

Equality and diversity data: Experimental statistics methodology

Overview

1. This document provides information on the methodology used in the Office for Students' (OfS) experimental equality and diversity statistics webpages.¹
2. As a result of limited data availability, the charts and tables are available only for UK-domiciled entrants at English higher education institutions. This group of students is defined below.
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 institutions or split by mode of study (i.e. full-time/part-time).
5. Finally, the statistics are shown by one the following student characteristics:
 - i. Gender identity
 - ii. Parental education
 - iii. Religion or belief
 - iv. Sexual orientation.

Detailed definitions of these student characteristics can be found in Table 2.

6. In 2012-13, equality and diversity questions on gender identity, religion or belief and sexual orientation were added as optional fields to the return of student data via the Higher Education Statistics Agency (HESA), to monitor equal opportunities issues in the higher education sector and support higher education institutions in meeting their obligations under the Equality Act 2010.² For the 2017-18 academic year, reporting of religion or belief became compulsory for all

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

² See <https://www.hesa.ac.uk/collection/c12051/summaryofchanges>

providers.³ As such, changes between this and previous academic years most likely reflect this change in reporting and should not be interpreted as time series trends.

7. In the early years of reporting, data was not returned by enough providers for it to be reported. This is because there was insufficient data to determine whether any patterns were representative of the whole population or only those providers reporting data. Since the 2015-16 academic year, data has been sufficiently reported to allow publication of high-level patterns for religion or belief and sexual orientation.⁴ Due to concerns around the quality of data collected on gender identity in 2015-16, we have decided to only publish this data from 2016-17 onwards for this field.
8. Parental education has been collected since 2008-09 as part of the UCAS application form and does not have the same data coverage issues associated with gender identity, sexual orientation and religion or belief, but is reported for the same time series here for consistency.
9. 'Information refused' is included in the proportions for these statistics as this is seen as an active response to the question. This is not the case for our official equality and diversity statistics⁵ because it is not possible to observe in that data the reason for non-response.
10. In future releases, we plan to include gender identity, religion or belief and sexual orientation data from the HESA alternative student record⁶, where data quality allows.
11. In addition, we hope to make it possible to split the data by broad subject of study, as with the official statistics⁷ associated with this release.

Methodology

12. The data presented in this release represents the English higher education institutions that returned data to HESA on the HESA student record⁸ for the academic years 2015-16 to 2018-19. The numbers of English providers and UK-domiciled students included in the HESA student record can be found in Table 1.
13. Table 1 includes the number of providers reporting data on gender identity, parental education, religion or belief and sexual orientation and the number of students for which data has been reported.
14. For a given academic year and student characteristic, a provider must have reported the profile of at least 50 of their students to be included in the sector calculations. This filtering of the data was performed to exclude providers reporting abnormally low amounts of data that on inspection appeared to be incorrect. For example, for gender identity, several large providers only reported the gender identity of students that had a gender different from assigned at birth, while the rest of their students were 'unknown'. For small categories like this, including data

³ <https://www.hesa.ac.uk/innovation/records/reviews/student-2017-18-changes>

⁴ See

<https://webarchive.nationalarchives.gov.uk/20180405120413/http://www.hefce.ac.uk/pubs/year/2017/201713>

⁵ www.officeforstudents.org.uk/data-and-analysis/equality-and-diversity/equality-and-diversity-data/

⁶ <https://www.hesa.ac.uk/collection/c18054>

⁷ www.officeforstudents.org.uk/data-and-analysis/equality-and-diversity/equality-and-diversity-data/

⁸ <https://www.hesa.ac.uk/collection/c18051>

that is not representative of the provider (i.e. in this example, the provider is reporting that 100 per cent of their students for which gender identity is known have a gender different from assigned at birth) could have a large impact on the proportions as data would only be included for this category and not the category that represents the majority of students at that provider. A cut-off value of 50 'knowns' was selected to reduce the risk of losing accurate data from providers with small numbers of enrolled students each year.

15. Further restrictions were made to the data concerning gender identity. Some providers have reported students with a gender different from assigned at birth at levels that are considerably higher than the general population. These instances most likely represent incorrect data reporting so restrictions were made to allow more robust sector-wide comparisons. Providers reporting more than 20 per cent of their students as having a different gender from that assigned at birth were removed from the data.
16. The impacts of this data limiting can be seen in Table 1.
17. Throughout this release, counts 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 less than 23 individuals, then neither the count nor proportion is presented.

Definitions

18. The student population consists of those in the standard higher education registration population who began their studies at English higher education institutions in the academic year specified. This population has been limited to only those students domiciled in England, Scotland, Wales or Northern Ireland.
19. The levels of study match those found in the official equality and diversity statistics associated with this release⁹ and were created using the HESA field COURSEAIM.¹⁰ The following detailed levels of study are used:
 - i. First degrees
 - ii. Undergraduate with postgraduate components
 - iii. Other undergraduate
 - iv. Postgraduate research
 - v. Postgraduate taught masters
 - vi. PGCE
 - vii. Other postgraduate
20. These detailed levels of study have been grouped into the following broad groups:
 - i. All levels of higher education – all
 - ii. All undergraduate – first degrees, undergraduate with postgraduate components, other undergraduate
 - iii. All postgraduate – postgraduate research, postgraduate taught masters, PGCE and other postgraduate

⁹ See www.officeforstudents.org.uk/data-and-analysis/equality-and-diversity/equality-and-diversity-data/

¹⁰ <https://www.hesa.ac.uk/collection/c18051/a/courseaim>

21. The data can be split by the mode of study (full-time or part-time), which is based on the HESA field XMODE01.¹¹ Alternatively, it can be displayed for all UK-domiciled entrants (Entire sector).
22. The student characteristics are defined in Table 2.

Changes

23. This release has some differences to the previous equality and diversity experimental data publication. Small changes have been made to the classification of level of study to align with changes made to the OfS level of study algorithm (IPLEVEL).¹² As a result, counts and proportions differ slightly from the 2019 publication.
24. In addition to the above change, counts and proportions differ slightly from the 2019 publication due to amendments and corrections to the data.
25. Where we previously published data on gender identity for the academic year 2015-16, after reviewing the data quality, we have decided to exclude this year from the release.

Contact

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

¹¹ <https://www.hesa.ac.uk/collection/c18051/derived/xmode01>

¹² The COVID-19 pandemic has resulted in a delay in publication of this supporting documentation. It will be made available as soon as possible.

Table 1: Numbers of English higher education providers and students before and after limiting

The 'All' dataset contains all the providers that returned data for the student characteristic in question and the number of students indicates the total for which information for that characteristic was reported. The 'Limited' dataset only contains data from providers returning information for at least 50 of their students and passed the data quality test for gender identity (paragraphs 11-13). *This is the proportion of students in the 'All' dataset rather than the proportion of students in the entire HESA student record.

HESA student record		Dataset	2015-16		2016-17		2017-18		2018-19	
			Count	Proportion	Count	Proportion	Count	Proportion	Count	Proportion
Entire record	Number of students	-	613,535		627,030		624,955		627,730	
	Number of providers	-	132		133		133		135	
Gender identity	Number of students	All	-	-	336,515	54%	353,245	57%	357,760	57%
	Number of providers	All	-	-	104	78%	108	81%	109	81%
	Number of students	Limited	-	-	327,570	52%	348,740	56%	357,690	57%
	Number of providers	Limited	-	-	89	67%	96	72%	101	75%
	Student data excluded	Limited	-	-	8,945	3%*	4,505	1%*	70	0%*
Parental education	Number of students	All	522,240	85%	530,840	85%	526,305	84%	529,915	84%
	Number of providers	All	124	94%	125	94%	126	95%	128	95%
	Number of students	Limited	522,240	85%	530,840	85%	526,305	84%	529,915	84%
	Number of providers	Limited	124	94%	125	94%	126	95%	128	95%
	Student data excluded	Limited	-	0%*	-	0%*	-	0%*	-	0%*
Religion or belief	Number of students	All	380,565	62%	413,115	66%	590,365	94%	589,150	94%
	Number of providers	All	106	80%	108	81%	133	100%	135	100%
	Number of students	Limited	380,520	62%	413,115	66%	590,365	94%	589,135	94%
	Number of providers	Limited	104	79%	107	80%	133	100%	134	99%
	Student data excluded	Limited	45	0%*	-	0%*	-	0%*	15	0%*
Sexual orientation	Number of students	All	395,305	64%	427,245	68%	444,425	71%	452,825	72%
	Number of providers	All	109	83%	111	83%	116	87%	118	87%
	Number of students	Limited	395,260	64%	427,245	68%	444,425	71%	452,820	72%
	Number of providers	Limited	107	81%	111	83%	116	87%	117	87%
	Student data excluded	Limited	45	0%*	-	0%*	-	0%*	5	0%*

Table 2: Definitions of student characteristics

Characteristic	Definition	Categories	Notes
Gender identity	The gender identity of the student	Same as the gender assigned at birth Not the same as the gender assigned at birth Information refused	Based on HESA field GENDERID ¹³
Parental education	Whether the parents of the student have a higher education qualification	Have an higher education qualification No higher education qualification Don't know Information refused	Based on HESA field PARED ¹⁴
Religion or belief	The religious belief of the student	Any other religion or belief Buddhist Christian Hindu Jewish Muslim No religion Sikh Spiritual Information refused	Based on HESA field RELBLF ¹⁵
Sexual orientation	The sexual orientation of the student	Bisexual Gay man Gay woman/lesbian Heterosexual Other Information refused	Based on HESA field SEXORT ¹⁶

¹³ <https://www.hesa.ac.uk/collection/c18051/a/genderid>

¹⁴ <https://www.hesa.ac.uk/collection/c18051/a/pared>

¹⁵ <https://www.hesa.ac.uk/collection/c18051/a/reblf>

¹⁶ <https://www.hesa.ac.uk/collection/c18051/a/sexort>