Appendix 2: Validation checks on HESES20 data

1. The Higher Education Students Early Statistics 2020-21 survey (HESES20) workbook contains a series of validation checks that help to ensure that incorrect data is not submitted. This appendix describes these validation checks. In addition, the workbook contains a series of credibility checks which will help providers check the credibility of their data prior to submission to Office for Students (OfS). The credibility checks are described in Appendix 3.

2. A number of validation checks are built into the Courses table and Tables 1 to 7 of HESES20. Validation cells above each table in the workbook return **'Validation: OK'** if the data has passed validation. If an error or inconsistency is detected in a table, **'Validation: Failure'** will appear in red above the column in which it is detected, and values associated with the failure will turn red. Errors and inconsistencies are described in more detail below or to the right of the table.

3. The 'Information' sheet in the HESES20 workbook contains more specific information about where validation checks have failed.

4. These errors **must be corrected** before the final submission of your HESES20 workbook. We will not accept workbooks containing validation failures. If an invalid workbook is uploaded to the portal, you will receive error messages in the comments section, and your upload will not be accepted. However, a results package will be generated, which you may find useful for checking purposes. Please see Appendix 1 for further information regarding the portal and submitting your workbook. For further guidance, please email heses@officeforstudents.org.uk.

Validation check for entire workbook

Data checked	Criterion
Date that the workbook has been uploaded to the extranet	Workbook has been uploaded before the census date. You will need to review and upload your HESES20 workbook on or after the census date in order to make a valid submission.

5. The census dates for HESES20 are as follows:

- For further education and sixth form colleges and academies: **1 November 2020**.

- For all other providers: **1 December 2020**.

Validation checks for Courses table

Data checked	Criterion
Learning aim reference	Must be valid. See the 'HESES20 Courses table information'[1] document for information as to which version of the Learning Aims Reference Service database has been used. Must be provided if there are values elsewhere in the row.

---

[1] Available at www.officeforstudents.org.uk/data-and-analysis/data-collection/heses/

Data checked	Criterion
Learning aim reference	Must not be entered in multiple rows with the same combination of level and length. This does not apply to temporary learning aims.
Price group proportions	Must be provided if price group is shown.
Price group proportions	Sum of proportions in row must equal 100 per cent.
Price group proportions	Must be a whole percentage.
Level	One of the following values must be entered if a valid code is entered in the learning aim reference field: 'UG'; 'PGT (UG fee)'; 'PGT (Masters' loan)'; 'PGT (Other)'; or 'PGR'.
Length	One of the following values must be entered if a valid code is entered in the learning aim reference field: 'Standard' or 'Long'.
Years countable	Must be a whole number.
Years countable	Must be ≥ 0.
All rows	Must not be blank if data entered in subsequent rows.
Years countable	Sum of years countable must be > 0 if a valid learning aim reference has been entered.

Validation checks for Table 1 Full-time

Data checked	Criterion
Totals automatically calculated for Columns 1, 2 and 3	Must be a whole number.
All values in Columns 1, 2 and 4	Must be ≥ 0.
All values in Column 3	Must be ≤ 0.
All values	Must be to no more than two decimal places.

Validation checks for Table 2 Sandwich year out

Data checked	Criterion
All values	Must be a whole number.
All values in Columns 1, 2 and 4	Must be ≥ 0.
All values in Column 3	Must be ≤ 0.

Validation checks for Table 3 Part-time

Data checked	Criterion
Totals automatically calculated for Columns 1, 2 and 3	Must be a whole number.
All values in Columns 1, 2, 4 and 4a	Must be ≥ 0.
All values in Column 3	Must be ≤ 0.
All values	Must be to no more than two decimal places.
All values in Column 4a	Must be ≤ corresponding value in Column 4.
All values in Column 4a	Must be ≥ 3% of corresponding value in Column 4.

Validation checks for Table 4 Year abroad

Data checked	Criterion
All values	Must be a whole number.
All values	Must be ≥ 0.
Total automatically calculated for…	Must be…
Column 1**(a)**(i)	≤ UG total (Standard + Long) in **Table 1**, Column 1(a)
Column 1**(a)**(ii)	≤ UG total (Standard + Long) in **Table 1**, Column 1(b)
Column 2**(a)**(i)	≤ UG total (Standard + Long) in **Table 1**, Column 2(a)
Column 2**(a)**(ii)	≤ UG total (Standard + Long) in **Table 1**, Column 2(b)
Column 1**(b)**(i)	≤ UG value in **Table 2**, Column 1(a)
Column 1**(b)**(ii)	≤ UG value in **Table 2**, Column 1(b)
Column 2**(b)**(i)	≤ UG value in **Table 2**, Column 2(a)
Column 2**(b)**(ii)	≤ UG value in **Table 2**, Column 2(b).

Validation checks for Table 5 Planning

Data checked	Criterion
All values	Must be a whole number.
All values	Must be ≥ 0.
For each of the following levels, the sum of the values in Section A, **Full-time and sandwich year out**, in each of Columns 2(a), 2(b) and 2(c): UG, PGT (UG fee), PGT (Masters' loan), PGT (Other), PGR	Must equal the sum of the corresponding values (Standard and Long) from **Tables 1 and 2**, in Column 2(a), 2(b) or 2(c) respectively.

Appendix 2: Validation checks on HESES20 data

Data checked	Criterion
For each of the following levels, the sum of the values in Section A, **Part-time**, in each of Columns 2(a), 2(b) and 2(c): UG, PGT (UG fee), PGT (Masters' loan), PGT (Other), PGR	Must equal the sum of the corresponding values (Standard and Long) from **Table 3**, in Column 2(a), 2(b) or 2(c) respectively.
All values in Section B	Must be ≤ corresponding total of Columns 1 and 2 in Section A.
The following checks do not apply to further education and sixth form colleges and academies because Column 1 in Tables 1, 2, 3 and 5 are pre-filled automatically from the Courses table data:	
For each of the following levels, the sum of the values in Section A, **Full-time and sandwich year out**, in each of Columns 1(a), 1(b) and 1(c): UG, PGT (UG fee), PGT (Masters' loan), PGT (Other), PGR	Must equal the sum of the corresponding values (Standard and Long) from **Tables 1 and 2**, in Column 1(a), 1(b) or 1(c) respectively.
For each of the following levels, the sum of the values in Section A, **Part-time**, in each of Columns 1(a), 1(b) and 1(c): UG, PGT (UG fee), PGT (Masters' loan), PGT (Other), PGR	Must equal the sum of the corresponding values (Standard and Long) from **Table 3**, in Column 1(a), 1(b) or 1(c) respectively.

Validation checks for Tables 6a, 6b, and 6c Health

Data checked	Criterion
All values in Columns 1-3 for: Nursing – adult; Nursing – children; Nursing – learning disability; Nursing – mental health	Must be a multiple of 0.5.
All values in Columns 1-3 except for: Nursing – adult; Nursing – children; Nursing – learning disability; Nursing – mental health	Must be a whole number.
All values except those in Column 3	Must be ≥ 0.

Appendix 2: Validation checks on HESES20 data

Data checked	Criterion
All values in Column 3	Must be ≤ 0.
For Tables 6a, 6b and 6c: Column 1(a) [OfS-fundable] Column 1(b) [non-fundable] Column 2(a) [OfS-fundable] Column 2(b) [non-fundable] Column 3(a) [OfS-fundable] Column 3(b) [non-fundable] Column 4(a) [OfS-fundable] Column 4(b) [non-fundable]	Must be … ≤ corresponding total in Column 1(a) of Table 1, 2 or 3 ≤ corresponding total in Column 1(b) of Table 1, 2 or 3 ≤ corresponding total in Column 2(a) of Table 1, 2 or 3 ≤ corresponding total in Column 2(b) of Table 1, 2 or 3 ≥ corresponding total in Column 3(a) of Table 1, 2 or 3 ≥ corresponding total in Column 3(b) of Table 1, 2 or 3 ≤ corresponding total in Column 4(a) of Table 1, 2 or 3 ≤ corresponding total in Column 4(b) of Table 1, 2 or 3.
For Table 6c only: Column 4a(a) [OfS-fundable] Column 4a(b) [non-fundable] For Table 6c only, all values in Column 4a For Table 6c only, all values in Column 4a For Table 6c only, all values in Column 4a	Must be … ≤ corresponding total in Column 4a(a) of Table 3 ≤ corresponding total in Column 4a(b) of Table 3 Must be to no more than two decimal places Must be ≤ corresponding value in Column 4 Must be ≥ 3% of corresponding value in Column 4.

Validation checks for Table 7 Location

Data checked	Criterion
All rows	Must not be blank if data entered in subsequent rows.
Location type	One of the following values must be entered: 'Other' or 'Subcontracted out'.
UKPRN	Must be a valid UK Provider Reference Number (UKPRN) or generic code, must not be left blank.
UKPRN	Must be the provider's own UKPRN when the location type has been set as 'Other'.
UKPRN	Must not be the provider's own UKPRN when the location type has been set as 'Subcontracted out'.
Name of location	Must not be blank.
Campus postcode	Must be valid and must not be left blank.
Section 1 and 2	Must be to no more than two decimal places.
Section 1 and 2	Must not be left blank and must be ≥ 0.
Estimated full-time equivalence (FTE) for:	Must be… < Estimated completed years for price group A in section 1

Data checked	Criterion
Price group A	< Estimated completed years for price group B in section 1
Price group B	< Estimated completed years for price group C1 in section 1
Price group C1	< Estimated completed years for price group C2 in section 1
Price group C2	< Estimated completed years for price group D in section 1
Price group D	
For each of the following price groups, the sum of the values in Section 1, **estimated completed years:** Price group A, Price group B, Price group C1, Price group C2, Price group D	Must equal the sum of the corresponding values (Standard and Long) from **Tables 1, 2 and 3**, in Column 4(a).
For each of the following price groups, the sum of the values in Section 2, **estimated FTE:** Price group A, Price group B, Price group C1, Price group C2, Price group D	Must equal the sum of the corresponding values (Standard and Long) from Column 4 in **Tables 1 and 2** and Column 4a(a) **in Table 3**.

Further information

6. Any questions about your data or validation checks should be emailed to heses@officeforstudents.org.uk.