

Office for
Students



HESA AP data checking tool

2017-18

HEAPES17 comparison technical document

Enquiries to APdata@officeforstudents.org.uk

Publication date August 2018

Contents

HEAPES17 comparison algorithms	3
Purpose	3
HEAPES17 comparison tables	3
Table 1 Excel workbook 'HEAP17_DCT_AXXXXXXXXX.xlsx'	3
Table 2 Fields used in the HEAPES17 re-creation	4
Description of derived fields for re-creating tables	6
Table 3 HEAPES comparison derived fields	6
Instance period data	7
HEAPES comparison context worksheets and coversheet	12
Recruitment against the 2017-18 student number controls (SNC worksheet)	12
Coversheet	12

HEAPES17 comparison algorithms

Purpose

1. This document describes the methods used to generate the data needed to re-create HEAPES from the HESA AP Student data.
2. This document is aimed at readers with in-depth knowledge of the data. Readers are advised to have a copy of the HESA AP Student Record Coding Manual 2017-18 and 'HEAPES17: Higher Education in Alternative Providers Early Statistics Survey' (HEFCE 2017/19) to hand when using this document.

HEAPES17 comparison tables

3. The HEAPES17 comparison tables can be accessed as an IRIS output on the HESA data collection system. The Excel workbook HEAP17_DCT_XXXXXXXX.xlsx (where XXXXXXXX denotes the UK Provider Reference Number (UKPRN)) contains the following worksheets.

Table 1 Excel workbook 'HEAP17_DCT_XXXXXXXX.xlsx'

Worksheet*	Title
Coversheet	Title page
Summary	Summary comparison of HEAPES17 and the HEAPES17 re-creation
FT_1	HEAPES17 re-creation Table 1: Students in the full-time HEAPES population (present on or before 1 December 2017)
FT_2	HEAPES17 re-creation Table 2: Students in the full-time HEAPES population
DLPT_1	HEAPES17 re-creation Table 3: Students in the DLPT HEAPES population (present on or before 1 December 2017)
DLPT_2	HEAPES17 re-creation Table 4: Students in the DLPT HEAPES population
SNC	Report on recruitment against the 2017-18 student number controls using the HEAPES17 re-creation
hFT_1	HEAPES17 Table 1: Students in the full-time HEAPES population (present on or before 1 December 2017)
hFT_2	HEAPES17 Table 2: Students in the full-time HEAPES population
hDLPT_1	HEAPES17 Table 3: Students in the DLPT HEAPES population (present on or before 1 December 2017)
hDLPT_2	HEAPES17 Table 4: Students in the DLPT HEAPES population
hSNC	HEAPES17 recruitment against the 2017-18 student number controls
FT_1Diff	Difference between HEAPES17 and HEAPES17 re-creation Table 1: Students in the full-time HEAPES population (present on or before 1 December 2017)

Worksheet*	Title
FT_2Diff	Difference between HEAPES17 and HEAPES17 re-creation Table 2: Students in the full-time HEAPES population
DLPT_1Diff	Difference between HEAPES17 and HEAPES17 re-creation Table 3: Students in the DLPT HEAPES population (present on or before 1 December 2017)
DLPT_2Diff	Difference between HEAPES17 and HEAPES17 re-creation Table 4: Students in the DLPT HEAPES population

* This worksheet reference corresponds to the spreadsheet tabs.

- The 'DIFF' sheets (see items ending in 'DIFF' in Table 1) will indicate where differences in cell totals between the HEAPES17 re-creation and HEAPES17 exceed a given threshold. The size of this threshold can be altered by entering the desired value where indicated on the worksheets. These sheets are provided to assist providers in reconciling differences between HEAPES17 and the HEAPES17 re-creation.

Using the individualised file

- When working through this document it is necessary to use the individualised file HEAP17_DCT_XXXXXXXXX_IND.csv where XXXXXXXX is the UKPRN. This will show the allocation of students to cells within the tables and, where relevant, details of why they were excluded.

HESA fields used in the re-creation

- Only certain fields, detailed in Table 2, were used to generate the HEAPES re-creation.
- Fields taken from the HESA return or derived as part of the comparison are shown in capitals using the names given in Tables 2 and 3 respectively.

Table 2 Fields used in the HEAPES17 re-creation

Entity	Field	Description	Data set
Course	COURSEAIM	General qualification aim of course	HESA
Course	CTITLE	Course title	HESA
Course	TTCID	Teacher training course identifier	HESA
Instance	COMDATE	Start date of instance	HESA
Instance	NUMHUS	Student instance identifier	HESA
Instance [†]	OWNINST	Provider's own instance identifier	HESA
Instance period	COURSEID	Course identifier	HESA
Instance period	HEAPESPOP	HEAPES population	HESA
Instance period	LOCATION	Location identifier	HESA
Instance period	LOCSDY	Location of study	HESA
Instance period	MODE	Mode of study	HESA
Instance period	PERIODSTART	Instance period start date	HESA

Entity	Field	Description	Data set
Student	HUSID	HESA unique student identifier	HESA
Student [†]	OWNSTU	Provider's own internal identifier for the student	HESA
	CDSAWARD	Awarding body as returned during specific course designation process	CDS [‡]
	CDSCTITLE	Course title as returned during specific course designation process	CDS [‡]
	CDSDISTANCE	Distance learning course identifier as returned during specific course designation process	CDS [‡]
	CDSLOCATION	Location as returned during specific course designation process	CDS [‡]
	CDSMODE	Mode as returned during specific course designation process	CDS [‡]

* The individualised file HEAP17_DCT_AXXXXXXXXX_IND.csv, downloadable as an IRIS output from the HESA data collection system.

† These fields are not used in the re-creation but are included in the individualised file to allow easy identification of students.

‡ The submission made when applying for specific course designation

Description of derived fields for re-creating tables

8. Here we give details of the derived fields in the individualised file. These fields are used to build the HEAPES re-creation tables.

Table 3 HEAPES comparison derived fields

Derived field name	Description	Paragraph
HEAPPERIODSTART1-3†	Instance period start date	11
HEAPCOURSEAIM1-3†	General qualification aim of course	12
HEAPCOURSEID1-3†	Course identifier	13
HEAPCTITLE1-3†	Course title	14
HEAPHEAPESPOP1-3†	HEAPES population	15
HEAPLOCATION1-3†	Location identifier	16
HEAPLOCSDY1-3†	Location of study	17
HEAPMODE1-3†	Mode of study	18
HEAPTTCID1-3†	Teacher training course identifier	19
HEAPCDSAWARD1-3†	Awarding body as returned during specific course designation process	20
HEAPCDSCTITLE1-3†	Course title as returned during specific course designation process	21
HEAPCDSDISTANCE1-3†	Distance learning course identifier as returned during specific course designation process	22 - 23
HEAPCDSLOCATION1-3†	Location as returned during specific course designation process	24
HEAPCDSMODE1-3†	Mode as returned during specific course designation process	25 - 26
HEAPPERIODINCLUDE	Whether the first selected instance period is included in the HEAPES17 re-creation	27
HEAPLEV1-3†	Qualification type	28
HEAPCAT1-3†	Whether the student is in the full-time or DLPT HEAPES population	29
HEAPREG1-3†	Column in Table 2 or 4 that the student will appear in	30

* The individualised file HEAP17_DCT_AXXXXXXXXX_IND.csv, downloadable as an IRIS output from the HESA data collection system.

† Due to file size restrictions the number of occurrences of this field is restricted to 3.

Instance period data

9. The number of instance periods that can be returned for a given instance is unbounded; however, due to file size considerations we restrict the number of separate instance periods displayed in the individualised file to 3. In the unusual case where an instance has more than 3 instance periods, all will be considered in our algorithms, although only 3 will be displayed in the individualised file. In selecting instance period data to show, we will first select those in the HEAPES population (HEAPESPOP = 1), then those that start in the academic year (PERIODSTART > 31 July 2017) and then those with the earliest period start date (PERIODSTART).
10. For providers that complete a HESA student return rather than an AP student return, all fields taken from the instance period entity are taken from instance instead. For example, HEAPCOURSEAIM1 will contain COURSEAIM.

HEAPPERIODSTART1-3

11. For providers making a HESA AP student return, these fields contain the instance period start dates (PERIODSTART) for the student. For providers making a HESA student return, HEAPPERIODSTART1 will contain the anniversary of the start date during the current academic year and HEAPPERIODSTART2 and HEAPPERIODSTART3 will be blank.

HEAPCOURSEAIM1-3

12. These fields contain the course aims (COURSEAIM) for the associated instance period. For example, HEAPCOURSEAIM1 will contain the general qualification aim of the course for HEAPPERIODSTART1. For providers making a HESA student return, HEAPCOURSEAIM1 will contain the general qualification aim of the course for the current academic year and HEAPCOURSEAIM2 and HEAPCOURSEAIM3 will be blank.

HEAPCOURSEID1-3

13. These fields contain the course identifiers (COURSEID) for the associated instance period. For example, HEAPCOURSEID1 will contain the course identifier for HEAPPERIODSTART1. For providers making a HESA student return, HEAPCOURSEID1 will contain the course identifier for the current academic year and HEAPCOURSEID2 and HEAPCOURSEID3 will be blank.

HEAPCTITLE1-3

14. These fields contain the course titles (CTITLE) for the associated instance period. For example, HEAPCTITLE1 will contain the course title for HEAPPERIODSTART1. For providers making a HESA student return, HEAPCTITLE1 will contain the course title for the current academic year and HEAPCTITLE2 and HEAPCTITLE3 will be blank.

HEAPHEAPESPOP1-3

15. These fields contain the HEAPES populations (HEAPESPOP) for the associated instance period. For example, HEAPHEAPESPOP1 will contain the HEAPES population for HEAPPERIODSTART1. For providers making a HESA student return, HEAPHEAPESPOP1

will contain the HEAPES population for the current academic year and HEAPHEAPESPOP2 and HEAPHEAPESPOP3 will be blank.

HEAPLOCATION1-3

16. These fields contain the location identifiers (LOCATION) for the associated instance period. For example, HEAPLOCATION1 will contain the location identifier for HEAPPERIODSTART1. For providers making a HESA student return, HEAPLOCATION1 will contain the location identifier for the current academic year and HEAPLOCATION2 and HEAPLOCATION3 will be blank.

HEAPLOCSDY1-3

17. These fields contain the location of study (LOCSDY) for the associated instance period. For example, HEAPLOCSDY1 will contain the location of study for HEAPPERIODSTART1. For providers making a HESA student return, HEAPLOCSDY1 will contain the location of study for the current academic year and HEAPLOCSDY2 and HEAPLOCSDY3 will be blank.

HEAPMODE1-3

18. These fields contain the mode of study (MODE) for the associated instance period. For example, HEAPMODE1 will contain the mode of study for HEAPPERIODSTART1. For providers making a HESA student return, HEAPMODE1 will contain the mode of study for the current academic year and HEAPMODE2 and HEAPMODE3 will be blank.

HEAPTTCID1-3

19. These fields contain the teacher training course identifier (TTCID) for the associated instance period. For example, HEAPTTCID1 will contain the teacher training course identifier for HEAPPERIODSTART1. For providers making a HESA student return, HEAPTTCID1 will contain the teacher training course identifier for the current academic year and HEAPTTCID2 and HEAPTTCID3 will be blank.

HEAPCDSAWARD1-3

20. These fields contain the awarding body (CDSAWARD) for the associated instance period as returned during specific course designation process. For example, HEAPCDSAWARD1 will contain the awarding body for the course for HEAPPERIODSTART1. This is obtained by linking the specific course designation submission to the AP student record using the course identifier (COURSEID). For providers with institutional level designation these fields will be blank.

HEAPCDSCTITLE1-3

21. These fields contain the course title (CDSCTITLE) for the associated instance period as applied for through specific course designation. For example, HEAPCDSCTITLE1 will contain the course title for HEAPPERIODSTART1. This is obtained by linking the specific course designation submission to the AP student record using the course identifier (COURSEID). For providers with institutional level designation these fields will be blank.

HEAPCDSDISTANCE1-3

22. For providers with specific course designation, these fields contain the distance learning course identifier (CDSDISTANCE) for the associated instance period as applied for through specific

course designation. For example, HEAPCDSDISTANCE1 will contain the distance learning course identifier for HEAPPERIODSTART1. This is obtained by linking the specific course designation submission to the AP student record using the course identifier (COURSEID).

23. For providers with institutional level designation, these fields allocate a distance learning course identifier for the course, where X denotes the associated instance period. For providers making a HESA student return, HEAPCDSDISTANCE1 will contain the distance learning course identifier for the current academic year and HEAPCDSDISTANCE2 and HEAPCDSDISTANCE3 will be blank

Value	Description	Definition
YES	Distance learning course	HEAPLOCSDYX = 6
NO	Otherwise	Otherwise

HEAPCDSLOCATION1-3

24. These fields contain the location identifiers (CDSLOCATION) for the associated instance period as applied for through specific course designation. For example, HEAPCDSLOCATION1 will contain the course title for HEAPPERIODSTART1. This is obtained by linking the specific course designation submission to the AP student record using the course identifier (COURSEID). For providers making a HESA student return these fields will be blank.

HEAPCDSMODE1-3

25. For providers with specific course designation, these fields contain the mode (CDSMODE) for the associated instance period as applied for through specific course designation. For example, HEAPCDSMODE1 will contain the mode for HEAPPERIODSTART1. This is obtained by linking the specific course designation submission to the AP student record using the course identifier (COURSEID).
26. For providers with institutional level designation, these fields allocate a mode for the course to the associated instance period, where X denotes the instance period. For providers making a HESA student return, HEAPCDSMODE1 will contain the mode for the current academic year and HEAPCDSMODE2 and HEAPCDSMODE3 will be blank

Value	Description	Definition
FT	Full-time course	HEAPMODEX ≠ 31, 44, 74*
PT	Otherwise	Otherwise

*This code is only relevant to providers making a full HESA student return

HEAPPERIODINCLUDE

27. This field indicates whether the first instance period selected using the approach in paragraph 9 is counted in the HEAPES17 re-creation.

Value	Description	Definition
1	Included in the HEAPES17 re-creation	HEAPHEAPESPOP1 = 1 and HEAPPERIODSTART1 > 31 July 2017
0	Otherwise	Otherwise

HEAPLEV1-3

28. These fields allocate students to a qualification type for the associated instance period, where X denotes the instance period. For example, HEAPLEV1 will contain the qualification type for HEAPPERIODSTART1.

Value	Description	Definition
PG_ITT	Postgraduate initial teacher training not leading to QTS	HEAPTTCIDX = 2 and HEAPCOURSEAIMX = M71, M73, H71
UG_ITT	Undergraduate initial teacher training not leading to QTS	HEAPTTCIDX = 2 and not above
HNC_HND	Higher National Certificate or Higher National Diploma	HEAPCOURSEAIMX = J30, C30 and not above
FIRST_DEGREE	First degree	HEAPCOURSEAIMX = M22, M26, M28, H00, H11, H16, H18, H22, H23, H50, I00, I11, I16 and not above
OTHER_UG	Other undergraduate	HEAPCOURSEAIMX = H41, H42, H43, H60, H70, H72, H76, H79, H80, H88, H90, H91, H99, I60, I70, I72, I74, I76, I80, I81, I90, I91, I99, J, C and not above
OTHER	Otherwise	Otherwise

HEAPCAT1-3

29. These fields identify whether the student is in the full-time or DLPT HEAPES population for the associated instance period, where X denotes the instance period. For example, HEAPCAT1 will identify which HEAPES population the student is in for HEAPPERIODSTART1.

Value	Description	Definition
FT	Student is in the full-time HEAPES population	HEAPCDSMODEX = FT and HEAPCSDSDISTANCEX = NO
DLPT	Student is in the DLPT HEAPES population	HEAPCDSMODEX = PT or HEAPCSDSDISTANCEX = YES

HEAPREG1-3

30. These fields identify whether the student will appear in Column 1 or 2 of Table 2 or 4 for the associated instance period, where X denotes the instance period. For example, HEAPREG1 will identify which column the student is in for HEAPPERIODSTART1.

Value	Description	Definition
1	Instance period starting on or before 1 December 2017	HEAPHEAPESPOPX = 1 and HEAPPERIODSTARTX < 2 December 2017
2	Instance period starting after 1 December 2017 and before 1 August 2018	HEAPHEAPESPOPX = 1 and HEAPPERIODSTARTX > 1 December 2017
0	Otherwise	Otherwise

HEAPES comparison context worksheets and coversheet

31. This section details how HEAPES re-creation data is used to inform the HEAPES comparison coversheet. It also describes how this data can be rebuilt from the HEAPES comparison individualised file.

Recruitment against the 2017-18 student number controls (SNC worksheet)

Full-time student number control for 2017-18: Recruited numbers

32. The students used to derive 'Recruited numbers' can be identified by selecting:

```
HEAPPERIODINCLUDE = 1  
HEAPHEAPESPOP1 = 1  
HEAPCAT1 = FT
```

Distance learning/part-time student number control for 2017-18: Recruited numbers

33. The students used to derive 'Recruited numbers' can be identified by selecting:

```
HEAPPERIODINCLUDE = 1  
HEAPHEAPESPOP1 = 1  
HEAPCAT1 = DLPT
```

Coversheet

34. The cover sheet consists of the following comparisons between HEAPES and the HEAPES re-creation:

Students in the full-time HEAPES population (present on or before 1 December 2017)

35. The difference is calculated by summing column 1 of FT_2 (HEAPES17 re-creation) and then subtracting the sum of column 1 of hFT_2 (HEAPES17) from this.

Students in the full-time HEAPES population

36. The difference is calculated by summing column 3 of FT_2 (HEAPES17 re-creation comparison) and then subtracting the sum of column 3 of hFT_2 (HEAPES17) from this.

Recruitment against the 2017-18 full-time student number control

37. The difference is calculated by taking the 'Full-time student number control for 2017-18: Recruited numbers' in the 'SNC' sheet and then subtracting the 'Full-time student number control for 2017-18: Recruited numbers' in the 'hSNC' sheet from this.

Students in excess of the full-time student number control

38. The difference is calculated by taking the 'Students in excess of the full-time student number control' in the 'SNC' sheet and then subtracting the 'Students in excess of the full-time student number control' in the 'hSNC' sheet from this.

Students in the DLPT HEAPES population (present on or before 1 December 2017)

39. The difference is calculated by summing column 1 of DLPT_2 (HEAPES17 re-creation) and then subtracting the sum of column 1 of hDLPT_2 (HEAPES17) from this.

Students in the DLPT HEAPES population

40. The difference is calculated by summing column 3 of DLPT_2 (HEAPES17 re-creation) and then subtracting the sum of column 3 of hDLPT_2 (HEAPES17) from this.

Recruitment against the 2017-18 distance learning/part-time student number control

41. The difference is calculated by taking the 'Distance learning/part-time student number control for 2017-18: Recruited numbers' in the 'SNC' sheet and then subtracting the 'Distance learning/part-time student number control for 2017-18: Recruited numbers' in the 'hSNC' sheet from this.

Students in excess of the distance learning/part-time student number control

42. The difference is calculated by taking the 'Students in excess of the distance learning/part-time student number control' in the 'SNC' sheet and then subtracting the 'Students in excess of the distance learning/part-time student number control' in the 'hSNC' sheet from this.



© The Office for Students copyright 2018

This publication is available under the Open Government Licence 3.0.

www.nationalarchives.gov.uk/doc/open-government-licence/version/3/