

HIGHER EDUCATION DATA WORKSHOP

6 JUNE 2007

Higher Education Data Workshop

Outline

- Data requirements
- Data quality analysis
- HESA/EYM & LLWR/EYM comparisons
- Possible use of HESA data for funding in future
- Non-completions, partial completions and transfers
- ITT reconfiguration
- Timetable
- Presentation by University of Wales Swansea
- New audit process
- WAG – update on data usage
- WAG - Welsh for Adults

Data Requirements – HESA Student Record 2007/08

Letter to HEIs (UK-wide)

- Confirmation that HEIs can meet the expectation to make a complete return for 2007/08 using the new system
 - Within normal timetables
 - Without a significant reduction in quality
- To ensure necessary steps are in place
- To be alerted to anticipated problems

Data Requirements – HESA Student Record 2007/08

Module entity

- MODLANG – Module is available through Welsh
- LANGPCNT – Percentage through Welsh

Data Requirements – HESA Student Record 2007/08

Module entity

- MODFTE – Module FTE
 - Guidance on calculating FTE
 - Related to STULOAD

Data Requirements – HESA Student Record 2007/08

Student on Module entity

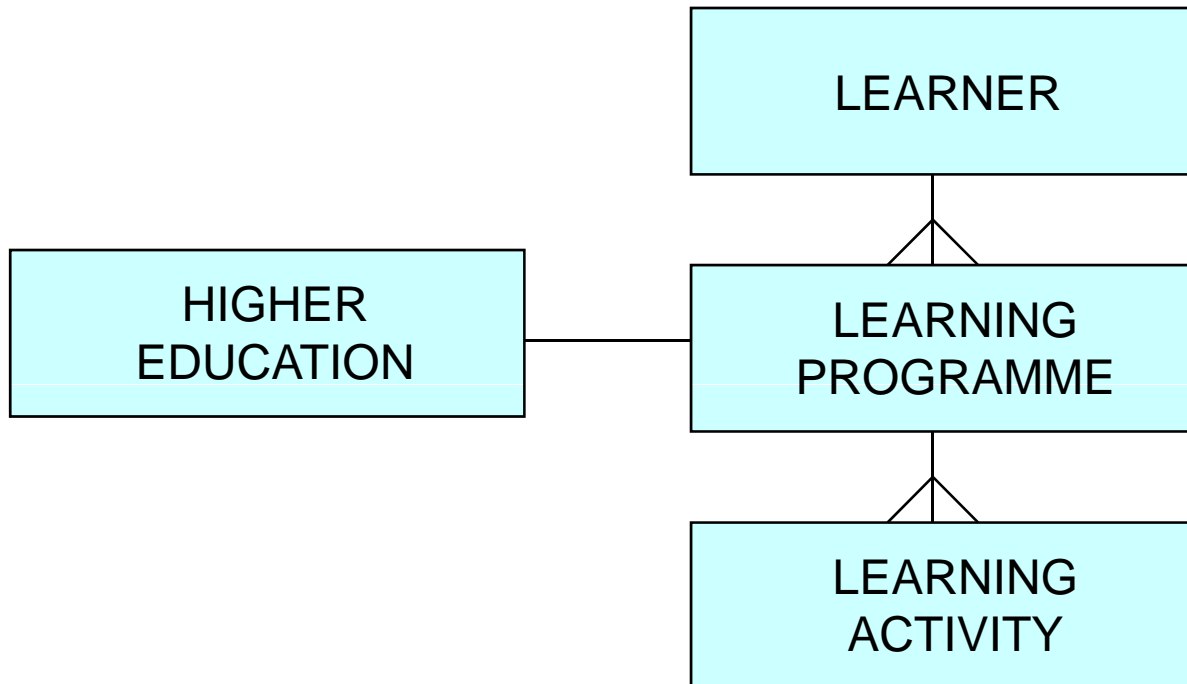
- MODSTAT – Module status
- MODOUT – Module outcome
 - 1 Completion - gained full credit
 - 2 Completion - did not gain credit
 - 3 Partial completion
 - 4 Student did not complete module
 - 5 Module taken on a not for credit basis
 - 6 Module outcome not yet known

Data Requirements – HESA Student Record 2007/08

- Other coding issues
 - RSNEND
 - PGR students
 - Fundability status
 - Subject of study (2006/07)
 - UoA (2007/08)
- 2007/08 data extraction and comparisons

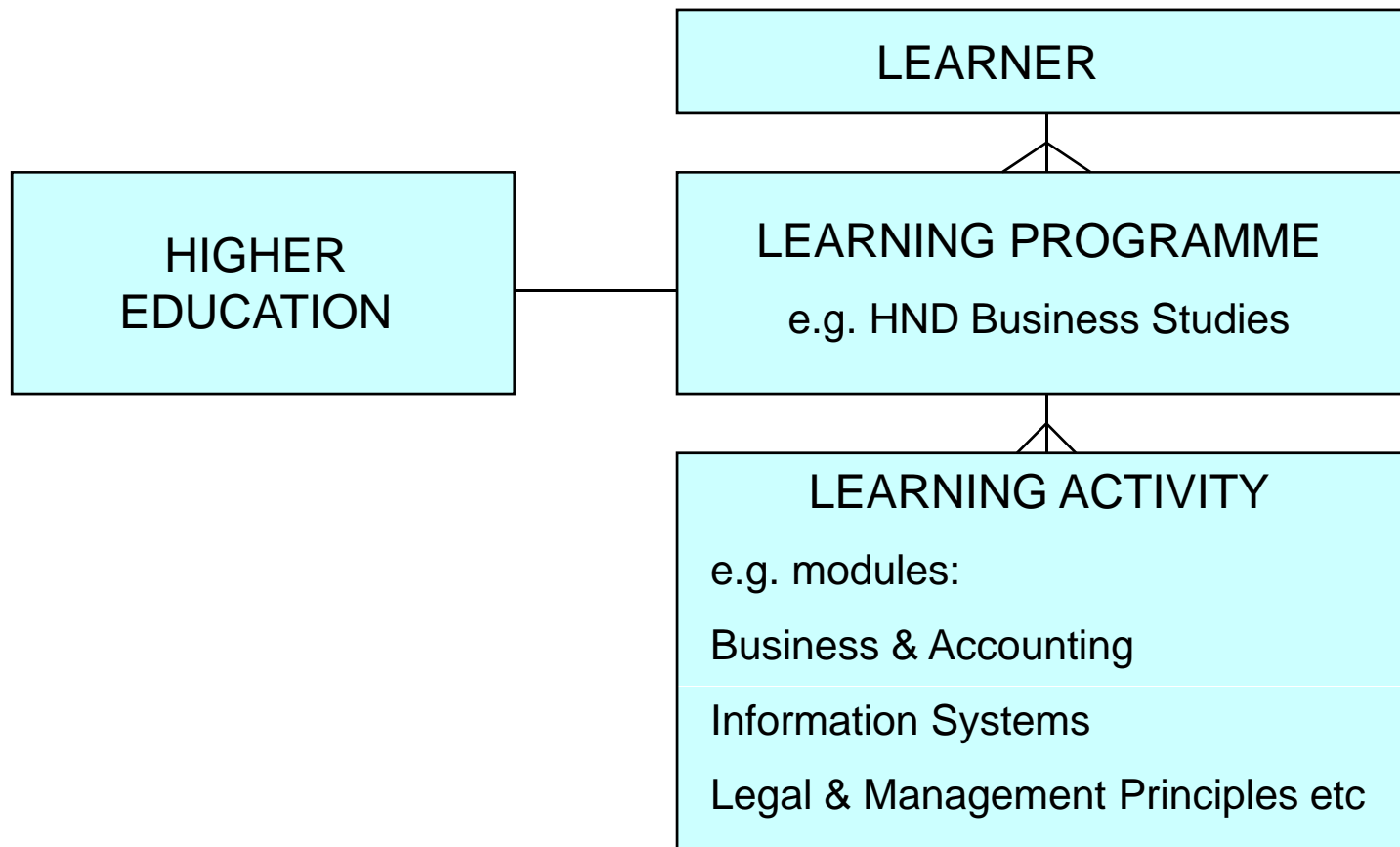
LLWR Coding Issues

Structure showing relationships between different datasets on Lifelong Learning Wales Record



LLWR Coding Issues (2)

Example structure of a HEFCW fundable programme of study



LLWR Coding Issues (3)

HE specific coding and guidance changes for 2007/08 data collection *(list not exclusive)*

LP07 (mode of attendance)

Coding: 07 (full-time), 08 (sandwich year out), 09 (part-time)

LP25 (financial support for learner)

Guidance: For HEFCW fundable HE learners in receipt of Disabled Students Allowance (DSA), this field must be coded as '10', irrespective of the source that provides the greatest amount in financial support. Essential for calculation of HEFCW Disability Premium Funding.

LA06 (learning aim reference)

Guidance: Partners delivering HE-level modules should record each module using the following generic codes: LDKKHE07 (general vocational) or LDLLHE07 (general non-vocational)

LLWR Coding Issues (4)

HE specific validation changes for 2007/08 data collection

(list not exclusive)

Validations have been introduced to ensure that HEFCW fundable Learning Programmes have specific fields relating to HE funding/analysis coded on a consistent basis.

For example:

- If HE mode (LP07='07'-'09') then HE level learning programme (LP17 = '51'-'59')
- If HE level learning programme then it maps to a HE dataset
- If HEFCW fundable learning activity (LA11='2') then HE level learning programme and HE mode of attendance
- If HEFCW fundable learning activity (LA11='2') then learning aim reference is coded generically (LA06 = 'LDKKHE07' / 'LDLLHE07')

LLWR Coding Issues (5)

**2007/08 data collection guidance issued by Welsh Assembly
Government in June/July 2007**

http://new.wales.gov.uk/topics/educationandskills/information_for_learning_provide/LLWR/?lang=en

The Support Desk for LLWR data collection can be contacted as follows:

E-mail: llwr@wales.gsi.gov.uk

Tel: 029 20 926046

HESA Data Quality Analysis 2005/06 – 1

Areas considered may have been included because they:

- have significance in relation to funding
- are areas receiving greater focus from policy teams or the Assembly
- impact on equal opportunities analysis
- relate to historic data problems

Changes from 2004/05 analysis:

- tables 1 – 9 are the same as last year
- basis of calculation has changed in tables 4 and 6
- tables 10 – 15 are new

HESA Data Quality Analysis 2005/06 – 2

Table 1 - Postcodes

- the level of valid postcodes is very high, with 1% of postcodes being invalid or missing

Table 2 - Ethnicity

- small improvement from 2004/05 in known ethnicity (94% to 95%)
- due to drop in ethnicity refused (from 3% to 2%)
- no change in proportion of enrolments with unknown ethnicity (3%)

HESA Data Quality Analysis 2005/06 – 3

Table 3 - Welsh subject provision

- Q500 Celtic studies used instead of Q560 Welsh or Q561 Welsh literature
- difficult to analyse Welsh provision

Table 4 – Proportion of module taught through Welsh

- very few modules where the number of credits taught through Welsh is less than 2

Table 5 – Coding of non-standard academic years

- academic years that are potentially non-standard coded as standard (e.g. masters courses expected to last more than 40 weeks)

HESA Data Quality Analysis 2005/06 – 4

Table 6 – Average FTE/credits per enrolment

Table 7 – Data linked by HUSID or HUSID and instance to 2004/05 data

- fewer linked records with different date of birth, ethnicity, commencement date or postcode than in 2004/05

Table 8 – Data linked to 2004/05 data using HEFCE linking

- linked by using name, date of birth etc.
- 8% of linked records with different HUSID
- students who are most likely to have different HUSID across years have changed institution

HESA Data Quality Analysis 2005/06 – 5

Table 9 – Students starting in 2005/06 also present in 2004/05

- some also present in 2004/05 (1% overall)
- proportions at individual institutions vary from 0% to 2% (0% to 5% in 2004/05)
- some large proportions for some mode/level combinations at a few institutions

HESA Data Quality Analysis 2005/06 – 6

Table 10 – Non-completion of current year

- majority of these students leave institution (78%)
- much greater proportions of students who have suspended studies or are dormant (varying between 19% and 44%) at some institutions than the sector overall (9%)

Table 11 - Non-completion of current year and reason for leaving

- 55% gave personal or other reasons for leaving
- 25% were written off after a lapse of time
- 10% reason for leaving was unknown

HESA Data Quality Analysis 2005/06 – 7

Table 12 – Welsh fluency

- 58% of Welsh domiciled students with known Welsh speaker status (increase of 5% on 2004/05)
- at five HEIs, more than 50% with unknown Welsh speaker status
- at seven HEIs, between 3% and 23% with unknown Welsh speaker status

Table 13 – Disability

- small improvement from 2004/05 in coding known disability status (92% to 93%)
- 100% known disability status at eight institutions

Table 14 – Disabled Student's Allowance (DSA)

- 20% of disabled student enrolments with information on DSA not known

HESA Data Quality Analysis 2005/06 – 8

Table 15 – Level of subject coding

JACS Code	JACS Title	Level of Coding
B100	Anatomy, Physiology and Pathology	Principal subject - coded to two places
B130	Pathology	Coded to three places
B131	Cellular Pathology	Coded to four places

- 75% of enrolments coded to principal subject level
- 25% of enrolments coded to three or more places
- at individual HEIs, the proportion of enrolments coded to three or more places varied between 1% and 49%

COMPARISONS BETWEEN HESA AND **EYM DATA**

Total number of assumed completed credit values extracted from HESA compared to verified EYM data

In 2004/05 HESA 2% higher than EYM data

In 2005/06 HESA 3% higher than EYM data

COMPARISONS BETWEEN HESA AND **EYM DATA**

**The range of % differences between HESA
and EYM total assumed completed credit
values range from:**

FT -7% to +9% (sector 1%)

SW -47% to +56% (sector -7%) (small numbers)

PT -19% to +61% (sector +11%)

Total -5% to 10% (sector +3%)

COMPARISONS BETWEEN HESA AND EYM DATA

Franchised Out (1):

2005/05 HEFCW fundable credit values

EYM = 427,833

HESA = 448,282

EYM 5% less than HESA

COMPARISONS BETWEEN HESA AND EYM DATA

Franchised Out (2):

- **No longer be collecting Table 5 in EYM.**
- **Will instead use HESA data for analysis of franchised out provision and publication.**
- **A summary of franchised out data from HESA will be sent to institutions each year for information and to alert us of any data errors before analysed/published.**

COMPARISONS BETWEEN HESA AND EYM DATA

Franchised Out (3):

Inconsistencies still occur:

- **HESA/EYM mapping advises franchised out is frnchact =3, should also include = 2 (mixed franchise arrangements)**
- **TINST1 'Other' institution coding**
- **TINST coded with FE colleges that no longer exist**

COMPARISONS BETWEEN HESA AND EYM DATA

Observations (Table 1a, 1b, 1c):

- **Some duplication of modules in a particular course in one institution**
- **Apportionment of modules to subject codes, especially between ASC 2 and 8**
- **Part-time PGT credits under reported at EYM**
- **Miscoding of sandwich year out on HESA**

COMPARISONS BETWEEN HESA AND EYM DATA

Observations (Regs and ITT):

- **Island and Overseas registrations not comparable**
- **ITT – some subjects of study coded at generic level on HESA e.g. Other MFL, Other Science**
- **Use of Y000 generic subject coding has been greatly reduced since 2004/05 HESA**

COMPARISONS BETWEEN LLWR AND EYM DATA

**LLWR month 5 data for 2005/06
= 1,100 registrations.**

**EYM 2005/06 HEFCW fundable at FE
colleges
= 1,560 registrations.**

COMPARISONS BETWEEN LLWR AND EYM DATA

Observations (1):

- **Inconsistencies of how FE colleges record LA on LLWR record. Course/Year/Module level.**
- **Credit value of year of LA for year/module not available.**
- **HEFCW fundable LAs not coded as HEFCW fundable (LA11=2) in error.**

COMPARISONS BETWEEN LLWR AND EYM DATA

Observations (2):

- **HE level LPs not coded using HE codes (LP17= 51-59) in error.**
- **HE level LP do not map to HE dataset entry**

**A number of validations and extra guidance
will be issued in the 2007/08 LLWR
guidance.**

COMPARISONS BETWEEN HESA/LLWR AND EYM DATA

**Will continue to make comparisons of the
HESA/LLWR data against EYM returns.**

Analysis will be sent to institutions in July (for
HESA) and December (for LLWR).

If anyone has any comments on the comparisons
or would like us to look at specific elements of
their data in more detail please email
hestats@hefcw.ac.uk

Possible Use of HESA data for Funding in Future

Reasons for Change

- HESSES and EYM burdensome for HEIs and HEFCW
- New fee arrangements
- With large numbers of 'fees only' students current model does not provide steers
- Likely to be good estimates of subject costs from TRAC-T work

Current HEFCW Thoughts

- Consider new teaching funding method based on HESA - dropping as many processes (eg Redistribution) as possible
- Consider funding by cost centre not subject area

Issues

- Need to adapt HESA Student Record slightly
- Need to improve data quality
- Need to get new mappings to cost centres
- Unlikely to be able to introduce before 2010/11 funding round

Non-completions

- Registrations
 - Where students do not take required assessment for the year
- Modules
 - Where students do not take required assessment for the module

Transfers

- Registrations
 - Not the same as non-completions – student remains at the institution and has not dropped out or withdrawn
 - Transfers should be returned in the ASC that they end the year in for the purposes of returning registrations

Transfers

- Modules
 - Transfers before or on 1 November – return as if always on the transferred to modules
 - Transfers after 1 November –
 - add credits to column 2 in the appropriate ASC of EYM tables for transferred to modules
 - subtract credits from column 2 in the appropriate ASC for transferred from modules
 - base column 3 figures on new column 2

Partial completions – part-time (1)

- Table 7 of HESES 2006/07
 - 80,287 partially completed credits in total (all HEIs/FEIs in Wales)
 - 28% of non-completed part-time credits
 - 5% of total completed + partially completed part-time credits
 - 1% of total completed + partially completed part-time and full-time credits

Partial completions – part-time (2)

- Credits equivalent to £2.3m over whole sector
- Count as fees-only in model due to overall level of recruitment
- Inclusion in teaching model does not affect recurrent grant outcome, but changes distribution between core and recovery (recovery reduced by 17%)

ITT Reconfiguration

- Numbers lost through ITT reconfiguration have been moved to ASC 11b and ringfenced (published as a separate category in the Recurrent Grant circular)
- Not included in redistribution
- Letter asks for proposals relating to this provision
- Will be monitored through HESES – likely to be additional table

Data Requirements – Widening Access Premium

- Alternative to the low affluence categorisation
- PISG have set up a technical group to consider widening participation and access performance indicators
- Technical group expected to report back later this year

2007/08 Timetable

6 June	Workshop for institutions
2 August	EYM return date
21 August	EYM verifications sent out
31 August	Return date for redistribution exercise
7 September	EYM verifications return date
Sept/October	Audit of data (selected institutions)
Mid September	Data submitted to HESA
Mid Sept/Oct	HESA data quality checking
Mid October	HESES survey issued
Mid November	HESES return date

University of Wales Swansea

New Audit Process (1)

- Proposals sent to heads of institutions in July 2006
- Responses received in October 2006
- Responses analysed and original proposals amended
- New process notified to institutions in April 2007

New Audit Process (2)

- New process:
 - Reliance placed on institutions' own internal audit reports in respect of systems and processes, with reports to be submitted in May each year
 - An external audit carried out at each institution at least once every five years
 - Additional external audits carried out where there is cause for concern
 - External audits will check the accuracy of data returns plus systems and processes
 - Auditors appointed by HEFCW will carry out all external audits

New Audit Process (3)

June 2007	Notification to 'at risk' institutions (11 June) Tender specification for external auditors issued Guidance for auditors issued
July 2007	Potential external auditors considered
August 2007	Appointment of external auditors Workshops for internal/external auditors
Sept/Oct 2007	Audit of 'at risk' institutions
May 2008	Internal audits reports returned
September 2008	First external audits on a sample basis

HE Data Workshop

Welsh Assembly Government use of HESA data

Susan Full

Higher Education & Lifelong Learning Statistics
Statistical Directorate
Welsh Assembly Government



Outline

- Who are we
- Examples of use of HESA data
- Some data quality concerns
- Other issues



Who are we?

- Part of the Welsh Assembly's Statistical Directorate
- Team of 7
- Very close links with the Assembly's Higher Education and Student Finance Divisions; HEFCW; HESA
- Also cover Student Finance (liaison with Student Loans company etc)
- Acting head of section/statistician: Susan Full
- Email HE.stats@wales.gsi.gov.uk



Our role in relation to HESA data

- Statutory Customer of HESA (along with HEFCW, DfES etc.)
- Represented on most HESA groups - e.g. Student Record Review, Statutory Customers technical group, DLHE Longitudinal Steering Group
- Represent the Assembly's views and requirements
- Interest in data on Welsh domiciles, Welsh institutions and UK comparisons



What use do we make of HESA data?

- National Statistics Publications
- Policy development and monitoring
- Internal briefing
- Assembly Questions
- Ad-hoc requests



National Statistics publications

- Series of First Releases and Statistical Bulletins intending to provide easy access to Welsh data, and to inform policy, for key stakeholders and the public in general:
 - Students in Higher Education Institutions, Wales
 - Staff and Finance of HEIs in Wales
 - Destination of Leavers from HE
 - Initial Teacher Training in Wales (& ITT destinations)
 - HE Performance Indicators
 - Student Flows in Wales, 2001/02 and 2002/03
 - Welsh in Higher Education (first published in 2006)



National Statistics publications

- Focus is on national trends of Welsh domiciles or students at Welsh institutions
- Some institutional data are published
- Can be found on www.wales.gov.uk/statistics
- Interactive tables available on www.statswales.wales.gov.uk



New output: Welsh in Higher Education

- Published for the first time in 2006
- Will be publishing 2005/06 version at end of June
- Data have been included in the Welsh Language Board's Vitality of Welsh Statistical Balance Sheet
- Some quality issues
- Student ability to a speak Welsh was published in an Annex as the rate of unknowns was high. This seems to have decreased for 2005/06



Welsh for Adults analysis

- A number of requests for WfA data
- Restructuring into WfA centres
 - Request for timeseries by institution and centre as far back as possible!
 - Age
 - Ethnicity
 - Disability
 - Domicile
 - Socio-economic class
 - Deprivation (used Community First areas)
 - Moved to Wales
 - Retention
 - Attainment
- Quality issues



Welsh for Adults and other FE data collection

- Have been working with DELLS and HESA to try and accommodate DELLS data needs in HESA student record
- Latest student record update included some changes to fields
 - Courseaim – WfA expanded to cover levels and specialist
 - Credit Transfer Scheme – GNVQ units/ OCN credits
 - Numunits –opened up to Welsh HEIs
- Will need detailed campus information
- Learn Direct subject classification not included
- Are considering a reduced record for FE in Welsh HEIs

Data quality issues

- Concern less on funding fields compared with HEFCW's interest
- Examples of recent concerns:
 - Year of study data
 - Postcodes
 - Subject coding
 - NS-SEC
 - Employer funding of part-time provision
 - Welsh speaking ability
 - Modules taught through Welsh
 - Staff teaching through Welsh



Some final thoughts for discussion

- Some specific Assembly requirements
 - e.g. Welsh medium ability, Welsh Bacc, additional consultation question on NS-SEC for part-time students
- Would welcome further direct liaison with sector and HEFCW
- Aware of managing burden on institutions
- Data collection issues
- Question bank

To sum up...

- Data collected are invaluable, not just for funding but for policy development, monitoring and the wider public interest.
- We never know which HESA data items we will be using next.....

Questions and discussion

