algorithms', available from www.officeforstudents.org.uk/data-and-analysis/institutional-performance-measures/technical-documentation/

- ii. For qualifiers, detailed level of study is now determined by the level of qualification with which they qualified. The field IPAWARDLEVEL is used instead of IPLEVEL (formerly B3MONLEVEL).
- iii. In Ethnicity (5 groups) the Gypsy or Traveller group is now within Other and no longer in White.

21. In addition to the above changes, counts and proportions differ from previous year's releases due to amendments and corrections to the data.

22. The CAH1 codes in this document have been updated to reflect Version 1.3, rather than Version 1.2. This has not contributed to any changes in counts and proportions.

23. We have made it possible to split the statistics by broad subject of study, as detailed in Table 2 and the paragraphs above.

24. We have added two new ethnicity groupings. Ethnicity (2 groups) provides a comparison of the counts and proportions of the White and minority ethnic students. Ethnicity (17 groups) allows for a view of counts and proportions by more detailed ethnic groups.

Contact

25. If you have any questions or feedback on these pages, please contact Stanley Rudkin at official.statistics@officeforstudents.org.uk.

Table 3: Definitions of study and student characteristics

Characteristic (and field)	Definition	Categories	Notes
Study characteristics:			
Entire sector	All data for English higher education providers	-	
Mode of study (IPMODE)	The mode of study being undertaken	Full-time Part-time	For the 'Qualifiers' population, the mode of study at qualification was used (B3MONEMPMODE)
Subject of study (broad)	Subject studied, grouped using an experimental classification created by the OfS. See Table 2 for a mapping of broad subject of study to detailed subject of study.	Medicine, dentistry and veterinary sciences (CAH01, CAH05) Nursing, allied health and psychology (CAH02, CAH04) Natural and mathematical sciences (CAH03, CAH07, CAH09) Engineering, technology and computing (CAH10, CAH11) Natural and built environment (CAH06, CAH13, CAH26) Law and social sciences (CAH15, CAH16) Business and management (CAH17) Humanities and languages (CAH19, CAH20, CAH23, CAH24) Education and teaching (CAH22) Design, creative and performing arts (CAH25)	CAH1 (Version 1.3) codes for each of these groups are given in parentheses

Characteristic (and field)	Definition	Categories	Notes
Subject of study (detailed)	Subject studied, grouped using the HESA Common Aggregation Hierarchy 1 (CAH1) Version 1.3 ⁹	<p>Medicine and dentistry (CAH01)</p> <p>Subjects allied to medicine (CAH02)</p> <p>Biological and sport sciences (CAH03)</p> <p>Psychology (CAH04)</p> <p>Veterinary sciences (CAH05)</p> <p>Agriculture, food and related studies (CAH06)</p> <p>Physical sciences (CAH07)</p> <p>Mathematical sciences (CAH09)</p> <p>Engineering and technology (CAH10)</p> <p>Computing (CAH11)</p> <p>Architecture, building and planning (CAH13)</p> <p>Social sciences (CAH15)</p> <p>Law (CAH16)</p> <p>Business and management (CAH17)</p> <p>Media, journalism and communications (CAH24)</p> <p>Language and area studies (CAH19)</p> <p>Historical, philosophical and religious studies (CAH20)</p> <p>Design, and creative and performing arts (CAH25)</p> <p>Education and teaching (CAH22)</p> <p>Combined and general studies (CAH23)</p> <p>Geography, earth and environmental studies (CAH26)</p>	

⁹ <https://www.hesa.ac.uk/innovation/hecos>

Characteristic (and field)	Definition	Categories	Notes
Student characteristics:			
Age at entry (IPAGEBAND)	Age of student as of 31 August in the academic year of entry	Under 21 years 21 to 25 years 26 to 30 years 31 to 40 years 41 to 50 years 51 years and over	
Disability – broad (IPDISABLE)	Indicates whether a student has reported as disabled	No disability reported Disability reported	
Disability – detailed (IPDISABLETYPE)	Provides detail of the type of disability a student has reported (if applicable)	No disability reported Cognitive or learning difficulties Mental health condition Multiple impairments Sensory, medical or physical impairments Social or communication impairment	Disability information is recorded on the basis of the student’s own self-assessment. Changes in the number of students in this category may occur as a result of changes in data reporting. For a small number of students, a disability has been reported but the type of disability is not known, as such these students are categorised as unknown.
Educational disadvantage – POLAR4 (IPPOLAR4)	A measure of the proportion of the population that participates in higher education ¹⁰	Quintile 1 – lowest participation areas Quintile 2 Quintile 3 Quintile 4 Quintile 5 – highest participation areas	POLAR4 is used for students under the age of 21 and all other students are assigned to ‘Not applicable’. Additionally, we have only included undergraduate students and postgraduate students under 21 have been assigned as ‘Not applicable’.
Ethnicity – 2 groups	Student ethnicity (2 groups)	White (including Irish and other) Minority ethnic groups	The black, Asian, mixed, and other ethnic groups of IPETHNIC are combined for this measure. It should be noted that the white ethnic group itself contains the ethnic minority groups,

¹⁰ Details of how POLAR4 quintiles are defined can be found here: www.officeforstudents.org.uk/data-and-analysis/polar-participation-of-local-areas/

Characteristic (and field)	Definition	Categories	Notes
			White – Irish and White - Other. For a small number of students each year, ethnicity is unknown. Ethnicity is only presented for UK- domiciled students and all other students are assigned to 'Not applicable'. ¹¹
Ethnicity – 5 groups (IPETHNIC)	Student ethnicity (5 groups)	Asian Black Mixed Other White	For a small number of students each year, ethnicity is unknown. Ethnicity is only presented for UK-domiciled students and all other students are assigned to 'Not applicable'.
Ethnicity – 17 groups (IPETHNICDETAIL)	Student ethnicity (17 groups)	Asian or Asian British - Chinese Asian or Asian British - Indian Asian or Asian British - Pakistani Asian or Asian British - other Black or black British - African Black or black British - Caribbean Black or black British - other Mixed - white and Asian Mixed - white and black African Mixed - white and black Caribbean Mixed - other White - English, Welsh, Scottish, Northern Irish, British White - Irish White - other Other ethnic group Gypsy or Traveller	For a small number of students each year, ethnicity is unknown. Ethnicity is only presented for UK-domiciled students and all other students are assigned to 'Not applicable'.

¹¹ <https://www.hesa.ac.uk/collection/c18051/a/ethnic>

Characteristic (and field)	Definition	Categories	Notes
Sex (IPSEX)	The sex of the student	Female Male Other sex	