

How the OfS uses your ILR data

Danny Shone
Christine Daniel
Myles Garrould

ilrdata@officeforstudents.org.uk

Session aims

In this session, we will cover:

- Outputs produced by the OfS using your ILR data
- ILR fields used in creating the outputs
- Other uses of ILR fields in OfS
- The OfS ILR data checking tool
- Other data sources that feed into our outputs.





What the OfS uses your ILR data for

Delivery of ILR data



- The Education & Skills Funding Agency (ESFA) deliver your ILR R14 data to the OfS in around November each year
- We received your final 2016-17 ILR R14 data in November 2017
- We start using your ILR data immediately once we have access to it

We also use your ILR data from the data checking tool before we have access to the final ILR R14 data for some purposes

Uses of ILR data

The OfS uses your ILR data in a number of ways. Below are some of our uses.

- Student premium funding allocations
- Teaching Excellence and Student Outcomes Framework (TEF)
- Unistats
- National Student Survey (NSS) target list
- Recreation of the HEIFES return in the same academic year
- Reconciliation of teaching funding
- Analysis
- Comparison data in HEIFES

Other data sources that we use with the ILR

To aid our use of your ILR data, we also use other data sources.

- **Your ILR data from previous years** – so we can link data across years and obtain Level 3 qualification for student premium
- **Other providers' ILR data** – for sector information, for example
- **National Pupil Database (NPD)** – student premium calculations
- **Learning Aims Reference Service (LARS) database** – course information



Fields from the ILR that we use

ILR fields that we use

We use a large amount of fields from the ILR record in our processes. A list of some of these is included in your packs.

OfS derived fields

- Generally, for the outputs that we produce, we calculate our own fields from the ILR data using algorithms
- These are known as derived fields and are used to in the calculations of specific OfS concepts, such as fundability status, for example
- We publish details of the algorithms used to create the OfS derived fields as part of the release of our outputs

HEFLEVEL (Column R)

.22 This field allocates students to level of study

Value	Description	Definition
UG	Undergraduate	FUNDLEV = 10, 11
PGT_UGF	Postgraduate taught and eligible under the undergraduate student support regime	FUNDLEV = 20, 21 and SPECFEE ≠ 9, BLANK and HEFQAIM = CERTED, DTLLS, DET, PGCE
PGT_ML	Postgraduate taught courses which are eligible for student finance support	FUNDLEV = 20, 21, 30, 31 and HEFQAIM = MASTER, HIGHER and ((HEFMODE = FTS, SWOUT and HEFCRSELGTH = 1, 2) or (HEFMODE = PT and HEFCRSELGTH = 1, 2, 3, 4))
PGT_OTH	Postgraduate taught courses that do not fall under the previous two categories	FUNDLEV = 20, 21, 30, 31 and not above

ILR fields used in OfS data outputs

The list of fields

The A3 table lists the ILR fields referenced in the algorithms used in our 2016-17 data outputs.

Some notes:

- This doesn't list the LARS fields that we also use. We link to Learning Aim information to determine certain aspects such as the type of course a learner is on.
- There are a number of other ILR fields which we don't specifically reference in our data output algorithms, but please remember that:
 - They may be used for statistical purposes
 - As our funding and assessment methods develop, they may be used more in the future.

ILR Field exercise

ILR fields used in OFS data checking tool outputs. (Based on 2016-17 ILR outputs).										
This document lists the ILR fields referenced in the algorithms used in the OfS's 2016-17 data outputs. These are listed by Entity, then a list of the fields that describe that entity.										
Looking through the fields below, we would like you to consider how and when you collect this information at your college, and to discuss your processes with others on your table. Some prompt questions are:										
- Is the information recorded when the student registers?								- Who collects the information in the college and how is it then recorded on your systems?		
- Is this updated each year?								- Is any of this information entered / updated by students themselves?		
- Is this information that needs to be updated within the year? (e.g. FUNDCOMP).								- Are there any fields that you do not collect, or use default values for?		
- How often is it updated?										
Is the field used by the OFS in our data outputs?										
Fields	XML Element Name	Any output	HEFES17 comparison	Student Premium	DHLE target list	NSS target list	2016 Unistats	TEF Yr2 Metrics	OFS notes	Your notes
Sex	Sex							✓	For TEF, this is used in benchmarking and high levels of unknowns leads to suppressed metrics. Will be used when calculating the disabled students' premium allocation. Also for TEF, this is used in benchmarking and high levels of unknowns leads to suppressed metrics.	
LLDD and health problem	LLDDHealthProb			✓				✓		
Learner Contact Entity	LearnerContact									
Postcode	PostCode			✓				✓	For TEF, the prior postcode is used in benchmarking and high levels of unknowns leads to suppressed metrics.	
Learner Contact Preference Entity	ContactPreference									
Contact preference code	ContPrefCode				✓					
LLDD and Health Problem Entity										
LLDD and health problem category	LLDDCat			✓				✓	Will be used when calculating the disabled students' premium allocation.	
Primary LLDD and health problem	PrimaryLLDD			✓				✓		
Learner Funding and Monitoring	LearnerFAM									
Learner funding and monitoring code	LearnerFAMCode			✓					LEARNFAM_DLA will be used when calculating the disabled students' premium allocation.	
Learner Provider Specified Monitoring Entity	ProviderSpecLearnerMonitoring									
Provider specified learner monitoring occurrence	ProvSpecLearMonOccur		✓		✓	✓	✓	✓		
Provider specified learner monitoring	ProvSpecLearMon		✓		✓	✓	✓	✓		
Learner Delivery Entity	LearnerDelivery									
Learner delivery reference	LearnerDelRef		✓	✓	✓	✓	✓	✓		

Running footer

Discussion

Looking through the fields below, we would like you to consider **how** and **when** you collect this information at your college, and to discuss these processes with others on your table.

For certain fields we've provided some extra notes; e.g. for fields that have caused issues with TEF metrics or which can cause students to be excluded from our algorithms.

Prompt questions:

- Is the information recorded when the student registers?
- Is this updated each year?
- Is this information that needs to be updated within the year? (e.g. FUNDCOMP).
- How often is it updated? How long does it take for updated information to get on to the system?
- Who collects / updates the information in the college and how is it then recorded on your systems?
- Is any of this information entered / updated by students themselves?
- Are there any fields that you do not collect, or use default values for?

Feedback



What themes and questions have come up in this exercise?



OfS outputs created using ILR data

OfS outputs (1)

We create outputs using your ILR data at various points in the year to show you how we have used the data. In the past, we have created the following outputs.

- HEIFES comparison
- Student premium funding data summary
- NSS target list
- Unistats preview
- Student characteristics data summary

HEIFES comparison (1)

The HEIFES comparison output compares the data from your HEIFES return with the data from your ILR return for the same academic year.

The aim of this output is to ensure that the HEIFES and ILR data match as closely as possible.

It can be used to identify discrepancies or errors in your HEIFES or ILR data, which then may need to be amended.

For the upcoming HEIFES17 comparison, we will be recreating the whole of the HEIFES workbook for the first time, rather than just those tables which are related to funding. This will help you identify any discrepancies across the entire return.

HEIFES comparison (2)

The HEIFES comparison workbook contains the following information:

- Your original HEIFES data from your HEIFES return
- Recreated HEIFES data using your ILR data
- Differences between HEIFES and the recreated HEIFES data
- Recalculations of funding using ILR data, comparing against original funding

Key differences between HEIFES17 and HEIFES17 recreated using 2017-18 ILR data

2018-19 Funding

	Allocations based on HEIFES17 return	Allocations based on HEIFES17 recreated using 2017-18 ILR data	Difference	Percentage difference
High-cost subject funding	£0	£0	£0	0.0%
Premium to support successful student outcomes: full-time	£0	£0	£0	0.0%
Premium to support successful student outcomes: part-time	£0	£0	£0	0.0%
Disabled students' premium	£0	£0	£0	0.0%
Postgraduate taught supplement	£0	£0	£0	0.0%
Accelerated full-time undergraduate provision	£0	£0	£0	0.0%
Intensive postgraduate provision	£0	£0	£0	0.0%
Students attending courses in London	£0	£0	£0	0.0%
Erasmus+ and overseas study programmes	£0	£0	£0	0.0%
Nursing and allied health supplement	£0	£0	£0	0.0%
Total 2018-19 funding (excluding fixed allocations)*	£0	£0	£0	0.0%

HEIFES comparison (3)

You should focus on the following areas:

- **Difference sheets** – these will display any differences for each table between the original and recreated data. You can toggle the threshold to highlight larger differences.
- **Rebuild information** – this shows a breakdown of students in different categories, showing differences between the original and recreated data, as well as the ILR fields which have been used to calculate the value shown.

Rebuild algorithm				
Summary sheet item	OfS Derived Fields	ILR fields	HEIFES17	HEIFES17 recreation
Students generating countable years				
Countable years including forecasts: Column 1 and 2	HEFEXCL = 0		0	0
Countable years excluding forecasts: Column 1	HEFEXCL = 0 and HEFREG = 1		0	0
Countable years forecasts: Column 2	HEFEXCL = 0 and HEFREG = 2		0	0
Forecast of years not completed				
Forecast of years not completed: Column 3	HEFEXCL = 0 and HEFCOMP = 3		0	0
Estimated countable years: Column 4	HEFEXCL = 0 and HEFCOMP = 4		0	0
Estimated countable years: undetermined completion status	HEFEXCL = 0 and HEFCOMP = 4 and FUNDCOMP = 3		N/A	0
Fundability				
Fundable countable years	HEFEXCL = 0 and HEFTYPE = HOMEF		0	0
Non-fundable countable years	HEFEXCL = 0 and HEFTYPE = HOMENF		0	0
Non-fundable countable years with no source of funding	HEFEXCL = 0 and HEFTYPE = HOMENF	and LEARNDEL FAM_SOF1 = BLANK	N/A	0

HEIFES comparison (4)

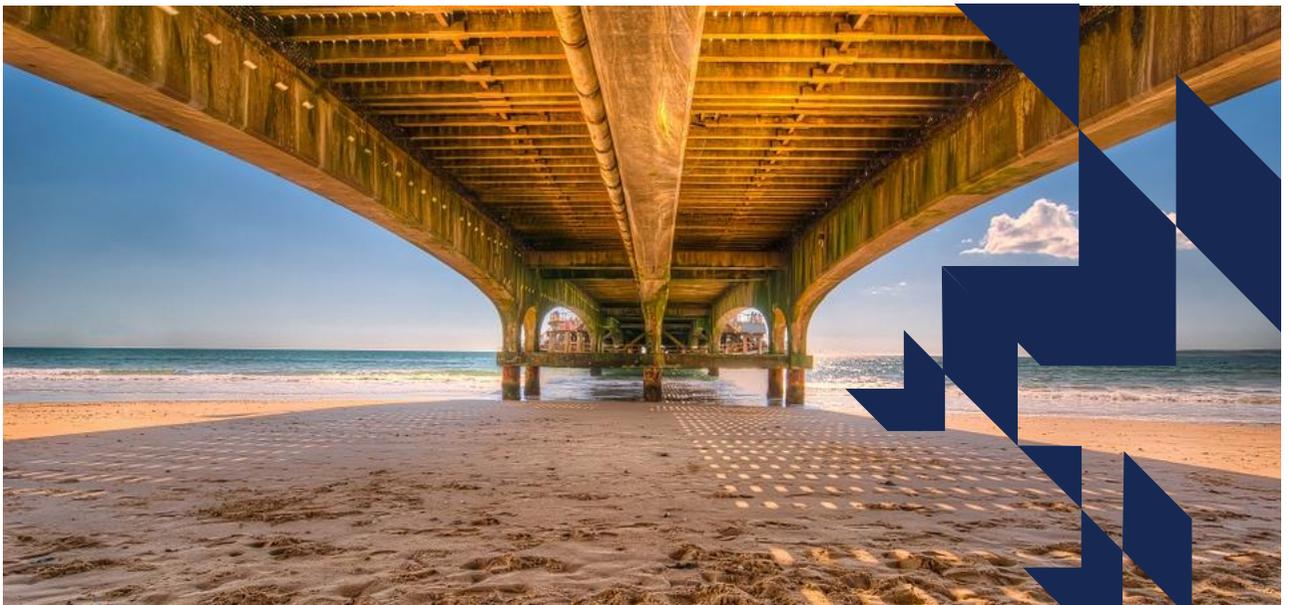
- You should also focus on the **Key differences** sheet. This shows key differences which we have identified in your ILR data, compared to your HEIFES data, as well differences in funding if we had used your ILR data in the calculations.
- These key differences form the basis of some of the queries you may receive as part of the verification process of the data checking tool.

2018-19 Funding					2017-18 Years countable					
	Allocations based on HEIFES17 return	Allocations based on HEIFES17 recreated using 2017-18 ILR data	Difference	Percentage difference			Years countable in HEIFES17 return	Years countable in HEIFES17 recreated using 2017-18 ILR data	Difference	Percentage difference
High-cost subject funding	£0	£0	£0	0.0%	By mode:	Full time	0	0	0	0.0%
Premium to support successful student outcomes: full-time	£0	£0	£0	0.0%		Sandwich year out	0	0	0	0.0%
Premium to support successful student outcomes: part-time	£0	£0	£0	0.0%		Part time	0	0	0	0.0%
Disabled students' premium	£0	£0	£0	0.0%	By level:	UG	0	0	0	0.0%
Postgraduate taught supplement	£0	£0	£0	0.0%		PG (UG fee)	0	0	0	0.0%
Accelerated full-time undergraduate provision	£0	£0	£0	0.0%		PG (Masters loan)	0	0	0	0.0%
Intensive postgraduate provision	£0	£0	£0	0.0%		PG (Other)	0	0	0	0.0%
Students attending courses in London	£0	£0	£0	0.0%	Total years countable		0	0	0	0.0%
Erasmus+ and overseas study programmes	£0	£0	£0	0.0%	Total new entrants		0	0	0	0.0%
Nursing and allied health supplement	£0	£0	£0	0.0%	Total apprenticeships		0	0	0	0.0%
Total 2018-19 funding (excluding fixed allocations)*	£0	£0	£0	0.0%	Total years countable subcontracted out to another provider		0	0	0	0.0%

HEIFES comparison (5)

What particular fields should you look out for?

- **FUNDCOMP** – to determine the completion status of the student. We assume a value of '3' is a completion, when it may actually be a non-completion. This could over inflate the difference in funding, leading to additional questions or selection for the reconciliation exercise.
- **HEFEXCL** – a derived field which excludes a student from the HEIFES population. Ensure that this value is '0' where you expect it to be, or you could be incorrectly excluding students from the population.
- **LEARNDELFAM_SOF** – the source of funding for the student. A value of '1' indicates OfS-fundable (previously HEFCE). You will need to determine fundability status yourself and set OfS non-fundable students to '998' or '999'. An exception is apprenticeships.



Apprenticeships

Apprentices: ILR fields used

ILR fields used in re-creating HEIFES Table 5 Section C

In addition to fields already described:

- **AIMTYPE**
3: Component learning aim within a programme
- **PROGTYPE**
10, 20, 21, 22, 23, 25
- **LEARNDELFAM_SOF**
105: Education and Skills Funding Agency (ESFA) - Adult

Apprentices: other work

Other work in the OfS on Higher and Degree Apprenticeships that uses ILR data includes:

- Quality assurance
- General analysis
- Extra NSS questions for Apprentices
- Exclusion from Unistats

Apprentices: Other ILR fields used

In addition to those used in re-creating Table 5 Section C of HEIFES, we use:

- **AIMTYPE**
 - Also look at **AIMTYPE** = 1: Programme aim
- Level of the Apprenticeship
 - Uses **PROGTYPE** for Frameworks
 - Uses notional end level field on LARS for Standards
- Level of the higher education component, and if it is prescribed or not
 - Uses information from LARS
- Possibly **STDCODE, FWORKCODE, PWAYCODE**
- Possibly **EMPID**

Apprentices: data quality issues

- Prior to 2016-17: missing information in the higher education fields: the ESFA have now added validation
- The higher education components sometimes being recorded as regular non-Apprenticeship OfS-fundable higher education
- Double counting across sub-contractual partners

Apprentices: other provider types

Higher education institutions or alternative providers generally need to record the higher education component on both HESA/HESA AP and ILR returns

- Currently we only use the HESA/HESA AP data
- A few cases where such providers are not filling in both returns

Fields used in HESA/HESA AP

- Only has the component aims (except where higher education providers have chosen to use HESA's ILR generation tool)
- **INITIATIVES** (HESA and HESA AP) and **PROGTYPE** (HESA) identify the Apprentices
- Higher education providers should complete **FUNDCODE** with respect to OfS funding rules, not ESFA funding rules
- **AIMTYPE** exists for higher education providers, but seems to be mixed practice in there

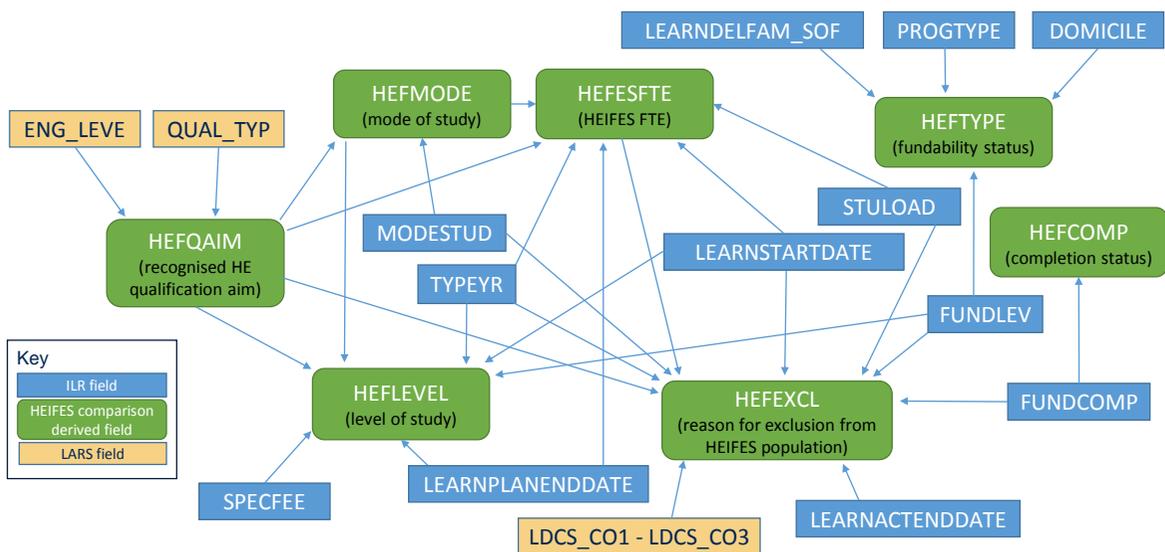


Student premium output

Student premium outputs: funding data summary

- The funding data summary shows how we have used a provider's data to calculate two targeted allocations:
 - Premium to support successful student outcomes: full-time
 - Disabled students' premium
- The download package contains:
 - Summary workbook
 - Individualised data file (.csv)
 - Technical document – contains descriptions and definitions of derived fields
 - Method document – explains funding methods and provides rebuild instructions
- We determine the headcounts in the summary using the following data:
 - ILR
 - National Pupil Database (NPD)
 - Young higher education participation rates by 2001 CAS ward (POLAR3)
 - Proportion of 16-74 year olds with a higher education qualification by 2001 CAS ward
- The individualised data file contains data from the datasets listed above. For data protection reasons, we exclude data related to qualifications on entry if it is sourced from ILR or NPD datasets.
- Some of the derived fields from the HEIFES comparison output are used to calculate the student premium (SP) derived fields.
- The algorithms for the derived fields in the SP and HEIFES comparison individualised files are included in the technical documents.

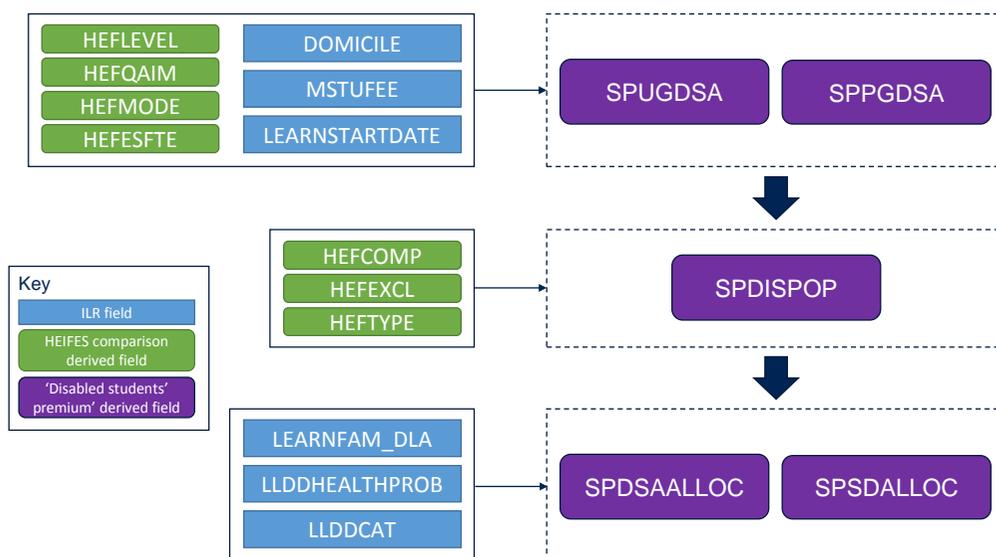
HEIFES comparison fields used in SP algorithms



Disabled students' premium

- Based on the number of students in a provider's ILR R14 data recorded as:
 - In receipt of the Disabled Students' Allowance (DSA)
 - Not in receipt of the DSA but having self-declared a disability
- These headcounts are recorded by the following derived fields:
 - SPDSAALLOC**
 - SPSDALLOC**
- The derived field **SPDISPOP** records the number of students in the recreated HEIFES population who were potentially eligible to receive the DSA and completed the year of their programme of study

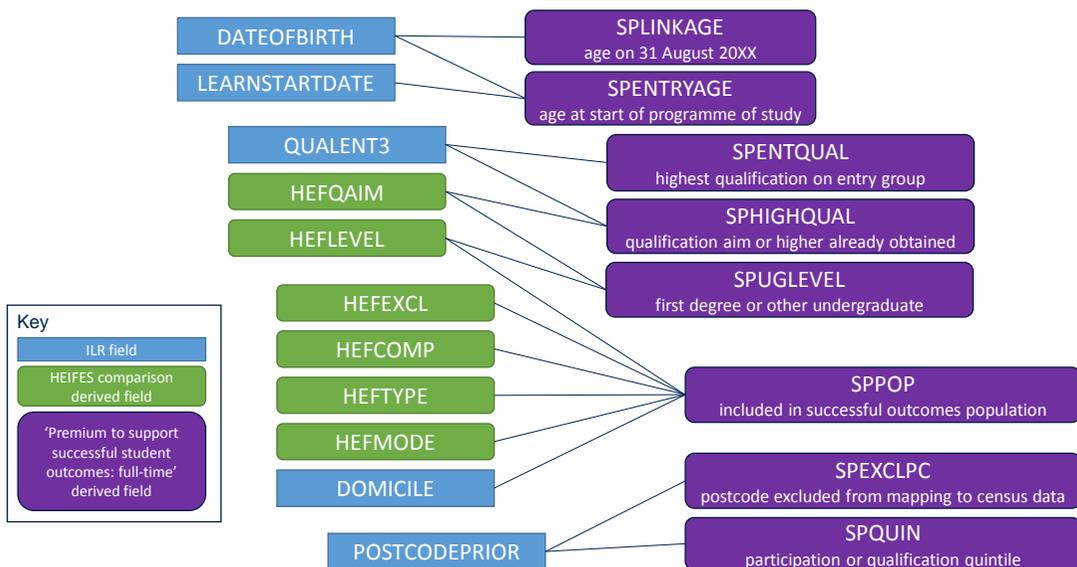
Key fields for the 'Disabled students' premium'



Premium to support successful student outcomes: full-time

- For full-time and sandwich year out undergraduates, the premium consists of:
 - A **main allocation** – based on the number of students assessed as being at risk of not completing their studies, using their Level 3 qualifications
 - A **supplement** – based on the number of these at-risk students who are assessed as being from the most disadvantaged backgrounds, using postcodes to determine participation or qualification quintiles
- The allocations are calculated using:
 - ILR R14 data
 - Level 3 qualifications from linked ILR and NPD data from 2002-03 to 2015-16
 - Young higher education participation rates by 2001 CAS ward
 - Proportion of 16-74 year olds with a higher education qualification by 2001 CAS ward
- We link ILR R14 data with other data sources (any provider’s ILR data and schools’ NPD data) to derive students’ UCAS tariff points and Level 3 grade combinations
- We assign each student to a higher education participation or qualification quintile using the 2001 CAS ward for their postcode. CAS wards in quintile 1 have the lowest young higher education participation rates or the lowest proportion of higher education qualified adults

Key fields for the ‘Premium to support successful student outcomes: full-time’



Student premium outputs: funding data comparison (1)

- The funding data comparison workbook generated by the data checking tool compares SP headcounts calculated using:
 - ILR data submitted to the data checking tool
 - ILR R14 data from the previous academic year
- It also compares the 'Disabled students' premium' allocations calculated using these datasets, using the previous academic year's funding method.
- The results package contains:
 - Comparison workbook
 - Individualised data file (.csv)
 - Technical document – contains descriptions and definitions of derived fields
 - Method document – explains funding methods and provides rebuild instructions
- In addition to ILR data, we use the following data to determine the headcounts in the comparison workbook:
 - Young higher education participation rates by 2001 CAS ward
 - Proportion of 16-74 year olds with a higher education qualification by 2001 CAS ward

Student premium outputs: funding data comparison (2)

- For data protection reasons, the data checking tool does not link to other data sources. The following fields are derived using data from other sources and are therefore excluded from the funding data comparison outputs:
 - **SPLINK, SPLINKAGE, SPTARIFF, SPGRADECOMB, SPENTQUAL**
- The fields listed above are used to inform the 'Premium to support successful student outcomes: full-time' allocations. Without them, it is not possible to assign students to risk categories using their Level 3 qualifications.
- Consequently, the funding data comparison calculates neither the main allocation nor the supplement for this premium, and compares only the number of students in particular participation or qualification quintile groups.
- The workbook is designed to enable the provider to check that the SP headcounts are as they would expect by highlighting significant year-on-year differences

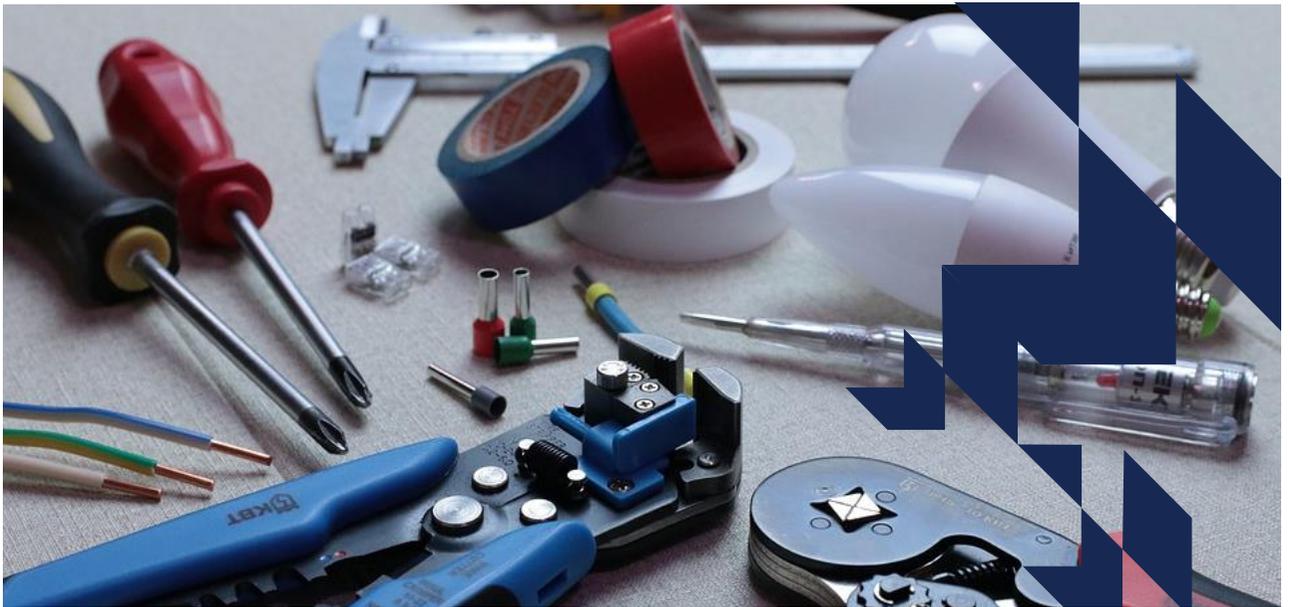
Learning Aims Reference Service (LARS)

- It is important that LARS is kept as up to date as possible throughout the year.
- LARS is maintained by the ESFA and contains all of the information regarding your courses.
- We use this for all outputs that we produce, therefore it is vital that it is kept up to date and that all of your courses are present on LARS.
- If not, it is likely that students could have the wrong qualification aim, be assigned to the wrong price group, or be excluded altogether.
- Get in touch with the ESFA service desk to add or amend learning aims on LARS.



HEIFES comparison exercise

Aim: how to use a data output and individualised file to identify errors or issues in your ILR data



Data checking tool

2017-18 ILR timeline

January 2018: HEIFES17 process completed

August 2018: Expected release of first outputs for 2017-18 ILR data checking tool

September 2018: Data verification of ILR data starts

18 October 2018: 2017-18 ILR R14 submission deadline to the ESFA

November 2018: OfS receive final 2017-18 ILR data

Spring 2019: OfS release outputs using final 2017-18 ILR data

Spring 2019: OfS 2019-20 Spring grant announcement

Spring 2019: Reconciliation exercise



What is the data checking tool?

The data checking tool is a process in which you can check the accuracy of your ILR data, before you submit the ILR R14 return to the ESFA

We create the outputs that we previously spoke about from your ILR data about which will allow you to check the data is correct

We strongly encourage you to use the data checking tool. Previous use has shown that many errors in ILR data have been identified through the use of the tool.

Using the data checking tool

- Colleges are able to upload their ILR returns to the OfS portal, which will return a set of outputs based on the data uploaded.
- Review these outputs as we suggested in the previous slides.
- Outputs are released in stages.
- You will also receive a list of queries based on the data you have uploaded, and will be contacted by our data verification team to help clarify the data. They will work with you to ensure that your data is correct.
- You can upload as many times as you like.



Welcome to the OfS portal

What we covered

- How the OfS uses the ILR
- Specific fields that we are interested in
- Outputs that we produce
- Data checking tool



Further information



- The OfS website will be updated in the run up to the release of the data checking tool with guidance regarding our use of ILR data
- You can look at the HEFCE website for guidance from last year.
 - Data and statistics >
 - How we use individualised student data
- You can email us:
ilrdata@officeforstudents.org.uk

Thank you for listening

Images used under a CC0 1.0 license:

- slide 2 by QuinceMedia (<https://pixabay.com/en/arrow-target-bullseye-goal-aim-2886223/>)
- slide 3 by geralt (<https://pixabay.com/en/matrix-technology-technique-data-3407242/>)
- slide 4 by aitoff (<https://pixabay.com/en/post-box-post-mail-box-wall-dry-2691454/>)
- slide 7 by ddzphoto (<https://pixabay.com/en/field-crop-storm-agriculture-farm-3416824/>)
- Slide 14 by qimono (<https://pixabay.com/en/question-mark-important-sign-1872634/>)
- slide 15 by Pexels (<https://pixabay.com/en/apple-caffeine-coffee-coffee-shop-1850613/>)
- Slide 17 by EsaRiutta (<https://pixabay.com/en/excel-spreadsheet-icon-mess-icons-1771393/>)

- Slide 18 by GDJ (<https://pixabay.com/en/binary-0-1-background-3322478/>)
- Slide 24 by Free-Photos (<https://pixabay.com/en/pier-jetty-wooden-beach-sand-801761/>)
- Slide 30 by Free-Photos (<https://pixabay.com/en/books-pages-story-stories-notes-1245690/>)
- Slide 41 by Sid74 (<https://pixabay.com/en/electrician-wiring-installation-3087536/>)
- Slide 42 by rawpixel (<https://pixabay.com/en/paper-composition-business-writing-3141341/>)
- Slide 45 by GDJ (<https://pixabay.com/en/thumbs-up-agree-approve-condone-3350838/>)
- Slide 46 by kreatikar (<https://pixabay.com/en/online-education-tutorial-3412473/>)

The copyright in this presentation is held either by the Office for Students (OfS) or by the originating authors.

Please contact info@officeforstudents.org.uk for further information and re-use requests.