

**Office for  
Students**



# **2021-22 Student Alternative data checking tool**

**2021-22 Pearson comparison technical  
document and rebuild instructions**

**Enquiries to** [hesasar@officeforstudents.org.uk](mailto:hesasar@officeforstudents.org.uk)

**Publication date** 4 August 2022

# Contents

|  |          |
|--|----------|
| <b>Summary</b>   | <b>2</b> |
| Changes from the 2020-21 HESA Student Alternative data checking tool                   | 3        |
| 2021-22 Pearson comparison workbook  | 3        |
| 2021-22 Pearson comparison individualised file   | 3        |
| <b>Student Alternative Record fields used to create the 2021-22 Pearson comparison</b> | <b>4</b> |
| <b>Description of derived fields</b>   | <b>5</b> |
| <b>Rebuild instructions</b>  | <b>8</b> |

# Summary

1. This document describes the Pearson comparison files generated by the 2021-22 Student Alternative Record (SAR) data checking tool, the algorithms used to generate the data and rebuild instructions. Throughout the document, fields taken or derived from the Student Alternative Record are shown in capitals.
2. The Pearson comparison can be used by higher education providers to verify and correct their 2021-22 individualised student data before signing off the final data with the Designated Data Body (DDB).
3. The purpose of this comparison is to ensure that the 2021-22 Student Alternative Record data is consistent with the Pearson qualification records. For each provider we check:
  - whether all students who received a qualification from Pearson during the 2021-22 academic year are included in the 2021-22 Student Alternative Record as having a qualification.
  - whether all students who are recorded as having achieved a Pearson qualification in the 2021-22 Student Alternative Record appear in the Pearson data.
  - whether the type of qualification recorded in the 2021-22 Student Alternative Record matches the type of qualification in the Pearson data.
4. The 2021-22 Pearson comparison is formed of two files. These are:
  - Pearson comparison workbook 'PEARCOM21\_AAA\_XXXXXXXX.xlsx',
  - Pearson comparison individualised file 'PEARCOM21\_AAA\_XXXXXXXX\_IND.csv'

where XXXXXXXX is your provider's eight-digit UK Provider Reference Number (UKPRN), and AAA identifies the output type as listed in the table below.

| File abbreviation | Source of outputs          |
|-------------------|----------------------------|
| DCT               | Data checking tool         |
| AST               | Amendments submission tool |

Details of how to use individualised files are given on the Office for Students (OfS) website at: [www.officeforstudents.org.uk/data-and-analysis/supplying-data/working-with-individualised-files/](https://www.officeforstudents.org.uk/data-and-analysis/supplying-data/working-with-individualised-files/).

5. These files can be accessed via the Information Reporting Interface Service (IRIS) report which is available on the DDB's data collection site at <https://datacollection.hesa.ac.uk/>. Access details will be sent to the appropriate contacts at providers when the data checking tool is available.
6. This document is aimed at readers with in-depth knowledge of the data. Readers are advised to refer to the Student Alternative Record coding manual 2021-22 (<https://www.hesa.ac.uk/collection/c21054>).

- We will document any changes or corrections to the algorithms in this document on the OfS website: [www.officeforstudents.org.uk/data-and-analysis/data-checking-tool/2021-22-student-alternative-data-checking-tool/](http://www.officeforstudents.org.uk/data-and-analysis/data-checking-tool/2021-22-student-alternative-data-checking-tool/) under 'Updates'.

## Changes from the 2020-21 HESA Student Alternative data checking tool

- There are no changes to the algorithms since the equivalent was published for the 2020-21 algorithms in '2020-21 HESA Student Alternative - PEARCOM20 Pearson comparison technical document and rebuild instructions' available from the data checking tool pages at [www.officeforstudents.org.uk/data-and-analysis/data-checking-tool/documentation-archive/](http://www.officeforstudents.org.uk/data-and-analysis/data-checking-tool/documentation-archive/).
- PEARCOMHESAPOP has been renamed to PEARCOMDDBPOP.

## 2021-22 Pearson comparison workbook

- The 2021-22 Pearson comparison workbook is the Excel workbook PEARCOM21\_AAA\_XXXXXXXX.xlsx which contains the following worksheets:

**Table 1: Worksheets in the 2021-22 Pearson comparison workbook**

| Worksheet* | Title                              |
|------------|------------------------------------|
| Coversheet | Pearson comparison summary figures |

\* This worksheet reference corresponds to the spreadsheet tabs.

## 2021-22 Pearson comparison individualised file

- When working through this document it is necessary to use the individualised file, PEARCOM21\_AAA\_XXXXXXXX\_IND.csv. The individualised file contains the values of all fields used to generate the Pearson comparison.
- Fields taken from the Student Alternative Record or derived as part of the 2021-22 Pearson comparison are shown in capitals using the names given in tables 2 and 3 respectively. Fields prefixed with 'PEARCOM' are derived, all others are taken directly from the Student Alternative Record or Pearson data as indicated in Table 2.
- Specific instructions for rebuilding the 2021-22 Pearson comparison workbook from the individualised file are provided in this document.
- The individualised file contains one record per instance (each unique combination of UKPRN, HUSID and NUMHUS) in the latest academic year. Additionally, the individualised file will contain one record per student as recorded in the Pearson data, where that record has not been matched to an instance in the latest academic year.

# Student Alternative Record fields used to create the 2021-22 Pearson comparison

15. The Student Alternative fields used to generate the data summary are listed in Table 2.

**Table 2: List of Student Alternative and Pearson fields used in the 2021-22 Pearson comparison**

| Name                 | Description   | Data source         | Column in individualised file |
|----------------------|---|---------------------|-------------------------------|
| UKPRN                | UK Provider Reference Number  | Student Alternative | A                             |
| HUSID                | Unique student identifier   | Student Alternative | B                             |
| OWNSTU <sup>†</sup>  | Provider's own student identifier                                     | Student Alternative | D                             |
| NUMHUS               | Student instance identifier   | Student Alternative | C                             |
| OWNINST <sup>†</sup> | Provider's own instance identifier                                    | Student Alternative | E                             |
| QUAL                 | Qualification awarded   | Student Alternative | N/A                           |
| PERIODSTART          | Instance period start date  | Student Alternative | N/A                           |
| CTITLE               | Course title  | Student Alternative | G                             |
| COURSEAIM            | Course aim  | Student Alternative | F                             |
| PEAR2122REGNO        | Learner's registration number (unique identifier in the Pearson data) | Pearson             | AA-AB                         |
| PEAR2122CTITLE       | Course title from Pearson data  | Pearson             | W-X                           |
| PEAR2122CMPLTDATE    | Learner's date of completion from the Pearson data                    | Pearson             | U-V                           |
| PEAR2122QUAL         | Learner's qualification aim from the Pearson data                     | Pearson             | AC-AD                         |
| PEAR2122STATUS       | Learner's outcome status  | Pearson             | Y-Z                           |
| PEAR2122GRADE        | Learner's outcome grade where a qualification has been awarded        | Pearson             | AE-AF                         |

<sup>†</sup> These fields are not used in the comparison but are included in the individualised file to allow for easy identification of students.

## Description of derived fields

16. The derived fields used to generate the 2021-22 Pearson comparison are listed in Table 3.

**Table 3: List of derived fields used in the 2021-22 Pearson comparison**

| Name                               | Description  | Paragraph | Column in individualised file |
|------------------------------------|--|-----------|-------------------------------|
| PEARCOMPERIODSTART1-3 <sup>†</sup> | Instance period start date in the Student Alternative Record   | 20        | K-M                           |
| PEARCOMQUAL1-3_1-2                 | Identifies the qualification awarded to the student in the Student Alternative Record                    | 21        | N-S                           |
| PEARCOMDDBPOP                      | Identifies students in the Student Alternative Record for the provider                                   | 22        | H                             |
| PEARCOMPEARQUAL1-2                 | Identifies the qualification awarded to the student in the Pearson data                                  | 23        | AG-AH                         |
| PEARCOMPEARPOP                     | Identifies students in the Pearson data for the provider   | 24        | I                             |
| PEARCOMQUALMATCH                   | Identifies students with matching qualification types in the Student Alternative Record and Pearson data | 25        | J                             |

<sup>†</sup> Due to file size restrictions the number of occurrences of this field is restricted to three.

### Instance period data

17. 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 three. In the unusual case where an instance has more than three instance periods, all will be considered in our algorithms, although only three will be displayed in the individualised file. In selecting instance period data to show, we will select those with a qualification ordered by:

- i. Those with an HND
- ii. Those with an HNC
- iii. Those whose period start date (PERIODSTART) is the earliest.

### Linking

18. The Student Alternative and Pearson records are linked using UKPRN, First Name(s), Last Name(s) and Date of Birth. A link is defined as an identical match of UKPRN, First Name(s), Last Name(s) and Date of Birth. Allowances are made for missing or incorrectly ordered names.

19. In the case that no link can be made between the 2021-22 Student Alternative Record and the Pearson data, we will attempt to link the Pearson data to the 2020-21 Student Alternative Record. If a link can be made, these students will be excluded from the output.

### PEARCOMPERIODSTART1-3

20. These fields contain the first three instance period start dates (PERIODSTARTs) for the student. See paragraph 17 for details on how these PERIODSTARTs are selected.

### PEARCOMQUAL1-3\_1-2

21. These fields contain the qualification awarded (QUALs) for the student as present in the Student Alternative Record. A maximum of two QUALs can be recorded per instance period, each qualification will be recorded in a separate column in the individualised file. For example, PEARCOMQUAL1\_1 will contain the first qualification awarded associated with PEARCOMPERIODSTART1.

### PEARCOMDDBPOP

22. This field identifies students who are in the Student Alternative Record for the provider and whether they have been awarded a Pearson qualification.

| Value | Description  | Definition   |
|-------|--|--|
| 2     | Student is in the Student Alternative Record for the provider and has been awarded a Pearson qualification     | A record was found in the Student Alternative Record for the provider and one of PEARCOMQUAL1_1-3_2 = J30, C30 |
| 1     | Student is in the Student Alternative Record for the provider but has not been awarded a Pearson qualification | A record was found in the Student Alternative Record for the provider and not above                            |
| 0     | Otherwise  | Otherwise  |

### PEARCOMPEARQUAL1-2

23. These fields identify students who have received a valid qualification in the Pearson data.

| Value | Description  | Definition   |
|-------|--|--|
| J30   | Student has received a HND qualification in the Pearson data | A record was found in the Pearson data where<br>PEAR2122QUAL = J30 and<br>PEAR2122STATUS = Complete and<br>PEAR2122GRADE = D, M, P |
| C30   | Student has received a HNC qualification in the Pearson data | A record was found in the Pearson data where<br>PEAR2122QUAL = C30 and<br>PEAR2122STATUS = Complete<br>PEAR2122GRADE = D, M, P     |
| 0     | Otherwise  | Otherwise  |

### PEARCOMPEARPOP

24. This field identifies students who are in the Pearson data for the provider and whether they have a qualification recorded.

| Value | Description  | Definition   |
|-------|--|--|
| 2     | Student is in the Pearson data for the provider, and they have a qualification recorded        | A record was found in the Pearson data for the provider and one of PEARCOMPEARQUAL1-2 = J30, C30 |
| 1     | Student is in the Pearson data for the provider, and they do not have a qualification recorded | A record was found in the Pearson data for the provider and not above                            |
| 0     | Otherwise  | Otherwise  |

### PEARCOMQUALMATCH

25. This field identifies students who have received the same qualification in the Pearson data and the Student Alternative Record.

| Value | Description  | Definition                        |
|-------|--|-----------------------------------|
| 1     | Student's recorded qualification type in the Student Alternative Record is the same as in the Pearson data | PEARCOMQUAL1_1 = PEARCOMPEARQUAL1 |
| 0     | Otherwise  | Otherwise                         |



## Rebuild instructions

25. The individualised file, PEARCOM21\_AAA\_XXXXXXXXX\_IND.csv, can be used to interrogate the data and our calculation of derived fields by filtering to certain groups of records. Full details of how to access and use individualised files are given on the OfS website ([www.officeforstudents.org.uk/data-and-analysis/supplying-data/working-with-individualised-files/](http://www.officeforstudents.org.uk/data-and-analysis/supplying-data/working-with-individualised-files/)).
26. Specific instructions to rebuild the 2021-22 Pearson comparison summary figures shown in the Pearson comparison workbook, PEARCOM21\_AAA\_XXXXXXXXX.xlsx, are provided here.

| Item in the workbook  | Derived field selection in individualised file                    |
|---|---|
| Students recorded in both the Student Alternative Record and Pearson data as receiving the same qualification                       | PEARCOMDDBPOP = 2 and PEARCOMPEARPOP = 2 and PEARCOMQUALMATCH = 1 |
| Students recorded with a qualification in the Pearson data but not appearing in the Student Alternative Record                      | PEARCOMDDBPOP = 0 and PEARCOMPEARPOP = 2                          |
| Students recorded with a qualification in the Student Alternative Record but that qualification does not appear in the Pearson data | PEARCOMDDBPOP = 2 and PEARCOMPEARPOP = 1                          |
| Students recorded with a qualification in the Pearson data but that qualification does not appear in the Student Alternative Record | PEARCOMDDBPOP = 1 and PEARCOMPEARPOP = 2                          |
| Students recorded in both Student Alternative Record and Pearson data as receiving a different qualification                        | PEARCOMDDBPOP = 2 and PEARCOMPEARPOP = 2 PEARCOMQUALMATCH = 0     |



© The Office for Students copyright 2022

This publication is available under the Open Government Licence 3.0 except where it indicates that the copyright for images or text is owned elsewhere.

[www.nationalarchives.gov.uk/doc/open-government-licence/version/3/](http://www.nationalarchives.gov.uk/doc/open-government-licence/version/3/)