

Office for
Students



2023-24 Student data checking tool

2023-24 Pearson comparison technical
document and rebuild instructions

Enquiries to DFAGateways@officeforstudents.org.uk

Publication date 19 September 2024

Contents

Summary	3
Changes from the 2022-23 Student data checking tool	5
2023-24 Pearson comparison data files	6
Workbook	6
Individualised file	6
Fields used to create the 2023-24 Pearson comparison	7
Description of derived fields	8
Rebuild instructions	11

Summary

1. This document describes the Pearson comparison files generated by the Student (23056) record, the algorithms used to generate the data and provides rebuild instructions. Throughout the document, fields taken or derived from the Student (23056) record or the Pearson data are shown in capitals.
2. The Pearson comparison files can be used by higher education providers to verify and correct their 2023-24 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 Student (23056) 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 2023-24 academic year are included in the Student (23056) record as having a qualification.
 - whether all students who are recorded as having achieved a Pearson qualification in the Student (23056) record also have a qualification in the Pearson data.
 - whether the type of qualification recorded in the Student (23056) record matches the type of qualification in the Pearson data.
4. The Student (23056) and Pearson records are linked using UK Provider Reference Number (UKPRN), First Name(s), Last Names(s), Date of Birth, and if possible, qualification aim and start date. A link is defined as a match of UKPRN, First Name(s), Last Name(s), Date of Birth, and if possible, qualification aim and start date. Allowances are made for missing or incorrectly ordered names and start dates not matching by less than 40 days.
5. The 2023-24 Pearson comparison is formed of two files. These are:
 - Pearson comparison workbook 'PEARCOM23_DCT_XXXXXXXX.xlsx'
 - Pearson comparison individualised file 'PEARCOM23_DCT_XXXXXXXX_IND.csv'

where XXXXXXXX is your provider's eight-digit UKPRN.

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/.

6. 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.

7. This document is aimed at readers with in-depth knowledge of the data. Readers are advised to refer to a copy of the Student (23056) record coding manual when using this document.¹
8. We will document in the data checking tool section of the OfS website any changes or corrections to the algorithms in this document.²

¹ See [Student record – 23056 coding manual](#).

² See [The OfS website – Data checking tool](#).

Changes from the 2022-23 Student data checking tool

9. This is the first publication of these algorithms relating to 2023-24 data. Any changes after initial publication will be listed here.
10. There have been no changes to the Pearson comparison algorithms since the 2022-23 Student data checking tool.

2023-24 Pearson comparison data files

Workbook

11. The 2023-24 Pearson comparison workbook is the Excel workbook PEARCOM23_DCT_XXXXXXXX.xlsx which contains the following worksheets:

Table 1: Worksheets in the 2023-24 Pearson comparison workbook

Worksheet*	Title
Coversheet	Pearson comparison summary figures

* This worksheet reference corresponds to the spreadsheet tabs.

Individualised file

12. When working through this document it is necessary to use the individualised file, PEARCOM23_DCT_XXXXXXXX_IND.csv. The individualised file contains the values of all fields used to generate the Pearson comparison.

13. Fields taken from the Student (23056) record, taken from the Pearson data or derived as part of the 2023-24 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 (23056) record or Pearson data as indicated in Table 2.

14. Specific instructions for rebuilding the 2023-24 Pearson comparison workbook from the individualised file are provided in this document.

15. The individualised file contains one row per engagement (each unique combination of UKPRN, SID 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 engagement in the latest academic year.

16. The number of qualifications awarded that can be returned for a given student engagement is unlimited; however, due to file size considerations we restrict the number of separate qualifications awarded per engagement displayed in the individualised file to three. In the unusual case where an engagement has more than three qualifications awarded, all will be considered in our algorithms, although only three will be displayed in the individualised file. In selecting qualifications awarded data to show, we will select those ordered by:

- a. Those with a HND.
- b. Those with a HNC.
- c. Those with the highest level of study (excluding HNDs and HNCs).

Fields used to create the 2023-24 Pearson comparison

17. The Student (23056) and Pearson fields used to generate the data summary are listed in Table 2.

Table 2: List of Student (23056) and Pearson fields used in the 2023-24 Pearson comparison

Source	Entity	Name	Description
Student	Engagement	UKPRN	UK Provider Reference Number
Student	Engagement	SID	Student identifier
Student	Engagement	OWNSTU [†]	Provider's own student identifier
Student	Engagement	NUMHUS	Engagement number
Student	Engagement	OWNENGID [†]	Provider's own engagement identifier
Student	Engagement	ENGSTARTDATE	Engagement start date
Student	Qualification	QUALCAT	Records the qualification category
Student	Engagement derived field	Z_LEVEL	Derived field records student's course level for the engagement
Student	Course*	COURSEID [†]	Course identifier
Student	Course*	COURSETITLE [†]	Course title
Pearson		PEAR2324CENTREREF [†]	Student identifier from the Pearson data
Pearson		PEAR2324REGNO	Learner's registration number (unique identifier in the Pearson data)
Pearson		PEAR2324COMMENCEMENTDATE	Learner's enrolment date
Pearson		PEAR2324CTITLE	Course title from Pearson data
Pearson		PEAR2324QUAL	Learner's qualification aim from the Pearson data
Pearson		PEAR2324STATUS	Learner's outcome status
Pearson		PEAR2324GRADE	Learner's outcome grade where a qualification has been awarded

[†] These fields are not used in the comparison but are included in the individualised file to allow for easy identification of students.

* The values of these fields are taken from the latest student course session associated with the student's engagement.

Description of derived fields

18. The derived fields used to generate the 2023-24 Pearson comparison are listed in Table 3.

Table 3: List of derived fields used in the 2023-24 Pearson comparison

Name	Description
PEARCOMQUAL1-3	Identifies the qualifications awarded to the student in the Student (23056) record
PEARCOMDDBPOP	Identifies students in the Student (23056) record for the provider
PEARCOMPEARQUAL1-2	Identifies the qualification awarded to the student in the Pearson data
PEARCOMPEARPOP	Identifies students in the Pearson data for the provider
PEARCOMQUALMATCH	Identifies students with matching qualification types in the Student (23056) record and Pearson data

PEARCOMQUAL1-3

19. These fields contain the qualifications awarded (QUALCATs) for the student as present in the Student (23056) record. See paragraph 6 for details on how these QUALCATs are selected.

PEARCOMDDBPOP

20. This field identifies students who are in the Student (23056) record for the provider and whether they have been awarded a Pearson qualification.

Table 4: PEARCOMDDBPOP – Value, description and definition

Value	Description	Definition
2	Student is in the Student (23056) record for the provider and has been awarded a Pearson qualification.	A record was found in the Student (23056) record for the provider and one of PEARCOMQUAL1-3 = J0003, C0001
1	Student is in the Student (23056) record for the provider but has not been awarded a Pearson qualification.	A record was found in the Student (23056) record for the provider and not above
0	Student is not in the Student (23056) record for the provider.	Otherwise

PEARCOMPEARQUAL1-2

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

Table 5: PEARCOMPEARQUAL1-2 – Value, description and definition

Value	Description	Definition
J0003	Student has received a HND qualification in the Pearson data	A record was found in the Pearson data where PEAR2324QUAL = J0003 and PEAR2324STATUS = Complete and PEAR2324GRADE = D, M, P
C0001	Student has received a HNC qualification in the Pearson data	A record was found in the Pearson data where PEAR2324QUAL = C0001 and PEAR2324STATUS = Complete PEAR2324GRADE = D, M, P
0	Student has not received a HND or HNC qualification in the Pearson data, or is not included in the Pearson data	Otherwise

PEARCOMPEARPOP

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

Table 6: PEARCOMPEARPOP – Value, description and definition

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= J0003, C0001
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	Student is not in the Pearson data for the provider	Otherwise

PEARCOMQUALMATCH

23. This field identifies students who have received the same qualification in the Pearson data and the Student (23056) record.

Table 7: PEARCOMQUALMATCH – Value, description and definition

Value	Description	Definition
1	Student's recorded qualification type in the Student (23056) record is the same as in the Pearson data	One of PEARCOMQUAL1-3 = one of PEARCOMPEARQUAL1-2

Value	Description	Definition
0	Student's recorded qualification type in the Student (23056) record if not the same as in the Pearson data	Otherwise

Rebuild instructions

24. The individualised file, PEARCOM23_DCT_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/.

25. Specific instructions to rebuild the 2023-24 Pearson comparison summary figures shown in the Pearson comparison workbook, PEARCOM23_DCT_XXXXXXXXX.xlsx, are provided here.

Table 8: Fields used in rebuilding the Pearson comparison summary figures

Item in the workbook	Derived field selection in individualised file
Student engagement recorded in both the Student (23056) record and Pearson data as receiving the same qualification.	PEARCOMDDBPOP = 2 and PEARCOMPEARPOP = 2 and PEARCOMQUALMATCH = 1
Student engagement recorded with a qualification in the Pearson data but not appearing in the Student (23056) record.	PEARCOMDDBPOP = 0 and PEARCOMPEARPOP = 2
Student engagement recorded with a qualification in the Student (23056) record but that qualification does not appear in the Pearson data.	PEARCOMDDBPOP = 2 and PEARCOMPEARPOP = 1
Student engagement recorded with a qualification in the Pearson data but that qualification does not appear in the Student (23056) record.	PEARCOMDDBPOP = 1 and PEARCOMPEARPOP = 2
Student engagement recorded in both Student (23056) record and Pearson data as receiving a different qualification.	PEARCOMDDBPOP = 2 and PEARCOMPEARPOP = 2 and PEARCOMQUALMATCH = 0



© The Office for Students copyright 2024

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/