European Tertiary Education Register (ETER)

[Contract No. EAC-2015-0280]

ETER Final Quality Report

Authors: ETER Quality Managers from Sapienza University of Rome: Cinzia Daraio, Monica Scannapieco, Tiziana Catarci and Léopold Simar.

Rome, 28 July 2017

Content

This document describes the Data Quality Methodology developed within the ETER project, the role of data quality in the overall ETER implementation and certifies the data quality of the four waves of data collected within the ETER contracts No. EAC-2013-038 (hereafter indicated as ETER, which refers to the 2014 and 2015 data collections) and EAC-2015-0280 (hereafter indicated as ETER II, which refers to the 2016 and 2017 data collections). The data quality indicators have been calculated on the ETER dataset as of May 30th 2017 and on those countries whose data were provided and validated by the National Statistical Authorities. For more general information on ETER, the report refers to the main documents on the ETER project namely the ETER Final Report and the ETER Handbook available on the project website (https://www.eter-project.com).[1]

Contents

TABLE 1. LIST OF ABBREVIATIONS AND COUNTRY CODES

Abbreviation	Full Name
DG EAC	Directorate General Education and Culture
DG RTD	Directorate General for Research and Innovation
EC	European Commission
EEA	European Economic Area
EFTA	European Free Trade Agreement
EHEA	European Higher Education Area
ERA	European Research Area
ETER	European Tertiary Education Register
EU	European Union
EUMIDA	European Microdata Project
EUROSTAT	European Statistical Office
FOE	Fields of Education
FTE	Full Time Equivalent
FYRM	Former Yugoslav Republic of Macedonia
HC	Head Count
HEI	Higher Education Institution
ISCED	International Standard Classification of Educational Degrees
NE	National Expert
NIFU	Nordic Institute for Studies in Innovation, Research and Education
NSA	National Statistical Authority
UAS	Universities of Applied Sciences
UOE	UNESCO OECD EUROSTAT handbook on educational statistics

ISO country code	Country	ISO country code	Country
AT	Austria	LI	Liechtenstein
BE	Belgium	LT	Lithuania
BG	Bulgaria	LU	Luxembourg
CH	Switzerland	LV	Latvia
CY	Cyprus	ME	Montenegro
CZ	Czech Republic	MK	Former Republic of Macedonia
DE	Germany	MT	Malta
DK	Denmark	NL	Netherlands
EE	Estonia	NO	Norway
GR	Greece	PL	Poland
ES	Spain	PT	Portugal
FI	Finland	RO	Romania
FR	France	RS	Serbia
HR	Croatia	SE	Sweden
HU	Hungary	SI	Slovenia
IE	Ireland	SK	Slovakia
IS	Iceland	TR	Turkey
IT	Italy	UK	United Kingdom

1 INTRODUCTION

The European Tertiary Education Register (ETER) is a project contracted by the European Commission's Directorate General Education and Culture (contract EAC-2013-0308 with study period 01.08.2013-31.07.2015 - hereafter ETER I, and contract EAC-2015-0280 with study period 01.08.2015-31.07.2017–hereafter ETER II), which aims to compile a register of European Higher Education Institutions (HEIs) and to collect a comparable set of data for the HEIs in the perimeter.

The study is a joint undertaking of five partners, Università della Svizzera italiana (USI), Lugano, Centre for Organizational Research; Joanneum Research, Graz; Nordic Institute for Studies in Innovation (NIFU), Research and Education, Oslo; University of Rome La Sapienza, Department of Computer, Control and Management Engineering, Antonio Ruberti, Rome; Department of Energy, Systems, Territory and Construction Engineering, University of Pisa. It is supervised by the Directorate General of Education and Culture of the European Commission, in cooperation with DG Research and Innovation and EUROSTAT.

It follows the pioneering Aquameth Project and the European MIcroData project (EUMIDA), a large-scale study supported by the European Commission from 2009 to 2011, which demonstrated the feasibility of collecting European-level data on individual HEIs.

This final Quality Report describes the quality of the data collected in the four waves of the ETER data collection process, i.e. the 2014, 2015, 2016 and 2017 data collections.

This Quality Report has been structured according to recent developments on quality in the European Statistical System (ESS) that include specific guidelines for the preparation of quality reports for a full range of statistical processes and their outputs[2]. According to such guidelines, a quality report may be user-oriented, producer-oriented or both. Producer-oriented reports have a specific focus on the process according to which statistics were obtained. User-oriented reports are, in general, a subset of the producer-oriented ones; users of final outputs can be advanced analysts and researchers, or the public at large.

With respect to practical implementations of quality reporting, recent developments indicate to use metadata structures for the purpose of quality reporting. The Euro-SDMX Metadata Structure (ESMS) was recommended as reference metadata report structure in Commission Recommendation 498/2009. This ESMS can be considered as the user-oriented format of quality reporting because it contains a basic level of quality information that is structured along the quality criteria as defined in the ES Code of Practice and Regulation 223/2009 – the information focuses more on the statistical output rather than on the underlying process itself. Instead, a more detailed quality reporting structure called ESS Standard for Quality Reports Structure (ESQRS) was developed in 2010 which is more addressed to the producers of statistics and which focuses more on the statistical process side.

The ESS handbook for quality reports[3] details practical guidelines that have been considered in ETER quality reporting. In particular, as detailed in the following of this document, ETER quality reporting has been designed in order to be (i) user-oriented and (ii) with a metadata structure that is conform to ESMS.

The main sections of the Quality Report are reported in Appendix 1 for purpose of being used as a reference for future data collections.

---

[2] http://ec.europa.eu/eurostat/web/quality/quality-reporting
[3] http://ec.europa.eu/eurostat/documents/3859598/6651706/KS-GQ-15-003-EN-N.pdf

2   QUALITY REPORT: INTRODUCTION

2.1   GENERAL DESCRIPTION OF THE PROCESS AND ITS OUTPUTS

The ETER project created a database collecting information on Higher Education Institutions (HEIs) in Europe, concerning their basic characteristics and geographical position, education activities, staff, finances and research activities.

The ETER database is targeted to include 36 countries composed by the 28 European Union Member States (the Belgium data collection is split between the Flemish and French part), plus EFTA countries (Iceland, Liechtenstein, Norway, Switzerland) and other four EU acceding countries (Former Yugoslav Republic of Macedonia, Montenegro, Serbia and Turkey). The current ETER coverage includes the perimeter (i.e. the list of HEIs) and descriptors for all countries, while quantitative data are provided for 33 countries, with the exclusion of Montenegro, Romania and Slovenia. The Belgian data collection is limited to the Flemish region. Data for Iceland and Turkey have been collected directly by the consortium based on official data published online: their coverage is partial and they have not been validated by the respective NSAs..   ETER I collected the data which refer to the years 2011 (academic year 2011/2012) and 2012 (academic year 2012/2013); while ETER II collected the data which refer to the year 2013 (academic year 2013/2014) and 2014 (academic year 2014/2015). For a few countries, quantitative data are missing in specific years: Denmark 2014, Hungary 2012, Serbia 2011, Turkey 2011 and 2012.

Overall, in the year 2014, the ETER database includes **2,624** HEIs with more than 22.7 million undergraduate and graduate students and around 715 thousands PhD students.

ETER data have been provided by the National Statistical Authorities (NSAs), Higher Education Ministries or Higher Education Agencies, based on national statistical databases or higher education information systems. They have been complemented by descriptors and geographical information mostly collected by the ETER consortium and validated by NSAs.

2.2   A BRIEF HISTORY OF THE STATISTICAL PROCESS AND OUTPUTS IN QUESTION

The ETER data should be considered as complementary to education and R&D statistics collected by EUROSTAT, as well as to the MULTIRANK project. EUROSTAT has provided national (regional) aggregates on educational activities at the tertiary level since 1998 – e.g. the number of students and of degrees – for both EU and some non-EU countries. The ETER data instead cover a subset of tertiary education activities, mostly those institutions that at minimum, graduate students at the bachelor level (ISCED-2011 level 6). By providing institutional-level data, ETER allows for the investigation of the diversity of HEI characteristics and the distribution of functions between HEIs (for example different orientations towards education or research). In terms of the number of students. In 2014, the ETER database includes in the considered countries, 94% of all students at the tertiary levels (ISCED 5-7), with reference to EUROSTAT national-level data.

2.3   THE BROAD STATISTICAL DOMAIN TO WHICH THE OUTPUTS BELONG AND RELATED STATISTICAL OUTPUTS

ETER includes the following main groups of variables:

- Institutional descriptors and geographical information on the HEIs included in the perimeter.
- Data on students and graduates, including breakdowns by ISCED-2011 level, gender, citizenship, mobility and field of education.
- Data on research, including students and graduates at the PhD level, as well as R&D expenditure.
- Financial data: expenditures and revenues.
- Staff data, distinguishing between academic and non-academic staff, including for the former breakdowns by gender, citizenship and field of education.

The ETER database also includes a small number of indicators that have been selected to characterize individual HEIs and their profile of activity.

When compared with the data provided by EUROSTAT about education and R&D statistics, ETER includes very similar variables and breakdowns for students and graduates, since ETER readily adopted the definitions from the UOE manual on education statistics, detailed at the HEI level.

Moreover, the ETER data provide substantial additional information concerning the other dimensions: descriptors are of the outmost importance in order to characterize types of HEIs and their history, while geographical information allows for an analysis of the distribution of HEI activities across the European space. ETER data also provide more detailed information on expenditures and revenues, including a breakdown of revenues by core budget and third-party funds, which is not foreseen in education statistics. Additional data have also been collected concerning staff, including the number of professors and breakdowns by gender and citizenship.

TABLE 2. ETER LIST OF VARIABLES

Dimension	Variables	Format
Identifiers	ETER ID	Text
	National identifier (optional)	Text
	Institution name (in own language)	Text
	English institution name (if available)	Text
	Acronym	Text
	Year	Integer
	Demographic event (past)	Nominal
	Affected HEIs (past)	Text
	Remarks (past)	Text
	Demographic event (future)	Nominal
	Affected HEIs (future)	Text
	Remarks (future)	Text
Basic institutional descriptors	Country Code	ISO code
	Legal status	Nominal
	Institution category, national definition (in own language)	Text
	Institution category, national definition (in English, if available)	Text
	Institution category standardized	
	Foreign campus	Nominal
	Foundation year	Binary
	Legal status year	Integer
	Ancestor year	Integer
	University hospital	Integer
	Institutional website	Nominal
		Text
Geographic information	Region of establishment, NUTS2 code	NUTS code
	Region of establishment, NUTS3 code	NUTS code
	Name of the city	Text
	Postcode	Text
	Multi-site institution	Binary
	Geographical coordinates	Numeric
Educational activities	Highest degree delivered	Nominal
	Lowest degree delivered	Nominal
	Number of enrolled students at ISCED levels 5, 6, 7, by fields of education, gender, citizenship and mobility	Integer
	Total number of students enrolled at ISCED 5-7	Integer
	Number of graduates at ISCED levels 5, 6, 7, by fields of education, gender, citizenship and mobility	Integer
	Total number of graduates at ISCED 5-7	Integer
	Distance education institution	Binary
Research activities	Research active institution	Binary
	Number of enrolled students at ISCED levels 8, by fields of education, gender, citizenship and mobility	Integer
	Number of graduates at ISCED levels 8 (doctorates), by fields of education, gender, citizenship and mobility	Integer

		Numeric
Expenditure	Currrent expenditure	Numeric
	Personnel expenditure	Numeric
	Non-personnel expenditure	Numeric
	Expenditure unclassified	Numeric
	Capital expenditure	Numeric
	Accounting system of capital expenditures	Nominal
	R&D expenditure	Numeric
Revenue	Current revenue	Numeric
	Core funding	Numeric
	Basic government allocation	Numeric
	Other core funding	Nominal
	Third party funding	Numeric
	Public third-party funding	Numeric
	Private third-party funding	Numeric
	Third-party funding from abroad	Numeric
	Third party funding unclassified	Numeric
	Tuition fees	Nominal
	Student fees funding	Nominal
	Revenues unclassified	Nominal
	Non-recurring revenues	Nominal
Staff	Number of academic staff in FTEs and headcounts	Numeric/Integer
	Number of academic staff by fields of education, gender and citizenship in headcounts	Integer
	Number of non academic staff in FTEs and headcounts	Numeric/Integer
	Number of professors by gender	Integer
	Inclusion of PhD students	Binary
	Number of total staff in FTE and HC	Numeric/Integer
Erasmus students	Number of incoming Erasmus students by ISCED level	Integer
	Number of outgoing Erasmus students by ISCED level	Integer

ETER includes also a range of indicators which are mostly related to students, graduates, PhD-students, PhD-graduates and staff in HEIs, variables for which most countries have delivered the data (see Table 3).

TABLE 3: LIST OF INDICATORS IN THE ETER DATABASE BY DIMENSION

Dimension	Variable	Note	N.
Gender in HEIs (share of women)	Students (ISCED 6 and 7)		2
	Graduates (ISCED 6 and 7)		2
	PhD-students		1
	PhD-graduates		1
	Academic staff		1
	Full professors	Included in 2012	1
Nationality in HEIs (share of foreigners)	Students (ISCED 6 and 7)		2
	Graduates (ISCED 6 and 7)		2
	PhD-students		1
	PhD-graduates		1
	Academic staff		1
Mobility in HEIs (Share of mobile individuals)	Students (ISCED 6 and 7)		2
	Graduates (ISCED 6 and 7)		2
	PhD-students		1
	PhD-graduates		1
Subject mix (Herfindahl index)	Students (ISCED5-7)		1
	PhD-graduates		1
	Academic staff	Included in 2012	1
Degree orientation	PhD-intensity		1
Staff	Full professors, as share of academic staff (HC)		1
	Academic staff (FTE), as share of total staff		1
Revenues	Total core budget as share of total current revenues		1

Dimension	Variable	Note	N.
	Total third party funding as share of total current revenues		1
	Tuition fees as share of total current revenues		1

2.4 ROLE OF DATA QUALITY WITHIN THE PROJECT ORGANIZATION

A complex project like ETER, which involves many different partners distributed across EUROPE, requires a well-thought management structure which needs to respect the following main principles:

- A clear allocation of tasks and responsibilities among partners in order to manage activities effectively, avoid duplications and reduce coordination costs.
- Integration and coordination among related tasks as the quality of the output will closely depend to the extent to which methodology, data collection and data quality will integrate smoothly throughout the whole project.
- Good and effective communication between all project partners in order to get a common understanding of the project tasks, to address emerging problems and issues and to motivate partners to provide high-quality work.

These principles were addressed through an organizational structure in three layers, which includes: i) a strongly integrated core team of people coordinating the whole project and taking care of the integration between the tasks they are supervising, ii) clearly defined work packages (located mostly at one consortium partner) where main executing activities will be implemented and iii) national experts in charge of data collection in their country (see Figure 1).
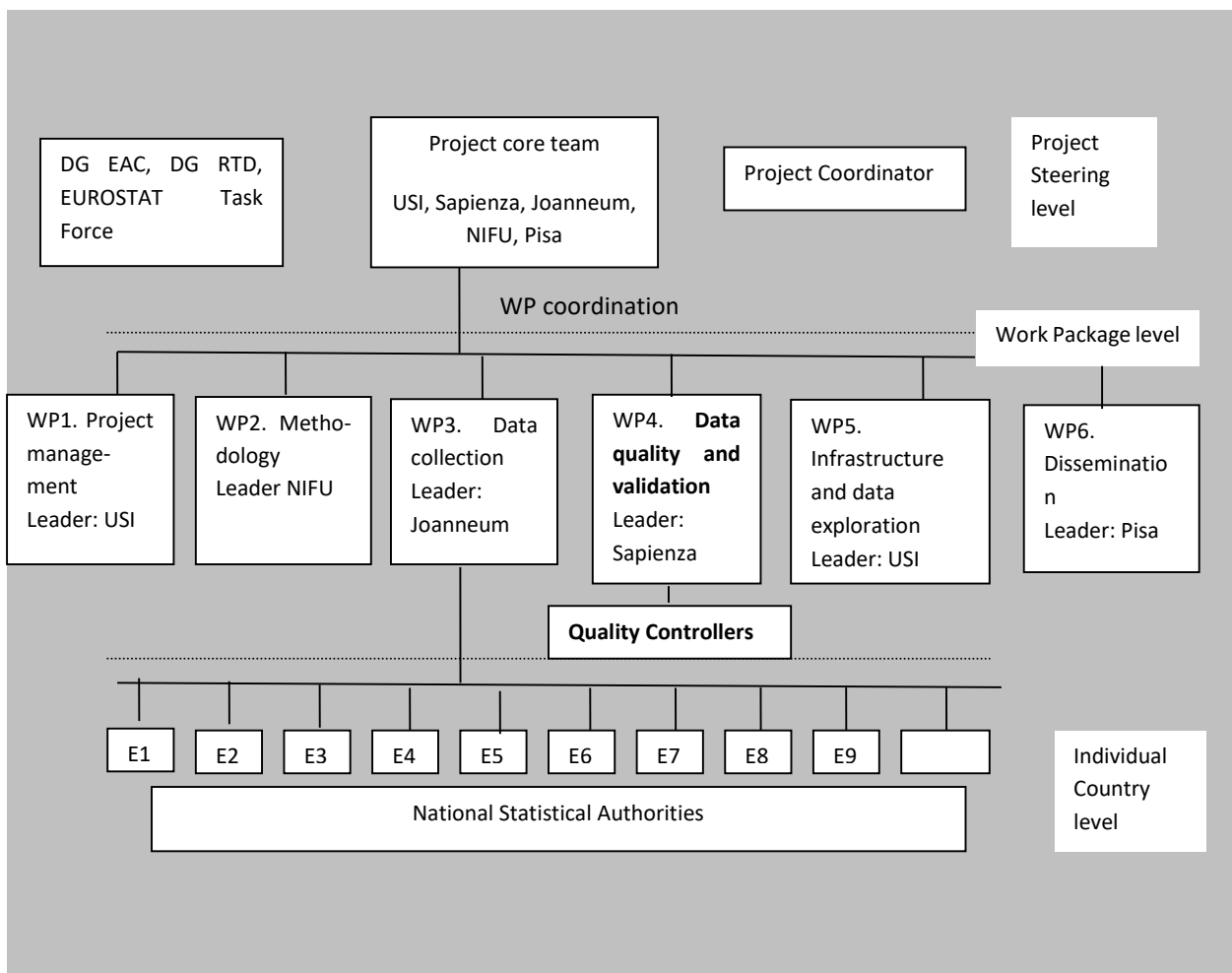
FIGURE 1 ORGANIZATIONAL STRUCTURE OF ETER

National experts (NE) are dedicated persons who manage the contacts to each individual country, finding the right contact points and solving practical problems with NSAs. National experts might also help for some parts of data collection, particularly for descriptors and geographic information. Most national experts are actually part of the core team members' organization ,only for a few countries, where particular issues were expected, ad hoc national experts have been nominated (for example because of their knowledge of the national language and higher education system).

National Statistical Authorities (NSAs) are the main source of data for ETER, with the partial exception of descriptors and geographic information. NSAs fill in the data in the ETER data collection sheets (in Excel), provide most of the metadata, as well as consent on the use of data, since they are the primary owners of most data.

2.5 OVERVIEW OF THE ETER DATA QUALITY METHODOLOGY

Data quality is a relevant issue in any data collection, but often underestimated in practical data collection procedures.

ETER data collection is a secondary data collection, meaning that data are not collected as primary data directly from respondents, but results from an elaboration done by NSAs starting from the data sources they in turn gathered.

As reported in the Handbook on Methodology for the Modern Business Statistics, with reference to secondary data collections (refs. https://ec.europa.eu/eurostat/cros/content/collection-and-use-secondary-data-theme_en):

"In order to be in a position to use data from secondary sources… the 'fitness for use' of the data source for official statistics needs to be determined. There are many ways to determine this. The most important approaches focus on the **metadata quality of the source**, *on the* **data quality of the input data**, *and on the* **data quality of the statistics produced***."*

Figure shows the secondary nature of ETER data collection. In particular, NSAs access to data collected either by means of surveys or as administrative data. As known, surveys do have the possibility of having quality metadata associated but these are not under ETER control, meaning that no "design-time" choice can be done at that level in order to measure quality of questionnaires answers. In addition, administrative data are by themselves secondary data sources, hence requiring a specific data quality evaluation.
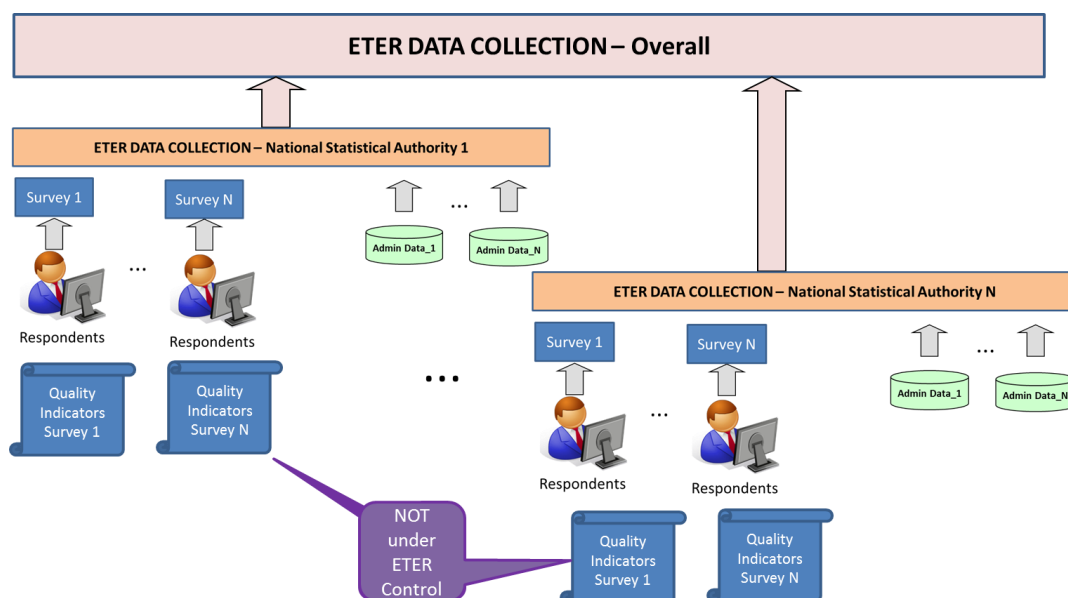
FIGURE 2. OVERALL ETER DATA COLLECTION

By taking into account these characteristics, the ETER Data Quality Methodology has been designed and implemented as an articulated process that spans from the very first stages of the collection, of which NSAs are in charge, to global checks and corrections done when the overall data collection is finished.
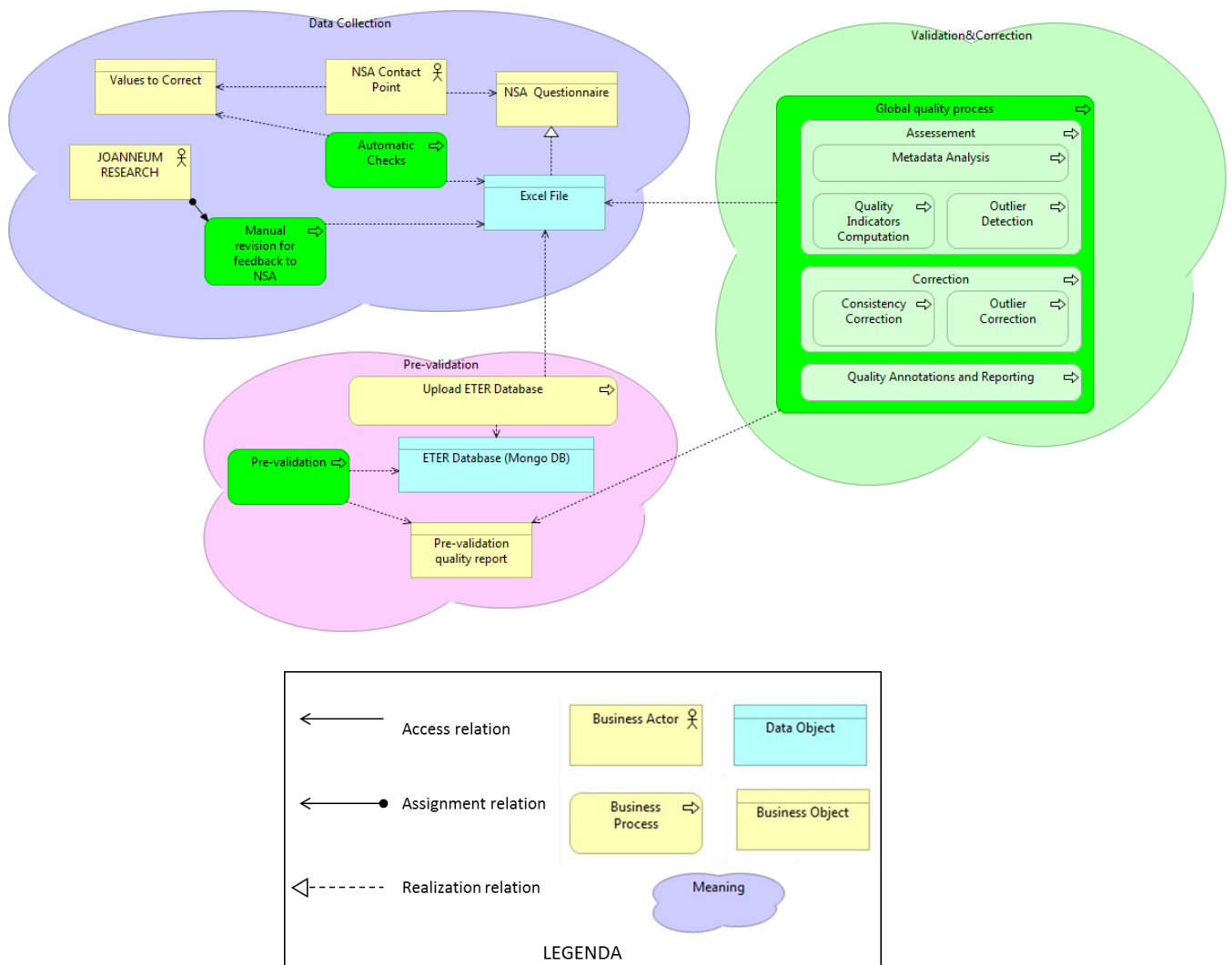
FIGURE 3: OVERALL ETER DATA QUALITY PROCESS

The overall process is described in Figure 3, where the quality activities are colored as green. Three main stages are highlighted: "data collection", "pre-validation" and "validation&correction".

In the early phases of data collection, NSAs collect questionnaires as Excel files. Some automatic checks have been implemented as automatic rules within such files, i.e. checks of (i) incomplete data and blank cells, (ii) format accuracy compliance, (iii) mistakes in sums and inconsistencies between variables. This step is represented in Figure 3 by the "Automatic Checks" process block. In addition, on the collected files, Johanneum Research performs a manual revision so that, if some feedbacks to improve the collected data are identified, these are conveyed to NSAs for following-up related corrections.

Later, a pre-validation of collected questionnaires is performed. As long as the questionnaires are provided by the countries, they are uploaded to a database that performs a pre-validation of data. Some indicators such as accuracy, completeness and consistency are computed country-by-country. Indeed, this phase is characterized for being *local-to-countries,* while the subsequent one is, instead *global to the whole data collection*. This stage produces a "Pre-validation quality report" provided as input to the subsequent and last stage.

The last stage is the final validation and correction stage. It consists of three main phases:

- a *quality assessment phase*, described in Section 2.5.1, with the main purpose of characterizing the "fitness for use" of ETER collected data;
- a *quality correction phase*, detailed in 2.5.2, that performs some correction actions when possible;
- a *quality reporting phase*, described in Section 2.5.3, with the purpose of documenting the results of the previous phases in order to enable quality-aware analyses on ETER data.

2.5.1 QUALITY ASSESSMENT

The quality assessment phase includes the following steps:

- Quality Metadata analysis. Together with the data collection, NSAs also collect a specific set of quality metadata. These include for instance the reference dates for the collected data that are relevant for the evaluation of data timeliness and for the data comparability.
- Quality Indicators Computation. After the described pre-validation phase, an in-depth quality assessment activity is performed on collected data in this step. Purposefully defined "Quality Indicators" are computed; these indicators have been defined for four data quality dimensions, namely: *format accuracy*, *completeness*, *consistency*, and *timeliness*.
- Outlier detection, consisting of a global data analysis performed according to advanced statistical methods aimed at identifying those outliers that can be labelled as erroneous data;
  - o Cross-section outliers detection: aimed at identifying outliers within a given data collection by comparing answers given by "similar" HEIs.
  - o Multi-annual outliers detection: aimed at identifying outliers across different years (4 years available altogether).

2.5.2 CORRECTION

The correction phase at this stage can be done on two main classes of identified errors:
- *Outliers* that emerge from the previous phase and that are (manually) labelled as erroneous.
- *Inconsistencies* that derive from the computed quality indicators and that can be reformulated in terms of deterministically defined corrections.

2.5.3 QUALITY ANNOTATIONS AND REPORTING.

One of the results of the data quality process is a documentation of the quality evaluation of the data set through the provision of specific flags and notations accompanying the data. Flags signal problems or specificities of data related to format accuracy, consistency, completeness and comparability. Flags can be attributed to:

- individual cells.
- one dimension or group of variables in a country (i.e. all variables concerning revenues).
- all variables for one or more HEIs in a specific country (i.e. all private HEIs).

The ETER flag system builds on a simplified and reduced version of the one adopted by EUROSTAT, but introduces a few additional codes for cases relevant at the level of individual HEIs.

To avoid blank cells in the database which may create ambiguity, *special codes* ("a", "m", "x", "xc", "xr", "nc", "c", "s") have been introduced, following the standard conventions used by Eurostat. Their description is reported in Table 4. Moreover, to warn the users about data and comparability issues *data flags* are introduced in the dataset, in separate columns alongside the data columns. A data flag "accompanies" an existing number and acts as a warning or an explanation. When relevant, a short explanation is included in a corresponding "Notes" column of the data set. This allows for a quick identification of the reason of the flag. Table 5 describes the flags.

TABLE 4. SPECIAL CODES FOR MISSING VALUES

Code	Description	Definition
m	Missing	The data is not available.
a	Not applicable	This variable is not applicable for the specific case. For example, number of PhD students in non-doctorate awarding HEIs is coded in this way.
x	Included in totals	The data are not available, but their values is included in the total. This applies for example when the data on total income is available but the distinction between core budget, third party funding and student fees funding is not..
xc	Included in another variable	The data are not available, but are included in the value of another variable. This applies for example when educational fields cannot be split.
xr	Included in another row	The missing data are actually included in another row. This happens when, for instance, there is an on-going merger process).
c	Confidential	When data are available, but restricted to public access.
s	Below threshold	This code is displayed in the public dataset when the count of a cell is so low that data protection issues might arise as individuals can be identified (for example number of students below 3).

TABLE 5: LIST OF FLAGS

Code	Description	Definition
b	break in time series	When changes in definitions or data collection procedures imply that the data are not comparable across years.
de	break in time series due to a demographic event	When data are not comparable across years because of demographic events (e.g. mergers, take-overs).
d	definition differs	Differences in definitions adopted for data collection imply that figures significantly differs from those complying with the ETER methodology and thus are not comparable across countries.
i	see metadata	There are specific conditions which imply that the value of a cell should be interpreted in a different way or not directly compared with others.
ic	inconsistent	Either when sum of break down differs from total or if another semantic rule is violated.
rd	rounded	When data have been rounded by the data provider and thus are included in this format in the database.
c	confidential	When data are available, but restricted to public access (this flag is relevant only for user with unrestricted access).
ms	missing subcategory	This flag is applied to totals in order to warn users that the total does not include one relevant subcategory (for example total expenditures not including capital expenditures).
p	provisional	This flag is used, if data have to be controlled and possibly revised at a later stage.

Summing up, flags are introduced as a result of the following steps:

- Computation of the data quality indicators.
- Analysis of the metadata in particular to identify important cases of departures from definitions, which should be notified to the users in order to enable *quality-aware* analyses.
- Pre-validation and quality assessment concerning deviant cases and inconsistencies.

2.6 REFERENCES TO OTHER DOCUMENTATION, ESPECIALLY ON METHODOLOGY.

The Documents produced in the ETER project are the following:

1. Project Handbook that reports all the details related to the methodology applied for the data collection. Available on the project website (https://www.eter-project.com/about/documents);

2. Final Report, available on the project website (https://www.eter-project.com/about/documents).

3 QUALITY REPORT: RELEVANCE, ASSESSMENT OF USER NEEDS AND PERCEPTIONS

3.1 USERS NEEDS AND SATISFACTION

International statistical agencies like EUROSTAT and OECD, as well as National Statistical Authorities, already perform data collections on higher education institutions based on existing methodological manuals. These include particularly:

- UNESCO-UIS/OECD/EUROSTAT data collection on education statistics (UOE): provides internationally comparable data on key aspects of education systems, specifically on the participation and completion of education programmes, as well as the cost and type of resources dedicated to education. Thus, UOE provides for data collection concerning students, degrees, education staff, finances and education expenditures; data are disaggregated by level of education (using the International Standard Classification of Education, ISCED) and by fields of education. The UOE manual is largely the reference also for ETER data collection and whenever possible definitions and classifications are adopted; however, as it shall be clear later, there are also differences which are related to the fact that the basic unit in UOE are educational programs, while in ETER are higher education institutions. Data collection is managed in most countries by national statistical institutes (NSI), which deliver summary tables of national aggregates to EUROSTAT.
- OECD/EUROSTAT joint data collection on research and development (R&D) statistics: provides data on R&D expenditure and R&D staff, based on the definitions, classifications and procedures for collecting R&D data provided in the OECD Frascati manual (OECD 2002) and, for EU Member States, following the requirements of Community Regulation CR/753/2004. R&D statistics is based on regular surveys of research performers, even if in some countries higher education R&D expenditures are derived from higher education statistics. The Frascati manual is relevant when looking to research expenditures of higher education institutions.

The basic approach of ETER is that, whenever the data collected through these methodologies reasonably correspond to the ETER framework, they are used with as less as possible changes. This concerns most specifically data on students, graduates, expenditures and revenues of HEIs. However, in order to meet users' needs and higher accuracy of data in individual cases, some slight departures have been decided to take into account some specific issues which emerged from the EUMIDA study: for instance, ISCED 7 long degrees (masters without an intermediate bachelor) will be singled out; nominal variables have been introduced to characterize how capital expenditures are accounted, etc.

Further, ETER introduces in two areas a number of variables not present in current statistical data (as they lack a definition of HEIs as such):

- First, non-statistical descriptors which characterize the legal status and the history of HEIs, like the name, the institutional category, the legal status, the foundation year, the presence of a university hospital, the institutional website.
- Second, geographical information like region of establishment, city, geographical coordinates. This allows interlinking with context data (like those from regional statistics), as well as to take into account the impact of geography (for example distance between HEIs).

Finally, in the specific case of staff data, it is considered that existing definitions in education statistics do not specifically consider HEI staff structures and issues like the specific status of PhD students – constituting a large part of researcher's workforce in some countries- as well as the status and role of permanent professors among the academic staff. In this area, ETER introduces a set of ad hoc definitions and performs a specific metadata collection in order to come to a robust methodology for the analysis of HEI staff.

While being largely a data collection study, a major goal of ETER is to lead to a further development of statistical foundations concerning Higher Education Institutions, taking into account their deep differences with other levels of schools.

3.2 COMPLETENESS

For each variable, dimension, and dataset, *completeness* evaluates the share of missing values (with the meaning relevant to completeness, i.e. unavailable or temporarily unavailable) which are present. In the following, completeness results are reported for all ETER data, with details by variable and by country.

The ETER dataset is fairly complete, with an overall completeness index ranging from 73% in 2011 and 2012 to 76% in 2013 and 78% in 2014. This indicator is computed excluding cases not applicable ("a" in the dataset, meaning that the variable is not relevant for that specific HEI).

Table 6 reports some statistics about missing values by typology.

TABLE 6: STATISTICS ON MISSING VALUES BY TYPOLOGY

Year	2011	2012	2013	2014
m	0.656	0.669	0.469	0.528
x	0.244	0.221	0.371	0.385
xc	0.087	0.095	0.138	0.067
xr	0.004	0.006	0.013	0.012
nc	0.001	0.001	0.002	0.000
c	0.008	0.008	0.007	0.007
Total count	**255711**	**259074**	**227147**	**217823**

The level of completeness largely varies by domain and variable as shown by Table 7. It is higher for data on students and graduates, although some breakdown by field and by mobility can be more problematic. However, the introduction of new standard information requested by Eurostat on this topic will improve the availability and comparability of mobility information in the coming years. Completeness is lower for financial data on income and expenditure (51% for totals and below 50% for financial categories breakdown). The lack of this information is due to the absence of standardized collection procedures at the national level for some countries, while the basic information is based on administrative compulsory data. An effort in this direction could significantly improve the situation. R&D expenditure is available in less than 40% of cases. Data on staff are in an intermediate position in the range 60%-70%, except for academic staff breakdowns. Data in FTE are more complete than data in HC, and data on academic staff more than technical and administrative staff, and even more complete than total staff.

The availability of descriptors is fairly complete, with the noticeable exception of information about the foundation years, especially legal status year (57%) and ancestor year (47%). The two latter descriptors were not mandatory in the data collection.

TABLE 7: COMPLETENESS BY VARIABLE AND REFERENCE YEAR

Variable code	Variable name	2011	2012	2013	2014
BAS.NATID	National identifier	1.00	1.00	1.00	1.00
BAS.ACRONYM	Institution Acronym	0.55	0.56	0.54	0.56
BAS.LEGALSTAT	Legal status	0.99	1.00	1.00	1.00
BAS.INSTCATSTAND	Institution Category standardized	0.99	1.00	1.00	1.00
BAS.FORCAMP	Foreign Campus	0.88	0.89	0.88	0.88
BAS.FOUNDYEAR	Foundation year	0.98	0.99	0.99	0.99
BAS.LEGALYEAR	Legal status year	0.58	0.58	0.58	0.65
BAS.ANCESTYEAR	Ancestor year	0.33	0.33	0.32	0.37
BAS.UNIHOSP	University hospital	0.99	1.00	0.98	0.98
BAS.WEBSITE	Institutional website	0.99	1.00	1.00	0.99
GEO.NUTS	Region of establishment (NUTS2-NUTS3)	0.99	1.00	1.00	1.00

Variable code	Variable name	2011	2012	2013	2014
GEO.CITY	Name of the city	0.99	1.00	1.00	1.00
GEO.POSTCODE	Postcode	0.99	1.00	1.00	1.00
GEO.MULTISITE	Multi-site institution	0.97	0.98	0.98	0.98
GEO.NUTS3MULTISITE	NUTS 3 codes of other campuses	0.89	0.92	0.93	0.93
EXP.CURRPERSON.NC	Personnel expenditure	0.44	0.52	0.58	0.57
EXP.CURRNONPERSON.NC	Non-personnel expenditure	0.44	0.48	0.54	0.57
EXP.CURRUNCL.NC	Expenditure unclassified	0.44	0.52	0.58	0.57
EXP.CURRTOTAL.NC	Total Current expenditure	0.48	0.52	0.58	0.57
EXP.CAPITAL.NC	Capital expenditure	0.39	0.43	0.49	0.49
EXP.ACCSYSTEM	Accounting system of capital expenditure	0.27	0.41	0.51	0.51
REV.COREBUDGETPUBLIC.NC	Basic government allocation	0.12	0.12	0.14	0.31
REV.COREBUDGETOTHER.NC	Other core budget	0.12	0.12	0.14	0.27
REV.CORETOTAL.NC	Total core budget	0.47	0.51	0.53	0.50
REV.THIRDPARTYPUBLIC.NC	Public third party funding	0.37	0.40	0.41	0.40
REV.THIRDPARTYPRIVATE.NC	Private third party funding	0.36	0.40	0.41	0.40
REV.THIRDPARTYABROAD.NC	Third party funding from abroad	0.12	0.12	0.14	0.41
REV.THIRDPARTYUND.NC	Third party funding undivided	0.12	0.12	0.14	0.42
REV.THIRDPARTYTOTAL.NC	Total third party funding	0.39	0.42	0.45	0.45
REV.TUITFEES	Tuition fees	0.42	0.45	0.54	0.53
REV.STUDFEES.NC	Student fees funding	0.39	0.45	0.48	0.44
REV.UNCL.NC	Revenue unclassified	0.45	0.50	0.51	0.49
REV.CURRTOTAL.NC	Total Current revenues	0.48	0.52	0.54	0.51
REV.NONRECURR.NC	Non-recurring revenues	0.03	0.03	0.03	0.19
STA.ACAFTETOTAL	Total academic staff (FTE)	0.69	0.66	0.70	0.70
STA.ACAHCMEN	Academic staff by sex	0.61	0.59	0.69	0.74
STA.ACAHCNAT	Academic staff by citizenship	0.38	0.38	0.51	0.55
STA.ACAHCFOE	Academic staff by FoE	0.27	0.29	0.30	0.35
STA.TOTACAHC	Total academic staff (HC)	0.59	0.57	0.67	0.72
STA.PROFMEN	Number of full professors by sex	0.64	0.62	0.74	0.83
STA.PROFTOTAL	Total number of full professors	0.68	0.66	0.76	0.83
STA.INCLPHDSTUD	Inclusion of PhD students	0.74	0.73	0.81	0.87
STA.NONACAFTE	Number of non-academic staff (FTE)	0.57	0.61	0.64	0.65
STA.NONACAHC	Number of non-academic staff (HC)	0.53	0.51	0.54	0.54
STA.TOTALFTE	Total staff (FTE)	0.68	0.65	0.68	0.68
STA.TOTALHC	Total staff (HC)	0.57	0.56	0.59	0.58
STUD.LOWDEG	Lowest degree delivered	0.89	0.87	0.97	0.97
STUD.HIGHDEG	Highest degree delivered	0.89	0.87	0.97	0.97
STUD.ISCED5MEN	Students enrolled at ISCED 5 by sex	0.65	0.59	0.88	0.87
STUD.ISCED5NAT	Students enrolled at ISCED 5 by citizenship	0.54	0.49	0.59	0.54
STUD.ISCED5RES	Students enrolled at ISCED 5 by mobility	0.59	0.54	0.55	0.52
STUD.ISCED5FOE	Students enrolled at ISCED 5 by FoE	0.57	0.51	0.57	0.56
STUD.ISCED5TOTAL	Total students enrolled at ISCED 5	0.65	0.59	0.88	0.87
STUD.ISCED6MEN	Students enrolled at ISCED 6 by sex	0.88	0.85	0.95	0.94
STUD.ISCED6NAT	Students enrolled at ISCED 6 by citizenship	0.82	0.80	0.84	0.81
STUD.ISCED6RES	Students enrolled at ISCED 6 by mobility	0.52	0.52	0.54	0.52
STUD.ISCED6FOE	Students enrolled at ISCED 6 by FoE	0.79	0.79	0.83	0.83
STUD.ISCED6TOTAL	Total students enrolled at ISCED 6	0.86	0.84	0.95	0.94
STUD.ISCED7MEN	Students enrolled at ISCED 7 by sex	0.86	0.84	0.95	0.95
STUD.ISCED7NAT	Students enrolled at ISCED 7 by citizenship	0.80	0.78	0.83	0.82
STUD.ISCED7RES	Students enrolled at ISCED 7 by mobility	0.53	0.53	0.55	0.53
STUD.ISCED7FOE	Students enrolled at ISCED 7 by FoE	0.80	0.80	0.83	0.84
STUD.ISCED7TOTAL	Total students enrolled at ISCED 7	0.86	0.84	0.96	0.95

Variable code	Variable name	2011	2012	2013	2014
STUD.ISCED7LONGMEN	Students enrolled ISCED 7 long degree by sex	0.80	0.77	0.85	0.85
STUD.ISCED7LONGNAT	Students enrolled ISCED 7 long degree by citizenship	0.80	0.77	0.84	0.84
STUD.ISCED7LONGRES	Students enrolled ISCED 7 long degree by mobility	0.56	0.53	0.59	0.56
STUD.ISCED7LONGFOE	Students enrolled ISCED 7 long degree dy FoE	0.74	0.73	0.80	0.82
STUD.ISCED7LONGTOTAL	Total students enrolled ISCED 7 long degree	0.80	0.77	0.85	0.85
STUD.TOTALISCED5.7	Total students enrolled ISCED 5-7	0.91	0.89	0.98	0.97
STUD.DISTEDUINST	Distance education institution	0.90	0.90	1.00	0.99
GRAD.ISCED5MEN	Graduates at ISCED 5 by sex	0.65	0.59	0.88	0.86
GRAD.ISCED5NAT	Graduates at ISCED 5 by citizenship	0.53	0.48	0.60	0.52
GRAD.ISCED5RES	Graduates at ISCED 5 by mobility	0.55	0.53	0.57	0.51
GRAD.ISCED5FOE	Graduates at ISCED 5 by FoE	0.56	0.51	0.57	0.64
GRAD.ISCED5TOTAL	Total graduates at ISCED 5	0.65	0.59	0.88	0.86
GRAD.ISCED6MEN	Graduates at ISCED 6 by sex	0.87	0.84	0.94	0.91
GRAD.ISCED6NAT	Graduates at ISCED 6 by citizenship	0.78	0.75	0.79	0.78
GRAD.ISCED6RES	Graduates at ISCED 6 by mobility	0.49	0.50	0.52	0.49
GRAD.ISCED6FOE	Graduates at ISCED 6 by FoE	0.78	0.78	0.82	0.80
GRAD.ISCED6TOTAL	Total graduates at ISCED 6	0.85	0.83	0.94	0.91
GRAD.ISCED7MEN	Graduates at ISCED 7 by sex	0.85	0.82	0.94	0.90
GRAD.ISCED7NAT	Graduates at ISCED 7 by citizenship	0.77	0.75	0.79	0.77
GRAD.ISCED7RES	Graduates at ISCED 7 by mobility	0.50	0.51	0.54	0.51
GRAD.ISCED7FOE	Graduates at ISCED 7 by FoE	0.79	0.78	0.81	0.79
GRAD.ISCED7TOTAL	Total graduates at ISCED 7	0.85	0.82	0.94	0.90
GRAD.ISCED7LONGMEN	Graduates at ISCED 7 long degree by sex	0.79	0.75	0.84	0.84
GRAD.ISCED7LONGNAT	Graduates at ISCED 7 long degree by citizenship	0.78	0.75	0.82	0.83
GRAD.ISCED7LONGRES	Graduates at ISCED 7 long degree by mobility	0.52	0.49	0.57	0.53
GRAD.ISCED7LONGFOE	Graduates at ISCED 7 long degree by FoE	0.73	0.71	0.79	0.81
GRAD.ISCED7LONGTOTAL	Total graduates at ISCED 7 long degree	0.79	0.75	0.84	0.84
GRAD.TOTALISCED5.7	Total graduates ISCED 5-7	0.87	0.87	0.97	0.94
RES.RESACTIVE	Research active institution	0.85	0.87	0.97	0.97
RES.STUDISCED8MEN	Students enrolled at ISCED 8 by sex	0.73	0.70	0.87	0.92
RES.STUDISCED8NAT	Students enrolled at ISCED 8 by citizenship	0.66	0.64	0.68	0.73
RES.STUDISCED8RES	Students enrolled at ISCED 8 by mobility	0.48	0.46	0.49	0.48
RES.STUDISCED8FOE	Students enrolled at ISCED 8 by FoE	0.58	0.58	0.56	0.62
RES.STUDISCED8TOTAL	Total students enrolled at ISCED 8	0.73	0.70	0.87	0.92
RES.GRADISCED8MEN	Graduates at ISCED 8 by sex	0.76	0.65	0.83	0.74
RES.GRADISCED8NAT	Graduates at ISCED 8 by citizenship	0.73	0.62	0.67	0.59
RES.GRADISCED8RES	Graduates at ISCED 8 by mobility	0.42	0.42	0.46	0.36
RES.GRADISCED8FOE	Graduates at ISCED 8 by FoE	0.63	0.62	0.61	0.46
RES.GRADISCED8TOTAL	Total graduates at ISCED 8	0.77	0.66	0.84	0.75
RES.R.DEXP.NC	R&D Expenditure	0.18	0.16	0.17	0.15

The level of completeness largely varies by country. In the analysed dataset there is a group of countries with a very high level of completeness (over 85%) including BE, BG, CH, CY, DE, ES, IE, IT, LI, LU, MT, PT, SE, a second group with an high level (50%-85%) including AT, CZ, DK, EE, FI, FR, GR, HR, HU, IS, LT, LV, NL, NO, PL, SK and a third group with a fair data availability (below 50%) including MK, UK.

TABLE 8: COMPLETENESS BY COUNTRY AND REFERENCE YEAR

Country	N.of HEIs	Min	Mean	Max	Std
AT 2011	68	0.54	0.55	0.58	0.010
AT 2012	68	0.50	0.54	0.56	0.008
AT 2013	68	0.54	0.55	0.57	0.009
AT 2014	68	0.56	0.58	0.61	0.012
BE 2011	31	0.30	0.80	0.89	0.168
BE 2012	31	0.30	0.79	0.89	0.166
BE 2013	26	0.28	0.79	0.90	0.185
BE 2014	26	0.30	0.84	0.96	0.204
BG 2011	52	0.49	0.76	0.79	0.052
BG 2012	52	0.49	0.76	0.79	0.052
BG 2013	52	0.49	0.76	0.78	0.053
BG 2014	52	0.49	0.76	0.79	0.053
CH 2011	33	0.84	0.98	0.99	0.035
CH 2012	35	0.84	0.98	0.99	0.034
CH 2013	35	0.85	0.98	0.99	0.034
CH 2014	35	0.85	0.98	0.99	0.040
CY 2011	25	0.51	0.85	0.97	0.085
CY 2012	25	0.52	0.84	0.94	0.078
CY 2013	25	0.45	0.84	1.00	0.115
CY 2014	23	0.77	0.86	1.00	0.049
CZ 2011	72	0.35	0.57	0.70	0.053
CZ 2012	71	0.66	0.72	0.78	0.044
CZ 2013	70	0.43	0.73	0.82	0.059
CZ 2014	71	0.73	0.78	0.82	0.036
DE 2011	374	0.50	0.93	0.95	0.052
DE 2012	386	0.41	0.93	0.95	0.062
DE 2013	390	0.41	0.93	0.95	0.061
DE 2014	396	0.34	0.94	0.96	0.081
DK 2011	34	0.27	0.79	0.91	0.167
DK 2012	34	0.27	0.79	0.92	0.166
DK 2013	33	0.35	0.84	0.93	0.132
DK 2014	33	0.30	0.42	0.53	0.070
EE 2011	30	0.69	0.74	0.80	0.047
EE 2012	29	0.69	0.74	0.79	0.045
EE 2013	26	0.69	0.74	0.79	0.045
EE 2014	25	0.69	0.74	0.79	0.045
ES 2011	77	0.75	0.79	0.79	0.013
ES 2012	80	0.71	0.79	0.79	0.020
ES 2013	80	0.75	0.79	0.80	0.012
ES 2014	81	0.74	0.80	0.81	0.014
FI 2011	44	0.53	0.65	0.67	0.031
FI 2012	44	0.63	0.70	0.72	0.021
FI 2013	42	0.76	0.83	0.84	0.018
FI 2014	41	0.79	0.85	0.87	0.020

Country	N.of HEIs	Min	Mean	Max	Std
FR 2011	285	0.24	0.62	0.80	0.190
FR 2012	286	0.25	0.66	0.89	0.218
FR 2013	316	0.24	0.66	0.89	0.207
FR 2014	323	0.25	0.66	0.95	0.216
GR 2011	50	0.39	0.68	0.73	0.062
GR 2012	50	0.40	0.67	0.72	0.066
GR 2013	47	0.39	0.70	0.75	0.067
GR 2014	49	0.39	0.56	0.63	0.046
HR 2011	32	0.35	0.43	0.53	0.064
HR 2012	33	0.40	0.53	0.67	0.085
HR 2013	36	0.54	0.65	0.72	0.036
HR 2014	37	0.75	0.87	0.92	0.051
HU 2011	52	0.79	0.88	0.89	0.016
HU 2012	52	0.31	0.31	0.31	0.001
HU 2013	52	0.75	0.83	0.88	0.022
HU 2014	52	0.87	0.87	0.88	0.002
IE 2011	27	0.76	0.85	0.86	0.019
IE 2012	27	0.76	0.85	0.86	0.024
IE 2013	27	0.76	0.85	0.86	0.023
IE 2014	27	0.80	0.91	0.93	0.037
IS 2011	7	0.50	0.52	0.53	0.012
IS 2012	7	0.42	0.46	0.47	0.019
IS 2013	7	0.47	0.49	0.50	0.011
IS 2014	7	0.50	0.52	0.55	0.018
IT 2011	176	0.39	0.81	0.90	0.078
IT 2012	176	0.44	0.82	0.91	0.075
IT 2013	176	0.59	0.80	0.88	0.042
IT 2014	215	0.71	0.80	0.92	0.044
LI 2011	1	0.96	0.96	0.96	-
LI 2012	1	0.96	0.96	0.96	-
LI 2013	1	0.96	0.96	0.96	-
LI 2014	1	0.96	0.96	0.96	-
LT 2011	43	0.36	0.67	0.77	0.104
LT 2012	43	0.43	0.67	0.77	0.098
LT 2013	44	0.18	0.83	0.93	0.198
LT 2014	43	0.17	0.89	0.97	0.211
LU 2011	1	0.89	0.89	0.89	-
LU 2012	1	0.30	0.30	0.30	-
LU 2013	1	0.30	0.30	0.30	-
LU 2014	1	0.30	0.30	0.30	-
LV 2011	49	0.13	0.63	0.68	0.075
LV 2012	49	0.13	0.64	0.70	0.076
LV 2013	46	0.13	0.66	0.71	0.087
LV 2014	45	0.14	0.80	0.87	0.133
MK 2011	10	0.44	0.63	0.72	0.081

Country	N.of HEIs	Min	Mean	Max	Std
MK 2012	10	0.39	0.67	0.75	0.112
MK 2013	10	0.41	0.65	0.75	0.105
MK 2014	10	0.39	0.68	0.75	0.113
MT 2011	1	0.94	0.94	0.94	-
MT 2012	1	0.90	0.90	0.90	-
MT 2013	2	0.88	0.88	0.89	0.007
MT 2014	2	0.91	0.93	0.95	0.023
NL 2011	57	0.56	0.75	0.81	0.075
NL 2012	55	0.58	0.75	0.81	0.075
NL 2013	56	0.56	0.77	0.82	0.068
NL 2014	56	0.46	0.79	0.84	0.079
NO 2011	50	0.30	0.84	0.89	0.104
NO 2012	50	0.33	0.84	0.88	0.081
NO 2013	50	0.33	0.85	0.89	0.079
NO 2014	48	0.41	0.89	0.94	0.078
PL 2011	286	0.57	0.70	0.76	0.025
PL 2012	272	0.41	0.70	0.76	0.038
PL 2013	280	0.54	0.70	0.78	0.032
PL 2014	281	0.58	0.74	0.82	0.037
PT 2011	113	0.72	0.82	0.94	0.070
PT 2012	106	0.73	0.81	0.93	0.077
PT 2013	94	0.73	0.82	0.93	0.077
PT 2014	91	0.75	0.85	0.96	0.084
RS 2011	17	0.26	0.26	0.26	0.000
RS 2012	17	0.30	0.30	0.30	0.001
RS 2013	17	0.38	0.55	0.62	0.070
RS 2014	17	0.38	0.54	0.59	0.064
SE 2011	39	0.89	0.89	0.90	0.002
SE 2012	39	0.92	0.92	0.92	0.002
SE 2013	40	0.45	0.91	0.96	0.076
SE 2014	37	0.97	1.00	1.00	0.005
SK 2011	32	0.70	0.79	0.83	0.061
SK 2012	32	0.70	0.79	0.84	0.063
SK 2013	32	0.70	0.82	0.86	0.069
SK 2014	32	0.70	0.84	0.90	0.085
TR 2011	177	0.30	0.30	0.31	0.002
TR 2012	181	0.30	0.30	0.31	0.002
TR 2013	188	0.31	0.45	0.56	0.038
TR 2014	189	0.31	0.44	0.55	0.035
UK 2011	162	0.86	0.98	0.99	0.012
UK 2012	161	0.32	0.98	0.99	0.053
UK 2013	160	0.69	0.98	0.99	0.025
UK 2014	161	0.62	0.99	1.00	0.031

Table A2 in Appendix 2 reports additional details about completeness by country and by variable.

4 QUALITY REPORT: FORMAT ACCURACY

Format accuracy evaluates the compliance of the value to the requested format, as defined in the data chapter of the ETER handbook, respectively in the definitions of each variable. This includes characteristics like being non-negative for all quantitative values, student and graduate data being integer variables, and so on (see ETER Handobook for details).

For a dataset like ETER, a central dimension of accuracy consists of monitoring the adherence to the rules concerning missing values, their coding, and the correct distinction between "0" values and missing values (respectively not applicable values).

This control has been performed systematically during the preliminary validation phase. After that, it has been performed again on the overall dataset. The identified deviant cases have been either directly corrected or reported back to the national experts for checking and correction. Only a few cases of format inaccuracy are remaining in the dataset, namely:

- cases where expenditures or revenues are negative. These comprise cases depending on national accounting rules (often reporting correction of previous data as negative variations) or on imperfect matching with ETER categories. The largest number of cases concerns core budget in Germany. All data are flagged;
- the number of students and graduates in AT are not integer since students enrolled in multiple types of programmes are pro-rated. The number of ISCED8 students in NL is not integer as well, since figures reported are in FTE; this impact also data on academic staff in headcount, which include PhD students and is not integer;
- in the metadata, the date of source release and reference period/date are reported in a variety of formats, even because often they are only a proxy of the exact date.

It is important to underline that the problems found in quantitative variables are present in national official data, and do not derive from the ETER collection process itself.

5 TIMELINESS

For each variable, *timeliness* evaluates the time lag between the ETER Collection publication date and the Source Release date. Ideally, it should be envisaged to reduce as much as possible this lag.

Information on the source release date at national level have been provided by most of the countries, but not for all; considerations below are derived from available information. As it could be expected the situation is diversified among countries and within them among different domains. Only in AT, BG, GR, LV, PT, SE all quantitative data were released at the same date, but it has to be noticed that for these countries information on financial variables are not provided. Differences in release dates among countries are wide, so that a trade-off should be defined when operating a European collection.

Published in June each year, the ETER data collection has an average lag of 12-18 months with respect to individual national releases. The choice of a more recent reference year (i.e. 2012 instead of 2011 for the 2014 data collection) to reduce the lag would have implied the exclusion of a number of countries because of temporary unavailability of data. It could be possible to reduce the time lag of several months changing the schedule of future waves of data collection, and/or trying to reduce the gap between the begin of ETER data collection and the publication of data (9 months now). This largely depends on responsiveness of NSAs and European Commission schedule.

TABLE 9: DATE OF DATA PUBLICATION

ETER wave of data collection	Reference year	Date of data publication
2014	2011 – A.Y. 2011/2012	02/07/2014
2015	2012 – A.Y. 2012/2013	03/07/2015
2016	2013 – A.Y. 2013/2014	20/06/2016
2017	2014 – A.Y. 2014/2015	19/06/2017

The table below provides insights on the average period of source release at national level for each ETER data dimension and country. The table contains the month of release and the lag with respect to the reference year (T). In the table "m" is used when the information on the date of source release is missing, while "-" refers to variables which are not available for the country.

Details by each year and variable are reported in the metadata available in the ETER website.

TABLE 10: DATES OF SOURCE RELEASE (MONTH YEAR, T=REFERENCE YEAR) **

Country	Expenditure	Revenue	Academic staff	Total staff	Enrolled students ISCED 5-7	Graduates ISCED 5-7	Enrolled students ISCED 8	Graduates ISCED 8	R&D Expenditure
AT	-	-	-	-	09-11 T+2	09-11 T+2	09-11 T+2	09-11 T+2	-
BE (FL)	03-06 T+1	03-06 T+1	02-12 T+1	02-12 T+1	01 T+1	02 T+2	01 T+1	02 T+2	-
BG	-	-	04 T+1	04 T+1	04 T+1	04 T+2	04 T+1	04 T+1	-
CH	11 T+1	11 T+1	12 T+1	12 T+1	03-06 T+1	06-08 T+1	03-06 T+1	06-08 T+1	11 T+1
CY	06-12 T+2	06-12 T+2	09 T+2	09 T+2	09 T+2	09 T+2	09 T+2	09 T+2	06-09 T+2
CZ	01-11 T+2	01-11 T+2	01-11 T+2	01-11 T+2	11-12 T+2	11-12 T+2	11-12 T+2	11-12 T+2	09 T+2
DE	05 T+2	05 T+2	10 T+1	10 T+1	09 T+1	09 T+1	09 T+1	09 T+1	-
DK	T+2	T+2	T+2	T+2	02-04 T+2	02-04 T+2	02-04 T+2	02-04 T+2	03 T+2
EE	m	m	m	m	11 T	11 T+1	11 T	11 T+1	-
ES	m	m	02 T+2	02 T+2	05 T+2	05 T+2	m	05 T+2	-
FI	m	m	m	m	m	m	m	m	m
FR	07 T+1	07 T+1	07 T+1	07 T+1	07 T+1	10 T+2	07 T+1	10 T+2	-
GR	-	-	10 T+2	10 T+2	10 T+2	10 T+2	10 T+2	10 T+2	-
HR	m	m	m	m	m	m	m	m	m
HU	m	m	m	m	m	m	m	m	-
IE	m	m	m	m	09 T+1	09 T+1	09 T+1	09 T+1	-
IS	-	-	m	m	m	m	m	m	-
IT	09 T+2	09 T+2	08 T+1 - 06 T+2	08 T+1 - 06 T+2	04 T+2	10 T+1	07 T+2	07 T+2	-
LI	-	-	03 T+2	03 T+2	03 T+2	03 T+2	03 T+2	03 T+2	-
LT	06 T+1	06 T+1	03-06 T+2	03-06 T+2	04-05 T+1	04-05 T+1	04-05 T+1	04-05 T+1	06 T+1
LU	m	m	m	m	m	m	m	m	m
LV	-	-	06 T+1	06 T+1	06 T+1	06 T+1	06 T+1	06 T+1	09 T+1
ME	-	-	-	-	-	-	-	-	-
MK	m	m	m	m	m	m	m	m	m
MT	06 T+2	06 T+2	05-06 T+2	05 T+2	06 T+2	01 T+2	01 T+2	01 T+2	06 T+2
NL	m	m	12 T+1	12 T+1	10 T+1	10 T+1	07 T+1	m	-
NO	m	m	m	m	m	m	m	m	-
PL	m	m	m	m	m	m	m	m	m
PT	03 T+1	03 T+1	03 T+1	03-10 T+1	02-11 T+2	02-11 T+2	02 T+2	02 T+2	-
RO	-	-	-	-	-	-	-	-	-
RS	m	m	m	m	m	m	m	m	m
SE	4 T+1	4 T+1	4 T+1	4 T+1	m	m	m	m	m
SI	-	-	-	-	-	-	-	-	-
SK	4 T+1	4 T+1	m	m	m	m	m	m	4 T+1
TR	m	-	m	m	m	m	m	m	-
UK	01 T+2	01 T+2	12 T+1 - 02 T+2	12 T+1 - 02 T+2	12 T+1 - 02 T+2	12 T+1 - 02 T+2	12 T+1 - 02 T+2	12 T+1 - 02 T+2	-

** "m" = data available but date of source release missing, "-" = variable missing.

6 ACCESSIBILITY AND CLARITY

6.1 ON-LINE DATABASE

The ETER database, which is the main output of the project, is made available through the ETER web application. The web application enables the user to retrieve data from the entire ETER data set in order to conduct research on micro data of the European higher education sector. This section will shortly outline the structure and functionalities of the web application, which are described in detail in chapter 9 of the ETER handbook.

The homepage includes a short description of the ETER project and the performing consortium members. Starting from the homepage of the ETER website, three paths are prepared for the user to define an individual query (depending on the information required):
- path1: the user wants to get an overview of the register and its contents and documentation,
- path2: the user wants to view and/or export data from the ETER micro data set, including country level metadata, or
- path3: the user wants to look for specific HEIs in the register.

1 – GETTING AN OVERVIEW ABOUT ETER
The selection field "ETER in a nutshell" offers the possibility of having detailed information about the register, its contents and methodology:
- obtain a general overview about the project itself, its targets and contents.
- gather detailed information on the data collected, the origin of the data and special codes and flags used in the data collection.
- retrieve all relevant documents accompanying the project. These include a report about the main conceptual and methodological choices of ETER, the data collection handbook with an in depth description of the data collection methodology, a technical report with detailed information on methodology, data collection, data management and data quality, etc.
- access results of analyses (ETER briefs) based on ETER data. These analyses were performed by the project team and individual analysis as well as figures, which can be downloaded separately.
- acquire information on publications based on the ETER data set.
- obtain answers to key questions by having a look at the frequently asked questions (FAQs) section on the website.

2 – VIEW AND EXPORT DATA AND METADATA
By selecting the label "Choose your HEI data", users can choose the required higher education institutions by year and country. Additional possibilities allow the user to refine the selection (Select your variables) and filter them. With the group of chosen institutions, the user can go a step forward and decide to download the requested data (Export data) or arrange and visualize them directly on the website (Tables and Visualizations). The default results mask includes some basic variables in order to get a first overview. Option Select your variables enables the user to select the required variables for export (an exporting of all variables at once is also possible without selecting them – see export function), tables and visualizations.
TABLE AND VISUALIZATION: After selecting variables, it is possible to change to the interactive Tables and Visualizations space. Using a table component, users are able to customize their data sets directly on the website. Users can drag and drop columns and arrange the data in order to retrieve the required data format. In interactive tables, different variables can be combined, displayed, and adapted to the specific user's needs. Users can additionally apply filter to all variables and use the table grid in order to calculate shares. Additionally, graphical options for the selected data are available on the web interface. This is the case for selected indicators and allows the comparison of single institutions in a broader context.
EXPORT FUNCTION: The most important feature for data analysts is the export function, which is dedicated to usability in order to support data analysis and the dissemination objectives of ETER. Data export will be possible in .xlsx and .csv format, where users have several options:
- choose the data to be exported by applying filter and selecting variables or export all variables at once.
- export the corresponding metadata.

3 – LOOK FOR SPECIFIC HEIs

In order to follow path3, the user needs to access "Look up your HEI" which leads to a filtering mask where it is possible to insert the name (or part of it) or the ETER code of the institutions of interest. The HEIs will be filtered in all available years and by clicking on them an overview of all available data will be showed for a quick consultation.

6.2 QUALITY DOCUMENTATION

Validation and data quality control are central tasks in ETER. They respond to the requirements of ensuring the best possible quality of the data collected and to limit inconsistencies as much as possible.

Institution-level databases, like ETER, raise particularly difficult issues concerning data quality given the lower level of aggregation, which makes them particularly sensible to quality issues. Further, given the fact that ETER is not only intended for research purposes, but also for providing transparency and visibility of the activities of individual HEIs, it is not sufficient that the dataset display a sufficient level aggregated data quality. Instead, care must also be taken to guarantee the correctness of individual data points and the comparability between individual observations.

A central issue for data quality is represented by the high level of heterogeneity in the HEI sample considered: not only are there systematic differences in the organization of national systems of higher education, which also translates to differences in indicators. Additionally, HEIs themselves are highly heterogeneous organizations, ranging from large and internationally reputed research universities to small-scale, mostly teaching oriented, specialized schools in arts, theology or teacher training. This significant level of heterogeneity makes the identification and interpretation of deviant observations highly problematic.

6.2.1 DATA FLAGGING

Based on the information provided with the data delivery, the accompanying metadata and the results of the data quality process, a number of flags are introduced in the dataset by the Consortium, with the aim to signal problems or specificities of data related to format accuracy, consistency, completeness and comparability.

Table 11 presents the number of flags introduced in the ETER database.

TABLE 11: NUMBER OF FLAGS BY CATEGORY AND BY YEAR

Year	b	de	d	i	ic	rd	ms	p	c
2011	0	0	4535	2945	395	1380	0	0	2048
2012	149	344	4233	2815	65	1307	1	2	2116
2013	730	752	4462	3362	78	1343	31	165	1567
2014	394	87	1204	3177	65	1377	30	0	1574

Two flags, namely "d" and "i", represent between 55% and 66% of total flags by year. They are the two flags specifically dealing with departures from ETER definitions and other specificities of data affecting comparability across countries. The following Table 12 shows the distribution of flags "d" and "i" by variable and year.

TABLE 12: NUMBER OF FLAGS 'D' AND 'I' BY VARIABLE AND YEAR

	d				i			
Variable code	2011	2012	2013	2014	2011	2012	2013	2014
GEO.FLAG	0	0	0	0	10	9	10	4
EXP.FLAGCURRPERSON	0	28	21	19	29	31	6	3
EXP.FLAGCURRNONPERSON	0	0	0	0	28	28	5	3
EXP.FLAGCURRUNCL	0	0	0	0	49	136	100	2
EXP.FLAGCURRTOTAL	0	0	0	0	0	2	163	5

Variable code	d				i			
	2011	2012	2013	2014	2011	2012	2013	2014
EXP.FLAGCAPITAL	27	27	26	19	1	2	43	2
REV.FLAGCOREBUDGETPUBLIC	0	0	0	0	0	0	0	2
REV.FLAGCOREBUDGETOTHER	0	0	0	0	0	0	0	57
REV.FLAGCORETOTAL	20	20	0	0	28	62	4	2
REV.FLAGTHIRDPARTYPUBLIC	0	0	0	0	0	0	0	2
REV.FLAGTHIRDPARTYPRIVATE	0	0	0	0	26	32	2	2
REV.FLAGTHIRDPARTYABROAD	0	0	0	0	0	0	0	50
REV.FLAGTHIRDPARTYUND	0	0	0	0	0	0	0	2
REV.FLAGTHIRDPARTYTOTAL	0	0	0	0	30	30	2	2
REV.FLAGSTUDFEES	0	0	0	0	29	55	7	2
REV.FLAGUNCL	0	0	0	0	168	133	97	2
REV.FLAGCURRTOTAL	0	0	0	0	185	82	119	29
REV.FLAGNONRECURR	0	0	0	0	0	0	0	2
STA.FLAGACAFTETOTAL	74	67	0	0	254	258	17	7
STA.FLAGACAGENDER	55	55	53	53	10	27	21	16
STA.FLAGACACITIZ	41	41	51	52	74	19	7	3
STA.FLAGACAFOE	0	0	0	39	0	4	52	3
STA.FLAGTOTACAHC	74	108	39	38	342	355	197	134
STA.FLAGPROF	35	37	13	14	4	1	289	307
STA.FLAGNONACAFTE	0	0	0	0	2	190	53	43
STA.FLAGNONACAHC	0	0	0	0	167	215	144	101
STA.FLAGTOTALFTE	0	0	0	0	156	239	54	42
STA.FLAGTOTALHC	0	0	0	0	344	248	275	433
STUD.FLAGISCED5GEN	0	0	0	0	0	0	3	1
STUD.FLAGISCED5CITIZ	0	0	0	0	39	0	182	184
STUD.FLAGISCED5MOB	27	27	27	0	0	0	3	1
STUD.FLAGISCED5FOE	66	66	61	0	6	6	9	1
STUD.FLAGISCED5TOTAL	0	0	38	0	23	24	29	19
STUD.FLAGISCED6GEN	32	33	0	0	1	1	3	3
STUD.FLAGISCED6CITIZ	0	0	214	183	40	1	3	3
STUD.FLAGISCED6MOB	60	58	52	23	1	1	5	3
STUD.FLAGISCED6FOE	503	483	588	0	18	18	20	13
STUD.FLAGISCED6TOTAL	48	22	217	206	43	50	15	3
STUD.FLAGISCED7GEN	38	36	0	0	1	1	3	1
STUD.FLAGISCED7CITIZ	0	0	0	0	40	1	182	176
STUD.FLAGISCED7MOB	83	79	39	10	1	1	5	1
STUD.FLAGISCED7FOE	477	448	530	0	18	18	20	11
STUD.FLAGISCED7TOTAL	55	44	31	10	35	56	54	42
STUD.FLAGISCED7LONGGEN	0	0	0	0	1	0	2	0
STUD.FLAGISCED7LONGCITIZ	0	0	0	0	40	0	14	0
STUD.FLAGISCED7LONGMOB	30	28	5	10	1	0	2	0
STUD.FLAGISCED7LONGFOE	274	253	224	0	8	1	3	0
STUD.FLAGISCED7LONGTOTAL	15	15	0	10	39	73	226	358
STUD.FLAGTOTALISCED5.7	0	0	0	23	5	9	11	5
GRAD.FLAGISCED5GEN	0	0	0	0	0	0	2	0

Variable code	d				i			
	2011	2012	2013	2014	2011	2012	2013	2014
GRAD.FLAGISCED5CITIZ	0	0	0	0	39	0	181	183
GRAD.FLAGISCED5MOB	27	27	27	0	0	0	2	0
GRAD.FLAGISCED5FOE	41	66	77	0	0	0	2	0
GRAD.FLAGISCED5TOTAL	0	0	0	0	72	72	71	84
GRAD.FLAGISCED6GEN	64	63	0	0	1	2	3	3
GRAD.FLAGISCED6CITIZ	32	30	179	183	40	2	3	3
GRAD.FLAGISCED6MOB	122	119	52	23	1	2	5	3
GRAD.FLAGISCED6FOE	471	484	584	0	11	12	13	13
GRAD.FLAGISCED6TOTAL	32	30	205	206	46	43	18	12
GRAD.FLAGISCED7GEN	36	36	0	0	1	2	3	1
GRAD.FLAGISCED7CITIZ	0	0	0	0	40	2	182	176
GRAD.FLAGISCED7MOB	81	79	39	10	1	2	5	1
GRAD.FLAGISCED7FOE	433	450	494	0	11	12	13	11
GRAD.FLAGISCED7TOTAL	36	38	8	10	41	53	54	88
GRAD.FLAGISCED7LONGGEN	0	0	0	0	1	1	2	1
GRAD.FLAGISCED7LONGCITIZ	0	0	0	0	40	1	2	1
GRAD.FLAGISCED7LONGMOB	30	28	5	10	1	1	2	1
GRAD.FLAGISCED7LONGFOE	214	199	191	0	1	1	2	1
GRAD.FLAGISCED7LONGTOTAL	15	15	0	10	42	72	226	398
GRAD.FLAGTOTALISCED5.7	0	0	0	23	13	13	26	19
RES.FLAGSTUDISCED8GEN	14	14	14	0	0	0	2	0
RES.FLAGSTUDISCED8CITIZ	0	0	0	0	39	0	2	28
RES.FLAGSTUDISCED8MOB	58	56	32	5	26	21	2	0
RES.FLAGSTUDISCED8FOE	107	96	145	0	12	10	11	7
RES.FLAGSTUDISCED8TOTAL	347	275	14	5	43	31	14	8
RES.FLAGGRADISCED8GEN	0	0	0	0	0	0	1	0
RES.FLAGGRADISCED8CITIZ	0	0	0	0	39	0	1	28
RES.FLAGGRADISCED8MOB	58	56	32	5	26	0	1	0
RES.FLAGGRADISCED8FOE	105	97	135	0	5	5	6	7
RES.FLAGGRADISCED8TOTAL	178	0	0	5	45	25	16	3
RES.FLAGR.DEXP	0	0	0	0	12	11	28	10
Total	4535	4233	4462	1204	2945	2815	3362	3177

Table 13 shows the number of most relevant flags by country and their evolution over time. Certain flags are concentrated in very few countries (i.e. "rd"-rounding in UK and DE), while others are more widespread (i.e. "i"- which means "see metadata" is present in almost all countries).

Country	b				d				i				ic				rd				p				c			
	2011	2012	2013	2014	2011	2012	2013	2014	2011	2012	2013	2014	2011	2012	2013	2014	2011	2012	2013	2014	2011	2012	2013	2014	2011	2012	2013	2014
AT	0	0	0	0	0	0	0	0	76	68	71	69	0	0	0	0	49	0	9	0	0	0	0	0	0	0	0	0
BE	0	0	419	84	0	0	0	0	117	36	14	69	27	0	2	2	1	0	0	0	0	0	5	0	0	0	0	0
BG	0	0	5	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	189	189	189	189
CH	0	0	101	0	21	23	0	0	7	42	79	14	0	0	0	0	5	6	0	0	0	0	2	0	0	0	0	0
CY	0	0	1	0	0	0	0	0	79	58	51	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
CZ	0	0	0	0	0	0	317	0	68	74	40	29	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0
DE	0	0	0	0	149	148	0	0	116	169	175	230	2	45	44	55	94	103	98	102	0	0	35	0	0	0	0	0
DK	0	0	0	0	0	0	0	0	0	0	2	0	76	0	0	0	0	0	17	0	0	0	1	0	0	0	0	0
EE	0	0	0	0	240	224	98	238	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ES	0	140	159	0	2	2	0	0	157	206	139	83	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0
FI	0	0	0	0	292	333	319	114	41	3	0	0	0	0	0	0	1	0	3	0	0	0	0	0	0	0	0	0
FR	0	0	0	0	362	141	21	19	27	36	64	245	3	7	0	1	0	0	0	0	0	0	25	0	52	52	52	49
GR	0	0	0	195	95	60	79	0	138	138	29	44	0	0	0	0	0	0	0	0	0	0	0	0	44	44	0	0
HR	0	0	0	0	96	99	0	0	128	132	73	149	0	0	0	0	0	0	0	0	0	0	0	0	576	592	640	666
HU	0	0	0	0	0	0	339	0	0	0	0	0	0	0	26	0	0	0	0	0	0	0	19	0	0	0	0	0
IE	0	0	0	0	459	459	450	19	187	240	78	82	0	0	0	0	0	0	0	0	0	0	6	0	0	0	0	0
IS	0	0	0	0	0	0	0	0	34	48	48	5	0	0	0	0	0	0	0	0	0	0	0	0	56	70	0	0
IT	0	0	0	0	0	0	0	0	411	703	0	1	175	1	3	4	0	0	0	0	0	0	26	0	0	0	0	0
LI	0	0	0	0	6	6	0	0	1	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
LT	0	0	0	0	0	0	0	39	1	1	0	0	0	1	0	0	0	0	0	0	0	0	2	0	236	274	335	335
LU	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
LV	0	0	0	0	192	188	46	0	50	51	47	2	0	0	0	0	0	0	0	0	0	0	4	0	895	895	351	335
MK	0	0	0	0	0	0	0	0	91	113	111	113	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
NL	0	0	0	0	134	121	85	43	60	137	101	205	16	1	0	0	0	0	1	1	0	0	0	0	0	0	0	0
NO	0	0	0	0	0	0	0	0	175	207	201	61	0	0	3	0	0	6	0	0	0	0	1	0	0	0	0	0
PL	0	0	45	0	1340	1279	1172	0	206	213	489	305	7	0	0	0	0	0	0	0	0	0	9	0	0	0	0	0
PT	0	0	0	115	826	763	486	0	276	47	23	142	0	0	0	0	0	0	0	0	0	0	6	0	0	0	0	0

Country	b				d				i				ic				rd				p				c			
	2011	2012	2013	2014	2011	2012	2013	2014	2011	2012	2013	2014	2011	2012	2013	2014	2011	2012	2013	2014	2011	2012	2013	2014	2011	2012	2013	2014
RS	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0
SE	0	0	0	0	301	367	341	0	468	80	111	56	89	0	0	1	0	0	0	0	0	0	6	0	0	0	0	0
SK	0	0	0	0	20	20	0	0	0	0	0	0	0	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0
TR	0	0	0	0	0	0	709	732	0	0	1251	1248	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
UK	0	9	0	0	0	0	0	0	27	12	165	20	0	0	0	2	1228	1192	1215	1270	0	1	12	0	0	0	0	0
ETER	0	149	730	394	4535	4233	4462	1204	2945	2815	3362	3177	395	65	78	65	1380	1307	1343	1377	0	2	165	0	2048	2116	1567	1574

7 QUALITY REPORT: COHERENCE AND COMPARABILITY

7.1 COHERENCE: DATA CONSISTENCY

Consistency verifies possible violations of semantic rules defined over the involved data, and specifically between different variables.

Given the nature of the ETER dataset, there is a high number of mutual dependencies between variables, which can be exploited for purposes of data quality analysis. In broad terms, they can be grouped in the following categories (for a complete list of consistency checks in ETER see Chapter 7 in the handbook):

- Logical dependencies between categorical variables and values. For example, when the highest degree delivered is ISCED 5, all numbers of students and graduates at levels 6-8 have to be coded as "not applicable." Similarly, if an HEI is non-research active, R&D expenditures should be "not applicable." Most of these rules are already stipulated in the definition of these variables.
- Sums of breakdowns of variables equal to the total, for example the sum of male, female, and gender unclassified students should be equal to the total.
- Relationships between valued variables. For example, R&D expenditures should be lower than total expenditures; the ancestor year should precede the foundation year of the actual HEI (which should, in turn, precede the legal status year).

A first round of checks was performed during the data collection phase in order to check with national respondents for possible mistakes. Consistency has been controlled again during the validation phase, through the calculation of a set of consistency indicators. Table 14 reports a list of consistency indicators identified as relevant for the assessment of the ETER data quality. Table 15 shows a list of detected inconsistencies identified through the calculation of consistency-like indicators.

TABLE 14: LIST OF CONSISTENCY INDICATORS

Consistency indicator		Value
1	Total Expenditure=SUM(personnel expenditure, non-personnel expenditure, capital expenditure, unclassified expenditures)	TRUE, FALSE
2	Total expenditure>0	TRUE, FALSE
3	Total Income=SUM(core budget, third party funding, tuition fees, revenues unclassified)	TRUE, FALSE
4	Total Income>0	TRUE, FALSE
5	Staff Total (HC and FTE)=SUM(academic staff, non-academic staff)	TRUE, FALSE
6	Staff Total>0	TRUE, FALSE
7	Academic staff total=SUM(female academic staff, male academic staff, unclassified)	TRUE, FALSE
8	Academic staff total=SUM(national academic staff, foreign academic staff, unclassified)	TRUE, FALSE
9	Academic staff total=SUM(academic staff by field of education)	TRUE, FALSE
10	Academic staff total-full professors>0	TRUE, FALSE

Consistency indicator		Value
11	Full professors=SUM(female full professors, male full professors, unclassified)	TRUE, FALSE
12	*If* lowest degree delivered=ISCED 8 *then* Enrolled Students, Graduates ISCED 5-7 ="a" *If* lowest degree delivered=ISCED 7 *then* Enrolled Students, Graduates ISCED 5-6 ="a" *If* lowest degree delivered=ISCED 6 *then* Enrolled Students, Graduates ISCED 5 ="a"	TRUE, FALSE
13	*If* highest degree delivered=ISCED 5 *then* Enrolled Students, Graduates ISCED 6-8 ="a" *If* highest degree delivered=ISCED 6 *then* Enrolled Students, Graduates ISCED 7-8 ="a" *If* highest degree delivered=ISCED 7 *then* Enrolled Students, Graduates ISCED 8 ="a"	TRUE, FALSE
14	Student Total=SUM(female students, male students, unclassified) (for each ISCED level)	TRUE, FALSE
15	Student Total=SUM(national students, foreigner students, unclassified) (for each ISCED level)	TRUE, FALSE
16	Student Total=SUM(resident students, mobile students, unclassified) (for each ISCED level)	TRUE, FALSE
17	Student Total=SUM(students by fields of education) (for each ISCED level)	TRUE, FALSE
18	SUM(Total students enrolled ISCED 5-7, Total students ISCED 8)>0	TRUE, FALSE
19	Graduates Total=SUM(female graduates, male graduates, unclassified) (for each ISCED level)	TRUE, FALSE
20	Graduates Total=SUM(national graduates, foreigner graduates, unclassified) (for each ISCED level)	TRUE, FALSE
21	Graduates Total=SUM(resident graduates, mobile graduates, unclassified) (for each ISCED level)	TRUE, FALSE
22	SUM(Total graduates ISCED 5-7, Graduates ISCED 8)>0	TRUE, FALSE
23	*If* Number of students=0 *then* number of graduates=0 (for each ISCED level)	TRUE, FALSE
24	*If* Non research active *then* R&D expenditure "a"	TRUE, FALSE
25	Total expenditure-R&D expenditure>0	TRUE, FALSE
26	Ancestor year ≤ foundation year ≤ legal status year	TRUE, FALSE

TABLE 15: LIST OF DETECTED INCONSISTENCIES

Country / HEI	Year	Explanation	Solution
Total current expenditure=SUM(personnel exp, non-personnel exp, unclassified exp)			
BE	2011	large inconsistency in expenditure breakdown, since negative "unclassified expenditure" are replaced by 'x'	flagged "ic"
FR, PT, other	2011-2014	minimal inconsistencies due to rounding	flagged "rd"
NL0016	2014	total current expenditure are € 1000 lower than the sum of breakdowns	rounded to € 1,000 in national figures

Country / HEI	Year	Explanation	Solution
Total current revenues=SUM(core budget, third party funding, student fees funding, revenues unclassified)			
BE0081	2011	large inconsistency in income breakdown, since negative "unclassified revenue" has been replaced by 'x'	flagged "ic"
BE0072	2012	large inconsistency in income breakdown, since negative "unclassified revenue" has been replaced by 'x'	flagged "ic"
FR, PT, other	2011-2014	minimal inconsistencies due to rounding	flagged "rd"
Total core budget=SUM(basic government allocation, other core budget)			
BE0061, PT0012	2014	minimal inconsistencies due to rounding	flagged "rd"
BE, CH, FR, PT	2013	minimal inconsistencies due to rounding	flagged "rd"
Personnel Total (HC and FTE)=SUM(academic staff, non-academic staff)			
BE, BG, ES, other	2011, 2012, 2013 or 2014	minimal inconsistencies in FTE fractional numbers	flagged "rd"
DE	2011-2014	minimal inconsistencies (+/- 1) due to rounding in FTE	flagged "rd"
DK	2013	minimal inconsistencies in HC due to rounding	flagged "rd"
NL	2011	some units of staff HC cannot be distributed to either academic or administrative staff	flagged "ic"
UK	2011-2014	minimal inconsistencies both FTE and HC due to UK rounding policy	flagged "rd"
Academic staff total=SUM(female academic staff, male academic staff, unclassified)			
DK	2013	minimal inconsistencies due to rounding	flagged "rd"
FI, NL	2012-2014	breakdowns only available in FTE, not comparable with aggregate academic staff in HC as explained in metadata	flagged
UK	2011-2014	minimal inconsistencies due to UK rounding policy	flagged "rd"
Academic staff total=SUM(national academic staff, foreign academic staff, unclassified)			
DK	2013	minimal inconsistencies due to rounding	flagged "rd"
FI	2011-2014	breakdowns only available in FTE, not comparable with aggregate academic staff in HC as explained in metadata	flagged
NL	2013-2014		
UK	2011-2014	minimal inconsistencies due to UK rounding policy	flagged "rd"
Academic staff total=SUM(academic staff by FoE)			
UK	2011-2014	minimal inconsistencies due to UK rounding policy	flagged "rd"
Personnel Total (HC and FTE)>0			
CZ	2014	several HEIs with staff=0 to be corrected with "m"	flagged[4]
Highest/Lowest degree consistency			
CY	2011	Some HEIs with highest degree=0 reports graduates ISCED6 pre-reform	flagged "ic"
DE	2011-2014	Several cases of students and graduates ISCED7=0 even if highest degree=1	flagged
NO	2011-2014	Different classification of students and graduated by levels (HEIs where lowest degree=0 and students ISCED5=a)	flagged

[4] After May 30th 2017 apart from small changes on descriptors (names, acronyms, websites and so on) the main changes introduced have been i) Flagging of Financial data for Germany; ii) Changes in Expenditures and Revenues of 2011 and 2012 for Ireland; iii) Changes in Expenditures and Revenues of 2013 for Hungary; iv) The coding of Staff data for some HEIs of the Czech Republic has been changed from 0 to 'm'.

Country / HEI	Year	Explanation	Solution
UK	2011-2014	HEIs without degree power at a certain level may offer programmes which are validated by other HEIs	flagged
Students/Graduates Total=SUM(female, male, unclassified)			
AT, UK	2011-2014	minimal inconsistencies due to rounding (not integer in AT and rounding policy in UK)	flagged "rd"
Students/Graduates Total=SUM(national, foreigner, unclassified)			
AT, UK	2011-2014	minimal inconsistencies due to rounding (not integer in AT and rounding policy in UK)	flagged "rd"
GR	2013	the inconsistency is explained referring to different coverage (breakdown by citizenship includes both ISCED6 and ISCED7long students + data refer to the end of academic year)	flagged "ic"
Students/Graduates Total=SUM(resident, mobile, unclassified)			
AT, UK	2011-2014	minimal inconsistencies due to rounding (not integer in AT and rounding policy in UK)	flagged "rd"
Students/Graduates Total=SUM(students/graduates by field of education)			
GR	2012	students by ISCED-F not consistent because of different reference date (students by FoE only available for the end of academic year)	flagged "ic"
SE	2011	breakdown by FoE refers to number of degrees awarded and not to 'graduates'	flagged "ic"
UK	2011-2014	minimal inconsistencies due to UK rounding policy	flagged "rd"
Total Students/Graduates > 0			
TR	2013, 2014	graduates for some HEIs are zero	none, newborn HEIs

The calculation of consistency-like indicators permits to detect and correct many inconsistencies of the ETER data.

There are however a few remaining inconsistencies due to:

- small-scale differences due to rounding errors;

- inconsistent coding of variables according to the rules defined in the ETER handbook, because of different rules and decisions taken at the national level (e.g. HEIs classified as non-research active which report small amount of research expenditure, or when the lowest/highest degree level is not consistent with the presence of students/graduates);

- breakdowns of students/graduates data non consistent due to the use of different data sources.

The list of problems which still remain is summarised in Table 15 according to which data have been flagged in the ETER dataset. It should be underlined that in most cases problems are not due to mistakes in data collection, but point to more fundamental problems with the available data. In any case they have a low or minimal impact on quality of data, referring to rounding or missing data coding rules.

7.2 COHERENCE: OUTLIER DETECTION

An outlier can indicate an observation or processing error, or a special element of the observed population which needs to be treated differently from the bulk in the subsequent processes.

In ETER, the outlier detection phase is part of the quality validation, and has been performed with two different objectives: (i) to identify possible errors in data collected; (ii) to better understand the collected data in order to perform a proper analysis.

The ETER outlier detection phase consists of a *basic phase,* based on the consistency-like indicators described in Table 14 and an *advanced phase*.

The advanced phase is based on the estimation of the distribution that better approximates the data. The method is applied on a set of ratios reported in Table 16.

TABLE 16. RATIOS USED IN THE ADVANCED OUTLIER DETECTION PHASE

Ratio	Breakdowns
Total expenditure / Total revenue	
Core budget / Total revenue	
Third party funding / Total income	
Student fees funding / Total income	
Tot expenditure / Total enrolled students ISCED 5-8	
Personnel expenditure / Total staff (FTE)	
Academic staff (FTE) / Number of administrative staff (FTE)	
Total staff (FTE) / Total staff (HC)	Academic staff, Administrative staff
Total enrolled students (ISCED 5-8) / Total academic staff	
Foreign enrolled students / Total enrolled students	by ISCED level
Mobile enrolled students / Total enrolled students	by ISCED level
Total students enrolled ISCED 5-7 / Total graduates ISCED 5-7	
Enrolled students ISCED 8 / Number of doctoral degrees	
Number of foreign students / Number of foreign graduates	by ISCED level
Number of mobile students / Number of mobile graduates	by ISCED level
R&D exp / Tot exp	

The method performs parameter estimation by regression on QQ plot positions[5].

For the purpose of outlier detection, the method assumes that the observations y_i are generated by a model probability density, with cumulative density function (cdf) $F(Y\,|\theta)$ where θ is a vector of parameters specifying F. The value of θ can be robustly estimated from the bulk of the observations (data) by minimizing the sum of squares:

$$\theta = \arg\min \sum_{i\epsilon\Lambda}\big[g(y_i) - g(F^{-1}(\hat{F}|\theta))\big]^2 \qquad\qquad (1)$$

Where:

- Λ indexes a subset of the observations y_i, i.e. all the observations that will be used by the fitting phase
- g is a monotonic function, differentiable on the range of Y.

---

[5] A Q–Q (Quantile-Quantile) is a probability plot, which is a graphical method for comparing two probability distributions by plotting their quantiles against each other.

- $\widehat{F_i}\widehat{F_i}$ are plot positions as used in QQ plots.

The method can be applied with different distributions. We used the implementation of the method proposed by the R package "extremevalues"[6]. In particular we found that most of the empirical data show their best fit with the *lognormal* distribution. In addition, the lognormal distribution solving Equation (1) with a suitable transformation g yields linear regression equations in the form:

$$b = (A'A)^{-1}A'x \tag{2}$$

Where:

- b is a 2-dimensional vector containing functions of the distribution parameters
- A is a $|\Lambda| \times 2$ matrix containing functions of $\widehat{F_i}\widehat{F_i}$ and
- x is a $|\Lambda|$ dimensional vector containing functions of yi.

By solving Equation (2), is it possible to estimate if the observed values are likely to be generated by the lognormal distribution or not.

Hence the method allows us to compare the real observations with the estimated ones, and thus to obtain a robust estimate for the bulk of the data. The list of HEIs detected as outliers by country and by ratio have been carefully examined by the consortium in order to check for reasons, respectively to correct the data when they are due to mistakes.

7.2.1 MULTIANNUAL CHECKS

The availability of data across years (4 years) allows us to investigate longitudinally the consistency of the collected data considering the impact of demographic events; revision of variable's categories and definitions, and their quality more generally. Indeed, multi-annual checks can help to detect suspect cases where the variation from one year to another is very large (as compared with what is expected, respectively with the average change in the whole sample). This type of check may be useful for detecting and reporting mistakes of respondents and/or changes in the methodology for data collection.

The availability of only 4 years of data, however, does not enable us to apply specific methods for time series analysis which require much more data. In addition the ETER dataset is composed by different typologies of variables (i.e. structural descriptors rather than quantitative variables) with a different propensity to change over time.

For these reasons the methodological approach developed for the multiannual checks is based on different techniques:

- *manual check* of the impact of demographic events (take-over, spin-off) on concerned institutions' figures and respective flagging (the code "b" for breakdown in time series was already foreseen);
- *analytic control* of descriptors and status variables supposed to be stable over time, i.e. legal status, foundation year, geographical information, lowest/highest degree awarded, etc.;
- *comparison of national aggregates* over time for a selected number of quantitative variables already during the validation phase, with alarm is the variation is over a pre-defined threshold;
- *use of measures of statistical dispersion* (interquartile range comparison over time) to assess the overall stability of the distribution of quantitative variables;
- *statistical analysis* to highlight the HEIs with annual growth (defined in Equation 3 and 4) outstanding from the overall distribution (outlier) with an adaptation of the methodology already used for yearly outlier detection.

---

[6]See URL: http://cran.r-project.org/web/packages/extremevalues/extremevalues.pdf

The statistical analysis allows us to robustly estimate[7] the distribution that generates our observations and to identify "outliers" as the observations that are "unlikely" to be generated from the empirically (robustly) estimated distribution. We analyse the distribution of the growth rate over a two year interval calculated "à la Malmquist" and weighted to take into account the absolute size of variation.

We are comparing the value of variable *x* over two points in time (year0 and year1). Since there is no reason to assume which one is the reference year (i.e. the correct value) the rate is calculated as the geometric average of the absolute values of the growth rates calculated as year1 over year0 and year0 over year1 respectively. In formula, the growth rate is given by:

$$\sqrt{|(x_1 - x_0)/x_0 \times 100| \times |(x_0 - x_1)/x_1 \times 100|} \qquad (3)$$

This methodological approach is inspired by the work of Malmquist (1953). In ETER we face another specific challenge due to the large heterogeneity of higher education institutions in the sample ranging from micro institutions (also below the size thresholds set for the ETER database) to huge institutions. In this context it is expected that small institutions register tremendous growth rates when the absolute numbers are small (i.e. enrolled students growing from 10 to 20 at ISCED level 6 lead to a growth rate of 100%). To correct for this possible source of distortion, the growth rate formula reported in Equation (3) is weighted for the difference of the levels of variable *x* in absolute terms. The *weighted growth rate* calculated is then:

$$\sqrt{|(x_1 - x_0)/x_0 \times 100| \times |(x_0 - x_1)/x_1 \times 100| \times |x_1 - x_0|} \qquad (4)$$

The method has been applied to compare the actual distribution of the growth rates with different known distributions. We follow a data driven approach to the selection of the fitting distribution, selecting the one with highest correlation coefficient and most suitable to the specific aim of the analysis. In the ETER data quality process indeed the focus of multiannual checks is not on observations with small or null growth rate (stability of levels over time) but rather on those with the highest variations. In other words we are more interested in *upper bound* outliers.

The statistical multiannual outlier detection is applied to the variables listed in Table 17, covering the main ETER dimensions. The comparison is always made between current year against the previous one, but whenever an outlier is detected the whole longitudinal set (starting from 2011) is manually analysed to better interpret its nature. The goal of the statistical method is therefore essentially related to the identification of cases to be checked by the experts of the consortium working team.

TABLE 17. LIST OF MULTIANNUAL CHECKS

Variable	Variable
Total expenditure (PPP)	Total staff (FTE)
Total revenues (PPP)	Total staff (HC)
Total academic staff (FTE)	Total students enrolled (by ISCED level)
Total academic staff (HC)	Total graduates (by ISCED level)
Number of administrative staff (FTE)	

A systematic control of descriptors (which are collected yearly) showed that they are quite stable over time (i.e. the institutional category or legal status of a HEI do not change unless a reform process is in place and in this case it should be registered also in the 'legal status year'). Some regional codes have been aligned with the last NUTS version applied since 2015.

---

[7] The analysis has been done by using the R package "extremevalues".

The statistical analysis described above on multiannual outliers has been applied on a selected subset of relevant variables. The distribution that generates the data has been estimated using the weighed growth rate (in Equation 4) for the years 2011-2012, 2012-2013 and 2013-2014.

The statistical analysis identified two categories of outliers for the last two waves of data collection:

- *Lower bound outliers*, which are cases of almost perfect stability over time (i.e. the value in 2014 is the same than in 2013). These cases are not suspect unless they involve all HEIs in a country or all variables for one specific HEIs. Such situations would rise the doubt of misreporting (no update) of the data collection, but none is found in ETER.

- *Upper bound outliers*, which are cases of large variation between one year and another. Around 500 cases are detected each year, including records with zero in one of the two years (more frequent for students/graduates). The statistical outliers may point to HEIs registering a strong growth or shrinking process, or to possible change of the classification systems with direct impact on the values collected. Detected cases have been analysed individually and flagged in the database when relevant.

To summarise, most of the outliers detected seem to be individual cases to be analysed one by one. Reporting mistakes for some variables have been found for a couple of countries, but easily solved in the current version of the dataset. There is only one situation that probably has a structural explanation which is the decrease of ISCED level 5 students and graduates in IE, especially in institutes of technology, and the corresponding increase at level 6. Other breaks in the time series, like the new coverage of researches under academic staff in PT where already described in the metadata.

Finally, there are also a few cases of HEIs fast growing because of demographic events (take-over of other existing institutions) and this is reflected in most of the variables investigated.

7.3 COMPARABILITY

In order to highlight problems of comparability of ETER values across countries, specific metadata have been collected together with quantitative variables. Although further efforts have been made to improve the completeness and readability of metadata collected, including the revision of their structure, the information provided by NSA is incomplete in several countries. In addition, some information are not updated and their level of reliability is uncertain. Nevertheless their reading is very useful to better understand the concrete situation, since metadata and notes often anticipate the explanation of problems emerging from data quality analysis. In this respect, data quality and metadata analysis are complementary.

Most of the results and problems are recurring in different waves of data collection, being related to national data survey design, but some progress have been made solving some issues in specific countries.

The main emerging issues are the following:

- The exact mapping of income categories with ETER breakdowns is still not perfectly clear in several countries, implying possible comparability problems (i.e. Negative core budget in DE);

- Specificities about the inclusion and classification of staff across countries and within countries among HEI categories (typically university vs. colleges) may impact full comparability. Figures reported might be incomplete excluding some categories (i.e. atypical staff in UK);

- Availability of FTE and HC measures of staff are jeopardised and sometimes calculated on different perimeters;

- Classification of students and graduates according to the new ISCED levels has is still not perfectly stable in every country, but it has been possible to solve with an ad hoc concordance table for ETER. In some cases figures reports exclude minor categories or programmes;

- Breakdowns of students and graduates by mobility status is not fully comparable among countries;

- The problems due to jeopardised application of the new ISCED classifications -by level of education and by field of education- have been finally solved in 2014 with the complete adoption in all countries. This had a positive impact especially for the comparability of data by field.

In addition, there are a number of departures from the reference period/date required or suggested by ETER, according to the UOE manual rules. Sometime these exceptions also apply to aggregate reporting to UOE, while in other cases there is a discrepancy due to different availability of data when going down to the individual micro data.

Table 18 summarizes departures from ETER definitions or national specifications for each variable. Cases have been flagged individually in the dataset.

TABLE 18. COMPARABILITY ISSUES AND DEPARTURES FROM ETER DEFINITION

ETER variable	2011	2012	2013	2014	Main departures from ETER definition by country
Total current expenditure	x	x	x	x	IE: Total pay expenditure from funding statements for the academic year
	x	x	x	x	LT: HEIs have large discrepancies between total revenues and total expenditures - this is because those HEIs spend only a part of their revenues in financial year
Personnel expenditure		x	x	x	CY: Financial data and personnel data come from different sources. Professors or other personnel may be paid for their services but not included under 'personnel expenditure' but as "other expenditure". Smaller colleges give in general lower salaries
Non-personnel expenditure		x	x	x	NL: Non-personnel expenditure is the sum of Housing Costs
Capital expenditure	x	x	x	x	IE: Not expenditures but total capital grants received
Total current revenues	x	x	x	x	CH, FI: partial figures, corresponding to current expenditure, excluding revenues for capital expenditure
	x	x	x		CZ: total revenue non available, only some breakdowns available at the HEI level
Core budget	x	x	x	x	DE: some core budgets are negative, because revenues are higher than expenditures
Third party funding	x			x	BE: 2011 data not comparable (estimates)
Student fees funding			x	x	BE: tuition fees (includes student fee funding) = operational revenue
Number of academic staff (FTE)		x	x	x	BE: for universities: professors, assistants, researchers, and scholarships; for university colleges: teaching and policy staff
		x	x	x	ES: Academic staff excluded personnel whose primary assignment is only research and graduate students employed for teaching assistance or research
	x	x	x	x	IT: estimation. FT is the standard for (tenured) academic staff
		x	x	x	NL: in the eight universities with affiliated hospital all staff of the universities' medical faculties has been formally transferred to the academic hospitals and therefore it is not anymore included in the staff (academic and non-academic) of the universities
	x				NO: All positions included, also II-positions
	x	x	x		PT: data do not include researcher for all private institutions due to lack of information
Number of academic staff (HC)		x	x	x	BE: for universities: professors, assistants, researchers, and scholarships; for university colleges: teaching and policy staff

ETER variable	2011	2012	2013	2014	Main departures from ETER definition by country
		x	x	x	ES: Academic staff excluded personnel whose primary assignment is only research and graduate students employed for teaching assistance or research
		x	x	x	FI: academic staff based on person-workyears (FTE) including breakdown
	x	x	x	x	IE: only Core-funded staff (perimeter narrower than FTE)
				x	IT: For some HEIs active also in non-tertiary education (Free academies of fine arts, Music Institutes) data refers to all academic staff and is not comparable with students figures
				x	LT: classification by field not corresponding to ISCED-F (national classification applied)
		x	x	x	LV: Personnel elected to academic position and visiting scholars are included
		x	x	x	NO: A person can only be registered once in the registry. II-positions are excluded, as only persons with an employment of minimum 40 per cent are included. On the contrary FTE includes also II-positions and research fellows
	x	x	x		PT: data do not include researcher for all private institutions due to lack of information
Number of full professors			x	x	ES: 'Full professors' includes all the staff officer with full teaching capacity and full research capacity (this might different from the She figures definitions).The category are not properly demarcated in private institutions
	x	x			IE: only FTE available
Number of administrative staff (FTE)	x	x	x	x	IT: estimation (assumed equal to HC)
Number of administrative staff (HC)			x	x	FI: Staff data are based on person-workyears
		x	x	x	GR: relevant data available only for the end of the academic year
	x	x	x	x	IE: only Core-funded staff (perimeter narrower than FTE)
		x	x	x	NO: Only persons with an employment of minimum 40% are included
Total staff (FTE)	x	x	x	x	IT: estimation (assumed equal to HC)
				x	MT: part time accounted as 0.5 FTE
				x	NL: for some universities, the number for staff in Health Faculties is based on an estimate
	x	x	x	x	UK: Atypical staff (both academic and non-academic) is never included. FTE is counted using a population of staff who were active during the reporting period 1 August to 31 July
Total staff (HC)				x	FI: academic staff based on person-workyears (FTE) including breakdown
				x	GR: data on administrative staff refer to the end of the academic year 2014-2015
	x	x	x	x	IE: only Core-funded staff (perimeter narrower than FTE)
				x	IT: For some HEIs active also in non-tertiary education (Free academies of fine arts, Music Institutes) data refers to all academic staff and is not comparable with students figures
				x	NL: for some universities, the number for staff in Health Faculties is based on an estimate
	x	x	x	x	NO: Only persons with an employment of minimum 40 per cent are included.
	x	x	x	x	UK: Figures reported in full person equivalent (FPE): individuals holding more than one contract with a provider are divided amongst the activities in proportion to the declared FTE for each activity. This may impact on the breakdown by academic/non-academic and by field

ETER variable	2011	2012	2013	2014	Main departures from ETER definition by country
Number of enrolled students at ISCED levels 5, 6, 7				x	AT: Students enrolled in multiple types of programmes are pro-rated
	x	x			AT: Not Available by ISCED level. Not classified and therefore missing: All "Universitätslehrgänge" of public universities (including "Master-Lehrgänge") as well as all "Lehrgänge" of private universities, universities of applied sciences and universities of teacher education that are not "Master-Lehrgänge"
		x	x	x	BE: only students actively enrolled with a diploma contract are included (students with a credit contract are excluded). For ISCED level 6, only Professional Bachelors, Academich Bachelors and Bachelor after Bachelor courses are included. Preparatory programs (schakelprogramma's en voorbereidingsprogramma's) or teacher training programs are not included. This is different from the UOE data collection where teacher training programs are included in the ISCED 6.
	x	x	x	x	CH: calculation of students/staff ratios and costs per students using these data are not reliable due to differences in the perimeter and data sources between students data, staff and expenditures data
	x	x	x		CZ: Number of students in headcounts, double-counting from different institutions is not avoided
		x			FR: Level 5 includes Capacité en droit and Préparation aux concours administratifs niveau bac
	x				GR: figures are not consistent with Eurostat aggregates
	x	x			HR: All levels 5-7 summarized under ISCED 6
	x	x	x	x	LV: ISCED 7 long degrees included in ISCED 7 (also graduates)
				x	NL: ISCED7 long degrees as well as the post-master students are included in ISCED7 because they have not been reported uniformly across all institutions
	x				NO: ad-hoc classification by levels. New ISCED non implemented yet
	x				PT: ad-hoc classification by levels. New ISCED non implemented yet
		x	x	x	SE: individuals are counted once per higher education institution (In UOE individuals are counted only once)
			x	x	TR: long-degree courses cannot be distinguished from bachelor degree courses
	x	x	x	x	UK: Writing-up and sabbatical students (applicable to HE level only) are not included in standard counts of students
breakdown by internationalisation / mobility			x	x	BE: breakdown by mobility according to the enrolment in a Flemish school in secondary education (also graduates)
		x	x	x	EE: mobility status based on country of permanent or usual residence (also graduates)
			x	x	GR: data on breakdown by citizenship available only for the end of the academic year (also graduates). ISCED7long foreigner students included under ISCED6
		x	x	x	IE: mobility status based on domiciliary data, permanent domicile prior to enrolment (also graduates)
		x	x	x	LI: mobility status refers to residence at the time of matriculation (also graduates)
			x	x	TR: data on foreign students cannot be broken down by ISCED level and are reported under ISCED6 (also graduates)
	x	x	x	x	UK: mobile students are those whose normal residence prior to commencing their programme of study was in countries other than UK (also graduates)
breakdown by field of education	x	x	x		CZ: according to ISCED97 (also graduates)
	x	x	x		EE: according to ISCED-97 (also graduates)
	x	x	x		FI: according to ISCED-97 (also graduates)

ETER variable	2011	2012	2013	2014	Main departures from ETER definition by country
	x	x	x		IE: according to ISCED-97 (also graduates)
	x	x	x		PL: according to ISCED97 (also graduates)
	x	x	x		SE: according to ISCED-97 (also graduates)
Number of graduates at ISCED levels 5, 6, 7	x	x	x	x	AT: graduates of in multiple types of programmes are pro-rated
		x	x	x	PT: students enrolled in an ISCED 7 Long Degree ("Mestrado Integrado") are accounted in ISCED 7 long degree student in terms of enrolments. They are included in graduates ISCED 6 whenever they complete 180 ECTS, then they are accounted as an ISCED 7 Long Degree graduate when they finish
		x	x	x	SE: The FOET-columns show degrees and not graduates (also ISCED8)
Students at ISCED level 8		x	x	x	BE: Includes students both enrolled for the academic degree of doctor and in doctoral education programmes
		x	x	x	DE: students cover around 50% of UOE doctoral students since ETER includes matriculated students and UOE data base on sample survey
		x			FR: includes HDR (Habilitation à Diriger des Recherches) (also graduates)
		x	x	x	NL: data in FTE
		x	x	x	NO: PhD students are counted by the number of PhD-agreements, with activity in the fall semester. The number includes research fellows (employed at the institutions), physicians that do their PhD as part of their R&D time etc. An institution can have research fellows without having PhD-students (the PhD students are enrolled at -and reported by- a HEI with a PhD-programme).
			x	x	SE: individuals are only counted once per higher education institution
Graduates at ISCED level 8			x	x	BE: Includes only graduates for the academic degree of doctor (not comparable with students)
				x	IS: ISCED8 graduates are classified by field of science
			x	x	SE: individuals are only counted once per higher education institution

7.3.1 CROSS YEAR COMPARISON

The comparability of data across different waves of data collections is affected by:

- *Demographic events*, which determine a change in the institutional perimeter of the involved HEIs. Indeed, according to ETER perimeter guidelines in case of take-over (i.e. incorporation of a HEI in another pre-existing one) the dominant HEI maintain its ID, but the institutional perimeter changes. Similarly in case of spin-off the original HEI maintain its ID but the perimeter shrinks;
- *Breakdown in time series* due to methodological changes in the data collection at national level (i.e. inclusion/exclusion of groups of staff in the collection) or institutional changes (i.e. the reform of educational curricula leading to the opening/closure of courses at a certain ISCED level).

Two specific flags have been introduced to track for these cases, respectively "de" –demographic event- and "b" –break in time series.

Table 19 summarises the cases of demographic events which determined a modification of the perimeter of specific HEIs in each wave of the data collection.

TABLE 19: LIST OF RELEVANT DEMOGRAPHIC EVENTS BY WAVE OF DATA COLLECTION

YEAR	DEMOGRAPHIC EVENT AND INVOLVED HEIs
2012	**CZ0047** College of Information Management and Business Administration took over in 2012 Central Bohemian Institute of Higher Education; **DE0094** Technische Universität Dresden took over in 2012 International Graduate School (IHI) Zittau; **EE0003** Tallinn University from 2012 includes also data for Tallinn Pedagogical College (take-over in 26-08-2012); NL0044 University of Applied Sciences Leiden took over Hogeschool Helicon (January 2013, data included already for 2012); **PL0276** Vistula University from 2012 took over Academy of Finance in Warsaw (01-11-2012); **UK0123** University College London took over the School of Pharmacy in January 2012; **UK0170** University of Wales Trinity Saint David took over in August 2013 Swansea Metropolitan University (data included starting from 2012/2013)
2013	**BE0056** Katholieke Universiteit Leuven took over BE0055 HUB-KU Brussel on 01/10/2013; **DK0001** Københavns Universitet took over DK0014 The Royal School of Library and Information Science on 01/04/2013; **EE0032** Institute of Theology of the Estonian Evagelical Lutheran Church took over EE0033 Tartu Academy of Theology on 01/07/2013; **GR0012** Panteion University of Social and Political Sciences took over the Department of Regional Economic Growth from GR0004 in June 2013; **GR0016** University of Patras took over GR0003 University of Western Greece in April/June 2013; **GR0023** University of Thessaly took over the Department of Informatics with Applications in Biomedicine from GR0004 in June 2013; **NL0058** Saxion University of Applied Sciences took over NL0039 Edith Stein University for Teacher Education on 01/01/2013; **NO0003** University of Tromsø - The arctic university of Norway took over NO0020 Finmark University College on 01/08/2013; **NO0045** NLA University College took over NO0060 Høgskolen i Staffeldtsgate on 01/01/2013; **SE0001** Uppsala universitet took over SE0025 Gotland University College on 01/07/2013; **UK0051** The University of Huddersfield's Barnsley campus (including the majority of its students), were transferred to Barnsley College in 2013/14, outside ETER perimeter (spin-out); **UK0170** University of Wales Trinity Saint David took over UK0141 Swansea Metropolitan University on 01/08/2013
2014	**DE0320** Bad Honnef-Bonn International University of Applied Sciences took over DE0374 Adam-Ries-Fachhochschule Erfurt (Priv.) on 01/09/2013; **EE0002** Tallinn University of Technology took over EE0018 Estonian Maritime Academy on 01/08/2014; **FR0282** Université des Antilles shrank after the spin-out of the Université de Guyane (new institution); **FR0285** University of New Caledonia shrank after the spin-out of the ESPE Nouvelle-Calédonie (new institution); **FR0286** University of French Polynesia shrank after the spin-out of the ESPE Polynésie Française (new institution); **IT0057** Scuola Normale Superiore di Pisa took over IT0024 Istituto Italiano di Scienze Umane (SUM) di Firenze in 2014; **NO0005** Norwegian University of Life Sciences took over NO0013 Norwegian School of Veterinary Science in 01/01/2014; **NO0003** University of Tromsø - The arctic university of Norway took over NO0020 Finmark University College on 01/08/2013 (first common data delivery in 2013); **PL0436** Higher School of Management and Law in Warsaw took over PL0265 Warsaw Customs and Logistics College in 2014; **UK01230** University College London took over UK0054 Institute of Education on 02/12/2014

Table 20 shows the distribution of the flag (b) which is breaks in time series, by country and by year. Overall eight countries have been interested by breaks in the series, and only two of them in more than one year.

TABLE 20. DISTRIBUTION OF BREAKS IN TIME SERIES BY VARIABLE, COUNTRY AND YEARAPHIC EVENTS BY WAVE OF DATA COLLECTION

Variable	ES 2012	UK 2012	BE 2013	BG 2013	CH 2013	CY 2013	ES 2013	PL 2013	BE 2014	GR 2014	PT 2014	TOT.
Total Current revenues	0	1	0	0	0	0	0	0	0	0	0	**1**
Total academic staff (FTE)	0	1	0	0	0	0	0	0	0	0	23	**24**
Academic staff (HC) - gender	0	0	0	0	22	0	0	0	0	0	23	**45**
Academic staff (HC) - citizenship	0	0	0	0	22	0	0	0	0	0	0	**22**
Academic staff (HC) - ISCED-FoE	0	0	0	0	22	0	0	0	0	0	0	**22**
Total academic staff (HC)	0	1	0	0	13	0	0	0	0	0	23	**37**
Number of full professors	0	1	0	0	0	0	0	0	0	0	0	**1**

Variable	ES 2012	UK 2012	BE 2013	BG 2013	CH 2013	CY 2013	ES 2013	PL 2013	BE 2014	GR 2014	PT 2014	TOT.
Number of non-academic staff (FTE)	0	1	0	0	0	1	0	0	0	0	0	**2**
Number of non-academic staff (HC)	0	1	0	0	0	0	0	0	0	0	0	**1**
Total staff (FTE)	0	0	0	0	0	0	0	0	0	0	23	**23**
Total staff (HC)	0	0	0	0	22	0	0	0	0	0	23	**45**
Students ISCED 6 - gender	7	0	22	0	0	0	8	6	0	45	0	**88**
Students ISCED 6 - citizenship	7	0	22	0	0	0	8	6	0	0	0	**43**
Students ISCED 6 - mobility	7	0	22	0	0	0	8	0	0	0	0	**37**
Students ISCED 6 - ISCED-FoE	7	0	22	0	0	0	8	0	23	0	0	**60**
Total students ISCED 6	7	0	22	0	0	0	8	6	0	45	0	**88**
Students ISCED 7 - gender	0	0	22	0	0	0	0	0	0	0	0	**22**
Students ISCED 7 - citizenship	0	0	22	0	0	0	0	0	0	0	0	**22**
Students ISCED 7 - mobility	0	0	22	0	0	0	0	0	0	0	0	**22**
Students ISCED 7 - ISCED-FoE	0	0	1	0	0	0	0	0	13	0	0	**14**
Total students ISCED 7	0	0	22	0	0	0	0	0	0	0	0	**22**
Students ISCED 7 ld - gender	7	0	0	0	0	0	8	0	0	20	0	**35**
Students ISCED 7 ld - citizenship	7	0	0	0	0	0	8	0	0	0	0	**15**
Students ISCED 7 ld - mobility	7	0	0	0	0	0	8	0	0	0	0	**15**
Students ISCED 7 ld - ISCED-FoE	7	0	0	0	0	0	8	0	0	20	0	**35**
Total students ISCED 7 ld	7	0	0	0	0	0	8	0	0	20	0	**35**
Total students ISCED 5-7	0	1	0	0	0	0	0	0	0	45	0	**46**
Graduates ISCED 6 - gender	7	0	22	0	0	0	8	9	0	0	0	**46**
Graduates ISCED 6 - citizenship	7	0	22	0	0	0	8	9	0	0	0	**46**
Graduates ISCED 6 - mobility	7	0	22	0	0	0	8	0	0	0	0	**37**
Graduates ISCED 6 - ISCED-FoE	7	0	22	0	0	0	8	0	23	0	0	**60**
Total graduates ISCED 6	7	0	22	0	0	0	8	9	0	0	0	**46**
Graduates ISCED 7 - gender	0	0	22	1	0	0	0	0	0	0	0	**23**
Graduates ISCED 7 - citizenship	0	0	22	1	0	0	0	0	0	0	0	**23**
Graduates ISCED 7 - mobility	0	0	22	1	0	0	0	0	0	0	0	**23**
Graduates ISCED 7 - ISCED-FoE	0	0	22	1	0	0	0	0	13	0	0	**36**
Total graduates ISCED 7	0	0	22	1	0	0	0	0	0	0	0	**23**
Graduates ISCED 7 ld - gender	7	0	0	0	0	0	8	0	0	0	0	**15**
Graduates ISCED 7 ld - citizenship	7	0	0	0	0	0	7	0	0	0	0	**14**
Graduates ISCED 7 ld - mobility	7	0	0	0	0	0	8	0	0	0	0	**15**
Graduates ISCED 7 ld - ISCED-FoE	7	0	0	0	0	0	8	0	0	0	0	**15**
Total graduates ISCED 7 ld	7	0	0	0	0	0	8	0	0	0	0	**15**
Total graduates ISCED 5-7	0	1	0	0	0	0	0	0	0	0	0	**1**
Students ISCED 8 - ISCED-FoE	0	0	0	0	0	0	0	0	6	0	0	**6**
Total students ISCED 8	0	1	0	0	0	0	0	0	0	0	0	**1**
Graduates ISCED 8 - ISCED-FoE	0	0	0	0	0	0	0	0	6	0	0	**6**
TOTAL	**140**	**9**	**419**	**5**	**101**	**1**	**159**	**45**	**84**	**195**	**115**	**1273**

8 QUALITY REPORT: COST AND BURDEN

The costs of ETER data collection can be considered with respect to the specific assets the project set up for the phases of:

- Data publication
- Data management and integration

In the following, we briefly detail the ICT assets made available for each phase.

8.1 DATA PUBLICATION

In order to increase transparency in European higher education, the ETER project aims to make data publicly available to interested users. ETER data can be accessed and searched on-line from the project website https://www.eter-project.com. The website has been designed as a unique entry point for all information in ETER. It provides users with the following information:

- Basic information on the project and methodological information in order to use the data.
- Searching and downloading subsets of the data or downloading the whole dataset in .xlsx, .csv or machine ready format.
- Downloading metadata information.
- Look up individual HEIs of interest.

The website provides open access for all users, as well as a restricted access (password-protected) through which users can also download data which are restricted to scholarly usage.

8.2 DATA MANAGEMENT AND INTEGRATION

The ETER project has developed an infrastructure that allows for the standardization and systematization of the data collection process, preliminary validation and data management. This is highly important in order to master the complexity of a data collection process that requires cooperation with the National Statistical Authorities and to ensure a sufficient level of uniformity in the data.

This infrastructure includes:

- Excel based data collection.
- A database for raw data.
- Statistical analysis software (mainly Stata and R) that have been ad-hoc developed.

8.3 BURDEN

The ETER project performs a secondary data collection, meaning that data are not collected as primary data directly from respondents, but results from an elaboration done by NSAs starting from the data sources they in turn gathered. In this way the burden on the final respondents is indeed not present. There is however a burden on the NSAs that was however minimized. Particular actions in this sense were:

- NSAs were asked only to **update** the country perimeter, indicating changes and demographic events.
- NSAs received the data collection sheets **pre-filled** with descriptors and most quantitative variables.

9 QUALITY REPORT: STATISTICAL PROCESSING

9.1 SOURCE DATA

The current ETER coverage includes 36 countries: EU-28 countries as well as Iceland, Liechtenstein, Norway, Switzerland, Republic of Serbia, the Former Yugoslav Republic of Macedonia, Montenegro, Turkey. For Montenegro, Romania and Slovenia only descriptors are available. Data are provided for the years from 2011 (academic year 2011/2012) to 2014 (academic year 2014/2015).

Table 21 presents an overview of the sources of data by country and by variable domain. It can be observed that in most cases data come from National Statistical Offices and/or from Ministry of Education and Research, according to national governance settlements. However, for very small countries with one or very few HEIs, the data come directly from their institutional database. Detailed information by country/variable is contained in country metadata available at the project website.

TABLE 21. DATA SOURCES BY COUNTRY

Country	Expenditure	Income	Academic and non academic staff	Students and graduates ISCED 5-7	Students and graduates ISCED 8	R&D Expenditures
AT	m	m	m	Statistics Austria	Statistics Austria	m
BE*	yearly financial reports		VLIR (2013). Statistische gegevens betreffende het personeel van de Vlaamse universiteiten. + administrative data for university colleges	-	-	m
BG	m		Republic of Bulgaria National Statistical Institute			m
CH	Federal Statistical Office					
CY	CYSTA: Services Survey for Private	CYSTAT: Annual Survey on Education				CYSTAT: R&D survey
CZ	Ministry of Education, Youth and Sports					m
DE	statistics of finances of universities		statistics of personnel of universities	statistics of students of universities	statistics of examinations of universities	m
DK	Public finances		Danske Universiteter + Uddannelses og Forskningsministeriet (UFM)	Student register		Annual questionnaire
EE	m	m	m	-	-	m
ES	m	m	-	-	-	m
FI	Ministry of Education and Culture/Finnish National Board of		Ministry of Education and Culture + Statistics Finland	Statistics Finland	Statistics Finland	Statistics Finland
FR	Ministère de l'Éducation nationale, de	m	Ministère de l'Éducation nationale, de l'Enseignement supérieur et de la			m
GR	m	m	Hellenic Statistical Autoritiy			National Documentation Center
HR	-	-	-	-	-	-
IE	Funding Statements of higher education		Quarterly staffing returns to HEA	HEA database (Student Record System)		m
IT	MIUR Ufficio Statistico		MIUR database + Survey of Contract academic staff and Technical and Administrative staff + Survey on AFAM	MIUR Ufficio Statistico		m
LI	m	m	-	-	-	m
LT	School report		Centre of Information Technologies of Education			School report
LV	m	m	Central Statistical Bureau data collection on Higher Educatiob Institutions			Central Statistical Bureau R&D data collection
MT	University of Malta					
NL	DUO (Dienst Uitvoering Onderwijs)		Vereniging Hogescholen + VSNU (Association of Universities in the Netherlands)	DUO (Dienst Uitvoering Onderwijs)	VSNU (Association of Universities in the Netherlands)	m
NO	DBH - Database on Higher education		NIFU/Register og research personnel + DBH	DBH - Database on Higher education	DBH + NIFU/Doctoral degree register	NIFU/R&D statistics
PL	Ministry of Science and HE		-	Ministry of Science and HE	Ministry of Science and HE	Ministry of Science and HE
PT	DGEEC- Direção-Geral de Estatísticas da Educação e Ciência					m
SE	Swedish Higher Education Authority		Swedish Higher Education Authority + Statistics Sweden	data not published	data not published	Statistics Sweden
SK	Ministry of Education, Science, Research and Sport of the Slovak Republic					m
UK	Higher Education Statistics Agency (HESA)					m

* Flanders part

9.2 FREQUENCY OF DATA COLLECTION

ETER performs yearly data collection. Four waves of data collection (2014, 2015, 2016 and 2017) have been concluded, for the baseline years 2011, 2012, 2013 and 2014.

9.3 DATA COLLECTION

ETER is a second level collection gathering information already available at National Statistical Offices and merging in a comparable and usable database at European level. The implications of this two stage process on data quality are recalled in Chapter 7 of ETER handbook.

The ETER project has developed an infrastructure which allows for standardization and systematization of the process of data collection, preliminary validation and data management. This is highly important in order to master the complexity of a data collection process that requires cooperation with the National Statistical Authorities and to ensure a sufficient level of uniformity in the data.

The infrastructure used in ETER includes:

- Templates for data collection, including documentation (e.g. flags and special values as commonly used for EUROSTAT-statistics), which guide national data providers (statistical offices, national authorities, other sources) and country experts addressing and supporting national data sources.
- Procedures for the preliminary data validation in order to detect mistakes and inconsistencies.
- A master database including upload interface and documentation of the database, which also constitutes the basis of the online tool for access to data by final users.

9.4 VALIDATION

The data collection is performed in Excel. The bases of the data collection templates for the reference year are the files from the preceding year.

Starting from the second wave of collection performed, the excel files already include a number of checks in order to call attention to NSAs/NEs and thus enable them to correct the data immediately.

The data collection files include prefilled cells with the previous year's values in order to support NSAs/NEs and reduce the burden for data collection. These include information that is not expected to change systematically from year to year, like descriptors, nominal variables, which are not likely to change, and the resulting not applicable values (e.g. no ISCED 5 level in a country, then "a" is prefilled in the respective categories), flags (except inconsistence "ic") and metadata.

As explained in Section 2.5, Data are validated through a two stages process including the followings:

1. Pre-validation, ensured by checks performed within the data collection phase on a country basis in order to allow for an easy return on the respondents and the correction of data before online integration. This step includes both automatic alerting rules embedded in the data collection files and format accuracy and consistency checks performed through an R script;

2. Global quality validation phase which has the role of performing more complex controls on data at both "global" and "local" levels. This step includes format accuracy and consistency checks, outlier detection, cross country and cross year comparability checks. When relevant, mainly to control for specific cases detected by previous methods, figures have been controlled against alternative source of data to solve possible doubts.

10 CONFIDENTIALITY

In principle, ETER data are public. In order to clarify in detail the status of data and possible restrictions with the original data providers, i.e. National Statistical Offices, they were asked to explicitly agree to the publication of the data they delivered.

Three possibilities are foreseen:

Public-data: data can be disclosed to the general public at the individual HEI-level, for example through the European Commission website. Use of these data is not subject to restrictions or to authorization, but users are requested to indicate the original source of data;

Restricted data: access can be granted for research purposes and carrying statistical analyses, on the condition that direct identification of statistical units in the final product is not possible;

Confidential data: data are available at the national provider, but due to confidentiality rules applying at national level they cannot be disclosed and therefore are not transferred to ETER.

As shown by Table 22, most NSAs informed the consortium that the delivered data were already public on national websites and therefore, no restrictions applied. Publication of data at the European level therefore largely reflects the on-going process of transparency in most European countries. In a few cases, NSAs asked for consent to publish data directly to the concerned HEIs.

For the countries that signed an agreement, two types of restrictions emerged:

•   On one hand, financial data (expenditures, revenues, R&D expenditures) is restricted in a few countries. This might apply to other countries that deliver these data in the future.

•   On the other hand, restrictions for a few individual HEIs apply in two cases: private HEIs in some countries (due to their legal status) and a few HEIs which did not give consent to deliver data in countries where NSAs had to ask permission directly, as there was no legal basis (Bulgaria, Lithuania).

In the public version of the dataset, data that was delivered, but restricted by a data disclosure agreement, were replaced by "c."

Moreover, ETER adopted usual statistical practices concerning data that are below some threshold, which would allow the identification of individuals, specifically for data on students and staff. To this aim, all cells below or equal to a count of 3 are set to "s" in the publicly available data. For breakdowns, the unclassified category is set to "s" in order to avoid the reconstruction of the concerned value by using the totals. The original data remain available for restricted purposes.

TABLE 22. STATUS OF DATA DISCLOSURE

Country	Data disclosure agreement signed	Specific issues by variable	Specific issue by HEI
Austria	yes	All data are already public, no data disclosure agreement required.	Confirmed by NSA.
Belgium (Flemish)	yes	Staff and all financial data are confidential	
Bulgaria	yes	None.	Data restricted for 3 HEIs which did not give consent.
Croatia	yes	All financial data (including R&D expenditures) are restricted.	
Cyprus	yes	All data are public	
Czech Republic	yes	None.	
Denmark	yes	All data are public	Confirmation by national expert.
Estonia	yes	All data are public.	Confirmed by NSA.
Finland	yes	All data are public	
Former Yugoslav Republic of Macedonia	no	All data are public.	All data were retrieved from public sources by the consortium.
France	yes	All data are public	Confirmed by NSA.
Germany	yes	All data are public	
Greece	yes	R&D expenditures are restricted only for research purposes.	
Hungary	yes	All data are public	
Iceland	no	Staff data are restricted access.	Data have been collected from the Website
Ireland	yes	No restrictions, all data are public.	Confirmed by NSA.
Italy		Financial data are confidential.	
Latvia	yes	Data on expenditures and revenues are restricted access only for scientific purposes.	One HEI did not provide consent to publication of data.
Liechtenstein	yes	None. All data are public.	
Lithuania	yes	None.	Data restricted for 4 HEIs which did not provide consent
Luxembourg	yes	All data are public.	Confirmed by NSA.
Malta		All data are public.	Confirmed by NSA.
Netherlands	yes	All data are public.	
Norway	yes	All data are already public, no data disclosure agreement required.	
Poland	yes	Financial data (including R&D) are restricted for the private HEIs	
Portugal	yes	All data are public.	
Romania	no		Data not delivered.
Serbia	yes	All data are public	Confirmed by NSA.
Slovakia	yes	All data are public.	Confirmed by NSA.
Slovenia	no		Data not delivered.
Spain	yes	All data are public.	Confirmed by NSA.
Sweden	yes	No restrictions, all data are public.	
Switzerland	yes	None. All data are public.	
Turkey	no		Data not delivered. Public data collected by the consortium
United Kingdom	yes	All data are public.	

11 OVERALL CONCLUSIONS AND RECOMMENDATIONS

Overall, the approach to data quality adopted by ETER has proved to be effective in managing a complex process of data collection from different countries and sources. The current version of the database, including the four waves of the ETER data collection, reaches good quality levels. In particular, internal accuracy, consistency and completeness of the data are overall good and most of the problems encountered found an explanation and have been systematically documented. The following practices have been particularly helpful in this respect:

- First, having a quality methodology spanning over all the data collection and processing steps. Indeed, the specific data quality methodology proposed for the ETER project has three main stages where quality activities have been identified, namely: "data collection", "pre-validation" and "validation and correction".
- Second, the combination of deterministic and statistical approaches to perform validation and correction. Indeed, ETER quality methodology does suggest the combination of simpler internal validity checks with computation of "quantitative" quality indicators, statistical testing and external checks. This makes it unlikely that problems are undetected.
- Third, the use of data flags to specifically annotate data problems and deviations from definitions is a suitable way to make users aware of the problems detected, but which cannot be corrected in the database.
- Fourth, the systematic use of metadata to explain observed problems, which are due to the underlying characteristics of data sources. This required in some cases an update to the collected metadata, since their quality and level of detail varies between countries.

Analyses on ETER data can be effectively performed in a quality-aware way. Indeed, the availability of quality information, as detailed throughout this report, does permit analysts to:

- Consider the level of completeness that may affect different variables and countries.
- As mentioned, conduct flag-aware analysis by filtering, for instance, uncorrected values.
- Figure out phenomena interpretation starting from outlier detection results, as well as carrying out new outlier detection analyses.

Short term Recommendations

In the light of the results summarised above, we briefly outline the following short term recommendations for the refinement of the Data Quality Process for the European Tertiary Education Register:

- Introducing a preliminary questionnaire to gather information from NSAs about the raw data they have available and their definition and contents in order to track and assess the match with ETER definition and categories (particularly relevant for financial variables).
- Improving the tracking of the changes introduced in the ETER database over time.
- Assessing internal consistency and stability of the flagging rules across countries and waves of data collection.
- Improving the user needs assessment providing measures of user satisfaction.

Long term Recommendations

In the following we briefly outline some long term recommendations on the development of the Data Quality Methodology for the European Tertiary Education Register:

- Investing on semantic modelling for industrializing the data quality framework. As better explained in the following, this investment would provide several benefits, including facing and overcoming the comparability problems

(through a Semantic Interoperability design and implementation analysis, along the lines of Catarci, Daraio and Scannapieco, 2015)[8];

- Development of the data quality framework (quality definition, quality metrics and quality assessment);
- Industrializing and standardizing quality reporting, by supporting a more accurate user analysis ("quality-aware" exploitation) of the European HEIs' microdata and by improving the opening and standardization of the quality reporting.

Investing on semantic modelling for industrializing the data quality framework

The ETER project deals with complex and heterogeneous metadata. We have remarked the several comparability problems that do exist in the data collection. The definition of an ontology as a shared representation of ETER concepts would greatly help to overcome these problems on one side, but also to manage the quality framework in a much more industrialized way.

Indeed, given that ontologies permit to *formally* represent concepts and variables it could be a valuable investment to invest on a shared "global" ontology for ETER concepts to which "mapping" the single files collected by NSAs. In other words, without impacting on how the collection is performed, an effort could be spent on how bring back to concepts and variables of the global ontologies those represented in the data collection files. This step can be done by using "mapping languages" to specify such mappings in a machine actionable way. This approach would univocally identify the comparability issues and would allow to engage a virtuous collaborative approach to progressively and systematically face with them.

On the other side, linking ETER data to an ontology would allow to perform more easily some data quality checks like the consistency and the completeness ones.

Finally, the data publishing phase could also greatly benefit from the use of an ontology. In particular, it could be expressed in terms of Web Semantics Standards, and in particular in OWL (Web Ontology Language). This would be a valuable result by itself: users of the ETER portal would be able to proper understand the concepts and variables relying the explicit representation of those that the ontology provides. In addition, it could be evaluated the possibility to publish ETER data as Linked Open Data (LOD). The advantages of this approach to data publication are several and range from the machine-to-machine data availability, to the navigational querying possibility that are inherent to the graph data model underlying LOD, to the adoption of state-of-the-art standards that allow to use bunch of tools for instance for data navigation and visualization.

Development of the Data Quality Framework

The ETER project is strongly founded on a proper data quality framework; however, such a framework, though being at a good stage of definition, still needs a proper industrialization.

As mentioned, the use of an ontology would allow to express consistency and completeness checks in terms of constraints over the ontology. In this way, part of the validation process should not be done "programmatically" but could rely on the data representation itself. This would permit obvious advantages in terms of maintainability of these quality checks in time.

In addition, the outlier detection approach could be engineered by developing a sets of reports that could be human controlled. A database could support this phase so that from one year to another the human controllers could access reports and interpret the detected outliers with a multi-year vision.

Industrializing and Standardizing Quality Reporting

---

[8] Catarci, T. Daraio C. Scannapieco M (2015), Implications of an OBDM approach for Data Quality, in *Efficiency, Effectiveness and Impact of Research and Innovation*, Proceedings of the Workshop of the 20 February 2015, DIAG Sapienza University of Rome, edited by C. Daraio, Efesto Edizioni, pp. 54-58

As described in Section 2.5.3, an important step of the data quality process is a documentation of the quality evaluation of the data set through the provision of specific flags and notations accompanying the data. A significant effort in identifying relevant flags and performing the data flagging process has been performed within ETER. A possible further investment could be made on a system supporting the whole annotation process as well as an easy access to annotated data. In particular, it would be relevant to have a system enabling both: (i) an easy annotation of the data according to the defined flags and (ii) a proper visualization of the flagged data in order to better support analyses taking them into account. The possibility of performing quality-aware analyses on ETER data is a real add-on of such data, and providing a specific support to make easier such analyses would be really helpful to the users.

As a further issue, we highlight the possibility of standardizing quality reporting and making them machine-actionable. This would imply a representation of such reports according to standard languages (e.g. SDMX) so to facilitate the accessibility and interpretability of the reports. It would be also an interesting issue to investigate the automatic generation of standard quality reports starting from the quality processing results.

APPENDIX 1: QUALITY REPORT TEMPLATE RECOMMENDED FOR FUTURE DATA COLLECTIONS

This section summarizes the main sections of the Quality Report developed for the ETER data collection and describes a template for ETER quality reporting that is conforming to ESS Guidelines. The resulting Quality Report template is reported in Table A1 with each composing section and related description. It is recommended for future data collection, to accompany the publication of data and certify their quality standards.

Table A1. ETER Quality Report's structure Recommended for future data collections

Quality Report Section Name	Description
Introduction	• General description of the process and its outputs • A brief history of the statistical process and outputs in question. • The broad statistical domain to which the outputs belong; related statistical outputs. References to other documentation, especially on methodology.
Relevance, assessment of user needs and perceptions	• User needs and satisfaction: Description of users and their respective needs with respect to the statistical data. Measures to determine user satisfaction • Completeness: The extent to which all statistics that are needed are available.
Accuracy	• Overall Accuracy: assessment of accuracy linked to a certain data set or domain, that is summarising the various components.
Timeliness and punctuality	Timeliness: describes the length of time between data availability and the event or phenomenon they describe.
Accessibility and clarity	**Accessibility** and **clarity** refer to the simplicity and ease, the conditions and modalities by which users can access, use and interpret statistics, with the appropriate supporting information and assistance. • On-line database: description of information about on-line databases in which the disseminated data can be accessed. • Documentation on methodology: Descriptive text and references to methodological documents available. • Quality documentation: Documentation on procedures applied for quality management and quality assessment.
Coherence and comparability	**Coherence** measures the adequacy of the statistics to be combined in different ways and for various uses. • Coherence – internal, in particular, measures the extent to which statistics are consistent within a given data set. Here we report the outcome of the outliers analysis. • Comparability – geographical, in particular, measures the extent to which statistics are comparable between geographical areas.
Cost and burden	**Cost and burden** is the cost associated with the collection and production of a statistical product and burden on respondents. • Costs: provide a summary of costs for production of statistical data. Describe efforts made to improve efficiency and the extent to which information and communications technology (ICT) is effectively used in the statistical process. • Response burden, where available, an estimate of respondent burden (in general measured in time used) should be reported as well as recent efforts made to reduce respondent burden (e.g. electronic means are used to facilitate data collection).

Quality Report Section Name	Description
Confidentiality	**Confidentiality** is a property of data indicating the extent to which their un-authorised disclosure could be prejudicial or harmful to the interest of the source or other relevant parties. • Confidentiality - data treatment: Rules applied for treating the data set to ensure statistical confidentiality and prevent un-authorised disclosure.
Statistical processing	**Statistical processing** refers to the operations performed on data to derive new information according to a given set of rules. • Source data: Characteristics and components of the raw statistical data used for compiling statistical aggregates. • Frequency of data collection: Frequency with which the source data are collected. • Data collection: Systematic process of gathering data for official statistics. • Data validation: Process of monitoring the results of data compilation and ensuring the quality of statistical results. • Data compilation: Operations performed on data to derive new information according to a given set of rules

APPENDIX 2 COMPLETENESS BY COUNTRY AND BY VARIABLE

Table A2: Completeness by country and by variable

Country	AT 2011	AT 2012	AT 2013	AT 2014	BE 2011	BE 2012	BE 2013	BE 2014	BG 2011	BG 2012	BG 2013	BG 2014	CH 2011	CH 2012	CH 2013	CH 2014	CY 2011	CY 2012	CY 2013	CY 2014	CZ 2011	CZ 2012	CZ 2013	CZ 2014	DE 2011	DE 2012	DE 2013	DE 2014
BAS.ETERIDYEAR	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BAS.ETERID	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BAS.NATID	-	-	-	-	-	-	-	-	1	1	1	1	-	-	-	-	-	-	-	-	1	1	1	1	1	1	1	1
BAS.INSTNAME	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BAS.REFYEAR	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BAS.ACRONYM	0.59	0.59	0.59	0.59	0.58	0.58	0.81	0.81	0.77	0.77	0.77	0.77	0.97	1	1	1	0.88	0.92	0.92	1	0.96	0.99	1	1	0.69	0.69	0.69	0.71
BAS.COUNTRY	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BAS.LEGALSTAT	1	1	1	1	0.90	0.90	0.88	0.88	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BAS.INSTCATSTAND	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BAS.FORCAMP	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BAS.FOUNDYEAR	0.97	0.97	0.97	0.97	1	1	1	1	1	1	1	1	0.73	0.77	0.77	0.77	1	1	1	1	1	1	1	1	1.00	1.00	0.99	0.99
BAS.LEGALYEAR	0.97	0.97	0.97	0.97	1	1	1	1	1	1	1	0.90	1	1	1	1	1	1	1	1	1	1	1	1	0.17	0.10	0.10	0.09
BAS.ANCESTYEAR	0.56	0.56	0.56	0.56	0.48	0.48	0.54	0.54	0.08	0.08	0.08	0.08	0.34	0.32	0.32	0.32	0.04	0.04	0.04	0.04	0.75	0.74	0.74	0.74	0.01	0.00	0	0
BAS.UNIHOSP	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BAS.WEBSITE	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.99	0.99	0.99	0.97
GEO.NUTS2	1	1	1	1	1	1	0.96	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
GEO.CITY	1	1	1	1	1	1	0.96	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
GEO.POSTCODE	1	1	1	1	1	1	0.96	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
GEO.MULTISITE	1	1	1	1	1	1	0.96	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
GEO.NUTS3MULTISITE	1	1	1	1	1	1	0.92	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.99	0.99	1	1
EXP.CURRPERSON	0	0	0	0	0.87	0.87	0.85	0.85	0	0	0	0	0.94	0.94	0.94	0.91	0.78	0.68	0.80	0.87	0.36	0.37	0.37	0.37	0.99	0.99	0.99	0.97
EXP.CURRNONPERSON	0	0	0	0	0.87	0.87	0.85	0.85	0	0	0	0	0.94	0.94	0.94	0.91	0.78	0.68	0.80	0.87	0.36	0.37	0.37	0.37	0.99	0.99	0.99	0.97
EXP.CURRUNCL	0	0	0	0	0.58	0.87	0.85	0.85	0	0	0	0	0.94	0.94	0.94	0.91	0.78	0.68	0.80	0.87	0.36	0.37	0.37	0.37	0.99	0.99	0.99	0.97
EXP.CURRTOTAL	0	0	0	0	0.87	0.87	0.85	0.85	0	0	0	0	0.94	0.94	0.94	0.94	0.78	0.68	0.80	0.87	0.36	0.37	0.37	0.37	0.99	0.99	0.99	0.97
EXP.CAPITAL	0	0	0	0	0.87	0.87	0.85	0.85	0	0	0	0	0	0	0	0	0.78	0.68	0.80	0.87	0	0	0	0	0.99	0.99	0.99	0.97
EXP.ACCSYSTEM	0	0	0	0	0	0	0	0.88	0	0	0	0	0.94	0.94	1	1	0.74	0.72	0.80	0.87	0.36	0.37	0.37	0.37	0	0	0	0
REV.COREBUDGETPUBLIC	0	0	0	0	0	0	0	0.85	0	0	0	0	0.91	0.91	0.91	0.91	0	0	0	0	0	0	0	0.85	0	0	0	0
REV.COREBUDGETOTHER	0	0	0	0	0	0	0	0.85	0	0	0	0	0.91	0.91	0.91	0.91	0	0	0	0	0	0	0	0.37	0	0	0	0
REV.CORETOTAL	0	0	0	0	0.87	0.87	0.85	0.85	0	0	0	0	0.94	0.94	0.94	0.91	0.14	0.13	0.13	0.14	0.83	0.86	0.86	0.37	0.99	0.99	0.99	0.97
REV.THIRDPARTYPUBLIC	0	0	0	0	0	0	0	0.54	0	0	0	0	0.94	0.94	0.94	0.91	0.14	0.13	0.13	0.14	0	0	0	0	0.99	0.99	0.99	0.97
REV.THIRDPARTYPRIVATE	0	0	0	0	0	0	0	0.19	0	0	0	0	0.94	0.94	0.94	0.91	0.14	0.13	0.13	0.14	0	0	0	0	0.99	0.99	0.99	0.97
REV.THIRDPARTYABROAD	0	0	0	0	0	0	0	0.54	0	0	0	0	0.91	0.91	0.91	0.91	0	0	0	0	0	0	0	0	0	0	0	0.97

Country	AT 2011	AT 2012	AT 2013	AT 2014	BE 2011	BE 2012	BE 2013	BE 2014	BG 2011	BG 2012	BG 2013	BG 2014	CH 2011	CH 2012	CH 2013	CH 2014	CY 2011	CY 2012	CY 2013	CY 2014	CZ 2011	CZ 2012	CZ 2013	CZ 2014	DE 2011	DE 2012	DE 2013	DE 2014
REV.THIRDPARTYUND	0	0	0	0	0	0	0	0.54	0	0	0	0	0.94	0.91	0.91	0.91	0	0	0	0	0	0	0	0	0	0	0	0.97
REV.THIRDPARTYTOTAL	0	0	0	0	0.87	0.00	0.85	0.85	0	0	0	0	0.94	0.94	0.94	0.91	0.14	0.13	0.13	0.14	0	0	0	0	0.99	0.99	0.99	0.97
REV.TUITFEES	0	0	0	0	0.87	0.90	0.88	0.88	0	0	0	0	1	1	1	1	1	1	1	1	0	0	0.94	1	0	0	0	0
REV.STUDFEES.NC	0	0	0	0	0.87	0.87	0.85	0.85	0	0	0	0	0.94	0.94	0.94	0.91	0.14	0.13	0.13	0.14	0	0	0	0	0.99	0.99	0.99	0.97
REV.UNCL.NC	0	0	0	0	0.84	0.84	0.85	0.85	0	0	0	0	0.94	0.94	0.94	0.91	0.14	0.13	0.13	0.14	0.36	0.37	0.37	0	0.99	0.99	0.99	0.97
REV.CURRTOTAL.NC	0	0	0	0	0.87	0.87	0.85	0.85	0	0	0	0	0.94	0.94	0.94	0.91	0.77	0.67	0.79	0.86	0.36	0.37	0.37	0	0.99	0.99	0.99	0.97
REV.NONRECURR.NC	0	0	0	0	0	0	0	0.88	0	0	0	0	0.91	0.94	0.94	0.91	0	0	0	0	0	0	0	0	0	0	0	0
STA.ACAFTETOTAL	0	0	0	0	0.87	0.87	0.85	0.85	0.98	0.98	0.98	0.98	1	1	1	1	0.96	0.96	0.92	1	0.51	0.51	0.66	1	1.00	1	1	0.99
STA.ACAHCMEN	0	0	0	0	0.87	0.87	0.85	0.85	0.98	0.98	0.98	0.98	1	1	1	1	0.96	0.96	0.92	1	0.51	0.51	0.66	1	1.00	1	1	0.99
STA.ACAHCNAT	0	0	0	0	0.19	0.16	0.19	0.19	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0	1.00	1	1	0.99
STA.ACAHCFOE05	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0	1.00	1	1	0.99
STA.TOTACAHC	0	0	0	0	0.87	0.87	0.85	0.85	0.98	0.98	0.98	0.98	1	1	1	1	0.96	0.96	0.92	1	0.51	0.51	0.66	1	1.00	1	1	0.99
STA.PROFMEN	0	0	0	0	0	0.19	0.19	0.85	0.98	0.98	0.98	0.98	1	1	1	1	0.87	0.92	0.88	1	0	0	0.06	1	1.00	1	1	0.99
STA.PROFTOTAL	0	0	0	0	0.19	0.19	0.19	0.85	0.98	0.98	0.98	0.98	1	1	1	1	0.87	0.92	0.88	1	0.51	0.51	0.66	1	1.00	1	1	0.99
STA.INCLPHDSTUD	0	0	0	0	0.87	0.90	0.88	0.88	0.98	1	1	1	1	1	1	1	0.96	0.96	0.92	1	0.96	0.94	0.66	1	1	1	1	0.99
STA.NONACAFTE	0	0	0	0	0.87	0.87	0.85	0.85	0.98	0.98	0.98	0.98	1	1	1	1	0.96	0.96	0.92	1	0.51	0.51	0.66	1	1.00	1	1	0.99
STA.NONACAHC	0	0	0	0	0.87	0.87	0.85	0.85	0.98	0.98	0.98	0.98	1	1	1	1	0.96	0.96	0.92	1	0.51	0.51	0.66	1	1.00	1	1	0.99
STA.TOTALFTE	0	0	0	0	0.87	0.87	0.85	0.85	0.98	0.98	0.98	0.98	1	1	1	1	0.96	0.96	0.92	1	0.51	0.51	0.66	1	1.00	1	1	0.99
STA.TOTALHC	0	0	0	0	0.87	0.87	0.85	0.85	0.98	0.98	0.98	0.98	1	1	1	1	0.96	0.96	0.92	1	0.51	0.51	0.66	1	1.00	1	1	0.99
STUD.LOWDEG	1	1	1	1	0.90	0.90	0.88	0.88	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
STUD.HIGHDEG	1	1	1	1	0.90	0.90	0.88	0.88	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.99
STUD.ISCED5MEN	1	1	1	1	0	0	0	0	-	-	-	-	-	-	-	-	1	1	0.92	1	-	-	-	-	-	-	-	-
STUD.ISCED5NAT	1	1	1	1	0	0	0	0	-	-	-	-	-	-	-	-	1	1	0.92	1	-	-	-	-	-	-	-	-
STUD.ISCED5RES	1	1	1	1	0	0	0	0	-	-	-	-	-	-	-	-	1	1	0.92	1	-	-	-	-	-	-	-	-
STUD.ISCED5FOE05	0	0	0	0	0	0	0	0	-	-	-	-	-	-	-	-	1	1	0.92	1	-	-	-	-	-	-	-	-
STUD.ISCED5TOTAL	1	1	1	1	0	0	0	0	-	-	-	-	-	-	-	-	1	1	0.92	1	-	-	-	-	-	-	-	-
STUD.ISCED6MEN	1	1	1	1	0.90	0.90	0.88	0.88	1	1	1	1	1	1	1	1	1	0.95	0.90	1	0.97	1	1	1	0.99	0.99	0.99	0.98
STUD.ISCED6NAT	1	1	1	1	0.90	0.90	0.88	0.88	1	1	1	1	1	1	1	1	1	0.95	0.90	1	0.97	1	0.97	1	0.99	0.99	0.99	0.98
STUD.ISCED6RES	1	1	1	1	0.90	0.90	0.88	0.88	1	1	1	1	1	1	1	1	1	0.95	0.90	1	0	0	0.03	0	0.99	0.99	0.99	0.98
STUD.ISCED6FOE05	0	0	0	0	0.90	0.90	0.88	0.88	1	1	1	1	1	1	1	1	1	0.95	0.90	1	0	1	0.97	1	0.99	0.99	0.99	0.98
STUD.ISCED6TOTAL	1	1	1	1	0.90	0.90	0.88	0.88	1	1	1	1	1	1	1	1	1	0.95	0.90	1	0.97	1	1	1	0.99	0.99	0.99	0.98
STUD.ISCED7MEN	1	1	1	1	0.89	0.89	0.82	0.81	1	1	1	1	1	1	1	1	0.94	1	1	1	0.96	1	1	1	0.99	0.99	0.99	0.98
STUD.ISCED7NAT	1	1	1	1	0.89	0.89	0.82	0.81	1	1	1	1	1	1	1	1	0.94	1	1	1	0.96	1	0.96	1	0.99	0.99	0.99	0.98
STUD.ISCED7RES	1	1	1	1	0.89	0.89	0.82	0.81	0	0	0	0	1	1	1	1	0.94	1	1	1	0	0	0.02	0	0.99	0.99	0.99	0.98
STUD.ISCED7FOE05	0	0	0	0	0.89	0.89	0.82	0.81	1	1	1	1	1	1	1	1	0.94	1	1	1	0	1	0.96	1	0.99	0.99	0.99	0.98

Country	AT 2011	AT 2012	AT 2013	AT 2014	BE 2011	BE 2012	BE 2013	BE 2014	BG 2011	BG 2012	BG 2013	BG 2014	CH 2011	CH 2012	CH 2013	CH 2014	CY 2011	CY 2012	CY 2013	CY 2014	CZ 2011	CZ 2012	CZ 2013	CZ 2014	DE 2011	DE 2012	DE 2013	DE 2014
STUD.ISCED7TOTAL	1	1	1	1	0.89	0.89	0.82	0.81	1	1	1	1	1	1	1	1	0.94	1	1	1	0.96	1	1	1	0.99	0.99	0.99	0.98
STUD.ISCED7LONGMEN	1	1	1	1	0	0	0	0	1	1	1	1	1	1	1	1	0.94	1	1	1	0.96	1	1	1	0.99	0.99	0.99	0.98
STUD.ISCED7LONGNAT	1	1	1	1	0	0	0	0	1	1	1	1	1	1	1	1	0.94	1	1	1	0.96	1	0.98	1	0.99	0.99	0.99	0.98
STUD.ISCED7LONGRES	1	1	1	1	0	0	0	0	1	1	1	1	1	1	1	1	0.94	1	1	1	0	0	0.41	0	0.99	0.99	0.99	0.98
STUD.ISCED7LONGFOE05	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	0.94	1	1	1	0	1	0.98	1	0.99	0.99	0.99	0.98
STUD.ISCED7LONGTOTAL	1	1	1	1	0	0	0	0	1	1	1	1	1	1	1	1	0.94	1	1	1	0.96	1	1	1	0.99	0.99	0.99	0.98
STUD.TOTALISCED5.7	1	1	1	1	0.90	0.90	0.88	0.88	1	1	1	1	1	1	1	1	0.95	0.96	0.96	1	0.97	1	1	1	0.99	0.99	0.99	0.98
STUD.DISTEDUINST	1	1	1	1	0.90	0.90	0.88	0.88	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
GRAD.ISCED5MEN	1	1	1	1	0	0	0	0	-	-	-	-	-	-	-	-	0.93	1	0.92	1	-	-	-	-	-	-	-	-
GRAD.ISCED5NAT	1	1	1	1	0	0	0	0	-	-	-	-	-	-	-	-	0.93	1	0.92	1	-	-	-	-	-	-	-	-
GRAD.ISCED5RES	1	1	1	1	0	0	0	0	-	-	-	-	-	-	-	-	0.93	1	0.92	1	-	-	-	-	-	-	-	-
GRAD.ISCED5FOE05	0	0	0	1	0	0	0	0	-	-	-	-	-	-	-	-	0.93	1	0.92	1	-	-	-	-	-	-	-	-
GRAD.ISCED5TOTAL	1	1	1	1	0	0	0	0	-	-	-	-	-	-	-	-	0.93	1	0.92	1	-	-	-	-	-	-	-	-
GRAD.ISCED6MEN	1	1	1	1	0.90	0.90	0.88	0.88	1	1	1	1	1	1	1	1	1	0.95	0.90	1	0.93	1	1	1	0.99	0.96	0.97	0.96
GRAD.ISCED6NAT	1	1	1	1	0.90	0.90	0.88	0.88	1	1	1	1	1	1	1	1	1	0.95	0.90	1	0.93	1	0.97	1	0.99	0.96	0.97	0.96
GRAD.ISCED6RES	1	1	1	1	0.90	0.90	0.88	0.88	0	0	0	0	1	1	1	1	1	0.95	0.90	1	0	0	0.06	0	0.99	0.96	0.97	0.96
GRAD.ISCED6FOE05	0	0	0	0	0.90	0.90	0.88	0.88	1	1	1	1	1	1	1	1	1	0.95	0.90	1	0	1	0.97	1	0.99	0.96	0.97	0.96
GRAD.ISCED6TOTAL	1	1	1	1	0.90	0.90	0.88	0.88	1	1	1	1	1	1	1	1	1	0.95	0.90	1	0.94	1	1	1	0.99	0.96	0.97	0.96
GRAD.ISCED7MEN	1	1	1	1	0.89	0.89	0.82	0.81	1	1	1	1	1	1	1	1	0.94	1	1	1	0.96	0.96	1	1	0.99	0.96	0.97	0.97
GRAD.ISCED7NAT	1	1	1	1	0.89	0.89	0.82	0.81	1	1	1	1	1	1	1	1	0.94	1	1	1	0.96	0.96	0.96	1	0.99	0.96	0.97	0.97
GRAD.ISCED7RES	1	1	1	1	0.89	0.89	0.82	0.81	0	0	0	0	1	1	1	1	0.94	1	1	1	0	0	0.04	0	0.99	0.96	0.97	0.97
GRAD.ISCED7FOE05	0	0	0	0	0.89	0.89	0.82	0.81	1	1	1	1	1	1	1	1	0.94	1	1	1	0	1	0.96	1	0.99	0.96	0.97	0.97
GRAD.ISCED7TOTAL	1	1	1	1	0.89	0.89	0.82	0.81	1	1	1	1	1	1	1	1	0.94	1	1	1	0.96	1	1	1	0.99	0.96	0.97	0.97
GRAD.ISCED7LONGMEN	1	1	1	1	0	0	0	0	1	1	1	1	1	1	1	1	0.94	1	1	1	0.96	0.96	1	1	0.99	0.96	0.97	0.97
GRAD.ISCED7LONGNAT	1	1	1	1	0	0	0	0	1	1	1	1	1	1	1	1	0.94	1	1	1	0.96	0.96	0.98	1	0.99	0.96	0.97	0.97
GRAD.ISCED7LONGRES	1	1	1	1	0	0	0	0	0	0	0	0	1	1	1	1	0.94	1	1	1	0	0	0.57	0	0.99	0.96	0.97	0.97
GRAD.ISCED7LONGFOE05	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	0.94	1	1	1	0	1	0.98	1	0.99	0.96	0.97	0.97
GRAD.ISCED7LONGTOTAL	1	1	1	1	0	0	0	0	1	1	1	1	1	1	1	1	0.94	1	1	1	0.96	1	1	1	0.99	0.96	0.97	0.97
GRAD.TOTALISCED5.7	1	1	1	1	0.90	0.90	0.88	0.88	1	1	1	1	1	1	1	1	0.95	0.96	0.96	1	0.96	1	1	1	0.99	0.96	0.97	0.96
GRAD.NOTESEDUCACT	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
RES.RESACTIVE	1	1	1	1	0.90	0.90	0.88	0.88	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.99
RES.STUDISCED8MEN	1	1	1	1	0.60	0.60	0.67	0.67	1	1	1	1	1	1	1	1	1	1	1	1	0.93	1	1	1	0.97	0.97	0.97	0.95
RES.STUDISCED8NAT	1	1	1	1	0.60	0.60	0.67	0.67	1	1	1	1	1	1	1	1	1	1	1	1	0.93	1	0.93	1	0.97	0.97	0.97	0.95
RES.STUDISCED8RES	1	1	1	1	0.60	0.60	0.67	0.67	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0.97	0.97	0.97	0.95
RES.STUDISCED8FOE05	0	0	0	0	0.60	0.60	0.67	0.67	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0.93	1	0.97	0.97	0.97	0.95

Country	AT 2011	AT 2012	AT 2013	AT 2014	BE 2011	BE 2012	BE 2013	BE 2014	BG 2011	BG 2012	BG 2013	BG 2014	CH 2011	CH 2012	CH 2013	CH 2014	CY 2011	CY 2012	CY 2013	CY 2014	CZ 2011	CZ 2012	CZ 2013	CZ 2014	DE 2011	DE 2012	DE 2013	DE 2014
RES.STUDISCED8TOTAL	1	1	1	1	0.60	0.60	0.67	0.67	1	1	1	1	1	1	1	1	1	1	1	1	0.93	1	1	1	0.97	0.97	0.97	0.95
RES.GRADISCED8MEN	1	1	1	1	0.70	0.70	0.67	0.67	1	1	1	1	1	1	1	1	1	1	1	1	0.93	1	1	1	0.97	0.97	0.97	0.95
RES.GRADISCED8NAT	1	1	1	1	0.70	0.70	0.67	0.67	1	1	1	1	1	1	1	1	1	1	1	1	0.93	1	0.93	1	0.97	0.97	0.97	0.95
RES.GRADISCED8RES	1	1	1	1	0.70	0.70	0.67	0.67	0	0	0	0	1	1	1	1	1	1	1	1	0	0	0	0	0.97	0.97	0.97	0.95
RES.GRADISCED8FOE05	0	0	0	0	0.70	0.70	0.67	0.67	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0.93	1	0.97	0.97	0.97	0.95
RES.GRADISCED8TOTAL	1	1	1	1	0.70	0.70	0.67	0.67	1	1	1	1	1	1	1	1	1	1	1	1	0.93	1	1	1	0.97	0.97	0.97	0.95
RES.R.DEXP	0	0	0	0	0	0	0	0	0	0	0	0	0.94	0.94	0.94	0.91	1	1	1	0.75	0	0	0	0	0	0	0	0

Country	DK 2011	DK 2012	DK 2013	DK 2014	EE 2011	EE 2012	EE 2013	EE 2014	ES 2011	ES 2012	ES 2013	ES 2014	FI 2011	FI 2012	FI 2013	FI 2014	FR 2011	FR 2012	FR 2013	FR 2014	GR 2011	GR 2012	GR 2013	GR 2014	HR 2011	HR 2012	HR 2013	HR 2014
BAS.ETERIDYEAR	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BAS.ETERID	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BAS.NATID	1	1	1	1	-	-	-	-	-	-	-	-	1	1	1	1	1	1	1	1	-	-	-	-	-	-	-	-
BAS.INSTNAME	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BAS.REFYEAR	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BAS.ACRONYM	0.91	0.91	0.94	0.94	0.83	0.86	0.96	1	1	1	1	1	0.55	0.55	0.57	0.59	0.38	0.38	0.35	0.36	1	1	1	1	0.31	0.33	0.31	0.3
BAS.COUNTRY	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BAS.LEGALSTAT	0.91	0.91	0.94	0.94	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BAS.INSTCATSTAND	0.97	0.97	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BAS.FORCAMP	0.97	0.97	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0.01	0.02	0.02	1	1	1	1	1	1	1	1
BAS.FOUNDYEAR	1	1	1	1	1	1	1	1	1	1	1	1	0.89	0.89	0.88	0.88	0.99	0.99	1	1	0.98	0.98	0.98	0.98	1	1	1	1
BAS.LEGALYEAR	0.97	0.97	1	1	0	0	0	0	0	0	0	0	0.91	0.91	0.9	0.9	0	0	0	0	0.98	0.98	0.98	0.98	1	1	1	1
BAS.ANCESTYEAR	0.97	0.97	1	1	1	1	1	1	0	0.01	0.01	0.01	0.89	0.89	0.86	0.83	0.08	0.08	0.07	0.08	0.26	0.26	1	1	1	1	1	1
BAS.UNIHOSP	0.97	0.97	1	1	1	1	1	1	1	1	1	1	0.98	0.98	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BAS.WEBSITE	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.98	0.98	1	1	1	1	1	1	1	1	1	1
GEO.NUTS2	0.97	0.97	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
GEO.CITY	0.97	0.97	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
GEO.POSTCODE	0.97	0.97	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
GEO.MULTISITE	0.91	0.91	0.94	0.94	1	1	1	1	1	1	1	1	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1
GEO.NUTS3MULTISITE	0.89	0.89	0.92	0.92	1	1	1	1	1	1	1	1	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1
EXP.CURRPERSON	0.85	0.85	0.88	0	0	0	0	0	0	0	0	0	0.93	0.93	0.93	0.93	0	0.41	0.34	0.36	0	0	0	0	0	0.85	0.94	0.78
EXP.CURRNONPERSON	0.85	0.85	0.88	0	0	0	0	0	0	0	0	0	0.93	0.93	0.93	0.93	0	0.41	0.34	0.36	0	0	0	0	0	0.85	0.94	0.78
EXP.CURRUNCL	0.85	0.85	0.88	0	0	0	0	0	0	0	0	0	0.93	0.93	0.93	0.93	0	0.43	0.36	0.36	0	0	0	0	0	0.85	0.94	0.78

Country	DK 2011	DK 2012	DK 2013	DK 2014	EE 2011	EE 2012	EE 2013	EE 2014	ES 2011	ES 2012	ES 2013	ES 2014	FI 2011	FI 2012	FI 2013	FI 2014	FR 2011	FR 2012	FR 2013	FR 2014	GR 2011	GR 2012	GR 2013	GR 2014	HR 2011	HR 2012	HR 2013	HR 2014
EXP.CURRTOTAL	0.85	0.85	0.88	0	0	0	0	0	0	0	0	0	0.93	0.93	0.93	0.93	0	0.41	0.34	0.36	0	0	0	0	0	0.91	0.94	0.78
EXP.CAPITAL	0.38	0.38	0.36	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.37	0.34	0.36	0	0	0	0	0	0.85	0.94	0.7
EXP.ACCSYSTEM	0.85	0.85	0.88	0.88	0	0	0	0	0	0	0	0	0	0	1	1	0	0.41	0.45	0.36	0	0	0	0	0	0.7	0.89	0.92
REV.COREBUDGETPUBLIC	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.84
REV.COREBUDGETOTHER	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.84
REV.CORETOTAL	0.85	0.85	0.88	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.41	0.34	0.32	0	0	0	0	0	0.73	0.94	0.84
REV.THIRDPARTYPUBLIC	0.24	0.24	0.24	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.41	0.34	0.32	0	0	0	0	0	0.55	0.91	0.84
REV.THIRDPARTYPRIVATE	0.24	0.24	0.24	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.41	0.34	0.32	0	0	0	0	0	0.55	0.91	0.84
REV.THIRDPARTYABROAD	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.32	0	0	0	0	0	0	0	0.84
REV.THIRDPARTYUND	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.32	0	0	0	0	0	0	0	0.84
REV.THIRDPARTYTOTAL	0.24	0.24	0.24	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.41	0.34	0.32	0	0	0	0	0	0.67	0.94	0.84
REV.TUITFEES	0.85	0.88	0.91	0.91	1	1	1	1	0	0	0	0	0	0	1	1	0	0.51	0.51	0.43	1	1	1	0.98	0	0.85	1	1
REV.STUDFEES.NC	0.71	0.62	0.67	0	0	0	0	0	0	0	0	0	0	0	-	-	0	0.42	0.34	0.32	0	0	0	0	0	0.88	0.94	0.89
REV.UNCL.NC	0.85	0.85	0.88	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.41	0.34	0.32	0	0	0	0	0	0.91	0.94	0.86
REV.CURRTOTAL.NC	0.85	0.85	0.88	0	0	0	0	0	0	0	0	0	0.93	0.93	0.93	0.93	0	0.41	0.34	0.32	0	0	0	0	0	0.91	0.94	0.86
REV.NONRECURR.NC	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.85
STA.ACAFTETOTAL	0.24	0.24	1	0	0	0	0	0	1	1	1	1	0.93	0.93	0.93	0.93	0	0	0	0	0	0	0	0	1	1	0.97	1
STA.ACAHCMEN	0	0	1	0	0	0	0	0	1	1	1	1	0.93	0.93	0.93	0.93	0	0	0	0.4	0.92	0.92	0.91	0.9	1	1	0.97	1
STA.ACAHCNAT	0	0	1	0	0	0	0	0	1	1	1	1	0.93	0.93	0.88	0.93	0	0	0	0.4	0	0	0	0	0	0	0	0
STA.ACAHCFOE05	0	0	0.06	0	0	0	0	0	0.85	0.87	0.87	0.85	0	0	0	0	0	0	0	0.4	0	0	0	0	0	0	0	0
STA.TOTACAHC	0	0	1	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0.4	0.92	0.92	0.91	0.9	1	1	0.97	1
STA.PROFMEN	0.24	0.62	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0.4	0.92	0.92	0.91	0.9	1	1	0.97	1
STA.PROFTOTAL	0.24	0.62	0	0	0	0	0	0	1	1	1	1	0.94	0.94	0	0	0	0	0	0.4	0.92	0.92	0.91	0.9	1	1	0.97	1
STA.INCLPHDSTUD	0.24	0.85	1	1	0	0	0	0	0.97	0.98	1	1	1	1	1	1	0	0	0	0.4	0.96	0.92	0.96	0.96	1	1	0.97	1
STA.NONACAFTE	0.24	0.24	1	0	0	0	0	0	1	1	1	1	0.93	0.86	0.93	0.93	0	0	0	0	0	0	0	0	0	0	0	0
STA.NONACAHC	0	0	1	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0	0.92	0.92	0.91	0	0	0	0	0
STA.TOTALFTE	0.24	0.24	1	0	0	0	0	0	1	1	1	1	0.93	0.86	0.93	0.93	0	0	0	0	0	0	0	0	0	0	0	0
STA.TOTALHC	0	0	1	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0	0.92	0.92	0.91	0	0	0	0	0
STUD.LOWDEG	0.91	0.91	0.94	0.94	1	1	1	1	1	1	1	1	1	1	1	1	0.76	0.76	0.78	0.78	1	1	1	1	1	1	1	1
STUD.HIGHDEG	0.91	0.91	0.94	0.94	1	1	1	1	1	1	1	1	1	1	1	1	0.76	0.76	0.78	0.78	1	1	1	1	1	1	1	1
STUD.ISCED5MEN	0.81	0.86	0.9	0	-	-	-	-	-	-	-	-	-	-	-	-	0.54	0.54	0.53	0.52	-	-	-	-	0	0	1	1
STUD.ISCED5NAT	0.81	0.86	0.9	0	-	-	-	-	-	-	-	-	-	-	-	-	0.54	0.54	0.53	0.52	-	-	-	-	0	0	1	1
STUD.ISCED5RES	0.81	0.86	0.9	0	-	-	-	-	-	-	-	-	-	-	-	-	0.54	0.54	0.53	0.52	-	-	-	-	0	0	0	1
STUD.ISCED5FOE05	0.81	0.86	0.9	0	-	-	-	-	-	-	-	-	-	-	-	-	0.54	0.54	0.53	0.52	-	-	-	-	0	0	0	1
STUD.ISCED5TOTAL	0.81	0.86	0.9	0	-	-	-	-	-	-	-	-	-	-	-	-	0.54	0.54	0.53	0.52	-	-	-	-	0	0	1	1

Country	DK 2011	DK 2012	DK 2013	DK 2014	EE 2011	EE 2012	EE 2013	EE 2014	ES 2011	ES 2012	ES 2013	ES 2014	FI 2011	FI 2012	FI 2013	FI 2014	FR 2011	FR 2012	FR 2013	FR 2014	GR 2011	GR 2012	GR 2013	GR 2014	HR 2011	HR 2012	HR 2013	HR 2014
STUD.ISCED6MEN	0.91	0.91	0.94	0	1	1	1	1	1	1	1	1	1	1	1	1	0.61	0.61	0.61	0.59	0.94	0.94	0.93	0.94	1	1	1	1
STUD.ISCED6NAT	0.91	0.91	0.94	0	1	1	1	1	1	1	1	1	1	1	1	1	0.61	0.61	0.61	0.59	0.94	0.94	0.93	0	0	0	1	1
STUD.ISCED6RES	0.91	0.91	0.94	0	1	1	1	1	1	1	1	1	0	1	1	1	0.61	0.61	0.61	0.59	0	0	0	0	0	0	1	1
STUD.ISCED6FOE05	0.91	0.91	0.94	0	1	1	1	1	1	1	1	1	1	1	1	1	0.61	0.61	0.61	0.59	0.94	0.94	0.93	0.94	0	0	0	1
STUD.ISCED6TOTAL	0.91	0.91	0.94	0	1	1	1	1	1	1	1	1	1	1	1	1	0.61	0.61	0.61	0.59	0.94	0.94	0.93	0.94	0	0	1	1
STUD.ISCED7MEN	0.81	0.8	0.86	0	1	1	1	1	1	1	1	1	1	1	1	1	0.71	0.72	0.74	0.72	0.95	0.95	0.94	0.94	0	0	1	1
STUD.ISCED7NAT	0.81	0.8	0.86	0	1	1	1	1	1	1	1	1	1	1	1	1	0.71	0.72	0.74	0.72	0	0	0.11	0	0	0	1	1
STUD.ISCED7RES	0.81	0.8	0.86	0	1	1	1	1	1	1	1	1	0	1	0.88	0.88	0.71	0.72	0.74	0.72	0	0	0.11	0	0	0	1	1
STUD.ISCED7FOE05	0.81	0.8	0.86	0	1	1	1	1	1	1	1	1	0.95	1	1	1	0.71	0.72	0.74	0.72	0.95	0.95	0.94	0.94	0	0	0	1
STUD.ISCED7TOTAL	0.81	0.8	0.86	0	1	1	1	1	1	1	1	1	1	1	1	1	0.71	0.72	0.74	0.72	0.95	0.95	0.94	0.94	0	0	1	1
STUD.ISCED7LONGMEN	0.7	0.63	0.86	0	1	1	1	1	1	1	1	1	0	0	1	1	0.72	0.73	0.71	0.7	0	0	1	1	0	0	1	1
STUD.ISCED7LONGNAT	0.7	0.63	0.86	0	1	1	1	1	1	1	1	1	0	0	1	1	0.72	0.73	0.71	0.7	0	0	0.3	0	0	0	1	1
STUD.ISCED7LONGRES	0.7	0.63	0.86	0	1	1	1	1	1	1	1	1	0	0	0.88	0.88	0.72	0.73	0.71	0.7	0	0	0.3	0	0	0	1	1
STUD.ISCED7LONGFOE05	0.7	0.63	0.86	0	1	1	1	1	1	1	1	1	0	0	1	1	0.72	0.73	0.71	0.7	0	0	1	1	0	0	0	1
STUD.ISCED7LONGTOTAL	0.7	0.63	0.86	0	1	1	1	1	1	1	1	1	0	0	1	1	0.72	0.73	0.71	0.7	0	0	1	1	0	0	1	1
STUD.TOTALISCED5.7	0.91	0.91	0.94	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0.99	1	1	0.92	0.92	0.91	0.92	1	1	1	1
STUD.DISTEDUINST	0.97	0.97	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
GRAD.ISCED5MEN	0.81	0.86	0.9	0	-	-	-	-	-	-	-	-	-	-	-	-	0.54	0.55	0.53	0.52	-	-	-	-	0	0	1	1
GRAD.ISCED5NAT	0.81	0.86	0.9	0	-	-	-	-	-	-	-	-	-	-	-	-	0.54	0.55	0.53	0.47	-	-	-	-	0	0	1	1
GRAD.ISCED5RES	0.81	0.86	0.9	0	-	-	-	-	-	-	-	-	-	-	-	-	0.54	0.55	0.53	0.52	-	-	-	-	0	0	1	1
GRAD.ISCED5FOE05	0.81	0.86	0.9	0	-	-	-	-	-	-	-	-	-	-	-	-	0.54	0.55	0.53	0.52	-	-	-	-	0	0	0	1
GRAD.ISCED5TOTAL	0.81	0.86	0.9	0	-	-	-	-	-	-	-	-	-	-	-	-	0.54	0.55	0.53	0.52	-	-	-	-	0	0	1	1
GRAD.ISCED6MEN	0.91	0.91	0.94	0	1	1	1	1	1	1	1	1	1	1	1	1	0.53	0.54	0.51	0.49	0.94	0.94	0.93	0	1	1	1	1
GRAD.ISCED6NAT	0.91	0.91	0.94	0	1	1	1	1	1	1	1	1	1	1	1	1	0.53	0.54	0.51	0.49	0	0	0	0	0	0	1	1
GRAD.ISCED6RES	0.91	0.91	0.94	0	1	1	1	1	1	1	1	1	0	1	1	1	0.53	0.54	0.51	0.49	0	0	0	0	1	1	1	1
GRAD.ISCED6FOE05	0.91	0.91	0.94	0	1	1	1	1	1	1	1	1	1	1	1	1	0.53	0.54	0.51	0.49	0.94	0.94	0.93	0	0	0	0	1
GRAD.ISCED6TOTAL	0.91	0.91	0.94	0	1	1	1	1	1	1	1	1	1	1	1	1	0.53	0.54	0.51	0.49	0.94	0.94	0.93	0	0	0	1	1
GRAD.ISCED7MEN	0.81	0.8	0.86	0	1	1	1	1	1	1	1	1	1	1	1	1	0.61	0.63	0.63	0.53	0.95	0.95	0.94	0	0	0	1	1
GRAD.ISCED7NAT	0.81	0.8	0.86	0	1	1	1	1	1	1	1	1	1	1	1	1	0.61	0.63	0.63	0.53	0	0	0.17	0	0	0	1	1
GRAD.ISCED7RES	0.81	0.8	0.86	0	1	1	1	1	1	1	1	1	0	1	0.88	0.88	0.61	0.63	0.63	0.53	0	0	0.17	0	0	0	1	1
GRAD.ISCED7FOE05	0.81	0.8	0.86	0	1	1	1	1	1	1	1	1	0.95	1	1	1	0.61	0.63	0.63	0.53	0.95	0.95	0.94	0	0	0	0	1
GRAD.ISCED7TOTAL	0.81	0.8	0.86	0	1	1	1	1	1	1	1	1	1	1	1	1	0.61	0.63	0.63	0.53	0.95	0.95	0.94	0	0	0	1	1
GRAD.ISCED7LONGMEN	0.77	0.63	0.86	0	1	1	1	1	1	1	1	1	0	0	1	1	0.67	0.66	0.68	0.69	0	0	1	0	0	0	1	1
GRAD.ISCED7LONGNAT	0.77	0.63	0.86	0	1	1	1	1	1	1	1	1	0	0	1	1	0.67	0.66	0.68	0.69	0	0	0.3	0	0	0	1	1
GRAD.ISCED7LONGRES	0.77	0.63	0.86	0	1	1	1	1	1	1	1	1	0	0	0.88	0.88	0.67	0.66	0.68	0.69	0	0	0.3	0	0	0	1	1

60

Country	DK 2011	DK 2012	DK 2013	DK 2014	EE 2011	EE 2012	EE 2013	EE 2014	ES 2011	ES 2012	ES 2013	ES 2014	FI 2011	FI 2012	FI 2013	FI 2014	FR 2011	FR 2012	FR 2013	FR 2014	GR 2011	GR 2012	GR 2013	GR 2014	HR 2011	HR 2012	HR 2013	HR 2014
GRAD.ISCED7LONGFOE05	0.77	0.63	0.86	0	1	1	1	1	1	1	1	1	0	0	1	1	0.67	0.66	0.68	0.69	0	0	1	0	0	0	0	1
GRAD.ISCED7LONGTOTAL	0.77	0.63	0.86	0	1	1	1	1	1	1	1	1	0	0	1	1	0.67	0.66	0.68	0.69	0	0	1	0	0	0	1	1
GRAD.TOTALISCED5.7	0.91	0.91	0.94	0	1	1	1	1	1	1	1	1	1	1	1	1	0.7	0.89	0.92	0.9	0.92	0.96	0.91	0	1	1	1	1
GRAD.NOTESEDUCACT	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
RES.RESACTIVE	0.91	0.91	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.44	0.77	0.81	0.82	1	1	1	1	1	1	1	1
RES.STUDISCED8MEN	0.79	0.79	0.83	0	1	1	1	1	0	0	0	1	1	1	1	1	0.63	0.63	0.63	0.62	0.96	0.96	0.95	0.95	1	1	1	1
RES.STUDISCED8NAT	0.79	0.79	0.83	0	1	1	1	1	0	0	0	1	1	1	1	1	0.63	0.63	0.63	0.62	0	0	0	0	0	0	1	1
RES.STUDISCED8RES	0.79	0.79	0.83	0	1	1	1	1	0	0	0	0	0	1	1	1	0.63	0.63	0.63	0.62	0	0	0	0	1	1	1	1
RES.STUDISCED8FOE05	0.79	0.79	0.83	0	1	1	1	1	0	0	0	1	1	1	1	1	0.63	0.63	0.63	0.62	0.96	0.96	0.95	0.95	0	0	0	1
RES.STUDISCED8TOTAL	0.79	0.79	0.83	0	1	1	1	1	0	0	0	1	1	1	1	1	0.63	0.63	0.63	0.62	0.96	0.96	0.95	0.95	1	1	1	1
RES.GRADISCED8MEN	0.79	0.79	0.83	0	1	1	1	1	1	1	1	1	1	1	1	1	0.55	0.54	0.55	0	0.96	0.96	0.95	0	1	1	1	1
RES.GRADISCED8NAT	0.79	0.79	0.83	0	1	1	1	1	1	1	1	1	1	1	1	1	0.55	0.54	0.55	0	0.09	0	0	0	0	0	1	1
RES.GRADISCED8RES	0.79	0.79	0.83	0	1	1	1	1	0	0	0.1	0	0	1	1	1	0.55	0.54	0.55	0	0.09	0	0	0	0	1	1	1
RES.GRADISCED8FOE05	0.79	0.79	0.83	0	1	1	1	1	1	1	1	0	0.94	1	1	1	0.56	0.54	0.55	0	0.96	0.96	0.95	0	0	0	0	1
RES.GRADISCED8TOTAL	0.79	0.79	0.83	0	1	1	1	1	1	1	1	1	1	1	1	1	0.55	0.54	0.55	0	0.96	0.96	0.95	0	1	1	1	1
RES.R.DEXP	0.57	0.36	0.4	0	0	0	0	0	0	0	0	0	0.36	0.36	0.98	0.98	0	0	0	0	1	0.98	0	0	1	1	1	1

Country	HU 2011	HU 2012	HU 2013	HU 2014	IE 2011	IE 2012	IE 2013	IE 2014	IS 2011	IS 2012	IS 2013	IS 2014	IT 2011	IT 2012	IT 2013	IT 2014	LI 2011	LI 2012	LI 2013	LI 2014	LT 2011	LT 2012	LT 2013	LT 2014	LU 2011	LU 2012	LU 2013	LU 2014
BAS.ETERIDYEAR	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BAS.ETERID	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BAS.NATID	-	1	1	-	1	1	1	1	-	-	-	-	-	-	-	1	-	-	-	-	1	1	1	1	-	-	-	-
BAS.INSTNAME	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BAS.REFYEAR	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BAS.ACRONYM	1	1	1	1	1	1	1	1	1	1	1	1	0.54	0.54	0.54	0.48	1	1	1	1	0.95	0.98	0.98	1	0	0	0	0
BAS.COUNTRY	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BAS.LEGALSTAT	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BAS.INSTCATSTAND	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BAS.FORCAMP	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BAS.FOUNDYEAR	0.85	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.99	1	1	1	1	1	1	1	1	1	1	1	1
BAS.LEGALYEAR	0.85	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BAS.ANCESTYEAR	1	1	1	1	1	1	1	1	1	1	1	1	0.97	0.97	0.99	0.99	1	1	1	1	1	1	1	1	1	1	1	1
BAS.UNIHOSP	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

Country	HU 2011	HU 2012	HU 2013	HU 2014	IE 2011	IE 2012	IE 2013	IE 2014	IS 2011	IS 2012	IS 2013	IS 2014	IT 2011	IT 2012	IT 2013	IT 2014	LI 2011	LI 2012	LI 2013	LI 2014	LT 2011	LT 2012	LT 2013	LT 2014	LU 2011	LU 2012	LU 2013	LU 2014
BAS.WEBSITE	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
GEO.NUTS2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
GEO.CITY	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
GEO.POSTCODE	1	1	1	1	-	-	-	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
GEO.MULTISITE	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
GEO.NUTS3MULTISITE	1	1	1	1	-	-	-	-	-	-	-	-	1	1	1	1	-	-	-	-	1	1	1	1	1	1	1	1
EXP.CURRPERSON	0.98	0	0.96	1	0.96	0.93	0.93	0.85	0	0	0	0	0.48	0.48	0.51	0.42	1	1	1	1	0.91	0.93	0.82	0.86	1	0	0	0
EXP.CURRNONPERSON	0.98	0	0.96	1	0.96	0.93	0.93	0.85	0	0	0	0	0.48	0.48	0.51	0.42	1	1	1	1	0.91	0.93	0.82	0.86	1	0	0	0
EXP.CURRUNCL	0.98	0	0.96	1	0.96	0.93	0.93	0.85	0	0	0	0	0.48	0.48	0.51	0.42	1	1	1	1	0.91	0.93	0.82	0.86	1	0	0	0
EXP.CURRTOTAL	0.98	0	0.96	1	0.96	0.93	0.93	0.85	0	0	0	0	0.48	0.48	0.51	0.42	1	1	1	1	0.91	0.93	0.82	0.86	1	0	0	0
EXP.CAPITAL	0.98	0	0.96	1	0.78	0.63	0.78	0.7	0	0	0	0	0.48	0.48	0.51	0.4	1	1	1	1	0.91	0.93	0.82	0.86	1	0	0	0
EXP.ACCSYSTEM	1	0	1	1	1	1	1	1	0	0	0	0	0.48	0.48	0.51	0.44	1	1	1	1	0.91	0.91	0.82	0.84	1	0	0	0
REV.COREBUDGETPUBLIC	0	0	0	1	0	0	0	0.85	0	0	0	0	0.48	0.48	0.51	0.42	1	1	1	1	0	0	0.82	0.86	0	0	0	0
REV.COREBUDGETOTHER	0	0	0	1	0	0	0	0.85	0	0	0	0	0.48	0.48	0.51	0.42	1	1	1	1	0	0	0.82	0.86	0	0	0	0
REV.CORETOTAL	0.98	0	0.96	1	0.96	0.93	0.93	0.85	0	0	0	0	0.48	0.48	0.51	0.42	1	1	1	1	0.91	0.93	0.82	0.86	0	0	0	0
REV.THIRDPARTYPUBLIC	0.98	0	0.96	1	0.96	0	0	0	0	0	0	0	0.38	0.38	0.38	0.33	0	0	0	0	0.91	0.93	0	0	0	0	0	0
REV.THIRDPARTYPRIVATE	0.98	0	0.96	1	0	0	0	0	0	0	0	0	0.38	0.38	0.38	0.33	0	0	0	0	0.91	0.93	0	0	0	0	0	0
REV.THIRDPARTYABROAD	0	0	0	0	0	0	0	0	0	0	0	0	0.38	0.38	0.38	0.33	0	0	0	0	0	0	0	0.86	0	0	0	0
REV.THIRDPARTYUND	0	0	0	0.44	0	0	0	0	0	0	0	0	0.48	0.48	0.51	0.42	0	0	0	0	0	0	0	0.86	0	0	0	0
REV.THIRDPARTYTOTAL	0.98	0	0.96	1	0.96	0.93	0.93	0.85	0	0	0	0	0.48	0.48	0.51	0.42	1	1	1	1	0.91	0.93	0.82	0.86	1	0	0	0
REV.TUITFEES	1	0	1	1	1	1	1	1	0	0	0	0	0.48	0.48	0.54	0.42	1	1	1	1	0.91	0.93	0.82	0.84	1	0	0	0
REV.STUDFEES.NC	0.46	0	0.96	0	0.96	0.93	0.93	0.85	0	0	0	0	0.47	0.47	0.5	0.41	1	1	1	1	0.9	0.93	0.81	0.86	1	0	0	0
REV.UNCL.NC	0.98	0	0.96	1	0.96	0.93	0.93	0.85	0	0	0	0	0.48	0.48	0.51	0.42	1	1	1	1	0.91	0.93	0.82	0.86	1	0	0	0
REV.CURRTOTAL.NC	0.98	0	0.96	1	0.96	0.93	0.93	0.85	0	0	0	0	0.48	0.48	0.51	0.42	1	1	1	1	0.91	0.93	0.82	0.86	1	0	0	0
REV.NONRECURR.NC	0	0	0	0	0	0	0	0.85	0	0	0	0	0	0	0	0.42	1	1	1	1	0	0	0	0.86	0	0	0	0
STA.ACAFTETOTAL	1	0	1	1	1	0.96	0.96	0.96	1	0	0	0	0.99	1	1	1	1	1	1	1	0.93	0.95	0.89	0.91	1	0	0	0
STA.ACAHCMEN	1	0	1	1	0.93	0.96	0.96	0.96	1	0	0	0	0.99	1	1	1	1	1	1	1	0.93	0.95	0.89	0.91	1	0	0	0
STA.ACAHCNAT	0	0	0	0	0	0	0	0	1	0	0	0	0.54	0.55	1	1	1	1	1	1	0	0	0.89	0.88	1	0	0	0
STA.ACAHCFOE05	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	0	0	0	0.91	0	0	0	0
STA.TOTACAHC	1	0	1	1	0.93	0.96	0.96	0.96	1	0	0	0	0.99	1	1	1	1	1	1	1	0.93	0.95	0.89	0.91	1	0	0	0
STA.PROFMEN	1	0	1	1	0	0.26	1	1	0	0	0	0	0.99	1	1	1	1	1	1	1	0	0	0.89	0.91	0	0	0	0
STA.PROFTOTAL	1	0	1	1	0	0.26	1	1	0	0	0	0	0.99	1	1	1	1	1	1	1	0.93	0.95	0.89	0.91	1	0	0	0
STA.INCLPHDSTUD	1	0	1	1	1	1	1	1	1	1	0	0	0.99	1	1	1	1	1	1	1	0.93	0.95	0.89	0.91	1	0	0	0
STA.NONACAFTE	1	0	1	1	0.96	0.96	0.96	0.96	0	0	0	0	0	1	1	1	1	1	1	1	0.93	0.95	0.89	0.91	1	0	0	0
STA.NONACAHC	1	0	1	1	0	0.96	0.96	0.96	0	0	0	0	0.99	1	1	1	1	1	1	1	0.93	0.95	0.89	0.91	1	0	0	0

Country	HU 2011	HU 2012	HU 2013	HU 2014	IE 2011	IE 2012	IE 2013	IE 2014	IS 2011	IS 2012	IS 2013	IS 2014	IT 2011	IT 2012	IT 2013	IT 2014	LI 2011	LI 2012	LI 2013	LI 2014	LT 2011	LT 2012	LT 2013	LT 2014	LU 2011	LU 2012	LU 2013	LU 2014
STA.TOTALFTE	1	0	1	1	0.96	0.96	0.96	0.96	1	0	0	0	0.99	1	1	1	1	1	1	1	0.93	0.95	0.89	0.91	1	0	0	0
STA.TOTALHC	1	0	1	1	0	0.96	0.96	0.96	1	0	0	0	0.99	1	1	1	1	1	1	1	0.93	0.95	0.89	0.91	1	0	0	0
STUD.LOWDEG	1	0	1	1	1	1	1	1	1	0.86	0.86	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0
STUD.HIGHDEG	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0
STUD.ISCED5MEN	1	0	1	1	1	1	1	1	1	0.86	0.86	1	-	-	-	-	-	-	-	-	-	-	-	-	-	0	0	0
STUD.ISCED5NAT	1	0	1	1	1	1	1	1	0	0	0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	0	0	0
STUD.ISCED5RES	1	0	0.06	0	1	1	1	1	0	0	0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	0	0	0
STUD.ISCED5FOE05	1	0	1	1	1	1	1	1	1	0.86	0.86	1	-	-	-	-	-	-	-	-	-	-	-	-	-	0	0	0
STUD.ISCED5TOTAL	1	0	1	1	1	1	1	1	1	0.86	0.86	1	-	-	-	-	-	-	-	-	-	-	-	-	-	0	0	0
STUD.ISCED6MEN	1	0	1	1	1	1	1	1	1	1	1	1	0.99	0.99	0.99	1	1	1	1	1	0.95	0.98	0.91	0.93	1	0	0	0
STUD.ISCED6NAT	1	0	1	1	1	1	1	1	0	0	0	0	0.98	0.99	0.99	1	1	1	1	1	0.95	0.98	0.89	0.88	1	0	0	0
STUD.ISCED6RES	1	0	0.02	0	1	1	1	1	0	0	0	0	0	0	0	0	1	1	1	1	0	0	0.89	0.88	1	0	0	0
STUD.ISCED6FOE05	1	0	1	1	1	1	1	1	1	1	1	1	0.98	0.99	0.99	1	1	1	1	1	0	0	0.91	0.93	1	0	0	0
STUD.ISCED6TOTAL	1	0	1	1	1	1	1	1	1	1	1	1	0.99	0.99	0.99	1	1	1	1	1	0.95	0.98	0.91	0.93	1	0	0	0
STUD.ISCED7MEN	0.98	0	1	1	1	1	1	1	1	1	1	1	0.98	0.99	1	1	1	1	1	1	1	1	0.95	0.95	1	0	0	0
STUD.ISCED7NAT	0.98	0	1	1	1	1	1	1	0	0	0	0	0.98	0.99	1	1	1	1	1	1	1	1	0.95	0.95	1	0	0	0
STUD.ISCED7RES	0.98	0	0.02	0	1	1	1	1	0	0	0	0	0	0	0	0	1	1	1	1	0	0	0.95	0.95	1	0	0	0
STUD.ISCED7FOE05	0.98	0	1	1	1	1	1	1	1	1	1	1	0.98	0.99	1	1	1	1	1	1	0	0	0.95	0.95	1	0	0	0
STUD.ISCED7TOTAL	0.98	0	1	1	1	1	1	1	1	1	1	1	0.98	0.99	1	1	1	1	1	1	1	1	0.95	0.95	1	0	0	0
STUD.ISCED7LONGMEN	-	0	1	1	-	-	-	-	1	1	1	1	0.99	0.99	1	1	-	-	-	-	1	1	0.88	0.88	-	0	0	0
STUD.ISCED7LONGNAT	-	0	1	1	-	-	-	-	0.86	0	0	0	0.99	0.99	1	1	-	-	-	-	1	1	0.88	0.88	-	0	0	0
STUD.ISCED7LONGRES	-	0	0.37	0	-	-	-	-	0.86	0	0	0	0	0	0	0	-	-	-	-	0	0	0.88	0.88	-	0	0	0
STUD.ISCED7LONGFOE05	-	0	1	1	-	-	-	-	1	1	1	1	0.99	0.99	1	1	-	-	-	-	0	0	0.88	0.88	-	0	0	0
STUD.ISCED7LONGTOTAL	-	0	1	1	-	-	-	-	1	1	1	1	0.99	0.99	1	1	-	-	-	-	1	1	0.88	0.88	-	0	0	0
STUD.TOTALISCED5.7	1	0	1	1	1	1	1	1	1	1	1	1	0.97	0.98	0.99	1	1	1	1	1	0.95	0.98	0.91	0.93	1	0	0	0
STUD.DISTEDUINST	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.95	0.98	0.91	0.93	1	0	0	0
GRAD.ISCED5MEN	1	0	1	1	1	1	1	1	0	0	0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	0	0	0
GRAD.ISCED5NAT	1	0	1	1	1	1	1	1	0	0	0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	0	0	0
GRAD.ISCED5RES	0	0	0.42	0	1	1	1	1	0	0	0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	0	0	0
GRAD.ISCED5FOE05	1	0	1	1	1	1	1	1	0	0	0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	0	0	0
GRAD.ISCED5TOTAL	1	0	1	1	1	1	1	1	0	0	0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	0	0	0
GRAD.ISCED6MEN	1	0	1	1	1	1	1	1	0	0	0	0	0.99	0.99	0.99	1	1	1	1	1	0.95	0.98	0.91	0.93	1	0	0	0
GRAD.ISCED6NAT	1	0	1	1	1	1	1	1	0	0	0	0	0.99	0.99	0.99	1	1	1	1	1	0.95	0.95	0.86	0.88	1	0	0	0
GRAD.ISCED6RES	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0	1	1	1	1	0	0	0.86	0.88	1	0	0	0
GRAD.ISCED6FOE05	1	0	1	1	1	1	1	1	0	0	0	0	0.99	0.99	0.99	1	1	1	1	1	0	0	0.91	0.93	1	0	0	0

Country	HU 2011	HU 2012	HU 2013	HU 2014	IE 2011	IE 2012	IE 2013	IE 2014	IS 2011	IS 2012	IS 2013	IS 2014	IT 2011	IT 2012	IT 2013	IT 2014	LI 2011	LI 2012	LI 2013	LI 2014	LT 2011	LT 2012	LT 2013	LT 2014	LU 2011	LU 2012	LU 2013	LU 2014
GRAD.ISCED6TOTAL	1	0	1	1	1	1	1	1	0	0	0	0	0.99	0.99	0.99	1	1	1	1	1	0.95	0.98	0.91	0.93	1	0	0	0
GRAD.ISCED7MEN	1	0	1	1	1	1	1	1	0	0	0	0	0.98	0.99	0.99	1	1	1	1	1	1	1	0.95	0.95	1	0	0	0
GRAD.ISCED7NAT	1	0	1	1	1	1	1	1	0	0	0	0	0.98	0.99	0.99	1	1	1	1	1	1	1	0.95	0.95	1	0	0	0
GRAD.ISCED7RES	0	0	0.04	0	1	1	1	1	0	0	0	0	0	0	0	0	1	1	1	1	0	0	0.95	0.95	1	0	0	0
GRAD.ISCED7FOE05	1	0	1	1	1	1	1	1	0	0	0	0	0.98	0.99	0.99	1	1	1	1	1	0	0	0.95	0.95	1	0	0	0
GRAD.ISCED7TOTAL	1	0	1	1	1	1	1	1	0	0	0	0	0.98	0.99	0.99	1	1	1	1	1	1	1	0.95	0.95	1	0	0	0
GRAD.ISCED7LONGMEN	-	0	1	1	-	-	-	-	0	0	0	0	0.99	0.99	0.99	1	-	-	-	-	1	1	1	0.88	-	0	0	0
GRAD.ISCED7LONGNAT	-	0	1	1	-	-	-	-	0	0	0	0	0.99	0.99	0.99	1	-	-	-	-	1	1	1	0.88	-	0	0	0
GRAD.ISCED7LONGRES	-	0	0.41	0	-	-	-	-	0	0	0	0	0	0	0	0	-	-	-	-	0	0	1	0.88	-	0	0	0
GRAD.ISCED7LONGFOE05	-	0	1	1	-	-	-	-	0	0	0	0	0.99	0.99	0.99	1	-	-	-	-	0	0	1	0.88	-	0	0	0
GRAD.ISCED7LONGTOTAL	-	0	1	1	-	-	-	-	0	0	0	0	0.99	0.99	0.99	1	-	-	-	-	1	1	1	0.88	-	0	0	0
GRAD.TOTALISCED5.7	1	0	1	1	1	1	1	1	0	0	0	0	0.99	0.99	0.99	1	1	1	1	1	0.95	0.98	0.91	0.93	1	0	0	0
GRAD.NOTESEDUCACT	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
RES.RESACTIVE	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.98	1	1	1	0	0	0
RES.STUDISCED8MEN	1	0	1	1	1	1	1	1	1	1	1	1	0.95	1	0.99	1	1	1	1	1	1	1	1	1	1	0	0	0
RES.STUDISCED8NAT	1	0	1	1	1	1	1	1	0	0	0	0	0.95	1	0.99	1	1	1	1	1	1	1	1	1	1	0	0	0
RES.STUDISCED8RES	1	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0	1	1	1	1	0	0	1	1	1	0	0	0
RES.STUDISCED8FOE05	1	0	1	1	1	1	1	1	1	1	1	1	0.95	1	0.01	0.01	1	1	1	1	0	0	1	1	1	0	0	0
RES.STUDISCED8TOTAL	1	0	1	1	1	1	1	1	1	1	1	1	0.95	1	0.99	1	1	1	1	1	1	1	1	1	1	0	0	0
RES.GRADISCED8MEN	1	0	1	1	1	1	1	1	0	0	0	1	0.94	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0
RES.GRADISCED8NAT	1	0	1	1	1	1	1	1	0	0	0	1	0.94	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0
RES.GRADISCED8RES	0	0	0.38	0	1	1	1	1	0	0	0	0	0	0	0	0	1	1	1	1	0	0	1	1	1	0	0	0
RES.GRADISCED8FOE05	1	0	1	1	1	1	1	1	0	0	0	1	0.94	1	0.01	0.01	1	1	1	1	0	0	1	1	1	0	0	0
RES.GRADISCED8TOTAL	1	0	1	1	1	1	1	1	0	0	0	1	0.94	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0
RES.R.DEXP	0.94	0	0	0	0.04	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.9	0.93	0.83	0.86	1	0	0	0

Country	LV 2011	LV 2012	LV 2013	LV 2014	MK 2011	MK 2012	MK 2013	MK 2014	MT 2011	MT 2012	MT 2013	MT 2014	NL 2011	NL 2012	NL 2013	NL 2014	NO 2011	NO 2012	NO 2013	NO 2014	PL 2011	PL 2012	PL 2013	PL 2014	PT 2011	PT 2012	PT 2013	PT 2014
BAS.ETERIDYEAR	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BAS.ETERID	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BAS.NATID	1	1	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	0.97	1	0.98	1	1	1	1
BAS.INSTNAME	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BAS.REFYEAR	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BAS.ACRONYM	0.76	0.76	0.8	0.82	1	1	1	1	1	1	1	1	0.65	0.67	0.66	0.66	0.92	0.92	0.92	1	0	0	0	0	0.34	0.37	0.4	0.43
BAS.COUNTRY	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BAS.LEGALSTAT	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BAS.INSTCATSTAND	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BAS.FORCAMP	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BAS.FOUNDYEAR	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.98	1	1	1	0.98
BAS.LEGALYEAR	1	1	1	1	1	1	1	1	1	1	1	1	0.88	0.87	0.88	0.88	1	1	1	1	0	0	0	-	1	0.99	0.99	0.97
BAS.ANCESTYEAR	1	1	1	1	1	1	1	1	1	1	1	1	0.33	0.35	0.35	0.35	0.88	0.88	0.98	0.98	0	0	0	-	1	1	0.99	0.97
BAS.UNIHOSP	1	1	1	1	0.8	0.8	0.8	0.8	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.98	1	1	1	1
BAS.WEBSITE	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.98	1	1	1	1	1	1	1	1
GEO.NUTS2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.98	0.98	1	1	1	1	1	1	1	1
GEO.CITY	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.98	0.98	0.98	0.98	1	1	1	1	1	1	1	1
GEO.POSTCODE	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.98	0.98	0.98	0.98	1	1	1	1	1	1	1	1
GEO.MULTISITE	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
GEO.NUTS3MULTISITE	-	-	-	-	1	1	1	1	1	1	1	1	1	1	1	1	0.94	0.94	0.94	0.93	1	1	1	1	1	1	1	1
EXP.CURRPERSON	0	0	0	0.64	0	0	0	0	1	1	1	1	0.96	0.98	0.96	0.98	0.98	0.98	0.98	0.98	0	0.35	0.34	0.33	0.28	0.33	0.36	0.37
EXP.CURRNONPERSON	0	0	0	0.64	0	0	0	0	1	1	1	1	0.96	0.98	0.96	0.98	0.98	0.98	0.98	0.98	0	0	0	0.33	0.28	0.33	0.36	0.37
EXP.CURRUNCL	0	0	0	0.64	0	0	0	0	1	1	1	1	0.96	0.98	0.96	0.98	0.98	0.98	0.98	0.98	0	0.35	0.34	0.33	0.28	0.33	0.36	0.37
EXP.CURRTOTAL	0	0	0	0.64	0	0	0	0	1	1	1	1	0.96	0.98	0.96	0.98	0.98	0.98	0.98	1	0.32	0.35	0.34	0.33	0.31	0.33	0.36	0.37
EXP.CAPITAL	0	0	0	0.64	0	0	0	0	1	1	1	1	0.96	0.98	0.96	0.98	0.96	0.98	0.98	0.96	0	0	0	0	0.28	0.33	0.36	0.37
EXP.ACCSYSTEM	0	0	0	1	0	0	0	0	1	1	1	1	1	1	0.98	1	0.98	0.98	0.98	1	0	1	1	0.98	0.31	0.33	0.36	0.37
REV.COREBUDGETPUBLIC	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0.96	0	0	0	0.96	0	0	0	0.33	0.31	0.33	0.36	0.37
REV.COREBUDGETOTHER	0	0	0	0	0	0	0	0	0	0	0.5	0	0	0	0	0	0	0	0	0.96	0	0	0	0.33	0.31	0.33	0.36	0.37
REV.CORETOTAL	0	0	0	0.64	0	0	0	0	1	1	1	1	0.96	0.98	0.95	0.96	0.94	0.96	0.98	0.96	0.36	0.35	0.34	0.33	0.31	0.33	0.36	0.37
REV.THIRDPARTYPUBLIC	0	0	0	0	0	0	0	0	1	1	0.5	0.5	0.96	0.98	0.93	0.96	0.76	0.74	0.76	0.96	0	0	0	0	0.31	0.33	0.36	0.37
REV.THIRDPARTYPRIVATE	0	0	0	0.5	0	0	0	0	1	1	0	0	0.96	0.98	0.93	0.96	0.76	0.74	0.76	0.96	0	0	0	0	0.28	0.33	0.36	0.37
REV.THIRDPARTYABROAD	0	0	0	0.64	0	0	0	0	0	0	0	0	0	0	0.93	0.96	0	0	0	0.96	0	0	0	0	0.28	0.33	0.36	0.37
REV.THIRDPARTYUND	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.93	0.96	0	0	0	0.96	0	0	0	0	0.31	0.33	0.36	0.37
REV.THIRDPARTYTOTAL	0	0	0	0.64	0	0	0	0	1	1	1	1	0.96	0.98	0.95	0.96	0.86	0.96	0.9	0.96	0	0	0	0	0.28	0.33	0.36	0.37
REV.TUITFEES	0	0	0	0.64	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.98	0.31	0.33	0.36	0.37

Country	LV 2011	LV 2012	LV 2013	LV 2014	MK 2011	MK 2012	MK 2013	MK 2014	MT 2011	MT 2012	MT 2013	MT 2014	NL 2011	NL 2012	NL 2013	NL 2014	NO 2011	NO 2012	NO 2013	NO 2014	PL 2011	PL 2012	PL 2013	PL 2014	PT 2011	PT 2012	PT 2013	PT 2014
REV.STUDFEES.NC	0	0	0	0.63	0	0	0	0	1	1	1	1	0.96	0.98	0.95	0.96	0.88	0.96	1	1	0	0	0	0	0.28	0.33	0.36	0.37
REV.UNCL.NC	0	0	0	0.64	0	0	0	0	1	1	1	1	0.96	0.98	0.95	0.96	0.96	0.96	0.98	0.96	0.32	0.35	0.34	0.33	0.28	0.33	0.36	0.37
REV.CURRTOTAL.NC	0	0	0	0.64	0	0	0	0	1	1	1	1	0.96	0.98	0.96	0.98	0.96	0.96	0.98	0.98	0.32	0.35	0.34	0.33	0.31	0.33	0.36	0.37
REV.NONRECURR.NC	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0.96	0	0	0	-	0	0	0	0	0.28	0.33	0.36	0.37
STA.ACAFTETOTAL	0	0	0	0	0	0	0	0	1	1	1	1	0.93	0.93	0.96	0.98	0.96	0.96	0.98	0.98	1	1	1	1	1	1	1	1
STA.ACAHCMEN	0.98	0.98	0.98	0.98	0.9	0.9	0.8	0.9	1	1	1	1	0.93	0.93	0.84	0.93	0.84	0.84	0.92	0.88	0	0	0	0	1	1	1	1
STA.ACAHCNAT	0	0	0	0	0	0	0	0	1	1	1	1	0	0	0.27	0.25	0	0	0	0	0	0	0	0	1	1	1	1
STA.ACAHCFOE05	0	0	0	0	0	0.8	0.7	0.7	1	1	0	0	0	0	0.02	0.05	0.84	0.84	0.92	0.88	0	0	0	0	0	0	0	0
STA.TOTACAHC	0.98	0.98	0.98	0.98	0.9	0.9	0.8	0.9	1	1	1	1	0.68	0.67	0.63	0.75	0.84	0.84	0.92	0.88	0	0	0	0	1	1	1	1
STA.PROFMEN	0.98	0.98	0.98	0.98	0	0	0	0	1	1	1	1	0.78	0.78	0.79	0.84	0.84	0.84	0.92	0.88	1	1	1	1	1	1	1	1
STA.PROFTOTAL	0.98	0.98	0.98	0.98	0	0	0	0	1	1	1	1	0.78	0.78	0.95	0.95	0.84	0.84	0.92	0.88	1	1	1	1	1	1	1	1
STA.INCLPHDSTUD	0.98	0.98	0.98	0.98	0	0	0	0	1	1	1	1	0.93	0.93	1	1	0.84	0.84	0.92	0.88	1	1	1	1	1	1	1	1
STA.NONACAFTE	0	0	0	0	0	0	0	0	1	1	1	1	0.93	0.93	0.96	0.98	0.96	0.96	0.98	0.98	1	1	1	1	0.31	0	0	0.37
STA.NONACAHC	0.98	0.98	0.98	0.98	0.9	0.9	0.8	0.9	1	1	1	1	0.68	0.67	0.63	0.73	0.84	0.84	0.9	0.88	0	0	0	0	0.31	0	0	0.37
STA.TOTALFTE	0	0	0	0	0	0	0	0	1	1	1	1	0.93	0.93	0.98	0.98	0.96	0.96	0.98	1	1	1	1	1	1	1	1	1
STA.TOTALHC	0.98	0.98	0.98	0.98	0.9	0.9	0.8	0.9	1	1	1	1	0.93	0.93	0.89	0.98	0.84	0.84	0.92	0.88	0	0	0	0	1	1	1	1
STUD.LOWDEG	0.98	0.98	0.98	0.98	1	1	1	1	1	1	1	1	1	1	1	1	0.98	0.98	0.98	0.98	1	1	1	1	1	1	1	1
STUD.HIGHDEG	0.98	0.98	0.98	0.98	1	1	1	1	1	1	1	1	1	1	1	1	0.98	0.98	0.98	0.98	1	1	1	1	1	1	1	1
STUD.ISCED5MEN	0.98	0.98	0.97	0.97	-	-	-	-	1	1	1	1	1	0.97	0.97	0.97	0	0	0	0	-	-	-	-	-	-	-	1
STUD.ISCED5NAT	0	0	0	0	-	-	-	-	1	1	1	1	0	0	0	0	0	0	0	0	-	-	-	-	-	-	-	1
STUD.ISCED5RES	0.98	0.98	0.97	0.97	-	-	-	-	1	0	0	1	0	0	0	0	0	0	0	0	-	-	-	-	-	-	-	1
STUD.ISCED5FOE05	0.98	0.98	0.97	0.97	-	-	-	-	1	1	1	1	1	0.97	0.97	0.97	0	0	0	0	-	-	-	-	-	-	-	1
STUD.ISCED5TOTAL	0.98	0.98	0.97	0.97	-	-	-	-	1	1	1	1	1	0.97	0.97	0.97	0	0	0	0	-	-	-	-	-	-	-	1
STUD.ISCED6MEN	0.97	0.97	0.96	0.96	0.9	0.9	0.8	0.9	1	1	1	1	0.98	0.98	0.98	0.96	0.96	0.98	0.98	0.98	1	1	1	1	1	1	1	1
STUD.ISCED6NAT	0	0	0	0	0.9	0.9	0.9	0.9	1	1	1	1	0	0	0	0	0.96	0.98	0.98	0.96	1	1	1	1	1	1	1	1
STUD.ISCED6RES	0.97	0.97	0.96	0.96	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0.99	1
STUD.ISCED6FOE05	0.97	0.97	0.96	0.96	0.9	0.9	0.9	0.9	1	1	1	1	0.98	0.98	0.98	0.96	0.96	0.98	0.98	0.98	1	1	1	1	1	1	1	1
STUD.ISCED6TOTAL	0.97	0.97	0.96	0.96	0.9	0.9	0.9	0.9	1	1	1	1	0.98	0.98	0.98	0.96	0.96	0.98	0.98	0.98	1	1	1	1	1	1	1	1
STUD.ISCED7MEN	0.96	0.96	0.96	0.96	1	0.78	0.89	0.9	1	1	1	1	0.98	0.98	0.96	0.96	0.96	0.98	0.98	0.98	1	1	1	1	1	1	1	1
STUD.ISCED7NAT	0	0	0	0	1	0.78	0.78	0.9	1	1	1	1	0	0	0	0	0.96	0.98	0.98	0.98	1	1	1	1	1	1	1	1
STUD.ISCED7RES	0.96	0.96	0.96	0.96	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0.99	1
STUD.ISCED7FOE05	0.96	0.96	0.96	0.96	1	0.78	0.78	0.9	1	1	1	1	0.98	0.98	0.98	0.96	0.96	0.98	0.98	0.98	1	1	1	1	1	1	1	1
STUD.ISCED7TOTAL	0.96	0.96	0.96	0.96	1	0.78	0.89	0.9	1	1	1	1	0.98	0.98	0.98	0.96	0.96	0.98	0.98	0.98	1	1	1	1	1	1	1	1
STUD.ISCED7LONGMEN	0	0	0	0.75	-	-	-	-	1	1	1	1	0	0	0	0	0.88	0.93	0.94	0.89	1	1	1	1	1	1	1	1
STUD.ISCED7LONGNAT	0	0	0	0	-	-	-	-	1	1	1	1	0	0	0	0	0.88	0.93	0.94	0.89	1	1	1	1	1	1	1	1

Country	LV 2011	LV 2012	LV 2013	LV 2014	MK 2011	MK 2012	MK 2013	MK 2014	MT 2011	MT 2012	MT 2013	MT 2014	NL 2011	NL 2012	NL 2013	NL 2014	NO 2011	NO 2012	NO 2013	NO 2014	PL 2011	PL 2012	PL 2013	PL 2014	PT 2011	PT 2012	PT 2013	PT 2014
STUD.ISCED7LONGRES	0	0	0	0.75	-	-	-	-	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1
STUD.ISCED7LONGFOE05	0	0	0	1	-	-	-	-	1	1	1	1	0	0	0	0	0.88	0.93	0.94	0.89	1	1	1	1	1	1	1	1
STUD.ISCED7LONGTOTAL	0	0	0	0.75	-	-	-	-	1	1	1	1	0	0	0	0	0.88	0.93	0.94	0.89	1	1	1	1	1	1	1	1
STUD.TOTALISCED5.7	0.98	0.98	0.98	0.98	0.9	0.9	0.9	0.9	1	1	1	1	0.98	0.98	0.98	0.96	0.96	0.98	0.98	0.98	1	1	1	1	1	1	1	1
STUD.DISTEDUINST	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.98	0.98	1	1	1	0.98	1	1	1	1
GRAD.ISCED5MEN	0.98	0.98	0.97	0.97	-	-	-	-	1	1	1	1	1	0.97	1	1	0.93	0.9	0.91	1	-	-	-	-	-	-	-	-
GRAD.ISCED5NAT	0	0	0	0	-	-	-	-	1	1	1	1	0	0	0.03	0.03	0	0	0	0	-	-	-	-	-	-	-	-
GRAD.ISCED5RES	0.98	0.98	0.97	0.97	-	-	-	-	1	0	0	1	0	0	0.03	0.03	0	0	0	0	-	-	-	-	-	-	-	-
GRAD.ISCED5FOE05	0.98	0.98	0.97	0.97	-	-	-	-	1	1	1	1	1	1	1	1	0.93	0.9	0.91	1	-	-	-	-	-	-	-	-
GRAD.ISCED5TOTAL	0.98	0.98	0.97	0.97	-	-	-	-	1	1	1	1	1	1	1	1	0.93	0.9	0.91	1	-	-	-	-	-	-	-	-
GRAD.ISCED6MEN	0.97	0.97	0.96	0.96	0.8	0.8	0.8	0.9	1	1	1	1	0.98	0.96	0.96	0.96	0.96	0.98	0.98	0.98	1	0.99	1	1	1	1	1	1
GRAD.ISCED6NAT	0	0	0	0	0.8	0.8	0.8	0.9	1	1	1	1	0	0	0	0	0	0	0	0	1	0.99	1	1	1	1	1	1
GRAD.ISCED6RES	0.97	0.97	0.96	0.96	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0.99	1
GRAD.ISCED6FOE05	0.97	0.97	0.96	0.96	0.8	0.8	0.8	0.9	1	1	1	1	0.98	0.98	0.98	0.96	0.96	0.98	0.98	0.98	1	0.99	1	1	1	1	1	1
GRAD.ISCED6TOTAL	0.97	0.97	0.96	0.96	0.8	0.8	0.8	0.9	1	1	1	1	0.98	0.98	0.98	0.96	0.96	0.98	0.98	0.98	1	0.99	1	1	1	1	1	1
GRAD.ISCED7MEN	0.96	0.96	0.96	0.96	0.78	0.78	0.78	0.8	1	1	1	1	0.98	0.96	0.96	0.96	0.93	0.93	0.98	0.98	1	0.99	1	1	1	1	1	1
GRAD.ISCED7NAT	0	0	0	0	0.78	0.78	0.88	0.8	1	1	1	1	0	0	0	0	0.02	0.07	0.06	0	1	0.99	1	1	1	1	1	1
GRAD.ISCED7RES	0.96	0.96	0.96	0.96	0	0	0	0	1	0	0	1	0	0	0	0	0.02	0.07	0.06	0	0	0	0	0	1	1	0.99	1
GRAD.ISCED7FOE05	0.96	0.96	0.96	0.96	0.78	0.78	0.78	0.8	1	1	1	1	0.98	0.98	0.98	0.96	0.93	0.93	0.98	0.98	1	0.99	1	1	1	1	1	1
GRAD.ISCED7TOTAL	0.96	0.96	0.96	0.96	0.78	0.78	0.78	0.8	1	1	1	1	0.98	0.98	0.98	0.96	0.95	0.96	0.98	0.98	1	0.99	1	1	1	1	1	1
GRAD.ISCED7LONGMEN	0	0	0	0.75	-	-	-	-	1	1	1	1	0	0	0	0	0.88	0.93	0.93	1	1	0.99	1	1	1	1	1	1
GRAD.ISCED7LONGNAT	0	0	0	0	-	-	-	-	1	1	1	1	0	0	0	0	0	0	0	0	1	0.99	1	1	1	1	1	1
GRAD.ISCED7LONGRES	0	0	0	0.75	-	-	-	-	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1
GRAD.ISCED7LONGFOE05	0	0	0	0.75	-	-	-	-	1	1	1	1	0	0	0	0	0.88	0.93	0.93	1	1	0.99	1	1	1	1	1	1
GRAD.ISCED7LONGTOTAL	0	0	0	0.75	-	-	-	-	1	1	1	1	0	0	0	0	0.88	0.93	0.93	1	1	0.99	1	1	1	1	1	1
GRAD.TOTALISCED5.7	0.98	0.98	0.98	0.98	0.8	0.8	0.8	0.9	1	1	1	1	0.98	0.98	0.98	0.96	0.96	0.98	0.98	0.98	1	0.99	1	1	1	1	1	1
GRAD.NOTESEDUCACT	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
RES.RESACTIVE	0.98	0.98	0.98	0.98	0.2	0.2	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.99	1	1	1	1
RES.STUDISCED8MEN	0.95	0.95	0.95	0.95	0.29	0.57	0.5	0.71	1	1	1	1	0.78	0.78	0.74	0.74	0.96	1	0.96	1	1	1	1	1	1	1	1	1
RES.STUDISCED8NAT	0	0	0	0	0.29	0.57	0.5	0.57	1	1	1	1	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
RES.STUDISCED8RES	0.95	0.95	0.95	0.95	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0.96	1
RES.STUDISCED8FOE05	0.95	0.95	0.95	0.95	0.29	0.57	0.5	0.57	1	1	1	1	0	0	0	0	0.96	1	0.96	1	0	0	0	0	1	1	1	1
RES.STUDISCED8TOTAL	0.95	0.95	0.95	0.95	0.29	0.57	0.5	0.71	1	1	1	1	0.78	0.78	0.74	1	0.96	1	0.96	1	1	1	1	1	1	1	1	1
RES.GRADISCED8MEN	0.95	0.95	0.95	0.95	0.43	0.57	0.43	0.57	1	1	1	1	0	0	0	0	0.92	1	1	1	1	0	0	0	1	1	1	1
RES.GRADISCED8NAT	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0.92	1	1	1	1	0	0	0	1	1	1	1

Country	LV 2011	LV 2012	LV 2013	LV 2014	MK 2011	MK 2012	MK 2013	MK 2014	MT 2011	MT 2012	MT 2013	MT 2014	NL 2011	NL 2012	NL 2013	NL 2014	NO 2011	NO 2012	NO 2013	NO 2014	PL 2011	PL 2012	PL 2013	PL 2014	PT 2011	PT 2012	PT 2013	PT 2014
RES.GRADISCED8RES	0.95	0.95	0.95	0.95	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0.16	0	0	0	0	0	1	1	0.96	1
RES.GRADISCED8FOE05	0.95	0.95	0.95	0.95	0.43	0.57	0.43	0.57	1	1	1	1	0	0	0	0	0.96	1	1	1	0	0	0	0	1	1	1	1
RES.GRADISCED8TOTAL	0.95	0.95	0.95	0.95	0.43	0.57	0.43	0.57	1	1	1	1	0.78	0.78	1	1	0.96	1	1	1	1	0	0	0	1	1	1	1
RES.R.DEXP	0.96	0.97	0.97	0.97	0	0	0	0	0	1	1	1	0	0	0	0	0.68	0	0.7	0	0.56	0.56	0.59	0.57	0	0	0	0

Country	RS 2011	RS 2012	RS 2013	RS 2014	SE 2011	SE 2012	SE 2013	SE 2014	SK 2011	SK 2012	SK 2013	SK 2014	TR 2011	TR 2012	TR 2013	TR 2014	UK 2011	UK 2012	UK 2013	UK 2014
BAS.ETERIDYEAR	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BAS.ETERID	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BAS.NATID	-	-	-	-	1	1	1	1	-	-	-	-	-	-	-	-	1	1	1	1
BAS.INSTNAME	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BAS.REFYEAR	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BAS.ACRONYM	0	0	0	0	0.97	0.97	0.98	1	1	1	1	1	0.41	0.4	0.4	0.4	0.31	0.32	0.03	0.33
BAS.COUNTRY	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BAS.LEGALSTAT	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BAS.INSTCATSTAND	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BAS.FORCAMP	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BAS.FOUNDYEAR	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BAS.LEGALYEAR	0	1	1	1	0.64	0.67	1	1	1	1	1	1	1	1	1	1	0.98	0.98	0.98	0.98
BAS.ANCESTYEAR	0	1	1	1	1	1	1	1	0	0	0	0	0.16	0.16	0.15	0.16	0.04	0.04	0.06	0.07
BAS.UNIHOSP	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BAS.WEBSITE	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
GEO.NUTS2	0	-	-	-	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
GEO.CITY	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
GEO.POSTCODE	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
GEO.MULTISITE	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
GEO.NUTS3MULTISITE	0	-	-	-	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
EXP.CURRPERSON	0	0	0	0	1	1	1	1	0.63	0.63	0.72	0.72	0	0	0.55	0.55	0.99	0.99	0.99	0.99
EXP.CURRNONPERSON	0	0	0	0	1	1	1	1	0.63	0.63	0.72	0.72	0	0	0.55	0.55	0.99	0.99	0.99	0.99
EXP.CURRUNCL	0	0	0	0	1	1	1	1	0.63	0.63	0.72	0.72	0	0	0.55	0.55	0.99	0.99	0.99	0.99
EXP.CURRTOTAL	0	0	0	0	1	1	1	1	0.63	0.63	0.72	0.72	0	0	0.55	0.55	0.99	0.99	0.99	0.99
EXP.CAPITAL	0	0	0	0	1	1	1	1	0.63	0.63	0.72	0.69	0	0	0.55	0.55	0.99	0.99	0.99	0.99
EXP.ACCSYSTEM	0	0	0	0	1	1	1	1	0.63	0.63	0.72	0.72	0	0	0.55	0.55	1	1	1	1

Country	RS 2011	RS 2012	RS 2013	RS 2014	SE 2011	SE 2012	SE 2013	SE 2014	SK 2011	SK 2012	SK 2013	SK 2014	TR 2011	TR 2012	TR 2013	TR 2014	UK 2011	UK 2012	UK 2013	UK 2014
REV.COREBUDGETPUBLIC	0	0	0	0	0	0	0	1	0	0	0	0.69	0	0	0	0	0.99	0.99	0.99	0.99
REV.COREBUDGETOTHER	0	0	0	0	0	0	0	1	0	0	0	0.69	0	0	0	0	0.99	0.99	0.99	0.99
REV.CORETOTAL	0	0	0	0	1	1	1	1	0.63	0.63	0.72	0.69	0	0	0	0	0.99	0.99	0.99	0.99
REV.THIRDPARTYPUBLIC	0	0	0	0	1	1	1	1	0	0	0.69	0.66	0	0	0	0	0.99	0.99	0.99	0.99
REV.THIRDPARTYPRIVATE	0	0	0	0	1	1	1	1	0	0	0.69	0.66	0	0	0	0	0.99	0.99	0.99	0.99
REV.THIRDPARTYABROAD	0	0	0	0	0	0	0	1	0	0	0	0.66	0	0	0	0	0.99	0.99	0.99	0.99
REV.THIRDPARTYUND	0	0	0	0	0	0	0	1	0	0	0	0.66	0	0	0	0	0.99	0.99	0.99	0.99
REV.THIRDPARTYTOTAL	0	0	0	0	1	1	1	1	0	0	0.69	0.66	0	0	0	0	0.99	0.99	0.99	0.99
REV.TUITFEES	0	0	0	0	1	1	1	1	0.63	0	1	1	0	0	0	0	1	1	1	1
REV.STUDFEES.NC	0	0	0	0	1	1	1	1	0.63	0.63	0.66	0.66	0	0	0	0	0.99	0.99	0.99	0.99
REV.UNCL.NC	0	0	0	0	1	1	1	1	0.63	0.63	0.72	0.69	0	0	0	0	0.99	0.99	0.99	0.99
REV.CURRTOTAL.NC	0	0	0	0	1	1	1	1	0.63	0.63	0.72	0.69	0	0	0	0	0.99	0.99	0.99	0.99
REV.NONRECURR.NC	0	0	0	0	0	0	0	-	0	0	0	0	0	0	0	0	0	0	0	0.99
STA.ACAFTETOTAL	0	0	0.88	0.88	1	1	0.97	1	0.63	0.63	0.63	0.63	0	0	0	0	1	0.99	1	0.99
STA.ACAHCMEN	0	0	0.88	0.88	1	1	0.97	1	0.63	0.63	0.63	0.63	0	0	0.98	0.96	1	0.99	1	0.99
STA.ACAHCNAT	0	0	0	0	1	1	0.97	1	0	0	0	0	0	0	0.98	0.96	1	0.99	1	0.99
STA.ACAHCFOE05	0	0	0	0	0	1	0.97	0.97	0	0	0	0	0	0	0	0	1	0.99	1	0.99
STA.TOTACAHC	0	0	0.88	0.88	1	1	0.97	1	0.63	0.63	0.63	0.63	0	0	0.98	0.96	1	0.99	1	0.99
STA.PROFMEN	0	0	0.88	0.88	1	1	0.97	1	0.63	0.63	0.63	0.63	0	0	0.98	0.96	0.99	0.99	0.99	0.99
STA.PROFTOTAL	0	0	0.88	0.88	1	1	0.97	1	0.63	0.63	0.63	0.63	0	0	0.98	0.96	0.99	0.99	0.99	0.99
STA.INCLPHDSTUD	0	0	0.88	0.88	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1
STA.NONACAFTE	0	0	0	0	1	1	0.97	1	0.63	0.63	0.63	0.63	0	0	0	0	1	0.99	1	1
STA.NONACAHC	0	0	0	0	1	1	0.97	1	0.63	0.63	0.63	0.63	0	0	0	0	1	0.99	1	1
STA.TOTALFTE	0	0	0.88	0.88	1	1	0.97	1	0.63	0.63	0.63	0.63	0	0	0	0	1	0.99	1	1
STA.TOTALHC	0	0	0.88	0.88	1	1	0.97	1	0.63	0.63	0.63	0.63	0	0	0	0	1	0.99	1	1
STUD.LOWDEG	0	0	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1
STUD.HIGHDEG	0	0	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1
STUD.ISCED5MEN	0	0	-	-	1	1	0.98	1	-	-	-	-	0	0	0.95	0.97	1	0.99	0.99	1
STUD.ISCED5NAT	0	0	-	-	1	1	0.98	1	-	-	-	-	0	0	0.16	0	0.99	0.99	0.99	1
STUD.ISCED5RES	0	0	-	-	1	1	0.98	1	-	-	-	-	0	0	0	0	1	0.99	0.99	1
STUD.ISCED5FOE05	0	0	-	-	1	1	0.98	1	-	-	-	-	0	0	0	0	1	0.99	0.99	1
STUD.ISCED5TOTAL	0	0	-	-	1	1	0.98	1	-	-	-	-	0	0	0.95	0.97	1	0.99	0.99	1
STUD.ISCED6MEN	0	0	0.88	0.88	1	1	0.98	1	1	1	1	1	0	0	0.95	0.97	1	0.99	0.99	1
STUD.ISCED6NAT	0	0	0.88	0.88	1	1	0.98	1	1	1	1	1	0	0	0	0	0.99	0.98	0.98	1
STUD.ISCED6RES	0	0	0	0	1	1	0.98	1	0	0	0	0	0	0	0	0	1	0.99	0.99	1

Country	RS 2011	RS 2012	RS 2013	RS 2014	SE 2011	SE 2012	SE 2013	SE 2014	SK 2011	SK 2012	SK 2013	SK 2014	TR 2011	TR 2012	TR 2013	TR 2014	UK 2011	UK 2012	UK 2013	UK 2014
STUD.ISCED6FOE05	0	0	0	0	1	1	0.98	1	1	1	1	1	0	0	0	0	1	0.99	0.99	1
STUD.ISCED6TOTAL	0	0	0.88	0.88	1	1	0.98	1	1	1	1	1	0	0	0.95	0.97	1	0.99	0.99	1
STUD.ISCED7MEN	0	0	0.88	0.88	1	1	0.98	1	1	1	1	1	0	0	0.94	0.97	1	0.99	1	0.99
STUD.ISCED7NAT	0	0	0.88	0.88	1	1	0.98	1	1	1	1	1	0	0	0	0	0.99	0.98	0.99	0.99
STUD.ISCED7RES	0	0	0	0	1	1	0.98	1	0	0	0	0	0	0	0	0	1	0.99	1	0.99
STUD.ISCED7FOE05	0	0	0	0	1	1	0.98	1	1	1	1	1	0	0	0	0.01	1	0.99	1	0.99
STUD.ISCED7TOTAL	0	0	0.88	0.88	1	1	0.98	1	1	1	1	1	0	0	0.95	0.97	1	0.99	1	0.99
STUD.ISCED7LONGMEN	0	0	-	-	1	1	0.98	1	1	1	1	1	0	0	0	0	1	0.99	1	0.99
STUD.ISCED7LONGNAT	0	0	-	-	1	1	0.98	1	1	1	1	1	0	0	0	0	1	0.99	1	0.99
STUD.ISCED7LONGRES	0	0	-	-	1	1	0.98	1	0	0	0	0	0	0	0	0	1	0.99	1	0.99
STUD.ISCED7LONGFOE05	0	0	-	-	1	1	0.98	1	1	1	1	1	0	0	0	0	1	0.99	1	0.99
STUD.ISCED7LONGTOTAL	0	0	-	-	1	1	0.98	1	1	1	1	1	0	0	0	0	1	0.99	1	0.99
STUD.TOTALISCED5.7	0	0	0.88	0.88	1	1	0.98	1	1	1	1	1	0	0	0.95	0.97	1	0.99	0.99	0.99
STUD.DISTEDUINST	0	0	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1
GRAD.ISCED5MEN	0	0	-	-	1	1	0.98	1	-	-	-	-	0	0	0.95	0.97	1	0.99	0.99	1
GRAD.ISCED5NAT	0	0	-	-	1	1	0.98	1	-	-	-	-	0	0	0.22	0	0.99	0.99	0.99	1
GRAD.ISCED5RES	0	0	-	-	1	1	0.98	1	-	-	-	-	0	0	0	0	1	0.99	0.99	1
GRAD.ISCED5FOE05	0	0	-	-	1	1	0.98	1	-	-	-	-	0	0	0	0	1	0.99	0.99	1
GRAD.ISCED5TOTAL	0	0	-	-	1	1	0.98	1	-	-	-	-	0	0	0.95	0.97	1	0.99	0.99	1
GRAD.ISCED6MEN	0	0	0.88	0.88	1	1	0.98	1	1	1	1	1	0	0	0.95	0.97	1	0.99	0.99	1
GRAD.ISCED6NAT	0	0	0.88	0.88	1	1	0.98	1	1	1	1	1	0	0	0.01	0	0.99	0.98	0.98	1
GRAD.ISCED6RES	0	0	0.35	0.06	1	1	0.98	1	0	0	0	0	0	0	0	0	1	0.99	0.99	1
GRAD.ISCED6FOE05	0	0	0.35	0.06	1	1	0.98	1	1	1	1	1	0	0	0	0	1	0.99	0.99	1
GRAD.ISCED6TOTAL	0	0	0.88	0.88	1	1	0.98	1	1	1	1	1	0	0	0.95	0.97	1	0.99	0.99	1
GRAD.ISCED7MEN	0	0	0.88	0.88	1	1	0.98	1	1	1	1	1	0	0	0.95	0.97	1	0.99	1	0.99
GRAD.ISCED7NAT	0	0	0.88	0.88	1	1	0.98	1	1	1	1	1	0	0	0	0	0.99	0.98	0.99	0.99
GRAD.ISCED7RES	0	0	0	0	1	1	0.98	1	0	0	0	0	0	0	0	0	1	0.99	1	0.99
GRAD.ISCED7FOE05	0	0	0	0	1	1	0.98	1	1	1	1	1	0	0	0	0	1	0.99	1	0.99
GRAD.ISCED7TOTAL	0	0	0.88	0.88	1	1	0.98	1	1	1	1	1	0	0	0.95	0.97	1	0.99	1	0.99
GRAD.ISCED7LONGMEN	0	0	-	-	1	1	0.98	1	1	1	1	1	0	0	0	0	1	0.99	1	0.99
GRAD.ISCED7LONGNAT	0	0	-	-	1	1	0.98	1	1	1	1	1	0	0	0	0	1	0.99	1	0.99
GRAD.ISCED7LONGRES	0	0	-	-	1	1	0.98	1	0	0	0	0	0	0	0	0	1	0.99	1	0.99
GRAD.ISCED7LONGFOE05	0	0	-	-	1	1	0.98	1	1	1	1	1	0	0	0	0	1	0.99	1	0.99
GRAD.ISCED7LONGTOTAL	0	0	-	-	1	1	0.98	1	1	1	1	1	0	0	0	0	1	0.99	1	0.99
GRAD.TOTALISCED5.7	0	0	0.88	0.88	1	1	0.98	1	1	1	1	1	0	0	0.95	0.96	1	0.99	0.99	0.99

Country	RS 2011	RS 2012	RS 2013	RS 2014	SE 2011	SE 2012	SE 2013	SE 2014	SK 2011	SK 2012	SK 2013	SK 2014	TR 2011	TR 2012	TR 2013	TR 2014	UK 2011	UK 2012	UK 2013	UK 2014
GRAD.NOTESEDUCACT	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
RES.RESACTIVE	0	0	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1
RES.STUDISCED8MEN	0	0	0.88	0.88	1	1	1	1	1	1	1	1	0	0	0.94	0.97	1	0.99	1	0.99
RES.STUDISCED8NAT	0	0	0.88	0.88	1	1	1	1	1	1	1	1	0	0	0.01	0	1	0.99	1	0.99
RES.STUDISCED8RES	0	0	0.06	0.06	1	1	1	1	0	0	0	0	0	0	0.01	0	1	0.99	1	0.99
RES.STUDISCED8FOE05	0	0	0.06	0.06	1	1	1	1	1	1	1	1	0	0	0.01	0	1	0.99	1	0.99
RES.STUDISCED8TOTAL	0	0	0.88	0.88	1	1	1	1	1	1	1	1	0	0	0.95	0.97	1	0.99	1	0.99
RES.GRADISCED8MEN	0	0	0.88	0.88	1	1	1	1	1	1	1	1	0	0	0.95	0.97	1	0.99	1	0.99
RES.GRADISCED8NAT	0	0	0.88	0.88	1	1	1	1	1	1	1	1	0	0	0.01	0	1	0.99	1	0.99
RES.GRADISCED8RES	0	0	0.29	0.29	1	1	1	1	0	0	0	0	0	0	0.01	0	1	0.99	1	0.99
RES.GRADISCED8FOE05	0	0	0.29	0.29	1	1	1	1	1	1	1	1	0	0	0.01	0	1	0.99	1	0.99
RES.GRADISCED8TOTAL	0	0	0.88	0.88	1	1	1	1	1	1	1	1	0	0	0.95	0.97	1	0.99	1	0.99
RES.R.DEXP	0	0	0	0	0.95	0.95	1	1	0	0.71	0.79	0.71	0	0	0	0	0	0	0	0