

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 webpages¹.
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:
 - All levels of higher education
 - All undergraduate
 - First degrees
 - Undergraduate with postgraduate components
 - Other undergraduate
 - 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:
 - Gender identity
 - Parental education
 - Religion or belief
 - 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 2010².
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 gender identity, religion or belief and sexual orientation³.

¹ <https://www.officeforstudents.org.uk/data-and-analysis/equality-and-diversity/experimental-data/>

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

³ See <https://webarchive.nationalarchives.gov.uk/20180405120413/http://www.hefce.ac.uk/pubs/year/2017/201713/>

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.

Methodology

9. The data presented in this release represents the English higher education institutions that returned data to HESA on the HESA student record⁴ for the academic years 2015-16 to 2017-18. The numbers of English providers and UK-domiciled students included in the HESA student record can be found in Table 1.
10. 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.
11. 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 that is not representative of the provider (i.e. in this example, the provider is reporting that 100% 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.
12. 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. Inspecting the data returned by English higher education institutions that have less than 500 enrolled students found they all returned 'known' data for at least 50 students for each year and student characteristic.
13. 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. These instances were especially common for 2015-16 and despite these additional exclusions, the values for gender being different from assigned at birth are almost certainly over estimations for 2015-16.
14. The impacts of this data limiting can be seen in Table 1.
15. 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.

⁴ <https://www.hesa.ac.uk/collection/c17051>

Definitions

16. 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.
17. The levels of study match those found in the non-experimental statistics associated with this release⁵ and were created using the HESA field COURSEAIM⁶. The following detailed levels of study are used:
 - a. First degrees
 - b. Undergraduate with postgraduate components
 - c. Other undergraduate
 - d. Postgraduate research
 - e. Postgraduate taught masters
 - f. PGCE
 - g. Other postgraduate
18. These detailed levels of study have been grouped into the following broad groups:
 - a. All levels of higher education – all
 - b. All undergraduate – first degrees, undergraduate with postgraduate components, other undergraduate
 - c. All postgraduate – postgraduate research, postgraduate taught masters, PGCE and other postgraduate
19. The data can be split by the mode of study (full-time or part-time), which is based on the HESA field XMODE01⁷. Alternatively, it can be displayed for all the institutions that return data to the HESA student record (Entire sector).
20. The student characteristics are defined in Table 2.

Changes

21. This release has some differences to the previous equality and diversity experimental data publication:
 - a. A different data cut-off method was utilised. In the previous release, providers were required to report 'known' data for at least 50% of their students to be included in calculations. Inspection of the data found limiting in this way had a small impact on the sector level proportions so the less conservative data limiting method detailed in paragraph 11 was used to allow more data to be included in the statistics.
 - b. For the 2017-18 academic year, reporting of religion or belief became compulsory for all providers⁸. 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. Because of this, only the 2017-18 data is represented in the graph.

⁵ See <https://www.officeforstudents.org.uk/data-and-analysis/equality-and-diversity/equality-and-diversity-data/>

⁶ <https://www.hesa.ac.uk/collection/c18051/a/courseaim>

⁷ <https://www.hesa.ac.uk/collection/c17051/derived/xmode01>

⁸ <https://www.hesa.ac.uk/innovation/records/reviews/student-2017-18-changes>

- c. Level of study is grouped in a different way to the previous release and now matches the levels used in the non-experimental equality and diversity statistics of this release⁹.
- d. For this year's release it is possible to choose a level of study and mode of study rather than one or the other. However, as a result of this increased detail, data is no longer split by subject of study as the HESA Common Aggregation Hierarchy 1 groupings used are too detailed and lead to very small numbers of students.
- e. All UK-domiciled students attending the Open University are included in the population and not only those that live in England to align with the OfS' access and participation data¹⁰.

If you have any feedback regarding these changes we encourage you to contact us using the details below.

Contact

22. If you have any questions or feedback on these pages, please contact William Rimington at william.rimington@officeforstudents.org.uk or the Foresight and Insight team at qapt@officeforstudents.org.uk

⁹ <https://www.officeforstudents.org.uk/data-and-analysis/equality-and-diversity/equality-and-diversity-data/>

¹⁰ <https://www.officeforstudents.org.uk/data-and-analysis/access-and-participation-data-dashboard/>

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		2015-16		2016-17		2017-18	
Dataset		Count	Proportion	Count	Proportion	Count	Proportion
Entire record							
Number of providers that returned data to the HESA student record	-	132		133		133	
Number of students on record	-	620,975		633,680		629,020	
Gender identity							
Number of providers	All	103	78%	104	78%	108	81%
Number of students	All	333,895	54%	344,255	54%	360,860	57%
Number of providers	Limited	80	61%	89	67%	96	72%
Number of students	Limited	303,225	49%	335,310	53%	356,350	57%
Student data excluded	Limited	30,670	9%*	8,945	3%*	4,510	1%*
Parental education							
Number of providers	All	124	94%	125	94%	126	95%
Number of students	All	526,515	85%	533,680	84%	528,165	84%
Number of providers	Limited	124	94%	125	94%	126	95%
Number of students	Limited	526,515	85%	533,680	84%	528,165	84%
Student data excluded	Limited	0	-	0	-	0	-
Religion or belief							
Number of providers	All	106	80%	108	81%	133	100%
Number of students	All	382,935	62%	415,950	66%	600,420	95%
Number of providers	Limited	104	79%	107	80%	133	100%
Number of students	Limited	382,885	62%	415,950	66%	600,420	95%
Student data excluded	Limited	50	0%*	0	-	0	-
Sexual orientation							
Number of providers	All	109	83%	111	83%	116	87%
Number of students	All	398,090	64%	430,515	68%	447,620	71%
Number of providers	Limited	107	81%	111	83%	116	87%
Number of students	Limited	398,040	64%	430,515	68%	447,620	71%
Student data excluded	Limited	50	0%*	0	-	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 ¹¹
Parental education	Whether the parents of the student have a higher education qualification	Have an HE qualification No HE qualification Don't know Information refused	Based on HESA field PARED ¹²
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 ¹³
Sexual orientation	The sexual orientation of the student	Bisexual Gay man Gay woman/lesbian Heterosexual Other Information refused	Based on HESA field SEXORT ¹⁴

¹¹ <https://www.hesa.ac.uk/collection/c17051/a/genderid>

¹² <https://www.hesa.ac.uk/collection/c16051/a/pared>

¹³ <https://www.hesa.ac.uk/collection/c16051/a/reblf>

¹⁴ <https://www.hesa.ac.uk/collection/c17051/a/sexort>