

HESA 1999-2000 derived statistics for funding allocations and monitoring

To	Heads of HEFCE-funded higher education institutions Heads of universities in Northern Ireland
Of interest to those responsible for	Student data, Audit, Finance
Reference	01/09
Publication date	February 2001
Enquiries to	For general enquiries contact: Anthony Ryan tel 0117 931 7297 Rhianne Cox tel 0117 931 7014 e-mail hesa_heses_stats@hefce.ac.uk For enquiries regarding franchising and campus data contact: Thomas Jackson tel 0117 931 7373 e-mail t.jackson@hefce.ac.uk

Executive summary

Purpose

1. This document describes how we will use the Higher Education Statistics Agency (HESA) 1999-2000 individualised student record to:

- monitor aggregate returns made to the HEFCE
- inform funding for 2001-02
- produce some regional analysis.

2. This document also outlines the procedures we will adopt where our monitoring process identifies significant differences between aggregate and individualised returns. We believe this exercise helps us, and institutions, to better understand the data supplied to HESA and ensures that higher quality more consistent data are delivered to both HESA and the HEFCE.

3. The data included in the reconstruction of the Higher Education Students Early Statistics survey 1999 (HESES99) from the HESA record will form the basis of calculating premiums based on the following:

- full-time mature undergraduate students
- the neighbourhood types of students
- students receiving a disabled students allowance.

These premiums will be used in the funding allocations for 2001-02.

4. Annex A gives full details of the methods used in generating the comparisons, including details of where assumptions have been made due to poor fit between the data collected on the HESES99 survey and that supplied to HESA. In addition, the code used to generate the comparisons can be found on our web-site, www.hefce.ac.uk under 'Learning and Teaching', 'Data collection'.

Franchised data

5. In HEFCE 00/55 we published details of franchised students elicited from the HESA 1998-99 return. We are providing information on franchised students from the HESA 1999-2000 return to enable institutions to check that it is accurate and suitable for publication. We expect to publish this information during 2001.

Campus data

6. We are providing summary campus information to enable institutions to check that it is accurate and suitable as the basis for use in regional analysis publications.

Performance indicators

7. A consultation document, including draft performance indicators (PIs) based on HESA data from 1999-2000 and earlier, will be sent to institutions after Easter. This is the third year we will publish PIs and we expect significant improvements in the quality of HESA data on previous years. In 'Consultation on performance indicators' (HEFCE 00/18) we clearly stated our expectation that the number of institutions amending their HESA data following this year's PI consultation would be minimal (see paragraphs 19-21 of HEFCE 00/18).

8. Details of the methodology and fields used in producing the PIs published in 2000 were given in HEFCE 00/18. The methodology and fields used in the 2001 PI publication are not expected to change. Institutions are therefore urged to take this opportunity to check data that will be used in the production of PIs to ensure that errors can be corrected in time to inform the publication of PIs. Where errors in HESA data affecting PIs are identified, institutions should contact Judy Akinbolu on 0117 931 7110 or e-mail j.akinbolu@hefce.ac.uk at the earliest opportunity.

Allocation of student load to academic cost centres

9. As part of our commitment to monitoring data used in funding we are currently developing a system for monitoring the allocation of student load to academic cost centres. Because of the way activity is allocated to academic cost centres different subjects may be reported to different cost centres in different institutions. However, over the institution as a whole we would expect the net effect of these variations to be small. Where we identify institutions whose allocations vary significantly from the norm we will include additional data with the data from this exercise seeking fuller details of the rationale used in assigning activity to cost centres. In future years we hope to include full details of this monitoring within this exercise.

Key points

10. Where we identify significant differences in the comparison of HESA data to HESES99, we will write to the institutions concerned and ask for an explanation.

11. If the data remain unreconciled, we may audit both sets of data to arrive at agreed figures. Any amendments to data following this exercise may result in retrospective adjustments to funding.

12. If institutions are concerned about the suitability for publication of franchised or campus data they should contact Thomas Jackson by 6 April 2001.

13. All institutions are advised to check the data quality of fields used in the production of PIs in conjunction with checks made as part of this exercise and ahead of the PI consultation which will be sent to institutions after Easter.

14. All institutions are invited to comment on the methods described in Annex A or in the SAS code.

Action required

15. Where we require a response it should be sent to Anthony Ryan no later than 6 April 2001.

Selection of institutions asked to respond

16. We will ask for a response from institutions if any of the following are true:

- a. The total difference in holdback calculated using HESES99 and HESA data exceeds £1,000,000.
- b. The difference in holdback is more than 10 per cent of teaching grant for 1999-2000.
- c. Average load factors vary by more than 0.1.

17. Holdback was chosen as the basis of selection because:

- a. Holdback is calculated at the highest level of aggregation used in operating our funding process.
- b. It shows major changes in funding while remaining neutral to variations in other institutions' data.

18. This comparison takes place after the data have been finalised with HESA. Changes will only be made to our copy of the HESA data if large discrepancies that can easily be corrected become apparent. If such changes are necessary we shall expect institutions to adopt procedures designed to ensure that, in future years, the final data submitted to HESA are correct. We may refuse to accept amendments where errors have previously been identified by HESA during collection.

Responses required

19. We will write separately to the institutions from which we require a response.

20. Responses should address one, or more, of the following problems and quantify the extent to which it contributes to the overall discrepancy.

- errors in HESES99
- errors in HESA
- problems of fit with the HEFCE algorithm.

Errors in HESES99

21. Where errors are found in the HESES return a revised return will be required. Institutions will be informed of the outcome of these changes by their HEFCE higher education adviser.

Errors in HESA

22. Institutions are required to submit timely and accurate data to HESA. However, it is recognised that HESA returns are necessarily complicated and that errors will occur in these returns. Therefore in previous years we have accepted a large number of amendments to HESA data as a result of this exercise and the annual 'Consultation on performance indicators'. This is the second year of data collected using the revised record structure and we would expect the number of errors to be greatly reduced.

23. The implications of processing and accepting amendments to HESA data are significant; so we want to keep amendments to a minimum. In particular we would not expect to make amendments where:

- a. Institutions have been notified of the error by HESA during the data collection process.
- b. Institutions have already identified a similar weakness in earlier years' data and we have not agreed to a continuation of amendments.
- c. The error does not significantly affect the outcome of this exercise, the PIs or other analysis.

24. If significant errors are identified in the data and we do not accept the amendment, this will be acknowledged in the performance indicator publication.

25. Where institutions are amending HESA data we require the corrections to be submitted to us in a standard format. Details of the standard format are given in Annex B. This is essential in order to establish an audit trail of data changes, and to ensure that corrections can be processed in a timely and accurate manner. If amendments to HESA data are received we will use this information to re-create HESES99 tables. Where we are content the amendments result in a reasonable comparison to the HESES99 return, we will ask the institution to confirm the accuracy of the amendments. A copy of the confirmation form is given in Annex C. The form should be photocopied and signed by the member of staff with responsibility for signing-off HESA data. Where we are not content that the amendments result in a reasonable comparison, we will ask for a further response. Details of this process are given in Figure 1.

Problems of fit with the HEFCE algorithm

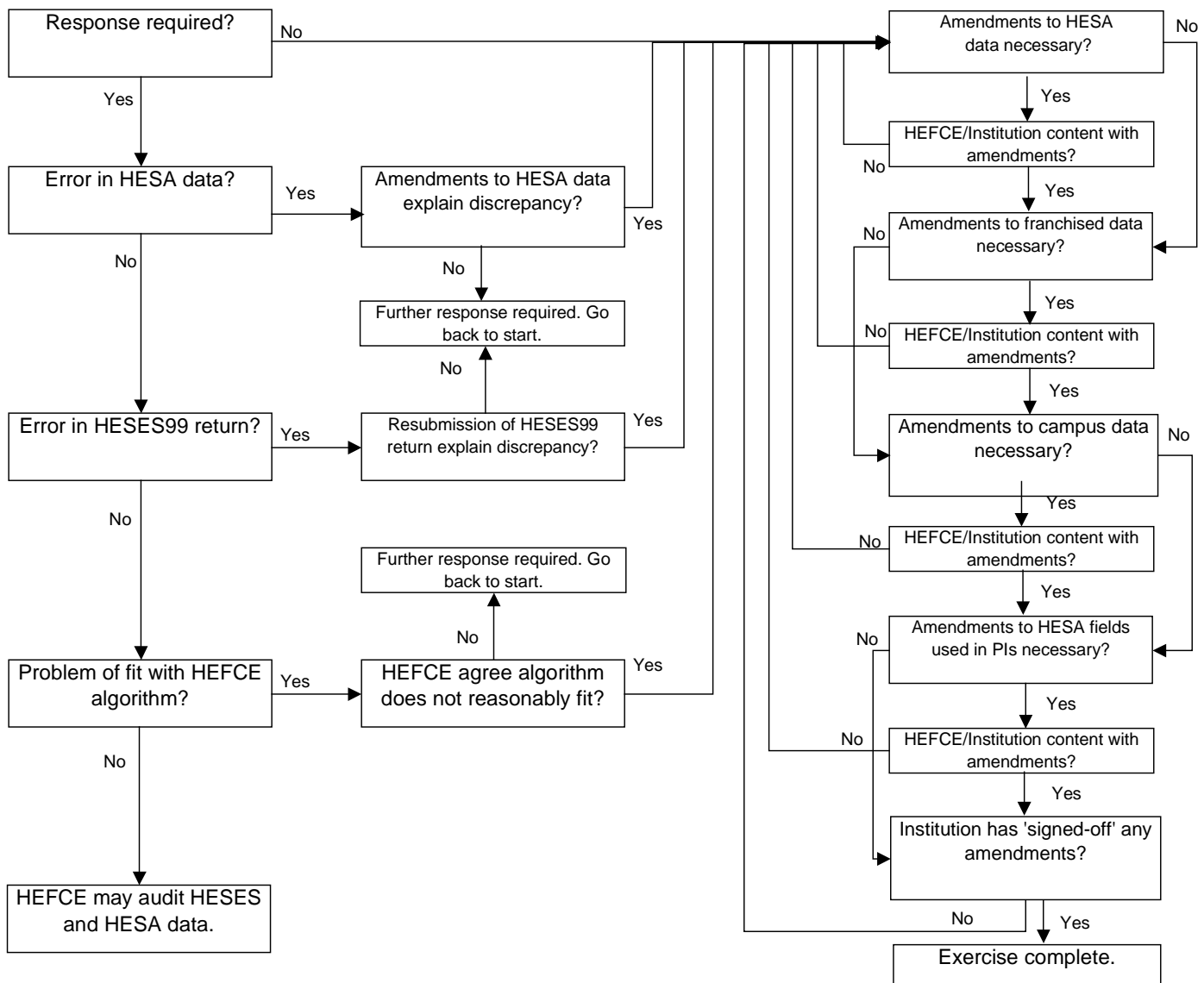
26. Where a difference of fit between our algorithm and data supplied explains the discrepancy, evidence of where this occurs will be required. To aid institutions in identifying such cases, the SAS code used to generate the comparison is available on our web-site under 'Learning and Teaching', 'Data collection'.

27. All institutions are invited to comment on the methods described in Annex A, and to suggest how they can be improved.

28. If differences between the data cannot be reconciled, we may audit both sets of data. If data are amended, we may make retrospective adjustments to funding.

29. We believe this exercise helps to improve the quality both of data supplied to HESA and of responses to the HESES survey.

Figure 1 Response process diagram



Information supplied

30. We will write to heads of institutions, copied to HESES2000 data contacts, by 28 February 2001 indicating whether a response is required and enclosing the following information:

- a. A copy of the HESES99 return, including all amendments made during the funding process.
- b. HESES99 as re-created using signed-off data supplied to HESA and the coding methods described in Annex A.
- c. The number of mature home and EC fundable, full-time undergraduates.
- d. The number of young, home and EC fundable, full-time undergraduates receiving some tuition-fee remittance.
- e. The number of home and EC, full-time, undergraduate entrants receiving a disabled students allowance.
- f. The numbers of young, English domiciled, home and EC fundable, full-time undergraduates in each of 160 neighbourhood types.
- g. The teaching institutions of franchised students and the number of such students.
- h. Summary information on campus details and activity.
- i. Summary information comparing HESES99 to HESA data including a summary of franchised students.
- j. A summary of HESA records excluded from the HESES re-creation.
- k. Details of the allocation of media studies and psychology to price groups.

31. All the above data are available electronically from our web-site. Details of how to obtain these data will be included with the letter referred to above.

32. Our web-site will also contain an individualised record. The record contains the following HESA fields:

- RECID
- CAMPID
- HUSID
- BIRTHDTE
- DISALL
- COMDATE
- SPCSTU
- DATELEFT
- QUALAIM
- SPLLENGTH
- UNITLGTH
- TTCID
- FUNDCODE
- FEEELIG
- MSTUFEE
- FEEBAND
- MODE
- LOCSDY
- YEARPRG
- STULOAD
- POSTCODE
- OWNSTU
- OWNPSD
- NUMHUS
- TYPEYR
- FUNDLEV
- FUNDCOMP
- INSTID

33. It also contains the following derived fields:

- a. A flag showing whether the student is counted in HESES99 and, if not, why the student was excluded.
- b. The HESES classification of the student.
- c. The classification of the student's neighbourhood type used to calculate the geodemographic premium. This field is completed for a restricted group of students only.
- d. The campus name of the student.
- e. The classification of our record of the method used to return FTE if the student is on a non-standard academic year.
- f. A flag indicating whether we allocate sport and leisure science cost centre activity to price group C.

34. We are conducting a similar exercise for further education colleges, using the Further Education Funding Council's Individualised Student Record.

contain any amendments or clarifications that need to be made after the circular and supplementary information are sent out. It can be found on the HEFCE

Frequently asked questions

35. There is a web page featuring answers to frequently asked questions. The web page will also

web-site under 'Learning and Teaching', 'Data collection'. The web page is updated regularly, and institutions are expected to look here for guidance in the first instance. We will use our e-mail list of HESES2000 contacts to notify institutions of any significant changes or updates. We will not use this simply to notify them of changes to the web-site.

Anthony Ryan
Analytical Services Group
HEFCE
Northavon House
Coldharbour Lane
BRISTOL
BS16 1QD

Deadline for responses

36. Responses should arrive **no later than 6 April 2001** and should be sent to:

Annex A

Technical supplement

Purpose

1. This annex describes the method used to generate the data distributed to institutions as part of this exercise. It also gives details of known discrepancies between the data sources.
2. This annex is aimed at expert readers with in-depth knowledge of the data. Readers are advised to have a copy of the 1999-2000 HESA coding manual and HESES99 (HEFCE 99/57) to hand when using this annex. In addition, users may wish to consult the SAS code used in the comparison which can be found on our web-site under 'Learning and Teaching', 'Data collection'.
3. The details listed below are similar to those in HESES99 Annex N, but some alterations have been made to improve the matching and clarity.

HESA fields used in comparisons

4. Only certain fields, detailed below, were used to generate the comparison between the HESES and HESA data. The field numbers shown relate to the combined record format of the HESA record. For institutions making a student module return, cost centre information is taken from the module portion of the return.
5. Throughout this annex, fields taken from the HESA return are shown in capitals using the names given below. Where we have used fields derived by HESA, a description of the algorithms used is given in Annex D.

Field number	Description	Name
1	Record type indicator	RECID
2	HESA institution identifier	INSTID
3	Campus identifier	CAMPID
4	Student identifier	HUSID
10	Date of birth	BIRTHDTE
15	Disability allowance	DISALL
26	Date of commencement of programme	COMDATE
28	Special students	SPCSTU
35	Date left institution or completed the programme of study	DATELEFT
41	General qualification aim of student	QUALAIM
43	Subject of qualification aim	SBJQA
49	Expected length of study programme	SPLENGTH
50	Units of length	UNITLGTH
53	Teacher training course identifier	TTCID
65	Fundability code	FUNDCODE
66	Fee eligibility	FEEELIG
67	Fee band	FEEBAND
68	Major source of tuition fees	MSTUFEE
70	Mode of study	MODE
71	Location of study	LOCSDY
72	Year of programme	YEARPRG

74	Student FTE	STULOAD
75	Postcode	POSTCODE
86	Other institution providing teaching 1	TINST1
87	Other institution providing teaching 2	TINST2
90	Proportion not taught by this institution	PCOLAB
100,103,106,109,112,115,118,1 21,124,127,130,133,136,139,14 2,145	Cost centre 1-16	COSTCN01-16
102,105,108,111,114,117,120,1 23,126,129,132,135,138,141,14 4,147	Proportion of subject 1-16	SBJPER01-16
151	Student instance number	NUMHUS
153	Type of programme year	TYPEYR
154	Level applicable to Funding Council HESES	FUNDLEV
155	Completion of year of programme of study	FUNDCOMP

6. The POSTCODE field was used to determine the neighbourhood type of full-time and sandwich, home and EC fundable undergraduates in the HESES column 4 population, who were young on entry and had an English postcode. The MSTUFEE field was used to determine those students who had some tuition fee remittance. The DISALL field was used to determine those students who received the Disabled Students' Allowance (DSA).

Linking programmes of study between years

7. Using the HUSID, INSTID, NUMHUS (HIN) triplet, we have in certain cases linked HESA data

between 1998-99 and 1999-2000 where we believe this will improve approximations arising from discrepancies between the two data sources. A link was attempted for the following groups of students:

- a. Students who are writing up a thesis or dissertation.
- b. Programmes of study generating two countable years in the final academic year.
- c. Non-standard academic year programmes of study of less than two years in duration where the FTE for each year of programme of study is split over two HESA returns.

Description of derived fields

8. This section contains details of the derived fields contained on the individualised data file. These fields are used to build the key dimensions of the HESES return.

Field name	Description	Paragraph
PRIKEY	Unique record identifier	9
FTE_TYPE	Method used to return FTE for non-standard academic years	11
HESMODE	Mode of study	16
HESLEVEL	Level of study	17
HESTYPE	Fundability status	18
HESFEELV	Fee level used in HESES	21
YEARONE	New entrant flag	22
LENGTH	Flag indicating long or standard length years of programme of study	24
XPRP101	Cost centre proportion indicator	25
Price groups	Proportion of countable year in each price group	26
XSBJA01	Principle subject of study	27
SPORTS	Flag indicating allocation of cost centre 38 to price groups	31
HESREG	HESES column 1 or 2 indicator	32
HESCOMP	HESES completion of year of programme of study flag	33
LOW_FTE	Flag indicating whether assumptions have been made where data are missing	34
WUP_LINK	Flag indicating whether linking was used for writing-up students	35
ATT_LINK	Flag indicating whether linking was used for course attributes	37
STUBID	Unique countable year of programme identifier	38
FTE_LINK	Flag indicating whether linking was used for FTE	40
STULOA98	STULOAD field from 1998-99 HESA July record	41
FTE_CASE	Indicator showing how HESESFTE was calculated	42
HESESFTE	FTE consistent with HESES definitions	45
HESEXCL	Reason for exclusion from the HESES population	49
HESCOL4	Flag indicating whether the student was included in HESES column 4	51
FRANCH	Flag indicating franchised students	52
FRNINST1	Franchised institution code(s)	53
FRNINST2		
TINSTNME	Name of institution(s) provision franchised to	53
XCOLAB01	Proportion of provision franchised	55
FRANFTE	Franchising FTE	56
FRANEXT	Extent of franchising	57
CAMPNME	Name of campus	58
CAMPPOST	Postcode of campus	59
SPC	Geodemographic grouping indicator	60

PRIKEY

9. This is a HESA derived field which uniquely identifies HESA records.

Method of reporting FTE

10. Information was used on the method chosen to return student load on the HESA student record, because this affects the way years of programme of study are counted.

FTE_TYPE

11. This field is used to identify the institution's method of returning FTE for students on non-standard academic years.

Value	Description
1	Standard academic year
2	Split FTE
3	100:0
4	0:100

Standard academic year

12. Where all the institution's activity for years of programme of study are within one academic year.

Split FTE

13. Where activity for a year of programme of study spans two academic years the FTE is split proportionally across them.

100:0

14. Where activity for a year of programme of study spans two academic years the whole of the FTE is reported in the academic year in which the year of programme of study begins.

0:100

15. Where activity for a year of programme of study spans two academic years the whole of the FTE is reported in the academic year in which the year of programme of study ends.

HESMODE

16. This field allocates students to mode of study.

Value	Description	Definition
FTS	Full-time and sandwich	Mode = 01, 52, 53 or (MODE = 23, 24 and FEEBAND ≠ 02, 42 and LOCSDY ≠4)
SWOUT	Sandwich year-out	MODE = 23, 24 and FEEBAND = 02, 42 and LOCSDY= 4, 8
PT	Part-time	Students not meeting the criteria above

HESLEVEL

17. This field allocates students to level of study.

Value	Description	Definition
UG	Undergraduate	FUNDLEV = 10, 11
PGT	Postgraduate taught	FUNDLEV = 20, 21
PGR	Postgraduate research	FUNDLEV = 30, 31

HESTYPE

18. This field allocates students to the four categories of fundability and residential status.

Value	Description	Definition
HOMEF	Home and EC HEFCE funded	FUNDCODE = 1
HOMEIF	Home and EC independently funded	FUNDCODE = 4
HOMENF	Home and EC non-fundable	FUNDCODE = 2, 5, 7 and FEELIG = 1 or 3.
ISOV	Island and overseas	Any not included above

Postgraduate research students

19. Full-time, home and EC postgraduate research students are assigned fundability status as follows:

Value	Description	Definition
HOMENF	Home and EC non-fundable	((FTE_TYPE=1,3 or TYPEYR=1) and COMDATE < 1 August 1999) or (FTE_TYPE=2,4 and TYPEYR=2,4,5 and COMDATE < 1 August 1998)
HOMEF, HOMEIF	See table in paragraph 18	Otherwise

20. Part-time, home and EC postgraduate research students are assigned fundability status as follows:

Value	Description	Definition
HOMENF	Home and EC non-fundable	((FTE_TYPE=1,3 or TYPEYR=1) and COMDATE < 1 August 1998) or (FTE_TYPE=2,4 and TYPEYR=2,4,5 and COMDATE < 1 August 1997)
HOMEF, HOMEIF	See table in paragraph 18	Otherwise

HESFEELV

21. This field contains the level of tuition fee charged to the student as in HESES99.

Value	Description	Definition
1025	Undergraduate full fee	FEEBAND = 01
510	Undergraduate half fee	FEEBAND = 02
2675	Postgraduate full fee	FEEBAND = 41
0	Regulated zero	FEEBAND = 99 and MSTUFEE = 98 and LOCSDY= 5
OTHER	Other fee charged or no fee level	Any not included above

YEARONE

22. This indicates whether a student is a new entrant as defined in HESES99.

Value	Description	Definition
1	New entrant	(FTE_TYPE = 1, 3 and YEARPRG = 1) or (FTE_TYPE = 2, 4, and TYPEYR = 1 and YEARPRG = 1) or (FTE_TYPE = 2, 4, and TYPEYR = 2, 4, 5 and YEARPRG = 2)
0	Otherwise	Any not included above.

23. For students on a course for which a year of programme is not a recognised concept, i.e. YEARPRG = 99, we calculated an indicative YEARPRG as one plus the number of elapsed years between COMDATE and 31 July 2000 for use in the above calculations.

LENGTH

24. This field indicates whether the student is on a standard or long year of programme of study.

Value	Description	Definition
L	Long (year of programme of study is over 45 weeks)	FUNDLEV = 11, 21, 31
S	Standard	Otherwise

XPRP101

25. XPRP101 is a field derived by HESA which evaluates the proportion of FTE to each cost centre / subject combination. A description of the algorithm used to derive this field is given in Annex D.

Price groups

26. The proportion of activity in each price group is contained in the eight price group fields given in the table below. The proportion of activity in each price group is calculated by mapping cost centre codes to price groups and summing the values of XPRP101 for each cost centre / subject combination over each price group. The table below shows the mapping of cost centre codes to price group fields and the value each field will take.

Field name	Cost centres	Value of field
PRGA	See paragraphs 28 and 29	
PRGB	01 [#] , 02 [#] , 03 [#] , 04, 08, 09, 10, 11, 12, 13, 14, 16, 17, 18, 19, 20, 21, 39	sum of XPRP101s
PRGC	05, 06, 23, 24, 25, 26, 28, 33, 36, 37	sum of XPRP101s
PRGD	27, 29, 30, 31, 32, 34, 35, 38*, 41	sum of XPRP101s
PRGMEDIA	30	sum of XPRP101s
PRGPSYCH	07	sum of XPRP101s
PRGITT	Courses of initial teacher training leading to QTS (TTCID=1)	1
PRGINSET	Courses of in-service education of teachers (TTCID=3), where the student has QTS	1

[#] Except those students identified as clinical medicine, dentistry and veterinary science in paragraphs 28 and 29.

* Except those institutions expected to have returned their sports science and leisure studies to price group C in the HESES99 tables.

XSBJA01

27. XSBJA01 is a field derived by HESA to identify primary subject of study. A description of the algorithm used to derive this field is given in Annex D.

Clinical medicine, dentistry and veterinary science

Undergraduates

28. Clinical medicine, dentistry and veterinary science undergraduates were assigned to price groups as follows:

Field	Description	Definition	Value of field
PRGA	Clinical medicine and dentistry	FUNDLEV= 10, 11 and QUALAIM = 19 and CRSELGTH* - YEARPRG = 0, 1, 2	1
PRGA	Veterinary science	QUALAIM = 19 and XSBJA01 = D1, D10	1
PRGA	Clinical dentistry	FUNDLEV = 10, 11 and QUALAIM=19 and XSBJA01 = A4, A40 and CRSELGTH* - YEARPRG = 3	One third
PRGB	Clinical dentistry	FUNDLEV = 10, 11 and QUALAIM=19 and XSBJA01 = A4, A40 and CRSELGTH* - YEARPRG = 3	Two thirds

* CRSELGTH is a calculation of the length of course using SPLENGTH and UNITLGTH. Where UNITLGTH=9, CRSELGTH=6 was assumed.

Postgraduates

29. Postgraduate activity in cost centres 01, 02 or 03 where XSBJA01=A3, A4, D1, A30, A40 or D10 is assigned to price group A.

Sports and leisure science

30. In 1998 we reviewed the mapping of the Sports Science and Leisure cost centre (cost centre 38) to price groups. As a result, a list of institutions was drawn up whose provision in this cost centre met threshold criteria for the use of well equipped sports science laboratories and/or sports facilities. This list was used in the allocation of students to price groups.

SPORTS

31. A flag to identify whether sports science and leisure studies activity is assigned to price group C at the institution.

Value	Description
1	Sports science allocated to price group C
0	Sports science allocated to price group D

HESREG

32. This field indicates whether the student will appear in column 1 or 2 of the appropriate HESES99 tables. It should be noted that if the student is excluded this field is not used to populate the tables.

Value	Description	Definition
1	Included in column 1	(FTE_TYPE=1,3 and start of year of programme of study < 2 December 1999) or (FTE_TYPE=2,4 and (TYPEYR=2,3,4,5 or (TYPEYR=1 and start of year of programme of study < 2 December 1999)))
2	Included in column 2	Otherwise

HESCOMP

33. This field indicates whether the student will appear in column 3 or 4 of the appropriate HESES99 tables. It should be noted that if the student is excluded this field is not used to populate the tables.

Value	Description	Definition
3	Included in column 3	((FTE_TYPE=1,3 or TYPEYR = 1) and FUNDCOMP=2) or (FTE_TYPE=2,4 and FUNDCOMP=2 and DATELEFT < anniversary of COMDATE and TYPEYR = 2,3,4,5)
4	Included in column 4	Otherwise

LOW_FTE

34. This field identifies students on low credit bearing courses. The following assumptions have been made for these students:

- SPCSTU = 9
- UNITLGTH = 1
- FEEBAND = 51
- LOCSDY = 1
- if DATELEFT is completed then SPLength = DATELEFT - COMDATE rounded up to the nearest year otherwise SPLength = 2.

Value	Description	Definition
1	Assumptions have been made	RECID = 99111, 99112, 99113
0	Assumptions have not been made	All other students.

WUP_LINK

35. This field indicates whether a link has been made to improve our estimate of MODE for writing-up students.

Value	Description	Definition
1	MODE from HESA 1998-99 assumed	<u>1999-2000 HESA data</u> MODE=41, 42, 43, 44 <u>1998-99 HESA data</u> MODE ≠ 41, 42, 43, 44
0	Otherwise	Any student not included above

Second countable years of programme of study

36. Non-standard academic years returned using the split FTE or 0:100 methods, where all activity for the final year of programme of study falls entirely within an academic year, generate two countable years of programme of study.

ATT_LINK

37. This field indicates whether a link has been made, when two years of programme of study are returned, to improve our estimate of attributes for the first countable year.

Value	Description	Definition
1	HESA record generates two countable years of programme of study	<u>1999-2000 HESA data</u> FTE_TYPE=2, 4 and TYPEYR=1 and COMDATE<1 August 1999 and DATELEFT<1 August 2000 and DATELEFT>anniversary of COMDATE + 14 days <u>1998-99 HESA data</u> FTE_TYPE=2, 4 and TYPEYR=2, 3, 4
0	Otherwise	Any not included above

STUBID

38. This field uniquely identifies years of programme of study when two years are generated.

Value	Description	Definition
1	First countable year of programme of study	ATT_LINK = 1
2	Second countable year of programme of study	ATT_LINK = 1
0	Otherwise	ATT_LINK = 0

39. When STUBID=1, we used HESA 1998-99 data in place of, or in the derivation of, the following fields:

- PRGA • PRGPSYCH • FEEBAND • FUNDCOMP • LOW_FTE • PRGMEDIA
- PRGB • FUNDCODE • FUNDEV • YEARPRG • FEEELIG • COSTCN
- PRGC • QUALAIM • FTE_TYPE • CAMPID • PRGITT • PRGINSET
- PRGD • SPCSTU • LOCSDY • XSBJA01 • XPRP101 • TYPEYR

FTE_LINK

40. This field indicates whether a link has been made to improve our estimate of FTE. The link has only been attempted for non-standard academic years of programme of study started during the 1998-99 academic year and completed during 1999-2000 using the split FTE method of returning STULOAD.

Value	Description	Definition
1	1998-99 STULOAD used to calculate HESESFTE	<u>1999-2000 HESA data</u> FTE_TYPE=2 and DATELEFT<1 August 2000 and COMDATE>31 July 1998 and COMDATE<1 August 1999 and (TYPEYR=2,5 or (TYPEYR=1 and DATELEFT>anniversary of COMDATE+14 days)). <u>1998-99 HESA data</u> FTE_TYPE=2 and TYPEYR=2,3
0	Otherwise	All students not included above.

STULOA98

41. This field contains the value of STULOAD from the HESA 1998-99 return.

FTE_CASE

42. This field contains the case description as given in the table below.

43. For non-standard academic years or when two years of programme of study are generated, the method used to calculate HESESFTE is dependent on the following factors:

- a. Method used to return FTE.
- b. Length of the programme of study.
- c. Number of countable years of programme of study generated in HESES99.
- d. Whether the year of programme of study is the last or not.

44. The table below shows how we identify different cases of non-standard academic years of programme of study.

Value	Description	Definition
1	100:0	FTE_TYPE = 3
2	0:100 and one year generated in HESES99	FTE_TYPE = 4 and ATT_LINK = 0
	0:100 and two years generated in HESES99	
3a	First year	FTE_TYPE=4 and ATT_LINK=1 and STUBID=1
3b	Second year	FTE_TYPE=4 and ATT_LINK=1 and STUBID=2
4	Split FTE, one year generated in HESES99 and the programme of study will only generate one year	FTE_TYPE=2 and FTE_LINK=1 and ATT_LINK=0
5	Split FTE, one year generated in HESES99, final year of programme of study	FTE_TYPE=2 and FTE_LINK=0 and ATT_LINK=0 and TAIL=1*
6	Split FTE, one year generated in HESES99, on a programme of study generating two, or more, years which is not the final year	FTE_TYPE=2 and FTE_LINK=0 and ATT_LINK=0 and TAIL=0*
	Split FTE, two years generated in HESES99 and the programme of study will only generate two years	
7a	First year	FTE_TYPE=2 and FTE_LINK=1 and ATT_LINK=1 and STUBID=1
7b	Second year	FTE_TYPE=2 and FTE_LINK=1 and ATT_LINK=1 and STUBID=2
	Split FTE, two years generated in HESES99 and the programme of study generates more than two years	
8a	First year	FTE_TYPE=2 and FTE_LINK=0 and ATT_LINK=1 and STUBID=1
8b	Second year	FTE_TYPE=2 and FTE_LINK=0 and ATT_LINK=1 and STUBID=2

* Where TAIL=1 if DATELEFT < 1 August 2000 and FTE_LINK = 0 and COMDATE < 1 August 1999 and (TYPEYR=2,5 or ATT_LINK=1), and TAIL=0 otherwise.

HESESFTE

45. This field contains the FTE we assume for the year of programme of study in column 4a of HESES99. When the year of programme of study is contained in a standard academic year and one year of programme of study is generated, HESESFTE is taken to be STULOAD. The table below shows the method of calculating HESESFTE for different groups of non-standard academic years of programme of study.

FTE_CASE	Definition
1	STULOAD
2	STULOAD
3a	STULOAD - STULOAD x PROP
3b	STULOAD x PROP
4	STULOAD + STULOA98
5	STULOAD + AVRGLOAD

6	STULOAD
7a	(STULOAD + STULOA98) - STULOAD x PROP
7b	STULOAD x PROP
8a	(STULOAD + AVRGLOAD) - STULOAD x PROP
8b	STULOAD x PROP

46. Where $PROP = (\text{DATELEFT} - \text{anniversary of COMDATE}) / (\text{DATELEFT} - \text{start of academic year})$ and AVRGLOAD is the arithmetic mean of STULOAD for all non-standard academic years of programme of study in their first academic year, with the same MODE and QUALAIM at the same institution.

47. STULOAD is capped at 100 for all calculations except where two years of programme of study are generated in HESES99. HESESFTE is also capped at 100.

48. HESESFTE is 50 for all sandwich year-out years of programme of study (HESMODE=SWOUT).

HESEXCL

49. This field indicates whether the student should be included in the HESES re-creation and, if not, the reason for the exclusion. The following table gives details of students excluded from HESES99 and the binary exclusion code used (HESEXCL).

Value	Description	Definition
1	Not active in academic year	COMDATE > 31 July 2000 or DATELEFT < 1 August 1999
2	FE, NVQ or QTS students	QUALAIM = 44, 45, 51, 52, 53, 54, 55, 71 to 78, 81, 82
4	Students with no qualification aim	QUALAIM = 97, 98, 99.
8	Students explicitly excluded from the HESES99 student population	FUNDLEV = 99 or FUNDCOMP = 9.
16	Students taught wholly outside the UK	LOCSDY = 7 and FUNDCODE ≠ 1.
32	Dormant, sabbatical or students writing-up	MODE = 51, 61, 62, 63, 64 or (MODE = 41, 42, 43, 44 and (COMDATE + CRSELGTH*) < 1 August 1999 and MODE98 = 41, 42, 43, 44)
64	Incoming exchange students	SPCSTU = 3, 4, 5, 6, 8.
128	Students with a FTE of less than 3%	HESESFTE < 3
256	Students with split FTE in first academic year	FTE_TYPE = 2 and COMDATE > 31 July 1999 and COMDATE < 1 August 2000 and TYPEYR = 2, 3.
512	Students on standard academic years who withdrew before 2 December 1999 or students on non-standard academic years who withdrew before the anniversary of their commencement date	DATELEFT < 2 December 1999 and ((TYPEYR = 1 and FUNDCOMP = 2) or (DATELEFT < anniversary of commencement date and TYPEYR = 2, 3, 4, 5 and FUNDCOMP = 2.))

* CRSELGTH is a calculation of the length of course using SPLENGTH and UNITLGTH. Where UNITLGTH=9, CRSELGTH=6 was assumed.

50. For example, if HESEXCL = 74, then subtracting figures from the above table starting at the bottom, we see that the student is an incoming exchange (HESEXCL = 64), explicitly excluded (HESEXCL = 8) FE student (HESEXCL = 2).

HESCOL4

51. This field indicates whether the student is assigned to column 4 of HESES99.

Value	Description	Definition
1	Included in column 4 of HESES	HESCOMP=4 and HESEXCL=0
0	Otherwise	Any students not included above

FRANCH

52. This field indicates whether the student is included in the supplementary franchising tables.

Value	Description	Definition
1	Student is included in the franchised tables	HESCOL4 = 1 and TINST1 or TINST2 are not empty
0	Student is not included in the franchised tables	All other students

FRNINST1 and FRNINST2

53. The teaching institution identifier(s) of franchised students included in the supplementary franchising tables.

TINSTNME

54. The full name of the teaching institution(s) of franchised students included in the supplementary franchising tables.

XCOLAB01

55. XCOLAB01 is a field derived by HESA which evaluates the proportion of franchising to other teaching institutions. A description of the algorithm used to derive this field is given in Annex D.

FRANFTE

56. The FTE assumed for students included in the supplementary franchising tables. $FRANFTE = HESESFTE \times XCOLAB01$.

FRANEXT

57. A flag to indicate whether the student is wholly franchised.

Value	Description	Definition
1	Wholly franchised	$XCOLAB01 = 100.0$
2	Not wholly franchised	Otherwise

CAMPNME

58. The name of the campus.

CAMPPOST

59. The postcode of the campus.

SPC

60. This holds a 3 digit code that identifies which of the 160 geodemographic clusters the record postcode (unaltered from POSTCODE) has been assigned to. Aggregates of these clusters may be used to identify low participation areas for funding purposes. Note that this assignment may not be consistent with that used in the PIs since different sources for the student postcode may be employed.

Differences between HESES and HESA data

61. Following the revision of the 1998-99 HESA record, the HESES data can be more easily re-created. However, there remain some data returned in HESES that cannot be re-created exactly using data supplied to HESA because of differences in definition. In such cases, reasonable approximations have been made. Listed below are areas where there may be some uncertainty about the correspondence of HESA records to HESES cells. Where possible, we have indicated the likely effects of the uncertainties.

Student load

62. The calculation of HESESFTE described in paragraphs 42 to 48 ensure that FTE, over the whole programme of study, is consistent with HESES, while attempting to minimise variance within years of programme of study.

63. Exceptionally, where the split FTE method is used, assumptions are made for the final year of programme of study. The average FTE of students in the first academic year returned to HESA, for all similar programmes of study at the same institution, is added to STULOAD. A constant FTE for the course has been assumed. Therefore HESESFTE will be deflated where the intensity of the course is increasing over time, and inflated when it is decreasing.

Assumptions affecting selection criteria

64. We made assumptions for the following groups of students excluded from the HESES99 student population.

- a. Students who are writing up (see paragraph 35 and HESEXCL = 32).
- b. Students with very low FTEs (see paragraphs 40 to 48 and HESEXCL = 128).

Assumptions not affecting selection criteria

Two countable years of programme of study - first countable year

65. Where two years of programme of study are generated we have assumed some programme of study attributes from 1998-99 HESA data for the first countable year. Data returned to HESA should reflect the year of programme of study at the end of the academic year, therefore 1999-2000 HESA data relates to the second countable year when two years are generated. We believe this provides a better estimate of attributes where the second countable year generated is short, and is unlikely to have an effect where the second countable year is substantial.

Mode of study

66. LOCSDY has codes that are not mutually exclusive. We may underestimate the numbers of sandwich year out students where another applicable code in LOCSDY has been returned in preference. Guidance from HESA (HESA circular 00/02) is as follows:

- 1=1 and not (4 or 5 or 6 or 7 or 8)
- 2=2 and not (7)
- 3=3 and not (7)
- 4=4 and not (2 or 3 or 7)
- 5=5 and not (2 or 3 or 7)
- 6=6 and not (2 or 3 or 7 or 8)
- 7=7
- 8=8 and not (2 or 3 or 7)

New entrants

67. Figures shown in column 5 of the re-creation may not accurately reflect the numbers of new entrants returned on HESES. Where a year of programme is not a recognised concept of a course, i.e. YEARPRG = 99, year of programme of study has been calculated using COMDATE (see paragraph 23).

Areas of uncertainty in completing HESES99

Forecasts of countable years of programme of study and non-completions

68. HESES99 required institutions to provide forecasts of countable years of programme of study between 2 December 1999 and 31 July 2000. Forecasts are by their nature inexact. Unless there are exceptional circumstances, it is expected that these figures will, when considered as a whole, be an accurate reflection of actual non-completions and forecasts.

Additional data derived from HESA

69. All additional data were derived for HESES column 4 students only.

Count of mature students

70. Undergraduate, full-time, home and EC fundable students included in the HESES99 column 4 population were counted as mature students if they were aged 25 or older on entry using COMDATE and BIRTHDTE.

Counts of young students in neighbourhood types

71. We enclose a table showing the number of young, English domiciled, home and EC fundable, full-time undergraduates in the HESES99 column 4 population in each of 160 neighbourhood types. An unknown/unclassified category is included because it was not always possible to determine the neighbourhood type. For the purpose of this count, students are 'young' if they are under 25 on entry to the programme of study; this is calculated using COMDATE and BIRTHDTE. For a general description of the geodemographic method refer to paragraphs 1-6 of annex A2 'Performance Indicators in higher education in the UK' (HEFCE 00/40).

Count of young students receiving some tuition fee remittance

72. Those young, home and EC fundable, full-time undergraduates in the HESES99 column 4 population with some award or financial backing for tuition fees, MSTUFEE ≠ 01, were counted.

Disabled students

73. Those home and EC, full-time undergraduate new entrants in the HESES99 column 4 population who have a disability and were in receipt of the student disability allowance, DISALL=4, were counted as disabled new entrants. It is assumed that the proportion of disabled in this restricted group of students is the same as the proportion of disabled students within the institution overall. New entrants were used as the changes made in 1998-99 to DISALL were not applied retrospectively to existing students.

Annex B

Correcting erroneous HESA data

1. If the institution's response is such that it needs to correct erroneous HESA data then, along with all fields which require amendments, please include the following fields, in the order given below, to enable identification of individual records:

Field 1, RECID

Field 2, INSTID

Field 4, HUSID

Field 149/134, OWNSTU

Field 150/135, OWNPSD

Field 151/136, NUMHUS

2. A single file containing all fields affected by the amendment should be sent for each reason for response. For example, if some records require amendments to the DATELEFT field which necessitates changes to RSNLEAVE, these fields should be returned only for those records where DATELEFT and RSNLEAVE need amending.

3. Institutions should return the file containing a header in the following form:

line 1 – amendment reference number in the form amdxxxxn where xxxx is the HESA institution identifier and n is a sequential number starting at 1

line 2 – date of amendment in the form ddmmyyyy

line 3 – brief description of change (such as changes to DATELEFT and RSNLEAVE)

line 4 – names of the variable(s) to be changed, comma separated (for example, DATELEFT, RSNLEAVE).

4. We require data to be sent as a comma-separated file on 3.5" floppy disk or CD-ROM. The fields returned should be correctly formatted according to HESA rules, for example HUSID should be numeric and 13 digits long, with zeros in front to pad the field to the necessary length.

5. The following check digits should be supplied with the data:

- the total number of records in the file excluding headers
- the sum of all the HUSIDs in the file.

6. We will then produce amended records, which will be available from our web-site as part of the individualised record (see paragraph 31 of the main document for further details on how to access these data). Institutions will then be asked to check that we have correctly made any amendments and, if so, to 'sign off' the changes using the form at Annex C.

7. These specifications are intended to improve the timeliness and accuracy of amending data. If institutions require advice or support in any of the technical aspects of our requirements please do not hesitate to contact Anthony Ryan, tel 0117 931 7297 or e-mail hesa_heses_stats@hefce.ac.uk.

Annex C

Confirmation sheet

Institution: _____

HESA code: _____

Data amendments supplied following HEFCE 01/09

File Name	Name of amended field(s)	Number of records amended

Please photocopy, complete and return the form to Rhianne Cox, Analytical Services Group, HEFCE, Northavon House, Coldharbour Lane, BRISTOL, BS16 1QD

Telephone 0117 931 7014

Fax 0117 931 7476

I confirm that the amendments made to the data file(s), as summarised above, are correct.

Signed _____

Name (please print) _____

Position in organisation _____

Annex D

Algorithms for HESA derived fields

1. This annex contains details of the algorithms used by HESA to derive fields we used in this exercise.

XPRP101

RECID = 99011

2. Validation on this record ensures that the proportions add up to 100 per cent (± 2).

Do for $n=1$ to 16, while COSTCN(n) NOT NULL:

COSTCN	XPRP101
COSTCN(n)	SBJPER(n)

RECID = 99111

3. There is only one subject and cost centre collected in this record.

COSTCN	XPRP101
COSTCN1	100.0

RECID = 99012 / 99013, 99112 / 99013

4. For each PRIKEY calculate:

$\text{totfte} = \text{SUM fte}(\text{module})$

Do for module = 1 to 16

and for COSTCN = 1 to 2

MODULE	TOTFTE	COSTCN	SBJ	XPRP101
NOT NULL	>0	NOT NULL	✓	$(\text{SBJPER} * ((\text{fte} / \text{totfte}) * 100)) / 100$
NOT NULL	>0	NULL	NULL	NULL
NOT NULL	=0	NOT NULL	✓	NULL
NOT NULL	=0	NULL	NULL	NULL
NULL		NULL	NULL	NULL

RECID = 99012 / 99113, 99112 / 99113

5. For each PRIKEY calculate:

$\text{totfte} = \text{SUM fte}(\text{module})$

Do for module = 1 to 16

MODULE	TOTFTE	XPRP101
NOT NULL	>0	$(\text{fte} / \text{totfte}) * 100$
NOT NULL	=0	NULL

RECID = 99211, 99311, 99411, 99611, 99212, 99312, 99412, 99612

6. None of these records contain any cost centre/subject information. Student load for these records cannot therefore be allocated to a cost centre and a NULL value is produced.

XSBJA01

C = character; n = number; Digits = character at that position in SBJQA1

SBJQA1 (CHAR 6)	SBJQA2 (CHAR 4)	SBJQA3 (CHAR 4)	SBJBID (INT 1)	XSBJA01 (CHAR 3)
null, #####				ZZZ
Cn or Cnn	blank	blank	blank	SBJQA1
Cn or Cnn	Cn or Cnn	blank	1	SBJQA1
X1 or X10	Cn or Cnn	blank	1	SBJQA1
Cn or Cnn	X1 or X10	blank	1	SBJQA2
Cn or Cnn	Cn or Cnn	blank	2	SBJQA1
X1 or X10	Cn or Cnn	blank	2	SBJQA1
Cn or Cnn	X1 or X10	blank	2	SBJQA2
Cn or Cnn	Cn or Cnn	Cn or Cnn	blank	SBJQA1
X1 or X10	Cn or Cnn	Cn or Cnn	blank	SBJQA1
Cn or Cnn	X1 or X10	Cn or Cnn	blank	SBJQA2
Cn or Cnn	Cn or Cnn	X1 or X10	blank	SBJQA3
CCnn	blank	blank	blank	1 3
CCnnnn	blank	blank	blank	1 3 4
CnCn	blank	blank	blank	1 2
CnnCnn	blank	blank	blank	1 2 3
CnCnCn	blank	blank	blank	1 2
X1Cn	blank	blank	blank	X1
X1Cnn	blank	blank	blank	X1
X10Cnn	blank	blank	blank	X10
X1CCnn	blank	blank	blank	X1
X1CnCn	blank	blank	blank	X1

XCOLAB01

RECID=99011, 99111, 99211, 99311, 99411, 99611

PCOLAB	XCOLAB01
✓	PCOLAB

RECID=99012, 99112, 99212, 99312, 99412, 99612

7. For each PRIKEY calculate:

totfte = SUM fte(module)

Do for module = 1 to 16

MODULE	TOTFTE	PCOLAB	FTE	TINST	XCOLAB01
NOT NULL	>0	✓	✓	TINST1	PCOLAB* fte/totfte
NOT NULL	=0	✓	NULL	NULL	0
NULL		NULL	NULL	NULL	NULL