ICCS

IEA

ICCS 2022 User Guide for the International Database

EDITORS:
Yasin Afana
Falk Brese
Hannah Kowolik
Diego Cortes
Wolfram Schulz

IEA INTERNATIONAL CIVIC AND CITIZENSHIP EDUCATION STUDY 2022

Yasin Afana
International Association for the Evaluation of Educational Achievement
Hamburg, Germany

Falk Brese
International Association for the Evaluation of Educational Achievement
Hamburg, Germany

Hannah Kowolik
International Association for the Evaluation of Educational Achievement
Hamburg, Germany

Diego Cortes
International Association for the Evaluation of Educational Achievement
Hamburg, Germany

Wolfram Schulz
The Australian Council for Educational Research
Camberwell, Victoria
Australia

Contributors: Renato Alves Coppi, Umut Atasever, Svenja Kalmbach, Julian Fraillon, Tim Friedman

IEA

The International Association for the Evaluation of Educational Achievement (IEA), with headquarters in Amsterdam, is an independent, international cooperative of national research institutions and governmental research agencies. It conducts large-scale comparative studies of educational achievement and other aspects of education, with the aim of gaining in-depth understanding of the effects of policies and practices within and across systems of education.

Yasin Afana • Falk Brese • Hannah Kowolik
Diego Cortes • Wolfram Schulz

ICCS 2022 User Guide for the International Database

IEA International Civic and Citizenship Education Study 2022

Contents

List of Tables and Figures

Tables

Figures

1 STUDY OVERVIEW: THE DATA AND THE IMPLICATIONS FOR ANALYSIS

Yasin Afana, Falk Brese, and Wolfram Schulz

1.1 Main objectives and scope

The International Civic and Citizenship Education Study (ICCS) 2022 investigated the ways in which young people are prepared to undertake their roles as citizens in a range of countries in the second decade of the 21st century. It studied students' knowledge and understanding of civics and citizenship as well as their attitudes, perceptions, and activities related to civics and citizenship. Based on nationally representative samples of students, the study further examined differences among countries in relation to these outcomes of civic and citizenship education, and explored how cross-national differences relate to student characteristics, school and community contexts, and national characteristics. As the third cycle of this study, ICCS 2022 is a continuation and an extension of ICCS 2009 and ICCS 2016. Some materials and variables are statistically linked and allow for changes over time to be investigated.

The International Association for the Evaluation of Educational Achievement (IEA) established ICCS in order to meet the need for continuing research on civic and citizenship education and as a response to widespread interest in conducting regular international assessments of this field of education. ICCS 2022 was intended as an exploration of enduring and emerging challenges of educating young people in a world where contexts of democracy and civic participation continue to change.

ICCS addressed research questions concerned with the following broader domains:

1. Students' knowledge and understanding of civics and citizenship and the factors associated with variations in this civic knowledge.

2. Students' current and expected future involvement in civic-related activities, their perceptions of their capacity to engage in these activities, and their perceptions of the value of civic engagement.

3. Students' beliefs about contemporary civil and civic issues in society, including those concerned with civic institutions, rules, and social principles (democracy, citizenship, and diversity), as well as their perceptions of their communities and threats to the world's future.

4. Contextual information that reflects the ways in which education systems and schools organize civic and citizenship education, with a particular focus on general approaches, the curriculum and its delivery, including processes implemented to facilitate future citizens' civic engagement and interaction within and across communities.

In each of these domains, ICCS 2022 investigated variations within and across countries, factors associated with those variations, and changes since ICCS 2009 and ICCS 2016.

ICCS gathered data from more than 82,000 students in their eighth year of schooling in about 3,400 schools from 22 countries and two German benchmarking entities (North Rhine-Westphalia and Schleswig-Holstein). Several countries had participated in either ICCS 2009, ICCS 2016, or both. The student data were augmented by data from more than 40,000 teachers in those schools and by contextual data collected from school principals and national research centers. An additional European student questionnaire in ICCS 2022 gathered data from almost 66,500 students in 18 European countries and the two benchmarking participants. The Latin American student questionnaire in ICCS 2022 gathered data from more than 10,000 students in the two participating Latin American countries.

1.2   The design in brief

The ICCS 2022 international database offers researchers and analysts a rich environment for examining students' civic knowledge in an international context. This includes:

- Comparable data for 22 countries and two benchmarking entities from around the world providing an international perspective from which to examine educational practices and student outcomes in civic and citizenship education.

- Comparable regional data for 18 countries and two benchmarking from the European region and two countries from the Latin American region that allow investigations on aspects of civic and citizenship education of specific relevance in each of these geographic regions.

- Students' civic knowledge linked to questionnaire information from students and school principals, providing policy-relevant contextual information on the antecedents of civic knowledge.

- Data from the teacher questionnaire that provide additional contextual information about the organization and culture of schools as well as data on general and civic-specific aspects of teaching.

- Student civic knowledge scores on the scale established in 2009 and 2016 to compare changes in civic knowledge across these three cycles of ICCS.

The ICCS 2022 main target population was defined as students in the grade that represented eight years of schooling, counting from the beginning of Level 1 of the International Standard Classification of Education (ISCED), provided that the average age of students in this grade was at least 13.5 years, so usually at grade 8 at the time of the assessment. If the average age of students in that grade was less than 13.5 years, the following grade (grade 9) became the target population. In Norway, the national population definition deviated from the international population definition as the survey was administered to grade 9 students and results are annotated accordingly in all ICCS 2022 reports.

The target population for the ICCS 2022 teacher survey was defined as all teachers teaching regular school subjects to the students of the target grade during the testing period and since the beginning of the school year. A specific segment of the teacher questionnaire collected information from teachers teaching subjects related to civic and citizenship education.

Random samples that involved multiple sampling stages, clustering, and stratification were selected for all target populations. In most participating countries, about 150 schools, one class per school, and 15 teachers per school were sampled. Minimum exclusion and target response rates were determined to secure high-quality data. Details about the sampling design are reported in Chapter 6 of the ICCS 2022 technical report (Schulz et al., 2024b).

1.3   Analyzing the data

The ICCS 2022 design and operations resembled procedures used in past and current educational surveys and student achievement studies, such as, for example, the IEA Trends in International Mathematics and Science Study (TIMSS), the IEA Progress in International Reading Literacy Study (PIRLS), and the IEA International Computer Information Literacy Study (ICILS). ICCS 2022 was an ambitious and demanding study, involving complex procedures for drawing samples, collecting data, and analyzing and interpreting findings. Researchers using the database need to understand the characteristics of the study to work effectively with the information it contains as well as the recommendations and advice provided in this user guide.

1.3.1 Resources and requirements

This user guide describes the organization, content, and use of the international database from a practical perspective. It is imperative that it be used in conjunction with the ICCS 2022 technical report (Schulz et al., 2024b), which provides a comprehensive account of the conceptual, methodological, and analytical implementation of the study. The ICCS 2022 international report (Schulz et al., 2024a), the European report (Damiani et al., 2024) are further key resources. Using all these publications in combination will allow analysts to understand and confidently replicate the procedures used, as well as to correctly undertake new analyses in areas of special interest.

At a minimum, an analyst carrying out statistical analysis will need to have a good understanding of the conceptual foundations of ICCS 2022, the themes addressed, the populations targeted, the samples selected, the instruments used, and the production of the international database. All this information is covered and explained in detail in the ICCS 2022 technical report and sketched in practical terms in this user guide. Researchers using the database also need to make themselves familiar with the database structure and its included variables (Chapter 2 in this guide). While it is not necessary to fully understand all the methods used to construct, validate, and compute the derived scales, analysts must be aware of their potential limitations (see Chapters 11 and 12 in the ICCS 2022 technical report).

Other important aspects to keep in mind when working with ICCS data are these:

- ICCS 2022 is an observational, nonexperimental study that collected cross-sectional data. For this reason, causal inferences, and language of the type "condition A caused effect B," "factor A influenced outcome B," and "variable A impacted on variable B" cannot and should not be established with ICCS 2022 data alone. The reports containing the international results of the study refrain from making such inferences or using causal language.

- The ICCS 2022 instruments included a variety of questions relating to factual information as well as to attitudes, beliefs, and perceptions. All this information was self-reported by the principals, teachers, and students. Furthermore, because population features were not observed but estimated using

7

sample data, wording such as "the estimated proportion of students with X is …" is preferable to writing "X percent of students are …"

- Nearly all variables in ICCS 2022 are categorical in nature (nominal or ordered). Analysts may therefore need to consider using categorical, nonparametric analysis methods for these types of variables. Techniques for continuous variables (provided that the required assumptions hold) should only be used on counts and on the derived scales obtained through data reduction or scaling methods such as factor analysis, structural equation modelling, or item response theory.

Analysts also need to have a working knowledge of SPSS (as the software of choice for this user guide) and knowledge of basic inferential statistics, such as estimating means, correlations, and linear regression parameters. Appropriate theoretical knowledge will be needed to conduct advanced analyses such as logistic regressions.

1.3.2 Estimation requirements

Researchers familiar with population estimation in large-scale education-survey databases such as TIMSS, PIRLS, and other IEA studies will have little difficulty analyzing ICCS 2022 data once they have familiarized themselves with the study's conceptual foundation and its methodological, operational, and analytical details. If, as a user of the ICCS 2022 international database, you are not accustomed to working with complex survey sample data, this guide should provide you with sufficient technical information to enable you to conduct correct basic analysis.

The three main design features of ICCS 2022 that you will need to consider during any secondary analysis of the study's data are the following:

1. The unequal selection probabilities of the sampling units that necessitate the use of weights during computation of estimates;

2. The complex multistage cluster sample design that was implemented to ensure a balance between the research goals and cost-efficient operations; and

3. The rotated design of the civic knowledge test, wherein students completed only samples of the test items rather than the full set of test items.

Chapter 3 of this user guide includes a brief account of the weights and variance estimation techniques intended for ICCS, whereas Chapters 6 and 10 of the ICCS 2022 technical report provide a more detailed description of the sample design and the weighting procedures. A detailed description of the ICCS 2022 scaling and how the civic knowledge scale was created is available in Chapter 11 of the ICCS 2022 technical report (Schulz et al., 2024b).

To obtain accurate and representative samples, ICCS 2022 used a two-stage sampling procedure whereby a random sample of schools is selected at the first stage and one or two intact target grade classes in the case of students or a random sample of teachers from the target grade are sampled at the second stage. This is an effective and efficient sampling approach given ICCS' purpose of describing population characteristics, but the resulting student sample has a complex structure that must be taken into consideration when analyzing the data. In particular, sampling weights need to be applied and a variance estimation technique adequate for complex

samples, in the case of ICCS the jackknife repeated replication (JK2) approach, needs to be used to estimate sampling variances correctly.[1]

Because each imputed score is an estimate based on limited information, it is subject to measurement error. To allow analysts to account for this error when analyzing the civic knowledge data, the international database provides five separate imputed scores for the civic knowledge scale. Any analysis involving civic knowledge scores needs to be replicated five times, using a different plausible value each time, with the results then combined into a single result that includes information on standard errors that incorporate both sampling and imputation error.[2]

This user guide is principally tailored to SPSS (IBM Corp., 2019), one of the most widely used statistical packages in the social sciences and educational research. Unfortunately, the base SPSS used in the current analyses (i.e., Version 26) does not directly support complex survey designs such as those used in ICCS 2022 and cannot be used "out of the box" for methodologically correct estimation of sampling errors and of test statistics. The base SPSS assumes that data come from a single-stage, simple random sample, which is not the case in ICCS 2022 and most if not all other large-scale assessments in education. A "complex samples" module for SPSS is available. However, it includes only one of many variance estimation approaches, namely Taylor Series Variance Estimation, and does not support jackknife replication for estimating sampling errors, which was the technique applied for the reporting of ICCS 2022 results.

This gap has been filled by IEA's International Database (IDB) Analyzer (IEA, 2024), available free of charge to analysts and researchers using the ICCS 2022 database, shorthand 'IEA IDB Analyzer'. The IEA IDB Analyzer employs SPSS, SAS (SAS Institute, 2016), and R (R Core Team, 2023) as an engine to compute population estimates and design-based standard errors using replication for a variety of international large-scale assessments. IEA developed the IEA IDB Analyzer in the context of its large-scale student assessments TIMSS and PIRLS and adapted it for use with data from ICCS 2022 and other studies. The IEA IDB Analyzer allows users to compute estimates of percentages, means, percentiles, correlations, and linear as well as logistic regression parameters, including their respective standard errors. It also simplifies the management of the ICCS 2022 international database by providing a module for selecting subsets of countries and variables, and for merging files for analysis. Chapter 4 of this current guide provides in-depth information about the IEA IDB Analyzer as well as examples illustrating its use.

If you are an occasional user of the database, you may not want to use one of the commercial statistical software packages with their associated costs. You can, however, access alternative packages suitable for analyzing complex sample data. Available in addition to the IEA IDB Analyzer are a growing number of software packages able to handle the jackknifing replication method implemented in ICCS 2022.

Commercial packages that include support for the weights and the replication method used in ICCS 2022 are SAS 9.4 and later editions, or Stata 13 and later editions (StataCorp, 2013). While these software packages support the variance estimation for complex samples in ICCS 2022, they do not always support these in combination with the multiple imputations' methodology that ICCS 2022 used for describing and representing students' civic knowledge

---

[1] More details on the sampling design and its implementation are provided in Chapter 6 of the ICCS 2022 technical report (Schulz et al., 2024b).
[2] More details on plausible values can be found in Chapter 11 of the ICCS 2022 technical report (Schulz et al., 2024b).

data. Third-party scripts and macros may exist to provide this support, for example more recently as packages for R.

1.3.3 Limitations of the international database

When analyzing ICCS 2022 data, researchers need to keep the following constraints in mind:

- Malta assessed grade 9 students given that the average age of grade 8 students is below 13.5.

- Norway (grade 9) deviated from the internationally defined population and surveyed the adjacent upper grade.

- Exclusion rates pertaining to the student population were above five percent in Croatia, Latvia, Norway (grade 9), and Sweden. The ICCS 2022 research team deemed this level of exclusion a significant reduction of target population coverage and researchers need to keep this feature in mind when interpreting results.

- Participation rates in the student survey were below ICCS 2022 standards in Brazil, Denmark, and the benchmarking participant Schleswig-Holstein (Germany), resulting in a separated presentation of the results in the ICCS 2022 reports. Population estimates derived with student-level data from these countries hold a higher risk of bias and therefore should be interpreted with caution and not compared with data from other countries.

- Participation rates in the teacher survey were below ICCS 2022 standards in Brazil, Colombia, Cyprus, Denmark, Estonia, France, Latvia, the Netherlands, Sweden, and the benchmarking participant North Rhine-Westphalia (Germany), resulting in a separated presentation of the results in the ICCS 2022 reports. Population estimates derived with teacher-level data from these countries hold a higher risk of bias and therefore should be interpreted with caution and not compared with data from other countries.

- Concerns about the substantial low response rates (less than 10%) for the teacher survey in Schleswig-Holstein (Germany) led to a decision not to include the corresponding data in the international database.

For details on population coverage and exclusion rates for countries participated in ICCS 2022, please refer to Chapter 6 of the ICCS 2022 technical report. For details on participation rates, please refer to Chapter 10 of the ICCS 2022 technical report (Schulz et al., 2024b).

1.4 Contents of this guide

This ICCS 2022 user guide describes the content and format of the data in the ICCS 2022 international database. In addition to this introduction, the ICCS 2022 user guide includes the following three chapters:

- Chapter 2 describes the structure and content of the ICCS 2022 international database.

- Chapter 3 introduces the use of weighting and variance estimation variables for analyzing the ICCS 2022 data.

- Chapter 4 introduces the IEA International Database (IDB) Analyzer software (IEA, 2024) and presents examples of analyses of the ICCS 2022 data using this software in conjunction with SPSS.

The ICCS 2022 user guide is accompanied by four appendices:

- Appendix A includes the international version of all international questionnaires administered in ICCS 2022, also containing the questionnaires of the regional student instruments. These serve as a reference guide to the questions asked and the variable names used to record the responses in the international database.

- Appendix B provides details on all national adaptations that were applied to the national version of all ICCS 2022 international questionnaires. When using the database, please refer to this appendix and check for any special adaptations made to the international versions of the ICCS 2022 questionnaires that could potentially affect the results of analyses.

- Appendix C describes how the derived questionnaire variables used for producing tables in the ICCS 2022 international and regional reports were computed.

- Appendix D contains all restricted use items of the ICCS 2022 assessment of civic knowledge along with their respective scoring guides. The restricted use items are made available to illustrate the content of ICCS 2022. Prior permission is required when using IEA data sources.[1]

---

[1] All publications and restricted use items by ICCS and other IEA studies, as well as translations thereof, are for non-commercial, educational and research purposes only. Prior permission is required when using IEA data sources for assessments or learning materials. IEA Intellectual Property Policy is inter alia included on the IEA Data Repository (http://www.iea.nl/data). IEA copyright must be explicitly acknowledged (© IEA 2024), and the need to obtain permission for any further use of the published text/material clearly stated in the requested use/display of this material.

Exploitation, distribution, redistribution, reproduction and/or transmitting in any form or by any means, including electronic or mechanical methods such as photocopying, information storage and retrieval system of these publications, restricted use items, translations thereof and/or part thereof are prohibited unless written permission has been provided by IEA.

References

Damiani, V., Agrusti, G., Losito, B., & Schulz, W. (2024). *Young citizens' views and engagement in a changing Europe. IEA International Civic and Citizenship Education Study 2022 European report*. International Association for the Evaluation of Educational Achievement (IEA).

IBM Corp. (2019). IBM SPSS Statistics for Windows, Version 26.0. [Computer Software]. Armonk, NY: IBM Corp

IEA (2024). IDB Analyzer (Version 5.0.23) [Computer Software]. International Association for the Evaluation of Educational Achievement. Available online at http://www.iea.nl/data.

R Core Team (2023) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

SAS Institute Inc. (2016). SAS® 9.4. [Computer Software]. Cary, NC: SAS Institute Inc.

Schulz, W., Ainley, J., Fraillon, J., Losito, B., Agrusti, G., Damiani, V., & Friedman, T. (2024a). *Education for Citizenship in Times of Global Challenge. IEA International Civic and Citizenship Education Study 2022 International Report*. International Association for the Evaluation of Educational Achievement (IEA).

Schulz, W., Friedman, T., Losito, B., & Fraillon, J. (Eds.) (2024b). *ICCS 2022 technical report.* International Association for the Evaluation of Educational Achievement (IEA).

StataCorp. (2013). Stata 13 [Computer software]. College Station, TX: StataCorp.

2 THE ICCS 2022 INTERNATIONAL DATABASE

Yasin Afana

2.1 Overview

The International Civic and Citizenship Education Study (ICCS) 2022 international database (IDB) contains student civic knowledge test data and international student, teacher, and school questionnaire data collected in the 22 countries and two benchmarking entities around the world that participated in the study. The database also includes data from the ICCS 2022 National Contexts Survey, providing information on the national contexts of civic and citizenship education for all participating countries. Additionally, for countries participating in one of the two regional student questionnaires included in ICCS 2022 (European and Latin American), the database contains regional questionnaire data.

Table 2.1 lists all ICCS 2022 countries, along with the operational codes used to identify them in the ICCS 2022 international database and includes information about country participation in the regional student questionnaires in ICCS 2022, 2016, and 2009. Some countries were not included in cross-national comparisons in the international reports (see Table 2.1), because of low participation in either ICCS 2022, ICCS 2016 or ICCS 2009, or any applicable combination(s) of the three cycles.

For details on population coverage and exclusion rates for countries that participated in ICCS 2022, please refer to Chapter 6 of the ICCS 2022 technical report; for details on participation rates, please refer to Chapter 10 of the ICCS 2022 technical report (Schulz et al., 2024b).

The database also contains materials that provide additional information on its structure and content. This chapter describes the content of the database and is divided into eight major sections corresponding to the different file types and materials included in the database.

Table 2.1: Countries participating in ICCS 2022

Countries	Operational codes		ICCS 2022 Regional Questionnaires		ICCS 2016	Regional Questionnaires		ICCS 2009	Regional Questionnaires	
	Alpha-3	Numeric	European	Latin American	Partici-pation in 2016	European	Latin American	Partici-pation in 2009	European	Latin American
Brazil[1,2]	BRA	76		•						
Bulgaria	BGR	100	•		•	•		•	•	
Chinese Taipei	TWN	158			•			•		
Colombia[2]	COL	170		•	•		•	•		•
Croatia	HRV	191	•		•	•				
Cyprus[2]	CYP	196	•					•	•	
Denmark[1,2]	DNK	208	•		•	•		•	•	
Estonia[2]	EST	233	•		•	•		•	•	
France[2]	FRA	250	•							
Italy	ITA	380	•		•	•		•	•	
Latvia[2]	LVA	428	•		•	•		•	•	
Lithuania	LTU	440	•		•	•		•	•	
Malta[3]	MLT	470	•		•	•		•	•	
Netherlands[2]	NLD	528	•		•	•		•	•	
Norway (9)[4]	NOR	578	•		•	•		•		
Poland	POL	616	•					•	•	
Romania	ROU	9642	•							
Serbia	SRB	688								
Slovak Republic	SVK	703	•					•	•	
Slovenia	SVN	705	•		•	•		•	•	
Spain	ESP	724	•					•	•	
Sweden[2]	SWE	752	•		•	•		•	•	

Benchmarking participants

Countries	Operational codes		ICCS 2022 Regional Questionnaires		ICCS 2016	Regional Questionnaires		ICCS 2009	Regional Questionnaires	
North Rhine-Westphalia (Germany)[2]	DNW	276001	•		•	•				
Schleswig-Holstein (Germany)[1,5]	DSH	276005	•							

2.2 ICCS 2022 database

The ICCS 2022 database comprises data from all instruments administered to the students, the teachers teaching in the target grade at their school, and their school principals. This includes the student responses to the international civic

---

[1] Country did not meet sample participation requirements for the student survey.

[2] Country did not meet sample participation requirements for the teacher survey.

[3] Country assessed Grade 9 students given that the average age of Grade 8 students is below 13.5.

[4] Country deviated from international defined population and surveyed adjacent upper grade.

[5] Concerns about the extremely low response rates (less than 10%) for the teacher surveys led to a decision not to include the corresponding data in the international database.

knowledge test items, the responses to the international student, teacher, and school questionnaires, and responses to the regional student questionnaires. The files also contain the civic knowledge test scores estimated for participating students and derived variables for reporting study findings in the ICCS 2022 international reports. National Research Coordinators' responses to the National Contexts Questionnaire are also part of the international database.

The ICCS 2022 data files are provided in SPSS format (.sav), SAS format (.sas7bdat), and R format (.Rdata) except for the National Contexts Survey data, which are only available in SPSS format (.sav). The files can be downloaded from the IEA Study Data Repository at http://www.iea.nl/data. The database contains files for each country that participated in ICCS 2022 for which internationally comparable data are available.

The file name identifies the type of data file and the country (Table 2.2). For example, ISGNORC4.sav is an SPSS file that contains Norway's ICCS 2022 target grade student questionnaire data. For each file type, a separate data file is provided for each participating country, with the exception of countries that did not meet sampling requirements for some types of respondents. For example: when a country did not meet sampling requirements for the teacher survey, then no teacher data were released. Files of the same type include the same uniformly defined set of variables across countries.

ICCS 2022 provided two assessment scenarios for participating countries. In the first scenario, the student assessment was administered in one mode, either as paper-based or computer-based. To provide an intermediate link (i.e., a bridge) between the paper-based and the computer-based administration, a so-called "bridge" scenario was implemented by those participating countries that changed from the paper-based in earlier cycles of ICCS to the computer-based mode in ICCS 2022. The aim was to investigate potential differences between the two different administration modes. In those countries, the study was administered to an additional "bridge" sample of students (and sometimes schools), with the "international" sample using the computer-based mode and the bridge sample using the paper-based mode. For all participating countries in ICCS 2022, the IDB data files are named with the two-digit suffix *C4*, referring to the 4th cycle of ICCS. For eleven countries that administered the bridge scenario, additional files are available containing the data used for checking for mode effects. These files are named with the suffix *B4*. Data from the bridge sample was not used for international reporting.

Table 2.2: ICCS 2022 data file names

File name (core/bridge sample)	Description
ISG•••C4/B4	International Student Questionnaire File
ISA•••C4/B4	Student Civic Knowledge Test File
ISR•••C4/B4	Student Reliability File
ISE•••C4/B4	European Student Questionnaire File
ISL•••C4/B4	Latin American Student Questionnaire File
ITG•••C4	Teacher Questionnaire File
ICG•••C4/B4	School Questionnaire File
NCQICSC4	National Contexts Questionnaire File

The SPSS files include full dictionary/meta information, namely, variable names, formats (type, width, and decimals), variable labels, value labels, missing values, and appropriately set measurement levels (nominal, ordinal, or scale). The dictionary information can be accessed through the *SPSS View → Variables* menu, or in output form through the *File → Display Data File Information* menu. SAS files include appropriate display formats and variable labels but do not permanently store value labels in data files. In R files, details about variables, mirroring the comprehensive metadata as in the SPSS files, are typically captured using the attributes' function. This displays variable label, missing values, class information, format specifics and value labels. In R (copied from SPSS), variable information is often stored in numeric format. Users may find it necessary to convert these numeric variables to a labeled format, especially when dealing with categorical variables.

Please note, the SPSS data files were created in Unicode mode. However, when saving SPSS data files in Unicode encoding in code page mode, defined string widths are automatically tripled. These format changes will then cause problems when merging data with the IDB Analyzer (i.e., when merging the data in SPSS). Researchers should take this into account when saving data files in SPSS.

2.2.1 Questionnaire data files

There are five types of ICCS 2022 questionnaire data files corresponding to the five types of questionnaires administered in ICCS 2022. The international, European, and Latin American student data files and the teacher and school data files contain the responses to the questions asked in the respective questionnaire.

All questionnaire data files feature a number of structure and design variables, sampling and weight variables, and derived variables from the respective questionnaire data that were used for analyses in the international reports including questionnaire scales. These variables are described later in this chapter (see section 2.4).

School questionnaire data files (ICG)

The school questionnaire data files contain responses from school principals to the questions in the ICCS 2022 school questionnaires.

Although school-level analyses where schools are the units of analysis can be performed, it is preferable to analyze school-level variables as attributes of students or teachers. To perform student- or teacher-level analyses with school data, the school questionnaire data files must be merged with the student or teacher questionnaire data files using the country and school identification variables. The merging procedure using the IEA IDB Analyzer is described in Chapter 4 of this user guide.

Teacher questionnaire data files (ITG)

The teachers that were sampled for ICCS 2022 were administered one questionnaire to collect information about school and classroom contexts, connections between schools and local communities, perceived objectives of civic and citizenship education, and approaches to teaching in this learning area.

It is important to note that, in contrast to other IEA surveys, the teachers in the teacher questionnaire data files constitute a representative sample of target grade teachers in a country. However, student and teacher data must

not be merged directly because these two groups constitute separate target populations. Chapter 4 of this user guide describes teacher-level analyses with school data using the IEA IDB Analyzer software.

International student questionnaire data files (ISG)

Students who participated in ICCS 2022 were administered a questionnaire with questions related to their home background, perceptions of their school context, their attitudes toward civic principles, institutions, and important topics in society, as well as aspects related to their civic engagement. The international student questionnaire data files contain students' responses to these questions. They also contain students' civic knowledge test scores (plausible values) to facilitate analyses of relationships between student background and student perceptions, characteristics, and achievement.

Regional student questionnaire data files (ISE; ISL)

Students from European and Latin American countries were administered regional student questionnaires in addition to the student test booklet and the international student questionnaire. The questions in the regional questionnaires were related to students' attitudes and perceptions relevant to the region. The questionnaire data files contain students' responses to these questions.

Questionnaire response code values

A series of conventions were adopted to code the data included in the ICCS 2022 questionnaire data files.

The values assigned to each of the questionnaire variables depend on the item format and the number of options available. For categorical questions, sequential numerical values were used to indicate the available response options. For example, the first response option was represented by a 1, the second response option by a 2, and so on. Check-all-that-apply questions were coded as "checked" if the corresponding option was chosen, otherwise, they were coded as "not checked". Open-ended questions, such as "the number of students in a school", were coded with the actual number given as the response.

2.2.2 Student civic knowledge test data files (ISA)

The ICCS 2022 student civic knowledge test data files contain the student responses to the individual test items in the ICCS 2022 assessments. The student test data files are best suited for performing item-level analyses. The civic knowledge test scores (plausible values) for the ICCS 2022 civic knowledge scale are available in both, the student test data files (ISA) and the international student questionnaire data files (ISG).

In ICCS 2022, three different booklet designs were provided depending on the sampling scenarios: For the paper-based test scenario (11 booklets), for the computer-based test (CBA) scenario (14 booklets), and for the so-called "bridging" scenario (8 booklets). Students who participated in ICCS 2022 were administered one of the test booklets depending on the sampling scenario, each including a series of items. Most of these items were multiple-choice items and some were constructed-response items. The student test data files contain the actual responses to the multiple-choice questions and the scores assigned to the constructed-response items.

Except for the items already presented in the international report and Appendix D of this user guide, the items administered in ICCS 2022, and associated materials (such as scoring guides) will remain secure for future use and hence are not available for secondary analysis.

Item response code values

A series of conventions also were adopted to code the data included in the civic knowledge test data files.

The values assigned to each of the test item variables also depend on the item format. For multiple-choice items, numerical values from 1 through 4 were used to correspond to the response options A through D, respectively. For these items, the correct response is marked with an asterisk (*) following the value label of the correct option.

Each of the 22 constructed-response items had its own scoring guide that used a one-digit scoring scheme. The constructed-response items in ICCS 2022 can be categorized into 3 categories depending on students' responses. Ten Items (CI3CPO1, CI3CPO2, CI4SLC3, CI4SLC4, CI4ZHC3, CI4ZHC4, CI4ZEC4, CI4ZSC4, CI4ZSC5Y, CI4ZSC5N) had a valid score range of 0 (= incorrect response) and 1 (= correct response). Eleven Items (CI2WFO2, CI4FKC1, CI4FWC2, CI2WFO1, CI4NHC2, CI2BIO1, CI4DVC1, CI3PRO1, CI4EDC2, CI4ZEC2, CI4ZEC5) followed a different scoring logic and scored so that responses that included two different described conceptual categories were scored as 2 and any response related to a single described conceptual category was scored as 1, resulting in codes 0 (= incorrect response), 1 (= partially correct response), and 2 (= correct response). One item (CI4ZHC1) had two values for partially correct responses, with codes 0 (= incorrect response), 1 (= partially correct response), 2 (= partially correct response), and 3 (= correct response). The "missing" code (9) was used when a student made no attempt to answer a question. This code was only allocated when the entire stimulus, question stem and question response area were left blank by the student.

2.2.3 Within-country scoring reliability data files (ISR)

The ICCS 2022 within-country scoring reliability data files contain data that can be used to investigate the reliability of the ICCS 2022 constructed-response item scoring. The scoring reliability data files contain one record for each booklet that was double-scored during the within-country scoring reliability exercise. For each constructed-response item in the civic knowledge test, the following three variables are included in the scoring reliability data files:

Original score (score assigned by the first scorer);
Second score (score assigned by the second scorer);
Score agreement (degree of agreement between the two scorers).

It should be noted that the second score data were used only to evaluate within-country scoring reliability and were not used when computing the test scores included in the database and presented in the international reports.

Reliability variable score values

The values contained in both the original score and second score variables are the one-digit diagnostic codes assigned following the ICCS 2022 scoring guides. The score agreement variable may have one of two values,

depending on the degree of agreement between the two scorers: code 0 was assigned if different scores were assigned. Code 1 was assigned in case of agreement between both scorers.

2.2.4 National Contexts Questionnaire data file

This data file contains the responses provided by National Research Coordinators of the participating countries to the ICCS 2022 National Contexts Questionnaire. The National Contexts Survey was designed to systematically collect relevant data on the structure of the education system, education policy, civic and citizenship education, teacher qualifications for civic and citizenship education, and the extent of current debate and reforms in this area. The survey also collected data on processes at the national level regarding assessment of and quality assurance in civic and citizenship education and in school curriculum approaches. The National Contexts Questionnaire was administered online using an assessment platform owned by the Australian Council for Educational Research (ACER).

The National Contexts Questionnaire data file (NCQICSC4.sav) is available in SPSS format and contains data for all countries participating in ICCS 2022.

2.3 Records included

The international database includes all records that satisfy the international sampling standards. Data from those respondents who either did not participate or did not pass adjudication because, for example, within-school participation was not sufficient, were removed from the final database.

More specifically, the database contains records for the following:

- All participating schools: any school where the school principal responded to the school questionnaire has a record in the school-level files. Participation in ICCS 2022 at the school level is independent of participation at the student and/or teacher levels for the same school.

- All participating teachers: any teacher who responded to the teacher questionnaire has a record in the teacher-level files, provided that at least 50 percent of the sampled teachers at that school participated in the study.

- All participating students: any student who responded to at least one item of the student test or the international student questionnaire has a record in the student-level files, but only if the respective school was regarded as participating in the student survey. A school was regarded as having participated in the student survey if, in its sampled class(es), at least 50 percent of the students participated, and all sampled classes participated. A class was regarded as having participated if at least 50 percent of its students participated.

Consequently, the following records were excluded from the database:

- Schools where the principal did not respond to the questionnaire;

- Teachers who did not respond to the questionnaire;

- Teachers from those schools where less than 50 percent of the sampled teachers participated;

- Students who could not or refused to participate, or did not respond to any items in the student test or the international student questionnaire;

- Students from those schools with sampled classes where less than 50 percent of the students participated;

- Students and/or teachers who were afterwards reported as not in scope, ineligible, or excluded;

- Students and/or teachers who participated but were not part of the sample; and

- Any other records that were considered unreliable, of undocumented origin, or otherwise in violation of accepted sampling and adjudication standards.

Any additional data collected by countries to meet national requirements were also excluded from the international database.

Further information on the ICCS 2022 participation and sampling adjudication requirements is available in Chapter 10 of the ICCS 2022 technical report (Schulz et al., 2024b).

2.4 Survey variables

The database contains the following information for each school that participated in the survey:

- The identification variables for the country and school;
- The school principal's responses to the school questionnaire;
- Additional structure and design variables;
- The school indices and scales derived from the original questions in the school questionnaires;
- Weights and variance estimation variables pertaining to schools; and
- The version and the scope of the database.

For each teacher who participated in the survey, the database contains:

- The identification variables for the country, school, and teacher;
- The teacher's responses to the teacher questionnaire;
- Additional structure and design variables;
- The teacher indices and scales derived from the original questions in the teacher questionnaire;
- The weights and variance estimation variables pertaining to teachers; and
- The version and the scope of the database.

For each student who participated in the survey, the following information is available:

- The identification variables for the country, school, class and student;

- The student's responses to the student questionnaire;
- The student's responses to the student civic knowledge test;
- Additional structure and design variables;
- The student civic knowledge test scores;
- The student indices and scales derived from the original questions in the student questionnaire;
- The weights and variance estimation variables pertaining to students; and
- The version and the scope of the database.

The next three sections of this chapter (sections 2.4.1–2.4.3) offer more detailed explanations of these variables.

2.4.1 Questionnaire variables

The questionnaire variable names consist of a 6- to 8-character string (e.g., IS4G04A). The variable names used in the database were assigned using a consistent and systematic naming convention:

- The first character indicates the reference level. The letter "I" is used for variables that are administered on an international level. The letter "E" is used for variables from the European student questionnaire, and the letter "L" is used for variables from the Latin American student questionnaire.

- The second character indicates the type of respondent. The letter "C" is used to identify data from school principals, the letter "T" is used for teacher data, and the letter "S" for student data.

- The third character indicates the study cycle: Number "4" identifies ICCS 2022 as the 4th cycle of an IEA study focusing on civic and citizenship education.

- The fourth character consists of the letter "G", which is used for all (background) questionnaire variables.

- The fifth, sixth, seventh and eighth characters indicate the question number. Their combination is unique to each variable within a questionnaire.

2.4.2 Student civic knowledge test item and scoring reliability variables

The names of the item variables pertaining to the international test are based on an alphanumeric code consisting of seven characters (e.g., CI4PRM1), which adheres to the following rules:

- The first character indicates the general study context. "C" stands for civic and citizenship education.

- The second character "I" indicates that the variable is originally a civic knowledge test variable.

- The third character indicates the assessment cycle when the item was first used in ICCS. The item names in the ICCS 2022 assessment consist of "2" for items used already in ICCS 2009, "3" for items used already in ICCS 2016, or "4" for items newly developed for ICCS 2022.

- The fourth and fifth characters indicate the unique item identifier.

- The sixth character is used for the item type. "M" represents multiple-choice items, and "O" stands for open-ended response items.

- The seventh digit represents the number of an item within a unit comprising the same content.

For example, CI4DDM1 is the first part of a multiple-choice item developed for ICCS 2022, and its unique (content) identifier is DD.

In the scoring reliability files, the variable names for the original score, second score, and score agreement variables are based on the same naming convention as for the international test item variables shown above. Only the second character in the variable name is used differently in order to differentiate between the three reliability variables:

- The original score variable has the letter "I" as the second character, in accordance with the test item naming convention (e.g., CI4ZHC4).

- The second score variable has the letter "R" as the second character (e.g., CR4ZHC4) and represents the score assigned by the reliability coder in the Reliability file.

- The score agreement variable has the letter "X" as the second character (e.g., CX4ZHC4).

2.4.3   Civic knowledge test scores

In ICCS 2022 a civic knowledge scale was derived from the test data. The ICCS civic knowledge reporting scale was developed in 2009, and the Rasch model (Rasch, 1960) was used to accomplish this work. The scale has a mean (the average score of countries participating in ICCS 2009) of 500 and a standard deviation of 100 for equally weighted national samples. Chapter 10 of the ICCS 2022 technical report (Schulz et al., 2024b) provides a detailed description of the scaling procedures used in ICCS 2022 and the creation of the civic knowledge scale. The ICCS 2022 international database provides five separate estimates of each student's score on that scale. These are included in the student test data files (ISA) and the international student questionnaire data files (ISG). The five estimates of students' civic knowledge are so-called "plausible values," and the variation between them reflects the uncertainty inherent in the measurement process.

The plausible values for the civic knowledge scale are the available measures of students' civic knowledge in the ICCS 2022 international database and should be used as the outcome measure in any study of students' civic knowledge. Plausible values can be readily analyzed using the IEA IDB Analyzer.

The test score variable names are based on a six-character alphanumeric code, where PV1CIV represents the first plausible value and PV5CIV represents the fifth plausible value.

2.4.4   Summary scales and derived variables from the questionnaires

In the ICCS 2022 questionnaires, typically, sets of items reflecting a number of different aspects were used to measure a single construct. In these cases, responses to the individual items were combined to create a derived

variable that provided a more comprehensive picture of the construct of interest than relying on individual item responses.

In the ICCS 2022 reports, a scale is a derived variable for which score values were assigned to students based on their responses to a set of indicator variables. In ICCS 2022, new scales were typically calculated as IRT WLE (weighted likelihood estimates) scores with a mean of 50 and a standard deviation of 10 for equally weighted countries that had met IEA sample participation guidelines. Student questionnaire scales based on (unmodified) item sets already included in the ICCS 2016 or ICCS 2009 questionnaires were equated, and their scale scores are comparable with the scales established in the previous ICCS cycles; in such cases, the metric reflects a mean of 50 and a standard deviation in the pooled ICCS 2016 or ICCS 2009 sample with equal weights given to each participating country that had met IEA sample participation guidelines. For student, teacher, and school questionnaire scaling, we only included records in the scale calculation if there were data for at least two of the corresponding indicator variables.

In addition to the scale indices, the ICCS 2022 international database also contains other (simple) indices that were derived by simple recoding or arithmetical transformation of original questionnaire variables.

Appendix C of this user guide provides a description of all derived variables (scale scores and indices) included in the international database. Chapter 12 of the ICCS 2022 technical report (Schulz et al., 2024b) provides further information about the scaling procedure for questionnaire items.

2.4.5   Weighting and variance estimation variables

To enable the calculation of the population estimates and correct jackknife variance estimates, sampling and weight variables are provided in the data files. Further details about weighting and variance estimation are provided in Chapter 3 of this user guide.

The following weight variables are included in the ICCS 2022 international database (see Table 2.3 for the availability of weight variables across file types).

TOTWGTS

This is the final student weight. It is computed as the product of WGTFAC1, WGTADJ1S, WGTFAC2S, WGTADJ2S, and WGTADJ3S. The final student weight must be applied when analyzing the students' data.

WGTFAC1

This is the school base weight. It corresponds to the inverse of the selection probability of the school.

WGTADJ1S

This is the school weight adjustment for students. It accounts for non-participating schools. The adjustment is done within explicit strata.

WGTFAC2S

This is the class weight factor. It corresponds to the inverse of the selection probability of the class within the school.

WGTADJ2S

This is the class weight adjustment. It accounts for the non-participating classes. The adjustment is done across schools but inside the explicit stratum.

WGTADJ3S

This is the student weight adjustment. It accounts for the non-participating students. The adjustment is done within classes.

TOTWGTT

This is the final teacher weight. It is computed as the product of WGTFAC1, WGTADJ1T, WGTFAC2T, WGTADJ2T, and WGTADJ3T. The final teacher weight must be applied when analyzing the teacher's data.

WGTADJ1T

This is the school weight adjustment for teachers. It accounts for non-participating schools. The adjustment is done within explicit strata.

WGTFAC2T

This is the teacher weight factor. It corresponds to the inverse of the selection probability of the teacher within the school.

WGTADJ2T

This is the teacher weight adjustment. It accounts for the non-participating teachers. The adjustment is done within schools.

WGTADJ3T

This is the teacher multiplicity adjustment. It accounts for teachers teaching in more than one school.

TOTWGTC

This is the final school weight for schools. It is computed as the product of WGTFAC1 and WGTADJ1C. The final school weight must be applied when analyzing the data from the school questionnaire.

WGTADJ1C

This is the school weight adjustment for schools. It accounts for the non-returned school questionnaires.

Table 2.3: Availability of weight variables in the ICCS 2022 international database

Weight variables	Data file types					
	ISA	ISG	ITG	ICG	ISE	ISL
TOTWGTS	•	•			•	•
WGTFAC1	•	•	•	•	•	•
WGTADJ1S	•	•			•	•

WGTFAC2S	•	•			•	•
WGTADJ2S	•	•			•	•
WGTADJ3S	•	•			•	•
TOTWGTT			•			
WGTADJ1T			•			
WGTFAC2T			•			
WGTADJ2T			•			
WGTADJ3T			•			
TOTWGTC				•		
WGTADJ1C				•		

A variance estimation method that considers the structure of the data is the jackknife repeated replication (JRR) method. The ICCS 2022 international database contains variables that support the implementation of this method (i.e., "jackknife zone," "jackknife replicate," "replicate weights"); we strongly encourage database users to use them. As the IEA IDB Analyzer automatically recognizes the data structure of ICCS 2022, it reports correct standard errors for all estimates using JRR with the respective variables.

The following variance estimation variables (or "jackknife variables") are included in the ICCS 2022 international database (see Table 2.4 for the availability of the different variance estimation variables across file types). The actual replicate weights are computed "on-the-fly" within the IEA IDB Analyzer, but they are also available in the data files for use with other analysis tools.

JKZONES

This variable indicates which sampling zone the student belongs to. The values of JKZONES can vary between 1 and 75. This variable is used to estimate sampling errors when analyzing student data.

JKREPS

This variable can take the values 0 or 1. It indicates whether the student should be disregarded, or its weight doubled when estimating sampling errors.

SRWGT01 to SRWGT75

These variables indicate the jackknife replicate weights variables (1–75) for the student survey.

JKZONET

This variable indicates which sampling zone the teacher belongs to. The values of JKZONET can vary between 1 and 75. This variable is used to estimate sampling errors when analyzing teacher data.

JKREPT

This variable can take the values 0 or 1. It indicates whether the teacher should be disregarded, or its weight doubled when estimating sampling errors.

TRWGT01 to TRWGT75

These variables indicate the jackknife replicate weights variables (1–75) for the teacher survey.

JKZONEC

This variable indicates to which sampling zone the school belongs. The values of JKZONEC can vary between 1 and 75. This variable is used to estimate sampling errors when analyzing school data.

JKREPC

This variable can take the values 0 or 1. It indicates whether the school should be disregarded, or its weight doubled when estimating sampling errors.

CRWGT01 to CRWGT75

These variables indicate the jackknife replicate weights variables (1–75) for the school survey.

Table 2.4: Availability of variance estimation variables in the ICCS 2022 international database

Variance estimation variables	Data file types					
	ISA	ISG	ITG	ICG	ISE	ISL
JKZONES	•	•			•	•
JKREPS	•	•			•	•
SRWGT01 to SRWGT75	•	•			•	•
JKZONET			•			
JKREPT			•			
TRWGT01 to TRWGT75			•			
JKZONEC				•		
JKREPC				•		
CRWGT01 to CRWGT75				•		

Notes:
ISA = Student Civic Knowledge Test File, ISG = International Student Questionnaire File, ITG = Teacher Questionnaire File, ICG = School Questionnaire File, ISE = European Student Questionnaire File, and ISL = Latin American Student Questionnaire File.

2.4.6   Structure and design variables

Besides the variables used to store responses to the questionnaires and test booklets, the ICCS 2022 data files also contain variables meant to store information used to identify and describe the respondents and design information that is required to properly analyze the data.

Identification variables

All ICCS 2022 data files contain several identification variables that provide information to identify countries and entries of students, teachers, or schools (see Table 2.5 for the availability of the different identification variables across file types). These variables are used to link variables for one case, clusters of cases (students and teachers pertaining to specific schools), and cases across the different data file types. However, the variables do not allow the identification of individual schools, students, or teachers in a country.

IDCNTRY

This variable indicates the country or participating education system the data refers to as an up to six-digit numeric code based on the ISO 3166 classification, with adaptations reflecting the education systems participating. This variable should always be used as the first linking variable whenever files are linked within and across countries.

COUNTRY

This variable indicates the participant's three-letter alphanumeric code, based on the ISO 3166-1 coding, with adaptations reflecting the education systems participating.

IDSCHOOL

IDSCHOOL is a four-digit identification code that uniquely identifies the participating schools within each country. The school codes are not unique across countries. Schools across countries can only be uniquely identified with the combination of IDCNTRY and IDSCHOOL.

IDCLASS

IDCLASS is a six-digit identification code that uniquely identifies the sampled classrooms within a country. The variable IDCLASS has a hierarchical structure and is formed by concatenating the IDSCHOOL variable and a two-digit sequential number identifying the sampled classrooms within a school. Classrooms can be uniquely identified across countries using the combination of IDCNTRY and IDCLASS.

IDSTUD

IDSTUD is an eight-digit identification code that uniquely identifies each sampled student within a country. The variable IDSTUD also has a hierarchical structure and is formed by concatenating the IDCLASS variable and a two-digit sequential number identifying all students within each classroom. Students can be uniquely identified across countries using the combination of IDCNTRY and IDSTUD.

IDTEACH

IDTEACH is a six-digit identification code that uniquely identifies the sampled teacher within a country. The variable IDTEACH has a hierarchical structure and is formed by concatenating the IDSCHOOL variable and a two-digit sequential number identifying the sampled teacher within a school. Teachers can be uniquely identified across countries using the combination of IDCNTRY and IDTEACH.

IDPOP

IDPOP identifies the grade and is set to "2" for the ICCS 2022 target grade for all participating countries (representing eight years of schooling).

IDGRADE

IDGRADE identifies the tested grade of the participating students. In ICCS 2022, the value is "8" (= Grade 8) for most countries.

IDBOOK

IDBOOK identifies the specific test booklet that was administered to each student. The test booklets are given a numerical value depending on the scenario (as previously mentioned): for the paper-based test scenario (from 1 to 11), for the computer-based scenario (from 1 to 14), and for the bridging scenario (from 1 to 8).

For reasons of confidentiality, the identification variables for the student (IDSTUD), teacher (IDTEACH), and school (IDSCHOOL) were scrambled, and hence do not match the identifiers used during data collection. For each country, unique matching tables were created and made available to authorized individuals. Still, with the scrambled identification variables, the structural link between the school and student/teacher level (the variable IDSCHOOL in the student and teacher files and the first four digits of any IDSTUD/ IDTEACH) has been maintained for all countries.

Table 2.5: Availability of identification variables in the ICCS 2022 international database

Identification variables	Data file types						
	ISA	ISR	ISG	ITG	ICG	ISE	ISL
IDCNTRY	•	•	•	•	•	•	•
COUNTRY	•	•	•	•	•	•	•
IDSCHOOL	•	•	•	•	•	•	•
IDCLASS	•	•	•			•	•
IDSTUD	•	•	•			•	•
IDTEACH				•			
IDPOP	•	•	•	•	•	•	•
IDGRADE	•	•	•	•	•	•	•
IDBOOK	•	•	•				

 Notes:
ISA = Student Civic Knowledge Test File, ISR = Student Reliability File, ISG = International Student Questionnaire File, ITG = Teacher Questionnaire File, ICG = School Questionnaire File, ISE = European Student Questionnaire File, and ISL = Latin American Student Questionnaire File.

Tracking variables

Information about students, teachers, and schools provided by the survey tracking forms or used otherwise in the process of within-school sampling is stored in the tracking variables (see Table 2.6 for the availability of the tracking variables across file types).

ITADMINI

Position of the test administrator of the test session as an attribute for each student. Code "1" is used for national center staff, code "2" is used for teachers from school but not from selected class, and code "3" is used for test administrators that did not fall into the groups coded as "1" or "2".

ITDATE

This variable indicates the date when the test was administered to a student.

ITLANG

This variable indicates the language used for student test administration. The two-digit alphanumeric language codes are based on the ISO 639-1 standard.

ITMODE_C

Administration mode of the school questionnaire in the data source: this variable indicates whether the principal completed the questionnaire online (code "1") or on paper (code "2").

ITMODE_T

Administration mode of the teacher questionnaire in the data source: this variable indicates whether the teacher completed the questionnaire online (code "1") or on paper (code "2").

ITMODE_SQ/ ITMODE_R

Administration mode of the student questionnaire (the international and regional) in the data source: this variable indicates whether the student completed the questionnaire online (code "1") or on paper (code "2").

ITMODE_SA

Administration mode of the student test in the data source: this variable indicates whether the student completed the civic knowledge test computer-based (code "1") or on paper (code "2").

STREAM

Stream of the class/student. In some countries, classes and/or students belong to or are organized in certain streams of, for example, different skill levels. This variable was derived from information provided by countries on the survey administration (using IEA's WinW3S software) and was recoded. The value scheme consists of the country's operational code and the number of the national category (last two digits).

TCERTAN

This variable indicates whether a teacher was sampled with certainty.

ILRELIAB

This variable indicates whether the student was sampled as part of the reliability sample.

Table 2.6: Availability of tracking variables in the ICCS 2022 international database file types

Tracking variables	Data file types

	ISA	ISR	ISG	ITG	ICG	ISE	ISL
ITADMINI			•				
ITDATE			•				
ITLANG			•				
ITMODE_C					•		
ITMODE_T				•			
ITMODE_SQ/ ITMODE_R			•			•	•
ITMODE_SA	•						
STREAM			•				
TCERTAN				•			
ILRELIAB	•	•	•				

Notes:
ISA = Student Civic Knowledge Test File, ISR = Student Reliability File, ISG = International Student Questionnaire File, ITG = Teacher Questionnaire File, ICG = School Questionnaire File, ISE = European Student Questionnaire File, and ISL = Latin American Student Questionnaire File.

2.4.7   Database creation variables

Information about the version number of the ICCS 2022 international database and the scope of the database is contained in the database creation variables. They are included in all data files.

VERSION

Throughout the data processing, a system of version numbers for the database was used. The version number of the ICCS 2022 final database is "32".

SCOPE

This variable identifies the scope of the data files. Code "2" indicates the Restricted Use Files (RUF) and code "3" is the Public Use Files (PUF).

2.5   Coding of missing data

A subset of the values for each variable type was reserved for specific codes related to different categories of missing data. We recommend that the user reads the following section with particular care, since the way in which these missing codes are used may have major consequences for analyses.

Omitted response codes (SPSS and R: 9, 99, 999, ...; SAS: .)

"Omitted" response codes are used for questions or items that a student, teacher, or school principal should have answered but did not; an omitted response code is thus given when an item is left blank. The length of the omitted response code given to a variable in the SPSS and R data files depends on the number of characters needed to represent the variable. For example, the omitted code for a one-digit variable is "9" whereas the omitted code for a three-digit variable is "999".

Invalid response codes (SPSS and R: 7, 97, 997, …; SAS: .I)

The response to a question is coded as "invalid" when the question was administered but an invalid response was given. This code is used for uninterpretable responses, for example when the respondent has chosen more than one option in response to a multiple-choice question. The length of the invalid response code in the SPSS and R data files depends on the number of characters needed to represent the variable. For example, the invalid code for a one-digit variable is "7" whereas the invalid code for a three-digit variable is "997". Invalid codes are not applicable for open-ended items of the international test instruments.

Not administered response codes (SPSS and R: 8, 98, 998, …; SAS: .A)

Specific codes were given to items that were "not administered" to distinguish these from data that were missing due to non-response. The not administered code was used in the following cases:

- *Civic knowledge test item was not assigned to the student*. All students participating in ICCS 2022 received only one of the test booklets. All variables corresponding to items that were not part of the booklet assigned to a student were coded as "not administered".

- *Student was absent from the test session*. When a student did not attend a particular testing session, for example, because of sickness, all variables relevant to that session were coded as "not administered".

- *Question or item misprinted*. When a particular question or item (or a whole page) was misprinted or otherwise not available to the respondent, the corresponding variable was coded as "not administered".

- *Question or item deleted or mistranslated*. If a question or item was identified during translation verification or item review as having a translation error, such that the nature of the question was altered or had poor psychometric properties, it was coded as "not administered" if it could not be recoded to match the international version.

- *A questionnaire or booklet was returned empty, was not returned, or was lost*. In such cases, all variables referring to that instrument and any derived variables were coded as "not administered".

- *A country chose, for cultural reasons, not to administer (include) a certain question in its national questionnaire*. The variables corresponding to the removed question were coded as "not administered". All national adaptations are provided in Appendix B of this user guide.

The length of the invalid response code in the SPSS and R data files depends on the number of characters needed to represent the variable. For example, the not administered code for a one-digit variable is "8," whereas the not administered code for three-digit variables would be "998".

Not reached response codes (SPSS and R: 6; SAS: .R)

An item was considered "not reached" in the test data files when the item itself and the item preceding it were not answered, and there were no other items completed in the remainder of the booklet. For scaling purposes, ICCS 2022

treated the not-reached items as incorrect responses, however, during the item calibration step of the IRT scaling, not-reached items were considered as not administered.

Logically not applicable response codes (SPSS and R: 6, 96, 996, …., SAS: .B)

"Logically not applicable" response codes were used for the questionnaire items for which responses were dependent on a filter question. If the filter question was answered such that the following questions would not apply, any follow-up question was coded as "logically not applicable".

The length of the logically not applicable response code in the SPSS and R data files depends on the number of characters needed to represent the variable. For example, the logically not applicable code for a one-digit variable is "6", whereas the logically not applicable code for three-digit variables would be "996".

2.6 Codebook files

All information related to the structure of the ICCS 2022 data files, as well as the source, format, descriptive labels, and response option codes for all variables, is contained in a codebook file in Excel format.

In the codebook file, there is a tab for each appropriate data file type in the ICCS 2022 international database. These tabs describe the contents and structure of the respective data file. Important codebook fields include "Label", which contains extended textual information for all variables, "Value Scheme Detailed", which lists the acceptable responses allowed for each variable, and "Missing Scheme Detailed", which lists all applicable missing codes in SPSS, SAS and R.

2.7 Program files

The ICCS 2022 international database includes SPSS, SAS and R programs that can be used to convert the response codes to individual items from the civic knowledge test data files to their corresponding score levels.

These SPSS, SAS and R programs are part of the ICCS 2022 international database and are available in the IEA Study Data Repository at http://www.iea.nl/data.

2.8 Two versions of the ICCS 2022 international database

Indirect identification of individuals was prevented by applying international disclosure risk edits, such as scrambling of identification variables and jackknife zone information. Some of the personal data variables that were needed only during field operations and data processing were removed; variables that were identified as highly identifying were suppressed or categorized.

The ICCS 2022 international database is available in two versions: Public Use Files (PUF) and Restricted Use Files (RUF). The public-use version is available for immediate access from the IEA Study Data Repository (https://www.iea.nl/data). A number of variables have been removed or categorized from the public use version in order to minimize the risk of disclosing confidential information or enabling re-identification. Users should be able to replicate all published ICCS 2022 results with this version of the ICCS 2022 international database. The restricted use files are an extended version for scientific use. Users who require any of the removed variables to conduct their analyses

should contact the IEA to obtain permission and access to the restricted-use version of the ICCS 2022 international database (see the IEA Study Data Repository at http://www.iea.nl/data).

Tables 2.7 to 2.9 list the variables that either have been scrambled, categorized, or removed in the restricted and public use version of the ICCS 2022 international database.

Table 2.7: Disclosure risk edits for sampling, identification, and tracking variables

Variables	Description	Data files	RUF	PUF
IDSCHOOL/IDCLASS/ IDSTUD/IDTEACH	ID variables	All files	Scrambled	Scrambled
JKZONEC/T/S	Jackknife zone	ICG, ITG, ISG, ISA, ISE, ISL	Scrambled	Scrambled
IDSTRATE	Explicit stratum code	ICG, ITG, ISG, ISA, ISE, ISL	Included	Suppressed
IDSTRATI	Implicit stratum code	ICG, ITG, ISG, ISA, ISE, ISL	Included	Suppressed
ITMODE_C	Administration mode school	ICG	Included	Suppressed
ITMODE_T	Administration mode teacher	ITG	Included	Suppressed
ITMODE_SQ, ITMODE_R, ITMODE_SA	Administration Mode Student	ISG, ISA, ISE, ISL	Included	Suppressed
ITDEV	Administration Device / Student	ISA	Included	Suppressed
T_BIRTHY	Teachers' year of birth	ITG	Included	Suppressed
S_BIRTHY, S_BIRTHM	Students' year/month of birth	ISG	Included	Suppressed
ITDATE	Date of testing	ISG	Included	Suppressed

Notes:
RUF = Restricted Use Files, PUF = Public Use Files, ISA = Student Civic Knowledge Test File, ISR = Student Reliability File, ISG = International Student Questionnaire File, ITG = Teacher Questionnaire File, ICG = School Questionnaire File, ISE = European Student Questionnaire File, and ISL = Latin American Student Questionnaire File.

Table 2.8: Disclosure risk edits for school questionnaire variables

Variables	Description	Data files	RUF	PUF
IC4G20	Public or private school	ICG	Included	Suppressed

C_PRIVATE	Public or private school-derived	ICG	Included	Suppressed
IC4G21A/IC4G21B	Total school enrollment	ICG	Included	Suppressed
C_SCSIZE	Total school enrollment - derived	ICG	Included	Categorized
IC4G22A/IC4G22B	Total enrollment <target grade>	ICG	Included	Suppressed
C_GENROL	Total enrollment <target grade> – derived	ICG	Included	Categorized

Notes:
RUF = Restricted Use Files, PUF = Public Use Files, ICG = School Questionnaire File.

Table 2.9: Disclosure risk edits for student questionnaire

Variables	Description	Data files	RUF	PUF
IS4G01A, IS4G01B	Date of birth (month, year)	ISG	Suppressed	Suppressed
IS4G06A, IS4G06B	Female guardian's job (open-ended)	ISG	Suppressed	Suppressed
IS4G08A, IS4G08B	Male guardian's job (open-ended)	ISG	Suppressed	Suppressed

Notes:
RUF = Restricted Use Files, PUF = Public Use Files, ISG = International Student Questionnaire File.

References

Damiani, V., Agrusti, G., Losito, B., & Schulz, W. (2024). *Young citizens' views and engagement in a changing Europe. IEA International Civic and Citizenship Education Study 2022 European report.* International Association for the Evaluation of Educational Achievement (IEA).

IBM Corp. Released 2019. IBM SPSS Statistics for Windows, Version 26.0. [Computer Software]. Armonk, NY: IBM Corp

IEA. (2024). IDB Analyzer (Version 5.0.23) [Computer Software]. International Association for the Evaluation of Educational Achievement. Available online at http://www.iea.nl/data.

Rasch, G. (1960). *Probabilistic models for some intelligence and attainments tests*. Danish Institute for Educational Research.

Schulz, W., Friedman, T., Losito, B., & Fraillon, J. (Eds.) (2024b). *ICCS 2022 technical report.* International Association for the Evaluation of Educational Achievement (IEA).

3 WEIGHTS AND VARIANCE ESTIMATION

Diego Cortes and Umut Atasever

3.1 Overview

This chapter introduces the use of weight and variance estimation variables in the International Civic and Citizenship Education Study (ICCS) 2022. The names and locations of these variables in the ICCS 2022 international database are described and their specific roles in student-, teacher-, and school-level analysis are explained. Examples demonstrating the importance of using the appropriate weighting and variance estimation techniques are given.

3.2 Estimation weights

Why weights are needed

A central objective of ICCS 2022 is to reach conclusions about an entire education system participating in the study. For this, all data in the ICCS 2022 international database were derived from randomly drawn samples of schools, classes, and teachers. This means that to make correct inference about population parameters the use of weighted estimators is necessary. Estimation weights in the publicly available database encapsulate two sets of information: 1) the selection probabilities that vary across sampling units; and 2) the unit nonresponse model implemented in ICCS 2022. Details about the sampling design are reported in Chapter 6 of the ICCS 2022 technical report (Schulz et al., 2024b), while details about the nonresponse model are reported in Chapter 10.

The ICCS 2022 sampling design leads to selection probabilities that vary across schools. The inverse of these probabilities (hereafter, sampling weights) reflects and compensates the disproportional selection probabilities of schools, students, and teachers. That is, a sampling unit with a small selection probability has a relatively large weight, and vice versa. Moreover, given that some sampled schools, students, or teachers refused to participate in the study, it was necessary to adjust the sampling weights such that participating units represent those that are nonparticipating. The adjustment is justified and described by the nonresponse model implemented in the study. Estimation (or total) weights are the product of sampling weights and adjustment factors at all levels of analysis. Details about weighting and adjustments are reported in Chapter 10 of the ICCS 2022 technical report (Schulz et al., 2024b).

3.2.1 Weight variables in the ICCS 2022 international database

Each record in the ICCS 2022 international database contains data for one or more variables that concern weighting. The last character of the variable name indicates the data type (S = Student, T = Teacher, C = School). The estimation weights, weighting factors, and weighting adjustments differ depending on the type of data. Only the value of the school base weight (variable WGTFAC1) is identical across all three types of datasets, as it is the sampling weight originated from the first sampling stage. Tables 3.1 – 3.3 list all the weighting factors and weighting adjustments for each population participating in ICCS 2022.

Student weight variables

Table 3.1 shows the student weight variables that are part of the ICCS 2022 international database.

Table 3.1: Weight variables in student data files

Variable	Description	Source Files
TOTWGTS	Total student weight	ISA, ISE, ISG, ISL
WGTFAC1	School base weight	ISG
WGTADJ1S	School non-participation adjustment for the student survey	ISG
WGTFAC2S	Class base weight	ISG
WGTADJ2S	Class non-participation adjustment	ISG
WGTADJ3S	Student non-participation adjustment	ISG

Teacher weight variables

Table 3.2 shows the weight variables in the teacher data files in the ICCS 2022 international database.

Table 3.2: Weight variables in teacher data files

Variable	Description	Source Files
TOTWGTT	Total teacher weight	ITG
WGTFAC1	School base weight	ITG
WGTADJ1T	School non-participation adjustment for the teacher survey	ITG
WGTFAC2T	Teacher base weight	ITG
WGTADJ2T	Teacher non-participation adjustment	ITG
WGTADJ3T	Teacher multiplicity adjustment	ITG

School weight variables

Table 3.3 shows the weight variables in the school data files of the ICCS 2022 international database.

Table 3.3: Weight variables in school data files

Variable	Description	Source Files
TOTWGTC	Total teacher weight	ICG
WGTFAC1	School base weight	ICG
WGTFAC1C	School non-participation adjustment for school level data analyses	ICG

3.2.2 Selecting the appropriate weight variable

For analyzing the ICCS 2022 data, the appropriate weights must be selected. The decision about which weight to use depends on the type of data used for analysis, the level of analysis and the number of countries involved.

Single level analysis

For analyses concerning one data type only, different weights must be applied depending on the type of data:

For student-level analyses, TOTWGTS should be used.

For teacher-level analyses, TOTWGTT should be used.

For school-level analyses, TOTWGTC should be used.

When the IEA IDB Analyzer is used for data analysis, the software automatically selects these variables.

Please note that ICCS 2022 is conceptually a student and teacher survey and was not designed as a school survey. Although it is possible to undertake analyses at the school level that generate unbiased results, the sampling precision of the estimates tends to be lower (with large standard errors and confidence intervals) at this level than it is for analyses at the student or teacher level. Therefore, results concerning school-level data tend to be associated with a high degree of uncertainty.

Merging files from different levels

If researchers plan to analyze data from more than one level and plan to merge data of different data types, they must choose the correct weight carefully.

The variable TOTWGTS should be used for analyzing student data with added school data. This type of analysis of disaggregated data is straightforward with the IEA IDB Analyzer. The software merges school-level data to the student data and selects the correct weight automatically. This way, school information becomes an attribute of the student, and the user can analyze information from both files. A sample research question could be: "What is the percentage of students studying at schools with a female headmaster?"

Analyzing combined teacher data and school data should be performed in the same way; TOTWGTT is the variable of choice. As for student data, the IEA IDB Analyzer takes care of the correct selection. In this type of analysis, school information becomes an attribute of the teacher. An example research question could be: "What is the percentage of teachers working at schools with a female headmaster?"

If student or teacher information is regarded as an attribute of school information, this cannot be handled easily with the IEA IDB Analyzer. The researcher must use other software (e.g. SPSS, SAS, or R) to aggregate the student or teacher data and to merge the resulting information to the school file.

For aggregating student data within schools, within-school weights (which are the product of class and student-level weight factors and adjustments WGTFAC2S x WGTADJ2S x WGTADJ3S), should be used.

Within-school teacher weights (defined as the product of teacher-level weight factors and adjustments WGTFAC2T x WGTADJ2T x WGTADJ3T), should be used for aggregating teacher data within the school.

After aggregation, the student or teacher file can be merged with the school file (with IDSCHOOL as the key variable). When this step is completed, the data can be processed further with the IEA IDB Analyzer. TOTWGTC should be used for school-level data analysis. A sample question is: "What is the percentage of schools in which more than 50 percent of the tested students do not speak the language of the test at home?"

It is neither possible nor meaningful to combine files of student and teacher data directly. These two groups constitute separate target populations. A sampled student may never have been taught by a sampled teacher, and a sampled teacher may never have taught a sampled student. However, it is possible to aggregate teacher data at the school level and then treat the result as a contextual attribute of the student data. Similarly, it is possible to aggregate student data at the school level and then treat the result as an attribute of the teacher data. This process is explained above.

Multi-level analysis

Working with aggregated or disaggregated data poses some methodological problems (for details, see Snijders & Bosker, 1999). It is possible to perform multi-level analysis with specialized software packages (e.g. HLM, Mplus, or R). For this type of analysis, users must compute the appropriate weights themselves. While the procedure outlined here may be applicable in a general sense, it is important to note that specific packages or software may require distinct sets of weights. Users are advised to consult the respective manuals for these packages or software to ensure the accurate application of weights.

At level 1 (student level), the analyst should apply a "within-school student weight" as the product of the class and student level weight factors (WGTFAC2S x WGTADJ2S x WGTADJ3S). If the teachers constitute level 1, the analyst should apply a "within-school teacher weight" as the product of the teacher-level weight factors (WGTFAC2T x WGTADJ2T x WGTADJ3T).

At level 2 (school level), the user should calculate a "school weight." For student data analysis, this is the product of the variables WGTFAC1 and WGTADJ1S; for teacher-level analysis, this is the product of WGTFAC1 and WGTADJ1T. Users should ensure that the software used for multi-level analysis normalizes the weights, that is, makes the sum of weights equal to the sample size. Users should not use the variable TOTWGTC from the school files, as non-response adjustments made to school questionnaire data may make these values slightly different from the correct ones.

Finally, note that there is no unique way of using estimation weights in multi-level analysis, as this depends on how the estimation model is set up to examine the relationship of interest. Mang et al. (2021) provides an in-depth discussion about the use of estimation weights in multi-level regression models.

Analyses of groups of countries

Thus far, the discussion has focused on analysis of data from one country at a time. However, all the above statements also hold when more than one country is analyzed. Some caution must be exercised when international averages are calculated, however. If an international average is computed using TOTWGTS, TOTWGTT or TOTWGTC, larger countries will contribute more to this average than smaller countries, which may not be the intention of the researcher.

Instead of performing weighted analyses across groups of countries, users must conduct weighted analyses separately for each country and calculate an average of these results afterwards. This is true regardless of whether single-level data, aggregated or disaggregated data, or multi-level data files are used for analysis.

Users of the IEA IDB Analyzer do not need to worry about the issue of international averages, since the software performs the correct calculations automatically. For calculating an international mean, the IEA IDB Analyzer first calculates national means using the TOTWGT variables and then averages the results over the countries that contribute to the international mean.

3.2.3   Example for analyzing weighted data

Not using weights in data analysis can lead to biased results. The following example illustrates the importance of using weights when conducting research with ICCS 2022 data.

A researcher may be interested in the average civic knowledge in Romania (variables PV1CIV, PV2CIV, PV3CIV, PV4CIV, and PV5CIV in ISG file). Using unweighted data (e.g., in SPSS), the mean of each plausible value is calculated and an average of the five values is calculated. Figure 3.1 shows that this average score is 461.03.

Figure 3.1: Example of unweighted analysis in SPSS

Descriptive Statistics

	N	Mean
Civic knowledge - 1st PV	2768	461.03
Civic knowledge - 2nd PV	2768	460.72
Civic knowledge - 3rd PV	2768	461.24
Civic knowledge - 4th PV	2768	460.59
Civic knowledge - 5th PV	2768	461.58
Valid N (listwise)	2768	

· → average: 461.03

But correctly using weighted data with the IEA IDB Analyzer, as in Figure 3.2, shows that in Romania, the correct estimate for civic knowledge is actually 470.

Figure 3.2: Example of weighted analysis using the IEA IDB Analyzer

N of Cases	Sum of TOTWGTS	Sum of TOTWGTS (s.e.)	Percent	Percent (s.e.)	PVCIV (Mean)	PVCIV (s.e.)
2768	170349	19938.05	100.00	.00	470.12	9.11

N of Cases	Sum of TOTWGTS	Sum of TOTWGTS (s.e.)	Percent	Percent (s.e.)	PVCIV (Mean)	PVCIV (s.e.)
2768	170349	19938.05	100.00	.00	470.12	9.11

The difference between the unweighted and the weighted result can be explained by the ICCS 2022 sampling design for Romania. The proportion of students from rural schools in the ICCS 2022 school sample is higher than the proportion of students in rural schools in the population. The sample was selected this way intentionally to meet ICCS minimum sample size requirements per explicit stratum. To balance out the disproportionate sample allocation, students from (some) rural schools were assigned relatively smaller weights. Since on average students from rural schools perform lower than students from urban schools, omitting weights leads to an under-estimate of the students' performance in Romania. The sampling weights compensate for that disproportional school sample.

3.3 Variance estimation

Since all information in ICCS 2022 is based upon sample data, analysts should report the uncertainty associated to the estimation of population parameters. Due to the complex sampling design used in ICCS 2022, it is not possible to calculate standard errors or to perform significance tests with standard software packages. While most of these programs implicitly assume that the data is derived from a simple random sample, the ICCS 2022 student and teacher data come from a two-stage stratified cluster sample (each school being regarded as a "cluster" of students or teachers). Any method for estimating sampling variance should account for this.

The ICCS 2022 international database contains variables that allow for the use of a variance estimation method known as the paired-jackknife repeated replication (JRR), which approximates the sampling distribution of estimated (population) parameters. These variables are referred to as "jackknife zones" and as "jackknife replicates." The JRR method is implemented in the IEA IDB Analyzer software (for details about the JRR technique used in ICCS 2022, please refer to Chapter 12 of the ICCS 2022 technical report (Schulz et al.,2024b).

3.3.1 Variance estimation variables in the ICCS 2022 international database

Table 3.4 shows student-level variance estimation variables (or "jackknife variables") that are included in the ICCS 2022 international database.

Table 3.4: Student-level variance estimation variables

Variable	Description	Source Files
JKZONES	Jackknife zone to which students of a school are assigned	ISA, ISE, ISG, ISL
JKREPS	Jackknife replicate to which students of a school are assigned	ISA, ISE, ISG, ISL

| SRWGT1 to 75 | Student jackknife replicate weight 1 to 75 | ISA, ISE, ISG, ISL |

Table 3.5 shows the jackknife variables included for teachers.

Table 3.5: Teacher-level variance estimation variables

Variable	Description	Source Files
JKZONET	Jackknife zone to which teachers of a school are assigned	ITG
JKREPT	Jackknife replicate to which teachers of a school are assigned	ITG
TRWGT1 to 75	Teacher jackknife replicate weight 1 to 75	ITG

Table 3.6 shows the school-level jackknife variables that are included in the ICCS 2022 international database.

Table 3.6: School-level variance estimation variables

Variable	Description	Source Files
JKZONEC	Jackknife zone to which a school is assigned for school-level data analyses	ITG
JKREPC	Jackknife replicate to which a school is assigned for school-level data analyses	ITG
CRWGT1 to 75	School jackknife replicate weight 1 to 75	ITG

3.3.2   Selecting the appropriate variance estimation variables

Different variance estimation variables must be applied depending on the type of data:

For all student level analyses, JKZONES and JKREPS should be used.

For all teacher level analyses, JKZONET and JKREPT should be used.

For all school level analyses, JKZONEC and JKREPC should be used.

Even for the same school, the variables at different levels of analysis can differ from each other and thus are not interchangeable. Just as with weights, data users should ensure to choose the correct jackknife variables when working with aggregated datasets. The level of analysis (student, teacher, or school) determines which variable to choose.

When calculations are performed with the IEA IDB Analyzer, the correct variables will be selected automatically. However, researchers may choose to use specialized software for types of data analysis that go beyond the range of the IEA IDB Analyzer's capabilities. In this case, researchers must specify the jackknife variables according to the requirements of the software. Usually, "-zone" variables must be specified as "variance stratum" or "variance strata" variables, while the "-rep" variables are commonly referred to as "cluster" variables.

3.3.3   Example for variance estimation

Not using the jackknife variables in data analysis will lead to incorrect measures to quantify the uncertainty in the estimation induced by the sampling design (i.e., standard errors of estimated parameters). The following example illustrates the importance of using the JRR technique for research and analysis with ICCS 2022 data.

A researcher may be interested in the average teacher age (variable T_AGE) in Romania. Using SPSS, he finds that the (weighted) average teacher age is about 45 years, and the standard error seems to be close to 0.2 years (see Figure 3.3).[1]

Figure 3.3: Example of incorrect variance estimation in SPSS

Descriptive Statistics

	N[2]	Mean	
	Statistic	Statistic	Std. Error
Teacher's age	**2241**	**44.71**	**0.214**
Valid N (listwise)	**2241**		

But using the JRR technique with the IEA IDB Analyzer, we find that the correct estimate for the standard error is larger than the SPSS estimate (see Figure 3.4).

Figure 3.4: Example of correct variance estimation using the IEA IDB Analyzer

N of Cases	Sum of TOTWGTT	Sum of TOTWGTT (s.e.)	Percent	Percent (s.e.)	T_AGE (Mean)	T_AGE (s.e.)
2241	66931.54	2346.76	100.00	.00	44.71	.31

The standard methods of the SPSS base version can neither handle weights correctly for sampling variance estimation, nor can it take the clustered data structure into account. This means that not only standard errors, but also all analyses that contain significance tests will be incorrect unless specialized software is used.

---

[1] In Figure 3.3, the standard error of the mean age is calculated assuming a simple random sample $SE(\hat{\mu}_{age}) = \frac{\hat{\sigma}_{age}}{\sqrt{n}}$.

[2] N = Number of cases

References

Mang, J., Küchenhoff, H., Meinck, S., Prenzel, M. (2021). Sampling weights in multilevel modelling: an investigation using PISA sampling structures. *Large-scale Assessments in Education, 9*(1). https://doi.org/10.1186/s40536-021-00099-0

Schulz, W., Friedman, Losito, B., & Fraillon, J. (Eds.) (2024b). *ICCS 2022 technical report.* International Association for the Evaluation of Educational Achievement (IEA).

Snijders, T.A.B., & Bosker, R.J. (1999). *Multilevel Analysis.* Sage Publications.

4 ANALYZING THE ICCS 2022 DATA USING THE IEA IDB ANALYZER

Falk Brese

4.1 Overview

The current chapter describes the use of the IEA International Database Analyzer software (IEA, 2024) for analyzing the ICCS 2022 international data files. Used in conjunction with either SPSS (IBM Corporation, 2019), SAS (SAS Institute, 2016), or R (R Core Team, 2023) the IEA IDB Analyzer provides a user-friendly interface to easily merge the various data file types of the ICCS 2022 International Database and seamlessly takes into account the sampling information and the multiple imputed civic knowledge scores to produce accurate statistical results.

Example analyses will illustrate the capabilities of the IEA IDB Analyzer to compute a variety of statistics, including percentages of students in specified subgroups, average civic knowledge in those subgroups, correlations, regression coefficients, and percentages of students reaching certain proficiency levels. The examples use student, teacher, and school questionnaire data to replicate some of the ICCS 2022 results included in the ICCS 2022 international report (Schulz et al., 2024a), as well as other useful analyses for investigating policy-relevant research questions.

Users should be able to perform statistical analyses with the IEA IDB Analyzer with a basic knowledge of the ICCS 2022 International Database. Chapter 2 gives a more detailed description of the data files contained in the International Database, including their structure and contents, along with a description of all the supporting documentation provided with the International Database.

Developed by the IEA, the IEA IDB Analyzer is an interface for SPSS, SAS, and R, all are well-known statistical analysis software packages. The IEA IDB Analyzer enables users to combine data files, either in SPSS, SAS, or R, from IEA's large-scale assessments and conduct analyses using either SPSS, SAS, or R, without actually writing programming code. The IEA IDB Analyzer generates SPSS, SAS, or R syntax that takes into account information from the sampling design in the computation of statistics and their standard errors. In addition, the generated syntax makes appropriate use of plausible values for calculating estimates of civic knowledge scores and their standard errors, combining both sampling variance and imputation variance.

The IEA IDB Analyzer consists of two main action modules—the merge module and the analysis module—which are integrated and executed in one common application. The merge module is used to create analysis datasets by combining data files of different types and from different countries, and selecting subsets of variables for analysis. The analysis module provides procedures for computing various statistics and their standard errors. A third module converts SPSS files to R format for merging or analyzing in R.

R and RStudio can be downloaded and installed at no cost from r-project.org and posit.co, respectively (R Core Team, 2023; Posit Team, 2023). Version 5.0 of the IEA IDB Analyzer requires R version 4.2.0 or later. When running a script produced by the IDB Analyzer, RStudio will print in the console a list of necessary packages.

The latest version of the IEA IDB Analyzer (version 5.0) is available for download from the IEA website: http://www.iea.nl/data. Once installed, the IEA IDB Analyzer can be accessed by using the START menu in Windows:

Start → IEA → IEA IDB Analyzer

When the IEA IDB Analyzer application is launched, the main window will appear, as shown in Figure 4.1. Users first need to select either SPSS, SAS, or R as their statistical software of choice. For the examples in this chapter, we will use the SPSS software and hence the SPSS data files from the ICCS 2022 International Database.

The main window will direct users to **Convert Files from SPSS to R**, the **Merge Module**, the **Analysis Module**, the **Help** manual, some example **Sample Files**, or to simply **Exit** the application. The IEA IDB Analyzer has an extensive manual, accessible from everywhere in the application through the **Help** button, which users are encouraged to consult for full details on all the functionalities and features of the IEA IDB Analyzer.

Figure 4.1: IEA IDB Analyzer main window
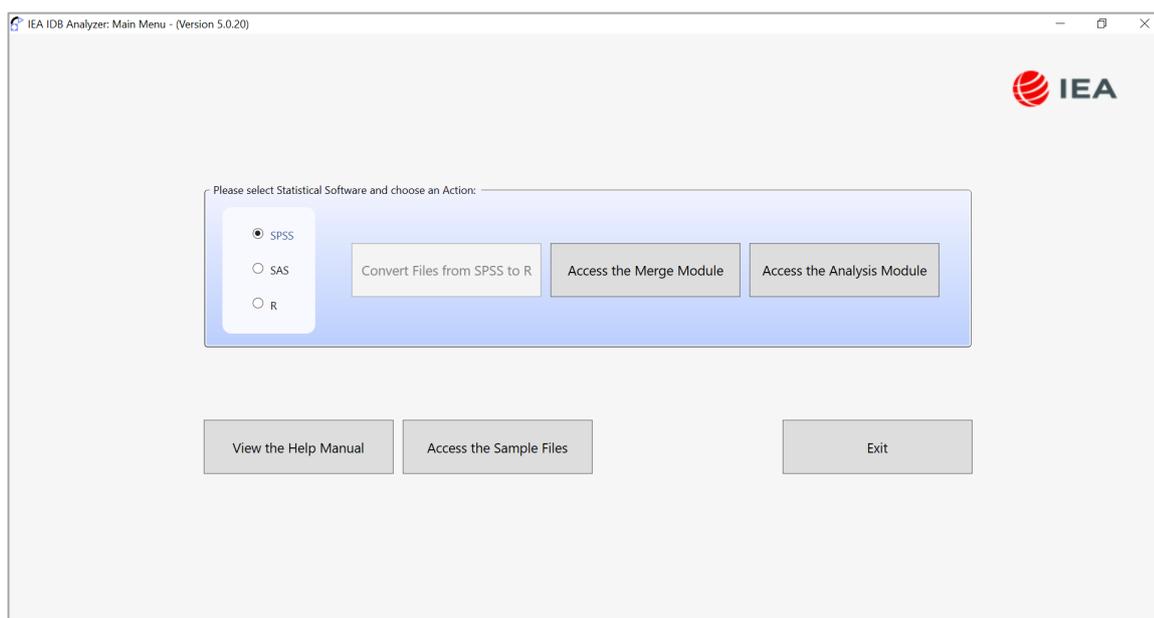
4.2   Scoring the individual civic knowledge items using SPSS, SAS, or R

The current section describes how the original answers from students to the civic knowledge assessment items can be scored. The original answers on multiple-choice items are located in the student test data files (ISA). The ICCS 2022 data already contains variables for each student's civic knowledge as a set of plausible values. Those are the preferred scores to be used for analysis. This section describes how to score the multiple-choice items in case item-level analysis is desired. Students' responses to the individual multiple-choice items need to be recoded into score points according to a scheme that specifies the correct option for each one of these items.

Two types of items were administered as part of the ICCS 2022 assessment. There were multiple-choice items, where students were asked to select one out of four options as the correct response. Numbers 1 through 4 are

used to represent response options A through D, respectively, in the student test data files (ISA). There also were constructed-response items, where students were asked to write a text response to a question, rather than choosing an answer from a list of options. Constructed-response items were worth a total of zero, one, two or three score points. Scorers from the national centers were trained to use the scoring guides to score the answers to these questions. The numbers 0 through 3 are used to represent the scored responses to these items and also represent their point values—"0" for incorrect response, "1" or "2" for partially correct response and "2" or "3" for correct response, depending on the item. For both types of items, special codes are set aside to represent missing data either as "Not Administered", "Omitted", "Not Reached", or "Invalid," Responses to multiple-choice items need to be converted to their appropriate score levels ("1" for correct and "0" for incorrect responses including missing responses except "not administered"), as well as responses coded to the special missing codes, in order to carry out specific item-level analyses. The user of the database can get an overview of the correct responses for this item type from the ISA codebooks and data files where the correct response option is marked with an asterisk "*". In case of constructed-response items, they have been already scored in advance by the national centers in each country.

The ICCS 2022 International Database includes a SPSS program (ISASCRC4.SPS), a SAS program (ISASCRC4.SAS), and an R program (ISASCRC4.R) which all allow you to recode the items from the civic knowledge data files to their score level. The program will convert the response option codes for multiple-choice items to dichotomous score levels (0 or 1) based on each item's scoring key. It will also convert the special missing codes as either incorrect (0) or missing. By default, the not administered response code is left as missing and the omitted and not reached response codes as incorrect. These default settings can be modified, depending on the requirements of your item-level analyses. For example, not reached responses were treated as missing for the purpose of calibrating the ICCS 2022 items, whereas they were treated as incorrect when scoring the results of individual countries and deriving civic knowledge scores for students.

4.2.1 SPSS scoring syntax

To use the SPSS program, you should do the following:

1. Open the SPSS data file containing the civic knowledge items data (e.g. ISABGRC4.sav);
2. Open the SPSS program file ISASCRC4.SPS;
3. Press CTRL+A followed by CTRL+R to submit the code for processing;
4. Save the SPSS data file.

The program recodes the items into new variables (using a new variable name with the original variable name the extension "_R", for example CI2ASM1_R) and saves the results in the same SPSS data file. If you would like not reached responses to be treated as missing rather than incorrect, you need to adapt the recode statement for each variable. For example, for variable CI2ASM1 replace the following statement:

```
recode CI2ASM1 (8=8) (4=1) (else=0) into CI2ASM1_R.
```

with this statement:

```
recode CI2ASM1 (6=6) (8=8) (4=1) (else=0) into CI2ASM1_R.
```

Adapt all recode statements accordingly. You can do that in one go for all recode statements by replacing "(8=8)" with "(6=6) (8=8)".

Figure 4.2 shows a condensed version of the SPSS program that scores the international civic knowledge items.

Figure 4.2: Example of ISASCRC4.sps SPSS program for converting item response codes to their score level

```
recode CI3ULM1 (8=8) (1=1) (else=0) into CI3ULM1_R.
recode CI3VGM1 (8=8) (3=1) (else=0) into CI3VGM1_R.
…
exe.

missing values CI2ASM1_R (6,8).
missing values CI3ULM1_R (6,8).
…
```

4.2.2 SAS scoring syntax

To use the SAS syntax, you will need to adapt the program code in the ISASCRC4.SAS program. You should do the following:

1. Open the SPSS program file ISASCRC4.SAS;
2. Specify the path to your data file in the "libname" statement;
3. Specify the file name of your data file in the "set" statement;
4. Submit the edited code for processing in SAS;
5. Save the resulting SAS data file.

The program recodes the items into new variables. If you would like not reached responses to be treated as missing rather than incorrect, you need to adapt add the following line directly below the "Ritems{i}=0;" statement:

if items{i}=".R" then Ritems{i}=.R;

Figure 4.3 shows a condensed version of the SAS program that scores the international civic knowledge items.

Figure 4.3: Example of ISASCRC4.sas SAS program for converting item response codes to their score level

```
libname test "D:\ICCS";

data ISAINTC4_SCR;
 set test.ISAINTC4;

format CI2ASM1_R CI2ASM2_R CI2BIO1_R CI2BPM1_R
…
array items{*} CI2ASM1 CI2ASM2 CI2BIO1 CI2BPM1
…
array Ritems{*} CI2ASM1_R CI2ASM2_R CI2BIO1_R CI2BPM1_R
…
do i=1 to dim(items);
 Ritems{i}=0;
 if items{i}=".A" then Ritems{i}=.A;|
end;

if CI2ASM1="4" then CI2ASM1_R = 1;
…
run;
```

In this example, the SAS data file of student achievement data files is called ISAINTC4, located in the folder D:\ICCS, and contains the civic knowledge test items. The resulting data file with the scored items will be called ISAINTC4_SCR and located in the same folder.

4.2.3 R scoring syntax

Figure 4.4 shows a condensed version of the ISASCRC4.R program to score the individual ICCS 2022 items. All R, SPSS, and SAS programs have a similar structure (see Figures 4.2 and 4.3). To score each individual ICCS 2022 item, the program code in the R score program must be adapted by completing the following steps:

1. Open ISASCRC4.R with R or RStudio (Posit Team, 2023)
2. At the end of the program, specify the folder where the merged R data file of student achievement data files is located in the argument indir
3. Specify the folder where the merged R data file of scored achievement items will be located in the argument outdir
4. Specify the name of the merged student achievement R data file in the argument infile
5. Submit the edited code for processing by R

In this example, the merged R data file of student achievement data files is called ISABGRC4, located in the folder D:\ICCS\, and contains the achievement item variables. The resulting data file of scored achievement items will be called ISABGRC4_SCR and saved in the location specified in the outdir argument. In this case, the same folder was used for both indir and outdir. Note that in R, (forward) slashes (/) should be used instead of backslashes in the file path, so the folder path in Figure 4.4 is given as: D:/ICCS/. If using SPSS or SAS, backslashes should be used.

Figure 4.4: Example of ISASCRC4.R R program for converting item response codes to their score level

```
scoreit <- function(data, item, type, right, NR, OM, other){

. . .

}
Doit <- function(indir=getwd(), outdir=getwd(), infile=""){

. . .

        # Score multiple-choice items with A key
        Aright <- c(<List of multiple-choice items where A is correct>)
        Data <- scoreit(data, item=Aright, type="MC", right=1, NR=6, OM=9, other=7)

        # Score multiple-choice items with B key
        Bright <- c(<List of multiple-choice items where B is correct>)
        Data <- scoreit(data, item=Bright, type="MC", right=2, NR=6, OM=9, other=7)

        # Score multiple-choice items with C key
        Cright <- c(<List of multiple-choice items where C is correct>)
        Data <- scoreit(data, item=Cright, type="MC", right=3, NR=6, OM=9, other=7)

        # Score multiple-choice items with D key
        Dright <- c(<List of multiple-choice items where D is correct>)
        Data <- scoreit(data, item=Dright, type="MC", right=4, NR=6, OM=9, other=7)

        # Score constructed-response items
        Constr <- c(<List of constructed-response items>)
        Data <- scoreit(data, item=constr, type="CR", right="", NR=6, OM=9, other=7)

. . .

}
doit(indir = "D:/ICCS/",
        outdir = "D:/ICCS/",
        infile = "ISABGRC4")
```

The data file of scored achievement items will have the same data structure as the student achievement data files (ISA), but with the score levels stored in the item variables instead of the student responses. The analysis module of the IEA IDB Analyzer will be able to conduct analyses with this data file.

By default, the score programs treat **Omitted** and **Not Reached** responses as missing. If **Omitted** and/or **Not Reached** responses are to be treated as incorrect rather than missing, users should replace the following lines of R code (which appear twice in the programs, once for multiple-choice items and once for constructed-response items):

```
== NR ~ NaN, # Not Reached
== OM ~ NaN, # Omitted
```

with these statements:

```
== NR ~ 0, # Not Reached
== OM ~ 0, # Omitted
```

4.3   Merging files with the IEA IDB Analyzer

The ICCS 2022 data files are disseminated separately for each country and by file type. In addition to allowing users to combine data from the same file type from more than one country for cross-country analyses, the merge module allows for the combination of data from different levels, e.g., merging student and school data into single SPSS dataset. This will allow analyzing the student data in relation with certain characteristics of the school they attend using the IEA IDB Analyzer Analysis Module later.

Table 4.1 provides an overview of possible combinations of data file types that the ICCS 2022 design allows to be merged at different levels. The grey-shaded cells on the diagonal represent the merges for the same file type. As the table shows,

- The school questionnaire file can be merged with every other file type.

- Teacher questionnaire files can be merged only with themselves (i.e., teacher questionnaire files from different countries) and with school questionnaire files, but merging teacher questionnaire files with any of the student files (student questionnaire, student test, and regional questionnaire data files) is not possible. The reason for this lies in the study's sample design—the ICCS 2022 teacher target population includes all teachers from the students' target grade. These are usually not only the teachers who teach the sampled students, so that teacher data cannot be directly linked to student data.[3]

---

[3] For more details on the ICCS 2022 sampling strategy and procedures see Chapter 5 of the ICCS 2022 technical report (Schulz et al., (2024b)).

- Also, the user will not be able to merge data from different regional questionnaires (European and Latin American) because students only answered the one questionnaire designed for the region they live (e.g., European students only completed the European questionnaire).

When merging a regional questionnaire file with some other file type, the IEA IDB Analyzer will always display a warning for files not found for specific countries. In general, it will list all countries that did not administer the specific regional questionnaire—European countries did not administer the Latin American regional questionnaire for example. Countries will also be listed if they belong to a region for which a regional questionnaire had been developed, but the country did not administer it. For example, there are European countries that decided not to use the European questionnaire. Such cases will be listed by the IEA IDB Analyzer in a warning message.

Table 4.1: Possible merging between different file types in ICCS 2022

	ISA	ISG	ITG	ICG	ISE	ISL
ISA	X	X		X	X	X
ISG	X	X		X	X	X
ITG			X	X		
ICG	X	X	X	X	X	X
ISE	X	X		X	X	
ISL	X	X		X		X

Notes:
ISA = Student Civic Knowledge Test File, ISR = Student Reliability File, ISG = International Student Questionnaire File, ITG = Teacher Questionnaire File, ICG = School Questionnaire File, ISE = European Student Questionnaire File, and ISL = Latin American Student Questionnaire File.

4.3.1   Merging data from different countries

Merging the files from different countries on a single level is simple. The same steps apply for merging school questionnaire, teacher questionnaire or any other file types. The following example will create an SPSS, SAS, or R data file[4] with student questionnaire data from all countries participating in ICCS 2022, not including the benchmarking participants North Rhine-Westphalia (Germany) and Schleswig-Holstein (Germany):

1. Start the IEA IDB Analyzer from the START menu and click the **Merge Module** button.

2. In the *Select Data Files and Participants* tab, click the **Select** button in the *Select Directory* section and browse to the folder where all ICCS 2022 data files are located. For example, in Figure 4.4, all SPSS data files are located in the folder titled "C:\ICCS2022\Data" The program will automatically recognize and

---

[4] Depending on the software selected in the main menu of the IEA IDB Analyzer.

complete the ***Select Study***, ***Select Year***, and ***Select Grade*** fields and list all countries available in this folder as possible candidates for merging. If the folder contains data from more than one IEA study, or from more than one grade, the IEA IDB Analyzer will prompt users to select files from the desired combination of study and grade for analyses. In Figure 4.4, ICCS 2022 eighth grade is selected.

3. Click a country of interest from the ***Available Participants*** list and click the **right arrow** button (▶) to move it to the ***Selected Participants*** panel. Individual countries can be moved directly to the ***Selected Participants*** panel by double-clicking on them. To select multiple countries, hold the **CTRL** key of the keyboard when clicking countries. Click the **tab-right arrow** button (▶I) to move all countries to the ***Selected Participants*** panel. In Figure 4.4, all available ICCS 2022 countries are selected.

Figure 4.5: IEA IDB Analyzer Merge Module: Selecting countries
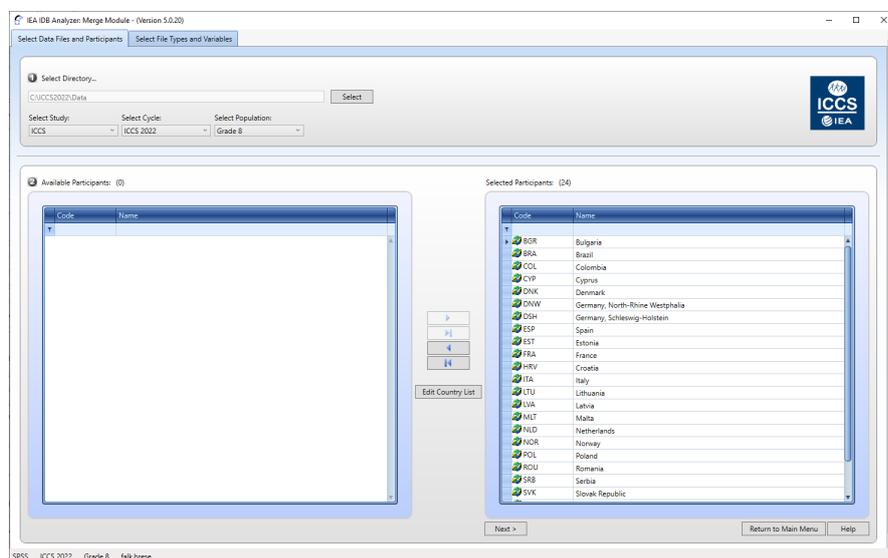
4. Click the **Next** > button to proceed to the next step. The software will open the ***Select File Types and Variables*** tab of the merge module, as shown in Figure 4.5, to select the file types and the variables to be included in the merged data file.

5. Select the file types for merging by **checking** the appropriate **boxes** to the left of the window. In the current example only the international student questionnaire file is selected (see Figure 4.5).

6. Select the variables of interest from the ***Available Variables*** list in the left panel. Variables are selected by clicking on them and then clicking the **right arrow** (▶) button. Clicking the **tab-right arrow** (▶I) button selects all variables. Note that there are two tabs: ***Background Variables and Scores*** and ***ID and Sampling Variables***. All civic knowledge scores and all identification and sampling variables are selected automatically by the IEA IDB Analyzer. Appendix A of this user guide provides the variable names for all questions in the questionnaires. In our example, we will select all student variables for merging.

7. Specify the desired name for the merged data file and the folder where it will be stored in the ***Output Files*** field by clicking the **Define/Modify** button. The IEA IDB Analyzer also will create a SPSS, SAS, or R syntax file (*.SPS, *.SAS, or *.R) of the same name and in the same folder with the SPSS, SAS, or R code used to perform the merge. In the example shown in Figure 4.5, the SPSS data file ISGALLC4.SAV as well as the syntax file ISGALLC4.SPS is stored in the "C:\ICCS2022\Work" folder. When using SAS, the resulting files would be ISGALLC4.sas7bdat (data) and ISGALLC4.sas (syntax). When using R, the resulting files would be ISGALLC4.RData (data) and ISGALLC4.R (syntax). The merged data file will contain student questionnaire data with the variables shown in the ***Selected Variables*** panel on the right.

Figure 4.6: IEA IDB Analyzer Merge Module: Selecting file types and variables
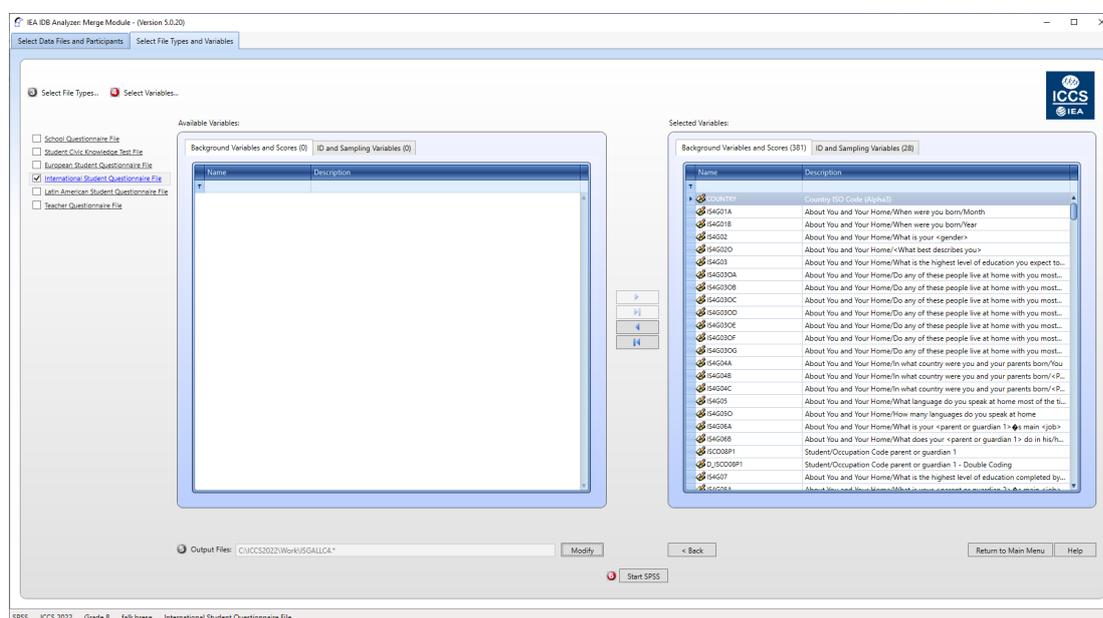
8. Click the **Start SPSS** button to create the SPSS syntax file with the specified name. The IEA IDB Analyzer stores the SPSS syntax file in the specified folder and opens it in a SPSS Syntax Editor window (Figure 4.6) ready for execution. The syntax file can be executed by opening the **Run** menu of SPSS and selecting the **All** menu option. In SAS, hitting the **Submit** option from the **Run** menu will execute the syntax. The IEA IDB Analyzer will display a warning if it is about to overwrite an existing file in the specified folder.

Figure 4.7: SPSS Syntax editor with merge syntax produced by the IEA IDB Analyzer Merge Module
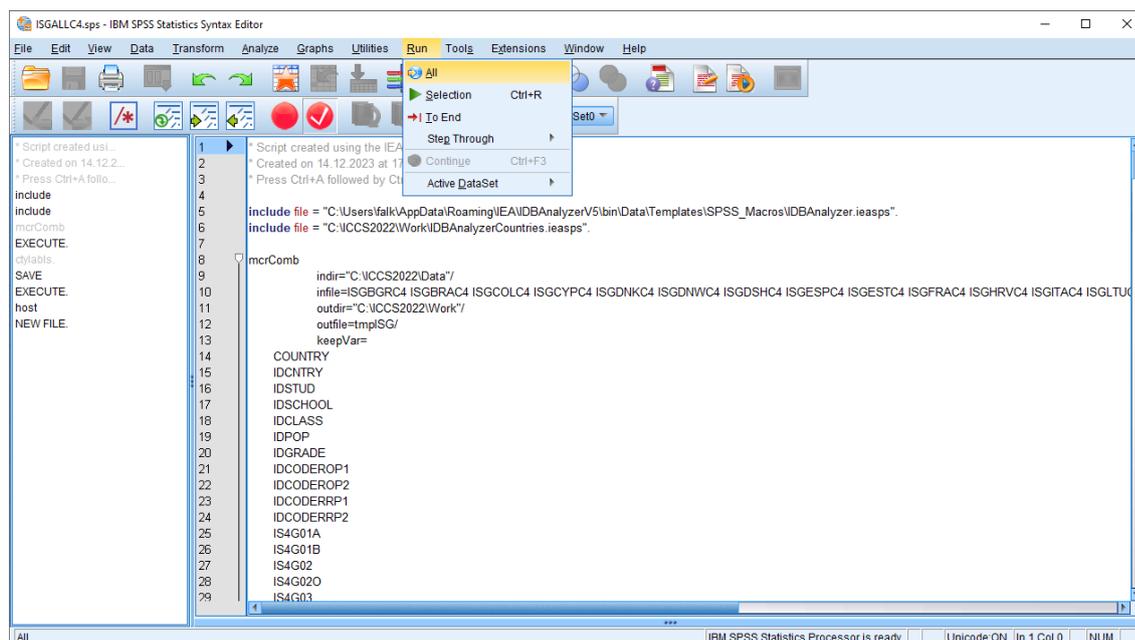
Once SPSS, SAS, or R has completed its execution, it is important to check the SPSS output window or the SAS or R log for possible warnings. If warnings appear, they should be examined carefully because they might indicate that the merge process was not performed properly and that the resulting merged data file might not be as expected.

4.3.2 Merging student questionnaire and regional questionnaire files

Student questionnaire files contain contextual variables related with students' background characteristics, their perceptions and behaviors. The regional questionnaire data files contain variables addressing specific regional issues and aspects in civic and citizenship education. As the use of the regional questionnaire instrument was optional, not all countries participated in that option. For example, some European countries decided not to use the European questionnaire.

Merging the student questionnaire data files with the regional data files can give researchers the chance to enrich the student-level analyses with variables that are specific for a certain region of the world.

To merge student questionnaire data with regional data, perform steps 1 to 4 as described in section 4.3.1. Then, simply select both file types in the second window of the IEA IDB Analyzer Merge Module. The variables of interest need to be selected separately for both file types, as follows:

- **Check the checkbox** next to the international **student questionnaire file type**. The *Background Variables and Scores* listed in the left panel will include all available variables from the student questionnaire data files. The plausible values, ID, and sampling variables are selected automatically and already listed in the right panel.

- Select the variables of interest from the left panel and click the **right arrow** (▶) button to move these variables to the *Selected Variables* panel on the right. Click the **tab-right arrow** (▶I) button to select all available variables.

- Next, **check the checkbox** next to the Latin American/European Student File (which one will depend on your interest). Based on your country selection, the IEA IDB Analyzer might display a warning that certain countries do not have data for the selected regional data. Close the warning message and select the variables of interest from the *Background Variables and Scores* panel in the same manner as described in step 1 and 2.

- **Define/Modify** the desired name of the merged data file and the folder where it will be stored in the *Output Files* field. The IEA IDB Analyzer will create an SPSS syntax file (*.SPS) of the same name and in the same folder with the code necessary to perform the merge.

- Click on the **Start SPSS** button to create the SPSS syntax file that will produce the required merged data file, which can then be run by opening the **Run** menu of SPSS and selecting the **All** option.

Note that when merging variables from more than one file type, the checkboxes of the desired file types need to remain to be checked throughout the selection of variables. Unchecking a file type will result a merged file without any variables from that file type—even if variables from that file type have been selected for merging in a previous step.

4.3.3   Merging school and student data files

The ICCS 2022 school samples were designed to optimize the student samples and the student-level estimates. Therefore, it is preferable to analyze school variables as attributes of students, rather than as elements in their own right. However, the school samples are representative probability samples of schools within each participating country and so it is possible to compute weighted numbers of schools with particular characteristics for providing reasonable estimates of percentages and means across the populations of schools in each country.

When merging school and student questionnaire data files, the school data will be disaggregated to student level, so that the school data become an attribute of the students in that school.

To merge the school and student questionnaire data files, select both the **School Questionnaire file** and **Student Questionnaire file** types. The variables of interest to be included in the merged data file need to be selected separately by file type using the same set of instructions as described in section 4.3.2. The ID and sampling variables will be selected automatically. Please note that when merging student and school data, only the Total Student Weight (TOTWGTS) variable will be included in the merged file, but not the Total School Weight (TOTWGTC). An analysis using school variables on student level together with student-level variables will use the Total Student Weight and will not allow the researcher to make inferences for the school themselves. The interpretation of the results will be about students who study in schools with certain characteristics. For example, if we used merged student and school data and used the type of school (public vs. private) as grouping variable, the Total Student Weight would be selected as weighting variable. The results then would be interpreted as

percentages of students who study in either public or private schools, e.g. "In Spain, 69% of the 8th grade students study in public schools, and 31% in private schools."

4.3.4 Merging school and teacher data files

Merging the school and teacher data files follows the same procedure as merging the school and student data files. School data will be disaggregated to the teacher level by adding the respective school level variables to each teacher record. To merge teacher questionnaire and school questionnaire data files, perform steps 1 to 4 as described in section 4.3.1. Then, simply select both file types in the second window of the IEA IDB Analyzer Merge Module. The variables of interest need to be selected separately for both file types, as follows:

- **Check the checkbox** next to the **Teacher Questionnaire File** type so that it appears checked. The ID and sampling variables are selected automatically and already listed in the right panel.

- Select the variables of interest from the left panel and click the **right arrow (▶)** button to move these variables to the *Selected Variables* panel on the right. Click the **tab-right arrow (▶|)** button to select all available variables.

- Next**, check the checkbox** next to **the School Questionnaire File** type. Based on your country selection, the IEA IDB Analyzer might display a warning that certain countries do not have data for the selected regional questionnaire. Close the warning message and select the variables of interest from the *Background Variables and Scores* panel in the same manner as described in step 1 and 2.

- **Define/Modify** the desired name of the merged data file and the folder where it will be stored in the *Output Files* field. The IEA IDB Analyzer will create an SPSS syntax file (*.SPS) of the same name and in the same folder with the code necessary to perform the merge.

- Click on the **Start SPSS** button to create the SPSS syntax file that will produce the required merged data file, which can then be run by opening the **Run** menu of SPSS and selecting the **All** option.

4.3.5   Merging data files for the example analyses

To carry out the analysis examples described in this chapter, the following merged data files including all available context and background variables and scores should be created:

- ISGALLC4.SAV   Merge the student questionnaire (ISG) data files for all countries

- ITGALLC4.SAV   Merge the teacher questionnaire (ITG) data files for all countries

- ICGISGALLC4.SAV   Merge the school questionnaire (ICG) and student questionnaire (ISG) data files for all countries

4.4   Performing analyses with the IEA IDB Analyzer

The IEA IDB Analyzer can perform statistical analyses on any files created using the **Merge Module** or on any original IEA study data file. Table 4.2 shows the statistical procedures available in the **Analysis Module** of the IEA IDB Analyzer.

Table 4.2: Statistical procedures available in the Analysis Module of the IEA IDB Analyzer

Statistic Type	Description
Percentages and Means	Computes percentages, means, and standard deviations for selected analysis variables by subgroups defined by grouping variable(s). Plausible values can be included as analysis variables.
Percentages only	Computes percentages by subgroups defined by grouping variable(s).
Linear Regression	Computes linear regression coefficients for selected independent variables to predict a dependent variable by subgroups defined by grouping variable(s). Plausible values can be included as dependent or independent variables.
Logistic Regression*	Computes logistic regression coefficients for selected independent variables to predict a dependent variable by subgroups defined by grouping variable(s). Plausible values can be included as dependent or independent variables. When used as a dependent variable, plausible values will be dichotomized using a specified cut-point, such as one of the ICCS international benchmarks.
Benchmarks	Computes percentages of students meeting a set of user-specified civic knowledge proficiency levels by subgroups defined by grouping variable(s).
Correlations	Computes means, standard deviations, and correlation coefficients for selected variables by subgroups defined by grouping variable(s). Plausible values can be included as analysis variables.
Percentiles	Computes the score points that separate a given proportion of the distribution of scores, by subgroups defined by the grouping variable(s). Plausible values can be included as analysis variables.
Group Differences by Performance*	Computes differences in an analysis variable, between groups, conditional on the civic knowledge or performance levels.

Notes:
The statistic types Logistic Regression and Group Differences by Performance will not be covered in this chapter. Refer to the IEA IDB Analyzer Help manual for details about these types of analysis.

All available features of the IEA IDB Analyzer are described extensively in its Help manual. All statistical procedures offered in the analysis module of the IEA IDB Analyzer make appropriate use of sampling weights, and standard errors are computed using the jackknife repeated replication (JRR) method (Schulz et al., 2024b). Percentages, means, linear regressions, correlations, and percentiles may be specified with or without civic knowledge scores. When civic knowledge scores are used the analyses are performed five times (once for each plausible value) and the results are aggregated to produce accurate estimates of civic knowledge and standard errors that incorporate both sampling and imputation errors. To conduct analyses using civic knowledge scores, select the **Use PVs** option

from the **Plausible Value Option** drop-down menu. The various variables required to perform an analysis need to be selected into specific variable fields according to their purpose, see Table 4.3 below.

Table 4.3: Fields for variable selection in the Analysis Module of the IEA IDB Analyzer

Field	Description
Grouping Variables	This is a list of variables to define subgroups of interest. The list must consist of at least one grouping variable. By default, the IEA IDB Analyzer includes the variable IDCNTRY used to distinguish the participating countries. Additional variables may be selected from the available list. If the Exclude Missing from Analysis option is checked, only cases that have non-missing values in the grouping variables will be used in the analysis. If it is not checked, missing values become reporting categories.
Analysis Variables	This is a list of variables for which means, percentages, correlations, or percentiles are to be computed. Usually, more than one analysis variable can be selected. To compute statistics based on civic knowledge scores, it is necessary to select the Use PVs option in the Plausible Value Option drop-down menu and select the civic knowledge scores of interest in the Plausible Values field.
Plausible Values	This section is used to identify the set of plausible values to be used when civic knowledge scores are the analysis variable for computing statistics. Select the Use PVs option in the Plausible Value Option drop-down menu before specifying the civic knowledge scores in the Plausible Values field.
Independent Variables	This is a list of variables to be treated as independent variables for a linear or logistic regression analysis. More than one independent variable can be selected. Categorical variables and continuous variables can be specified as independent variables. When specifying categorical variables as independent variables, they can be treated either by "effect coding" or "dummy coding" using the Contrast drop-down menu. Civic knowledge scores also can be included as an independent variable. To specify civic knowledge scores as an independent variable, it is necessary to select the Use PVs option in the Plausible Value Option drop-down menu and select the civic knowledge scores of interest in the Plausible Values field.
Dependent Variable	This is the variable to be used as the dependent variable when a linear or logistic regression analysis is specified. Only one dependent variable can be listed and can be either a context or background variable or civic knowledge scores. To use civic knowledge scores as the dependent variable, select the Use PVs option in the Plausible Value Option drop-down menu, click on the Plausible Values radio button in the Dependent Variable section, and select the civic knowledge scores of interest in the Plausible Values field.
Achievement Benchmarks	These are the values that will be used as cut points on the civic knowledge scale, selected in the Plausible Values section, for computing the percentages of students meeting the specified benchmarks. Multiple cut points can be specified, each separated by a blank space. A drop-down menu is available to select the ICCS international benchmarks.

Percentiles	These are the percentiles that will be calculated from the distribution of a continuous analysis variable selected in the Analysis Variables section. Civic knowledge scores can also be selected as an analysis variable. Select the Use PVs option in the Plausible Value Option drop-down menu and select the civic knowledge scores of interest in the Plausible Values field. Multiple percentiles can be specified, each separated by a blank space.
Weight Variable	This is the sampling weight variable that will be used in the analysis. The IEA IDB Analyzer automatically selects the appropriate weight variable for analysis based on the file types included in the merged data file. Generally, this will be TOTWGTS for analysis on student level. When analyzing teacher data TOTWGTT will be used. Chapter 3 of this user guide provides more information on the ICCS 2022 sampling weights.

4.5   Performing analyses with student-level variables

Many analyses of the ICCS 2022 data can be undertaken using student-level data only. This section presents examples of analyses used to produce tables for the ICCS 2022 international report—examples of percentages only, percentages and means, linear regression analyses and computing percentages of students reaching proficiency levels are included in this section. Additionally, conducting correlation analysis will be described.

4.5.1   Student-level analysis without civic knowledge scores

In our first example, we will replicate an analysis of students' reported age at the time of testing. The results, presented in Table 3.9 of the ICCS 2022 international report, are reproduced here in Figure 4.7. This example will focus on the results presented in the third column—the average age at the time of testing. We want to report the average ages (with their appropriate standard errors) and therefore we will be computing means without civic knowledge scores.

Figure 4.8: Table of example student-level analysis without civic knowledge scores taken from the ICCS 2022 international report (Table 3.9)

Table 3.9: Distributions of civic knowledge

Country	Grade	Average age	Civic knowledge scale scores						Average scale score			HDI
			250	350	450	550	650	750				
Chinese Taipei	8	14.2							583	(2.3)	▲	0.93
Sweden¹	8	14.8							565	(3.5)	▲	0.95
Poland	8	14.4							554	(2.5)	▲	0.88
Estonia	8	15.0							545	(5.5)	▲	0.89
Croatia¹	8	14.7							531	(2.6)	▲	0.86
Norway (9)¹	9	14.9							529	(2.8)	▲	0.96
Italy	8	13.8							523	(3.6)	▲	0.90
Spain	8	14.0							510	(3.3)		0.91
Lithuania	8	14.8							509	(4.0)		0.88
Netherlands†	8	14.1							508	(4.1)		0.94
France	8	13.9							508	(3.3)		0.90
Slovenia	8	13.9							504	(2.3)		0.92
Slovak Republic	8	14.3							501	(3.3)	▼	0.85
Latvia¹	8	14.8							490	(2.8)	▼	0.86
Malta	8	13.6							490	(7.4)	▼	0.92
Romania	8	15.0							470	(9.1)	▼	0.82
Serbia	8	14.6							464	(3.4)	▼	0.80
Cyprus	8	13.9							459	(2.5)	▼	0.90
Bulgaria	8	14.8							456	(4.6)	▼	0.80
Colombia	8	14.1							452	(3.8)	▼	0.75
ICCS 2022 average		**14.4**	Below D	D	C	B	A		**508**	**(0.9)**		
				Proficiency Level								
Countries not meeting sample participation requirements												
Brazil	8	14.1							457	(3.3)		0.75
Denmark	8	14.8							556	(3.5)		0.95
German benchmarking participant meeting sample participation requirements												
North Rhine-Westphalia	8	14.3							524	(2.6)	▲	0.94
German benchmarking participant not meeting sample participation requirements												
Schleswig-Holstein	8	14.5							544	(4.4)		0.92

Notes:
() Standard errors appear in parentheses.
(9) Country deviated from international defined population and surveyed adjacent upper grade.
† Nearly met guidelines for sampling participation rates only after replacement schools were included.
¹ National defined population covers 90% to 95% of national target population.
² Country surveyed target grade in the first half of the school year.
HDI = Human Development Index.

Percentiles of performance
5th 25th   75th 95th
Mean and confidence
interval (±2SE)

▲ Achievement significantly higher than international average
▼ Achievement significantly lower than international average

We need to undertake a number of steps to replicate the results in this table. After reviewing the student questionnaire data codebook, we identify the student questionnaire data variable S_AGE as the numeric variable reporting the age of students at the time of testing.

After creating the merged data file for the analysis, we use the analysis module of the IEA IDB Analyzer to perform our analysis in the following steps:

1. Open the **Analysis Module** of the IEA IDB Analyzer.

2. Select the merged data file ISGALLC4.SAV as the **Analysis File** by clicking the **Select** button.

3. Select **ICSS (Using Student Weights)** as the Analysis Type.

4. Select **Percentages and Means** as the Statistic Type.

5. Select **None Used** as the Plausible Value Option.

6. Leave **2** as the Number of Decimals and **Yes** for the Show Graphs Option as some additional settings of the output.

7. The variable IDCNTRY is selected automatically as **Grouping Variables**. No additional grouping variables are needed for this analysis. Leave the default option **Exclude Missing From Analysis** selected.

8. Click the **Analysis Variables** field to activate it. Select S_AGE from the list of available variables and move it to the **Analysis Variables** field by clicking the **right arrow (▶)** button in this section.

9. The **Weight Variable** is automatically defined by the software. As this example analysis uses student questionnaire data, TOTWGTS is selected by default.

10. Specify the name and the folder of the output files in the **Output Files** field by clicking the **Define/Modify** button. The IEA IDB Analyzer will use this name and folder to create six output files: an SPSS syntax file that contains the code for performing the analysis, and after running the syntax file an SPSS data file with the results, an SPSS output file and an Excel file with these same results, and another two files with significant test results, one as an SPSS data file, and the other on as an Excel file.

11. Press the **Start SPSS** button to create the SPSS syntax file. The file will open in an SPSS syntax window. The syntax file will be executed by opening the **Run** menu of SPSS and selecting the **All** option. If necessary, the IEA IDB Analyzer will prompt you to confirm overwriting already existing files.

Figure 4.8 shows what the IEA IDB Analyzer analysis module window looks like if all information is completed for our example. The results are displayed in Figure 4.9 though only the first countries are displayed to save space (this will be done for all analysis examples). Note that the IEA IDB Analyzer also displays an average of statistics for all countries included in the analysis at the end of the results tables ("Table Average").

In this example, each country's average for variable S_AGE is reported for all sampled students. The countries are identified in the first column. The second column reports the number of valid cases. The third column reports the sum of weights of the sampled students, followed by the percent, mean, and standard deviation, each accompanied by its jackknife standard error. The last but one column contains the percentage of missing cases, and the last column shows the number of variance strata contributing cases.

From the first line shown in Figure 4.9, we see that in Bulgaria valid data were available for 3,108 students and these sampled students represent a population of about 55,905 students. The estimate for the mean age of Bulgarian students was 14.83 years at the time they took the ICCS 2022 test, with a standard error of 0.01 years.

Figure 4.9: IEA IDB Analyzer setup for example student-level analysis without plausible values
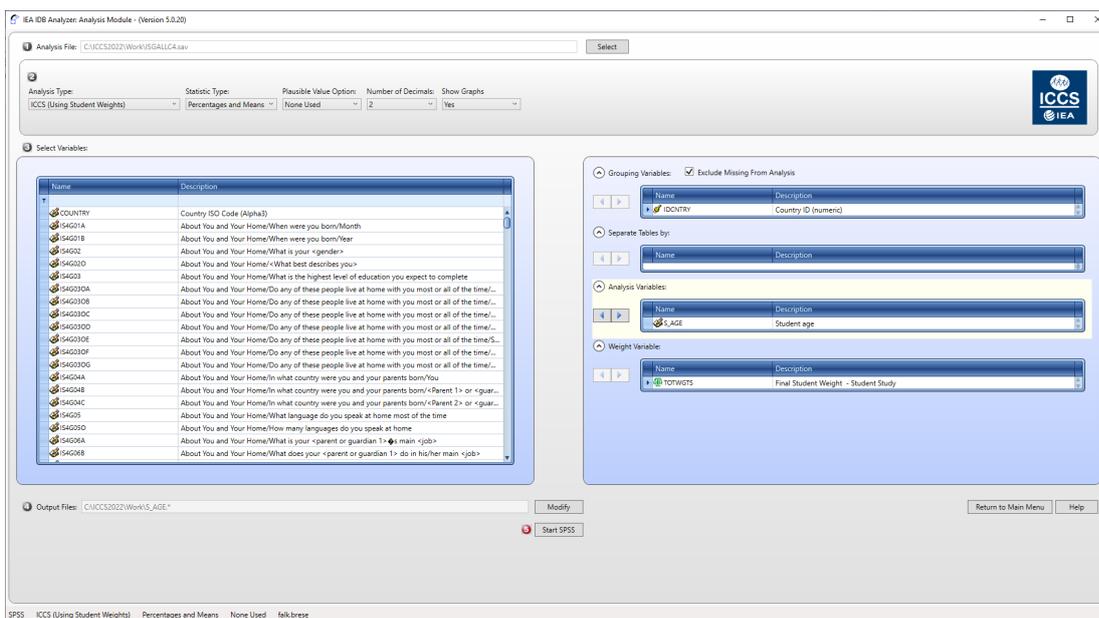
Figure 4.10: Output for example student-level analysis without civic knowledge scores

```
Average for S_AGE by (IDCNTRY )
PAGE     1

                                    Sum of
                         N of    Sum of    TOTWGTS            Percent   S_AGE    S_AGE             Std.Dev. Percent
Country ID (numeric)    Cases    TOTWGTS   (s.e.)    Percent  (s.e.)    (Mean)   (s.e.)  Std.Dev.  (s.e.)   Missing  nvarstrat
_____

Brazil                  5885    2764272,59  70665,26  38,56   ,63      14,10    ,02      ,73       ,02      ,00      75
Bulgaria                3108      55904,59   1280,08    ,78   ,02      14,83    ,01      ,42       ,02      ,18      73
Chinese Taipei          3852     185537,84   3234,70   2,59   ,06      14,21    ,01      ,37       ,01      ,00      75
Colombia                4192     741240,91  25486,39  10,34   ,35      14,10    ,04     1,02       ,02      ,04      75
Croatia                 2759      36519,71    562,71    ,51   ,01      14,67    ,01      ,33       ,01      ,00      72
…
Table Average             .           .          .     4,17   ,04      14,41    ,00      ,46       ,00      .        .
```

4.5.2   Student-level analysis with civic knowledge scores

In our second example, we want to replicate another set of results presented in the ICCS 2022 international report. We are interested in investigating the relationship between students' gender and civic knowledge, the latter being represented by a set of five plausible values. These results, presented in Table 3.14 of the ICCS 2022 international report, are repeated here in Figure 4.10. Since the results in this table are based on plausible values, we need to make sure we include them when we create the file using the merge module, and also to indicate that our analysis will make use of civic knowledge scores when we specify the analysis type.

Figure 4.11: Table of example student-level analysis with civic knowledge scores taken from the ICCS 2022 international report (Table 3.14)

Table 3.14: Gender differences in civic knowledge

Country	Mean scale score females	Mean scale score males	Gender difference (female−male)	Gender difference (female−male) -50 0 50 100
Bulgaria	477 (5.3)	436 (5.9)	**41** (6.5)	
Sweden[1]	583 (4.5)	546 (4.0)	**37** (4.8)	
Lithuania	527 (4.3)	491 (4.5)	**35** (4.1)	
Norway (9)[1]	550 (3.0)	514 (3.4)	**36** (3.3)	
Romania	487 (8.7)	454 (10.0)	**33** (4.9)	
Cyprus	476 (3.4)	443 (3.2)	**33** (4.3)	
Latvia[1]	507 (3.2)	474 (3.4)	**33** (3.4)	
Croatia[1]	547 (3.2)	515 (3.8)	**31** (4.7)	
Slovenia	519 (2.8)	490 (3.1)	**29** (3.7)	
Malta	503 (8.5)	476 (8.7)	**26** (8.7)	
Serbia	476 (3.5)	451 (4.4)	**25** (4.1)	
Italy	537 (4.5)	510 (3.7)	**27** (3.9)	
Poland	566 (2.7)	542 (3.4)	**24** (3.5)	
Estonia	558 (5.6)	533 (6.2)	**24** (4.9)	
Slovak Republic	511 (3.6)	492 (4.0)	**19** (3.9)	
Chinese Taipei	594 (2.5)	574 (3.3)	**20** (3.8)	
Spain	519 (3.6)	502 (4.0)	**17** (3.9)	
France	515 (3.6)	502 (4.2)	**13** (4.2)	
Netherlands†	514 (5.8)	504 (4.7)	10 (6.5)	
Colombia	456 (4.9)	449 (4.3)	6 (5.2)	
International average	521 (1.0)	495 (1.1)	**26** (1.1)	
Countries not meeting sample participation requirements				
Brazil	467 (3.4)	446 (4.0)	**22** (3.5)	
Denmark	575 (4.1)	543 (4.2)	**33** (4.9)	
German benchmarking participant meeting sample participation requirements				
North Rhine-Westphalia	525 (3.7)	524 (3.0)	2 (4.1)	
German benchmarking participant not meeting sample participation requirements				
Schleswig-Holstein	547 (5.7)	541 (5.8)	6 (7.7)	

Notes:
Statistically significant differences (*p* < 0.05) are displayed in **bold**.
Because results are rounded to the nearest whole number, some aggregate statistics may appear inconsistent.
( )   Standard errors appear in parentheses.
(9)   Country deviated from international defined population and surveyed adjacent upper grade.
†   Nearly met guidelines for sampling participation rates only after replacement schools were included.
[1]   National defined population covers 90% to 95% of national target population.

■ Gender difference statistically significant at *p* > 0.05
☐ Gender difference not statistically significant at *p* > 0.05

After reviewing the codebooks, we observe that the variable SD_GENDER contains (dichotomous) categorical information on the gender of the student, and this variable is found in the student questionnaire data files.

The **Percentages and Means** analysis type with the **Use PVs** option activated computes percentages and mean civic knowledge scores based on plausible values and their respective standard errors. Once you have opened the analysis module and selected the ISGALLC4.SAV data file, the steps in the IEA IDB Analyzer are as follows:

1. Open the **Analysis Module** of the IEA IDB Analyzer.

2. Select the merged data file ISGALLC4.SAV as the **Analysis File** by clicking the **Select** button.

3. Select **ICCS (Using Student Weights)** as the Analysis Type.

4. Select **Percentages and Means** as the Statistic Type. By default, the program will exclude records with missing grouping variables from the analysis.

5. Select **Use PVs** as the Plausible Value Option.

6. Add the variable SD_GENDER as a second **Grouping Variable**. The variable IDCNTRY is always selected automatically as a grouping variable.

7. Specify the civic knowledge scores to be used for the analysis by clicking the **Plausible Values** field to activate it. Select PVCIV01-05 from the list of available variables and move it to the **Plausible Values** field by clicking the **right arrow (▶)** button in this section.

8. The **Weight Variable** is automatically selected by the software. As this example analysis uses student questionnaire data, TOTWGTS is included in the analysis by default.

9. Specify the name and the folder of the output files in the **Output Files** field by clicking the **Define/Modify** button.

10. Click the **Start SPSS** button to create the SPSS syntax file. The file will open in an SPSS syntax window. The syntax file will be executed by opening the **Run** menu of SPSS and selecting the **All** menu option. If necessary, the IEA IDB Analyzer will prompt you to confirm or cancel overwriting already existing files.

Figure 4.11 below displays the analysis module with the proper settings for this example analysis. The output for this setup is shown in Figure 4.12.

Figure 4.12: IEA IDB Analyzer setup for example student-level analysis with civic knowledge scores
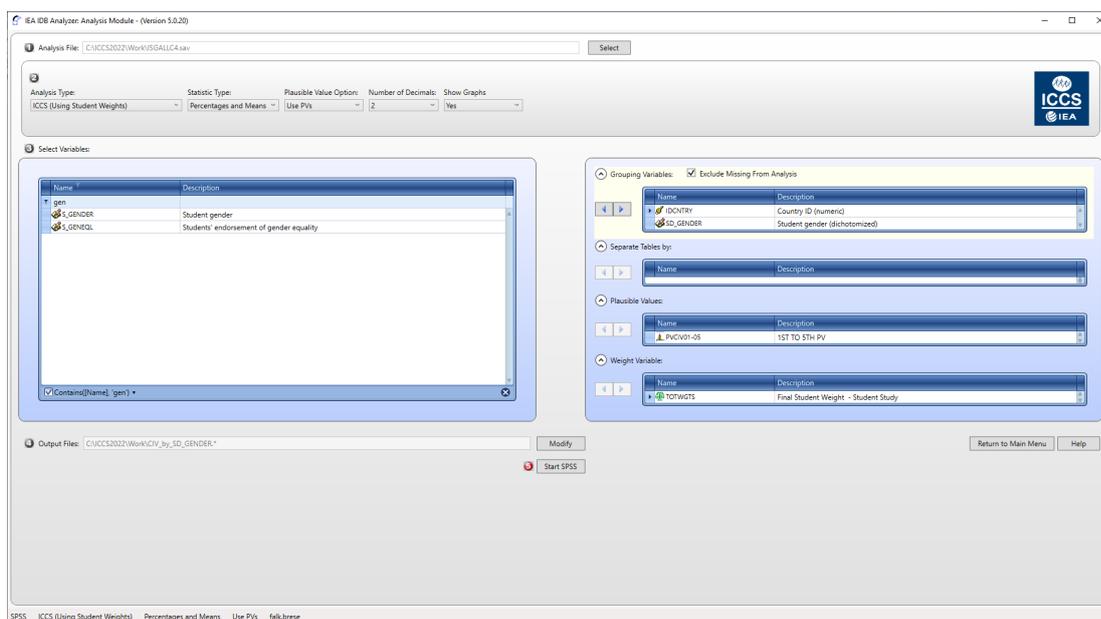
In this example, each country's results are presented on two lines, one for each of the dichotomized gender, which are the values of the SD_GENDER variable. The countries are identified in the first column and the second column describes the category of SD_GENDER being reported. The third column reports the number of valid cases and the fourth and fifth the sum of weights of the sampled students and the standard error of that sum. The next two columns report the percentage of students in each category and its standard error, followed by the estimated mean civic knowledge and its standard error. The standard deviation of the civic knowledge scores and its standard

error are reported in the next two columns. The last but one column contains the percentage of missing cases, and the last column shows the number of variance strata contributing cases to the corresponding group.

From Figure 4.12, we see that in Bulgaria 50.49 percent of the target population students are boys, and 49.51 percent are girls (standard error of 1.81%). We also see that the mean civic knowledge is 435.57 (standard error of 5.91) for boys and 467.80 (standard error of 5.32) for girls.

Figure 4.13: Output for example student-level analysis with civic knowledge scores

```
Average for PVCIV by IDCNTRY SD_GENDER
PAGE      1

                                             Sum of
                     Student gender  N of  Sum of  TOTWGTS         Percent  PVCIV   PVCIV            Std.Dev.
Country ID (numeric) (dichotomized)  Cases TOTWGTS (s.e.)  Percent (s.e.)   (Mean)  (s.e.)  Std.Dev  (s.e.)   pctmiss nvarstrat
_____

Brazil               Boy            2931 1363970 36650,56  49,34   ,75      445,61  3,98    102,27   2,38     ,00     75
                     Girl           2954 1400303 45055,93  50,66   ,75      467,45  3,40    100,65   2,77     ,00     75

Bulgaria             Boy            1510 28265   1319,79   50,49   1,81     435,57  5,91    101,17   2,76     ,00     73
                     Girl           1601 27712   1049,69   49,51   1,81     476,80  5,32    101,86   2,65     ,00     73

Chinese Taipei       Boy            1965 94526   2160,45   51,40   ,78      573,62  3,30    91,32    1,90     ,00     75
                     Girl           1850 89376   2152,15   48,60   ,78      594,12  2,48    82,62    1,60     ,00     75

Colombia             Boy            2159 386713 14592,53   52,20   1,37     449,17  4,34    89,46    2,28     ,00     75
                     Girl           2030 354092 17705,76   47,80   1,37     455,54  4,88    89,76    1,89     ,00     75

Croatia              Boy            1347 17749   481,03    48,60   1,09     515,26  3,83    86,01    2,41     ,00     72
                     Girl           1412 18771   493,42    51,40   1,09     546,50  3,18    76,70    2,26     ,00     72

...


Table Average        Boy            .     .       .         50,80  ,28      497,86  ,99     97,35    ,51      .       .
                     Girl           .     .       .         49,20  ,28      522,22  ,94     91,98    ,52      .       .
```

The IEA IDB Analyzer also provides output files for all analysis types in Excel format. These files are located in the folder specified in the **Output Files** field. For the Percentages & Means analysis type (and as well for the Benchmark analysis type) there is also an Excel file with significance test results for differences of percentages as well as differences in means. Figure 4.13 shows the first rows of this file.

Figure 4.14: Excel output including significance test results for example student-level analysis with civic knowledge scores

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1	IDCNTRY	dvar	groupvar	refgroup	compgrou	pct	pct_se	cpct	cpct_se	pctdiff	pctdiff_se	pctdiff_t	mnpv	mnpv_se	cmnpv	cmnpv_se	mnpvdiff	mnpvdiff	mnpvdiff_	wei
2	Brazil	PVCIV	SD_GENDI	Boy	Boy	49.34	0.75	49.34	0.75	0.00	0.00	#NULL!	445.61	3.98	445.61	3.98	0.00	0.00	#NULL!	TOT
3	Brazil	PVCIV	SD_GENDI	Boy	Girl	49.34	0.75	50.66	0.75	1.31	1.51	0.87	445.61	3.98	467.45	3.40	21.84	3.51	6.21	TOT
4	Brazil	PVCIV	SD_GENDI	Girl	Boy	50.66	0.75	49.34	0.75	-1.31	1.51	-0.87	467.45	3.40	445.61	3.98	-21.84	3.51	-6.21	TOT
5	Brazil	PVCIV	SD_GENDI	Girl	Girl	50.66	0.75	50.66	0.75	0.00	0.00	#NULL!	467.45	3.40	467.45	3.40	0.00	0.00	#NULL!	TOT
6	Bulgaria	PVCIV	SD_GENDI	Boy	Boy	50.49	1.81	50.49	1.81	0.00	0.00	#NULL!	435.57	5.91	435.57	5.91	0.00	0.00	#NULL!	TOT
7	Bulgaria	PVCIV	SD_GENDI	Boy	Girl	50.49	1.81	49.51	1.81	-0.99	3.63	-0.27	435.57	5.91	476.80	5.32	41.23	6.46	6.38	TOT
8	Bulgaria	PVCIV	SD_GENDI	Girl	Boy	49.51	1.81	50.49	1.81	0.99	3.63	0.27	476.80	5.32	435.57	5.91	-41.23	6.46	-6.38	TOT
9	Bulgaria	PVCIV	SD_GENDI	Girl	Girl	49.51	1.81	49.51	1.81	0.00	0.00	#NULL!	476.80	5.32	476.80	5.32	0.00	0.00	#NULL!	TOT
10	Chinese T	PVCIV	SD_GENDI	Boy	Boy	51.40	0.78	51.40	0.78	0.00	0.00	#NULL!	573.62	3.30	573.62	3.30	0.00	0.00	#NULL!	TOT
11	Chinese T	PVCIV	SD_GENDI	Boy	Girl	51.40	0.78	48.60	0.78	-2.80	1.57	-1.79	573.62	3.30	594.12	2.48	20.50	3.76	5.45	TOT

The file provides significance test results for all combinations of grouping variables that were specified in the analysis module. In this example, this comprises all combinations of students' gender—boys with boys, boys with girls, girls with boys, and girls with girls. The interesting comparison is girls with boys (or vice versa). In Bulgaria, the difference in civic knowledge between girls and boys we got already from the SPSS output (see Figure 4.12; girls score 435.57, boys score 476.80). Row 7 in Figure 4.13 shows that in Bulgaria, boys score 41.23 score points higher on the civic knowledge scale (with a standard error of 6.46). Column S provides the t-value of this difference,

which in this case is 6.38. Opting for a 95 percent confidence level, this difference is statistically significant, as the absolute value of that t-value (6.38) is ≥1.96.

4.5.3   Student-level linear regression analysis

The IEA IDB Analyzer is able to calculate multiple linear regressions between dependent variables and a set of independent variables. This section demonstrates an example for a regression analysis with civic knowledge scores using student-level variables selected in the example merged data file ISGALLC4.SAV.

The IEA IDB Analyzer can also be used to compute regression analysis without civic knowledge scores, The steps for such an analysis would be similar to those described for a regression analysis with civic knowledge scores.

In our third example, we do not replicate any results presented in the ICCS 2022 international report, instead we will examine the language spoken at home as a predictor of civic knowledge. The linear regression analysis will use the variable S_TLANG—indicating if the language the ICCS assessment was taken is the language spoken at home—as the predictor of the five plausible values for civic knowledge (PVCIV01 through PVCIV05), using the weighting variable TOTWGTS. The data will be read from the merged data file ISGALLC4.SAV and the standard errors will be computed based on 75 sets of replicate weights.

The example regression analysis is performed by the analysis module of the IEA IDB Analyzer with the completed analysis window shown in Figure 4.14, using the following steps:

1. Open the **Analysis Module** of the IEA IDB Analyzer.

2. Specify the data file ISGALLC4.SAV as the **Analysis File** by clicking the **Select** button and selecting the file from the location it has been saved.

3. Select **ICCS (Using Student Weights)** as the Analysis Type.

4. Select **Linear Regression** as the **Statistic Type**.

5. Select **Use PVs** as the Plausible Value Option.

6. The variable IDCNTRY is selected automatically as **Grouping Variables**. No additional grouping variables are needed for this analysis.

7. Click the **Categorical Variables** field in the **Independent Variables** section to activate it and select the variable S_TLANG as the independent variable. This is done by selecting S_TLANG from the list of available variables and moving it to the **Categorical Variables** field by clicking the **right arrow (▶)** button in this section. By clicking the **Contrast** field of S_TLANG, a drop-down menu will appear, from which **Dummy Coding** should be selected. By default, the IEA IDB Analyzer will assume that independent variables selected as categorical variables have two categories and it will select the first category according to the coding of the variable (i.e., the code with the lowest value) as the reference category. In our example, S_TLANG has two valid categories and code 0 (language of test is not equal to the language spoken at home) is selected as reference category. These settings fit our needs. This

way, the intercept or constant will be the estimated average civic knowledge for students with "other language[s]" spoken at home than the language of test, whereas the regression coefficient (S_TLANG (estimate)) shows the estimated difference in civic knowledge score points of students speaking the language of test at home and those students who speak another language at home. This also will allow us to perform a t-test to determine if the average civic knowledge is significantly different between students speaking and not speaking the language of test at home.

8. Click the **Plausible Values** radio button in the **Dependent Variable** section and select PVCIV01-05 as the Plausible Values. This is done by selecting PVCIV01-05 from the list of available variables and moving it to the **Plausible Values** field by clicking the **right arrow** (▶) button in this section.

9. The **Weight Variable** is automatically selected by the software. As this example analysis uses student questionnaire data, TOTWGTS is selected by default.

10. Specify the name and the folder of the output files in the **Output Files** field by clicking the **Define/Modify** button.

11. Click the **Start SPSS** button to create the SPSS syntax file. The file will open in an SPSS syntax window. The syntax file will be executed by opening the **Run** menu of SPSS and selecting the **All** option. If applicable, the IEA IDB Analyzer will prompt you to confirm overwriting already existing files.

Figure 4.15: IDB Analyzer setup for example student-level regression analysis with civic knowledge scores
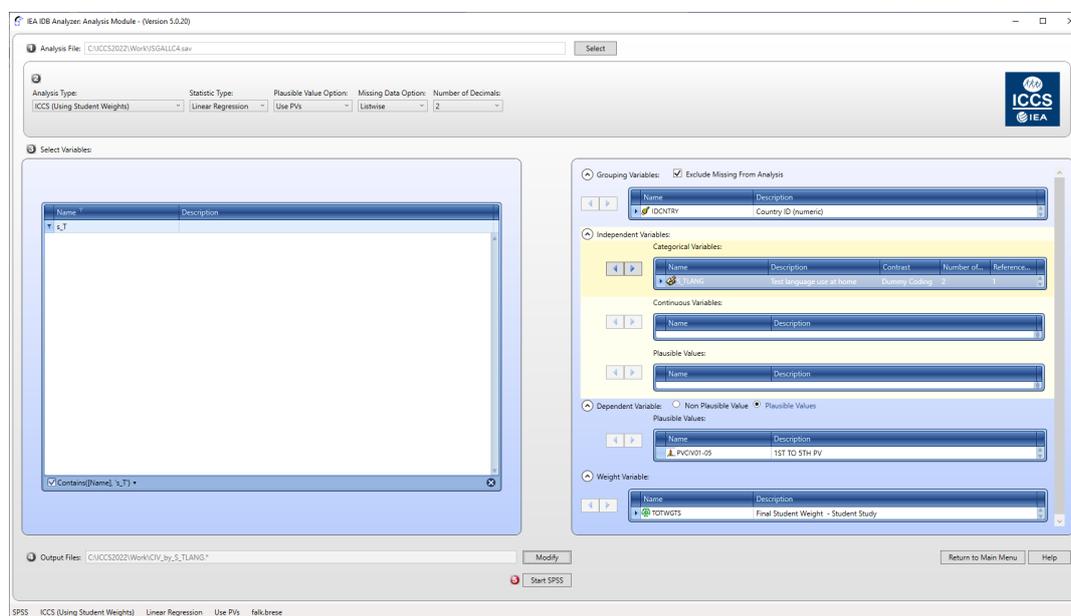
The results of this analysis are presented in Figure 4.15. We see that in Bulgaria the estimated mean civic knowledge of target-grade students speaking the language of test at home (Constant) is 386.12, with a standard error of 9.13. The Bulgarian target-grade students speaking the language of test at home have an estimated mean civic knowledge score 80.24 points (S_TLANG_D2) higher than Flemish students not speaking the language of test

at home with standard error of 9.65. The estimated t-test value is 7.52 (Stndrdzd. Coefficient (t-value)) which in absolute value is greater than 1.96, indicating that this difference is statistically significant at a 95 percent confidence level. In the model statistics section of the results output, we can see that in Bulgaria about 6 percent of the variance in student civic knowledge is explained by whether the language the ICCS test was taken was the language spoken at home.

Figure 4.16: Output for example student-level regression analysis with civic knowledge scores

```
Regression Coefficients
PAGE    1

                                   Regression  Regression              Stndrdzd.   Stndrdzd.
                        Regression  Coefficient Coefficient Stndrdzd.   Coefficient Coefficient
IDCNTRY        Variable Coefficient (s.e.)      (t-value)   Coefficient (s.e.)      (t-value)
_____       _____ _____ _____ _____ _____ _____ _____

Brazil         (CONSTANT)   428,36     30,41       14,08         .           .           .
               S_TLANG_D2    29,13     30,23         ,96        ,03         ,03         1,11

Bulgaria       (CONSTANT)   386,12      9,13       42,30         .           .           .
               S_TLANG_D2    80,24      9,65        8,32        ,25         ,03         7,53

Chinese Taipei (CONSTANT)   535,50      6,00       89,19         .           .           .
               S_TLANG_D2    52,08      5,73        9,09        ,16         ,02         8,66

Colombia       (CONSTANT)   433,44     13,55       32,00         .           .           .
               S_TLANG_D2    19,43     12,70        1,53        ,03         ,02         1,51

Croatia        (CONSTANT)   494,47     17,54       28,19         .           .           .
               S_TLANG_D2    38,14     17,34        2,20        ,06         ,03         2,20

...

Table Average  (CONSTANT)   467,13      2,38      196,05         .           .           .
               S_TLANG_D2    49,46      2,39       20,70        ,16         ,01        26,34
```

```
Model Statistics
PAGE    1

                                            Adjusted
                           R-Square  Adjusted R-Square
IDCNTRY        R-Square    (s.e.)    R-Square (s.e.)
_____       _____    _____  _____ _____

Brazil             ,00        ,00       ,00      ,00

Bulgaria           ,06        ,02       ,06      ,02

Chinese Taipei     ,03        ,01       ,03      ,01

Colombia           ,00        ,00       ,00      ,00

Croatia            ,00        ,00       ,00      ,00

...

Table Average      ,04        ,00       ,04      ,00
```

4.5.4   Calculating percentages of students reaching proficiency levels

This section describes the IEA IDB Analyzer's ability to perform benchmark analyses, which will compute the percentages of students reaching specified proficiency levels on the civic knowledge scale and within specified subgroups, along with appropriate standard errors.

As an example, we will compute the percentages of students (not) reaching the four ICCS 2022 international proficiency levels of civic knowledge (Level D - 311 to 394 score points; Level C – 395 to 478 score points; Level B – 479 to 562 score points; Level A – 563 score points and above) using the merged ISGALLC4.SAV data file. These results, presented in Table 3.10 of the *ICCS 2022 international report*, are provided in Figure 4.16.

Figure 4.17: Example table of proficiency levels analysis taken from the ICCS 2022 international report (Table 3.10)

Table 3.10: Percentages of students at each proficiency level of civic knowledge

Country	Below Level D	Level D	Level C	Level B	Level A	
Chinese Taipei	0.3 (0.1)	2.8 (0.4)	9.3 (0.7)	25.5 (1.1)	62.1 (1.2)	
Sweden[1]	1.2 (0.3)	6.4 (0.8)	14.3 (0.8)	25.2 (0.9)	52.8 (1.4)	
Poland	0.5 (0.2)	4.3 (0.5)	15.4 (0.8)	32.2 (1.1)	47.6 (1.3)	
Estonia	0.8 (0.3)	5.5 (0.7)	18.9 (1.3)	31.5 (1.5)	43.4 (2.4)	
Croatia[1]	2.6 (0.4)	9.3 (0.7)	18.1 (0.8)	29.4 (0.8)	40.6 (1.2)	
Norway (9)[1]	^ ^	5.0 (0.6)	20.9 (1.1)	37.6 (1.3)	36.1 (1.3)	
Italy	1.1 (0.3)	7.0 (0.8)	22.3 (1.6)	34.8 (1.4)	34.7 (2.0)	
Spain	2.0 (0.4)	12.0 (1.2)	23.7 (1.3)	31.5 (1.6)	30.8 (1.8)	
Lithuania	1.8 (0.3)	10.6 (0.9)	24.9 (1.2)	32.5 (1.1)	30.1 (1.8)	
Netherlands†	1.6 (0.3)	9.9 (0.8)	25.5 (0.9)	33.9 (0.9)	29.0 (1.2)	
France	1.5 (0.3)	9.6 (0.9)	24.1 (1.2)	36.1 (1.1)	28.8 (1.3)	
Slovenia	2.9 (0.6)	11.4 (1.1)	25.2 (1.1)	32.5 (1.2)	28.0 (1.3)	
Slovak Republic	4.4 (1.4)	15.2 (1.6)	25.1 (1.3)	29.3 (1.7)	25.9 (2.1)	
Latvia[1]	1.2 (0.2)	9.5 (0.6)	27.8 (1.0)	36.1 (1.2)	25.4 (1.2)	
Malta	1.8 (0.3)	11.8 (0.9)	30.8 (1.2)	35.6 (1.5)	20.0 (1.1)	
Romania	5.5 (1.1)	18.8 (2.1)	26.8 (2.1)	30.6 (2.5)	18.3 (3.1)	
Serbia	7.9 (1.0)	22.7 (1.6)	27.5 (1.4)	24.7 (1.4)	17.3 (1.3)	
Cyprus	5.8 (0.6)	21.9 (0.9)	31.0 (1.2)	25.9 (1.0)	15.4 (0.8)	
Bulgaria	4.2 (0.6)	19.4 (1.8)	32.7 (1.7)	29.6 (1.2)	14.0 (1.1)	
Colombia	5.2 (0.6)	23.6 (1.5)	32.4 (1.4)	26.7 (1.4)	12.2 (1.1)	
ICCS 2022 average	2.6 (0.1)	11.8 (0.2)	23.8 (0.3)	31.1 (0.3)	30.6 (0.4)	

Countries not meeting sample participation requirements						
Brazil	6.8 (0.6)	22.7 (0.9)	30.7 (1.0)	24.2 (1.0)	15.7 (1.1)	
Denmark	1.3 (0.3)	6.1 (0.6)	15.3 (0.9)	27.3 (1.1)	49.9 (1.5)	
German benchmarking participant meeting sample participation requirements						
North Rhine-Westphalia	1.5 (0.2)	9.0 (0.7)	22.8 (1.1)	30.5 (1.2)	36.3 (1.1)	
German benchmarking participant not meeting sample participation requirements						
Schleswig-Holstein	1.1 (0.4)	7.8 (1.2)	17.9 (1.6)	27.9 (1.6)	45.2 (1.9)	

Notes:
We have presented the countries in descending order according to the percentage of students with scores that positioned them at Proficiency Level A on the scale.
Because results are rounded to the nearest whole number, some aggregate statistics may appear inconsistent.
() Standard errors appear in parentheses.
(9) Country deviated from international defined population and surveyed adjacent upper grade.
† Nearly met guidelines for sampling participation rates only after replacement schools were included.
^ Number of students too small to report percentages at the level

Below level D ■ Level D □ Level C
■ Level B □ Level A

This example is performed by the analysis module of the IEA IDB Analyzer with the completed analysis window shown in Figure 4.17, using the following steps:

1. Open the **Analysis Module** of the IEA IDB Analyzer.

2. Specify the data file ISGALLC4.SAV as the **Analysis File** by clicking the **Select** button.

3. Select **ICCS (Using Student Weights)** as the Analysis Type.

4. Select **Benchmarks** as the **Statistic Type**.

5. Select the **Discrete** option under the **Benchmark Option** drop-down menu to get discrete percentages of students reaching the ICSS international benchmarks.

6. The variable IDCNTRY is selected automatically as **Grouping Variables**. No additional grouping variables are needed for this analysis.

7. Specify the civic knowledge scores to be used for the analysis by clicking the **Plausible Values** field. Select PVCIV01-05 from the list of available variables and move it to the **Plausible Values** field by clicking the **right arrow (▶)** button in this section.

8. Specify the ICCS 2022 International Benchmarks—311, 395, 479 and 563 (Level D, Level C, Level B and Level A respectively). These values can be entered manually in the **Achievement Benchmarks** field, each separated by a blank space, or they can be selected by clicking on the drop-down menu available for this field.

9. The **Weight Variable** is automatically defined by the software. As this example analysis uses student questionnaire data, TOTWGTS is selected by default.

10. Specify the name and the folder of the output files in the **Output Files** field by clicking the **Define/Modify** button.

11. Click the **Start SPSS** button to create the SPSS syntax file. The file will open in an SPSS syntax window. The syntax file will be executed by opening the **Run** menu of SPSS and selecting the **All** option. If necessary, the IEA IDB Analyzer will prompt you to confirm overwriting already existing files.

Figure 4.18: IDB Analyzer setup for example benchmark analysis
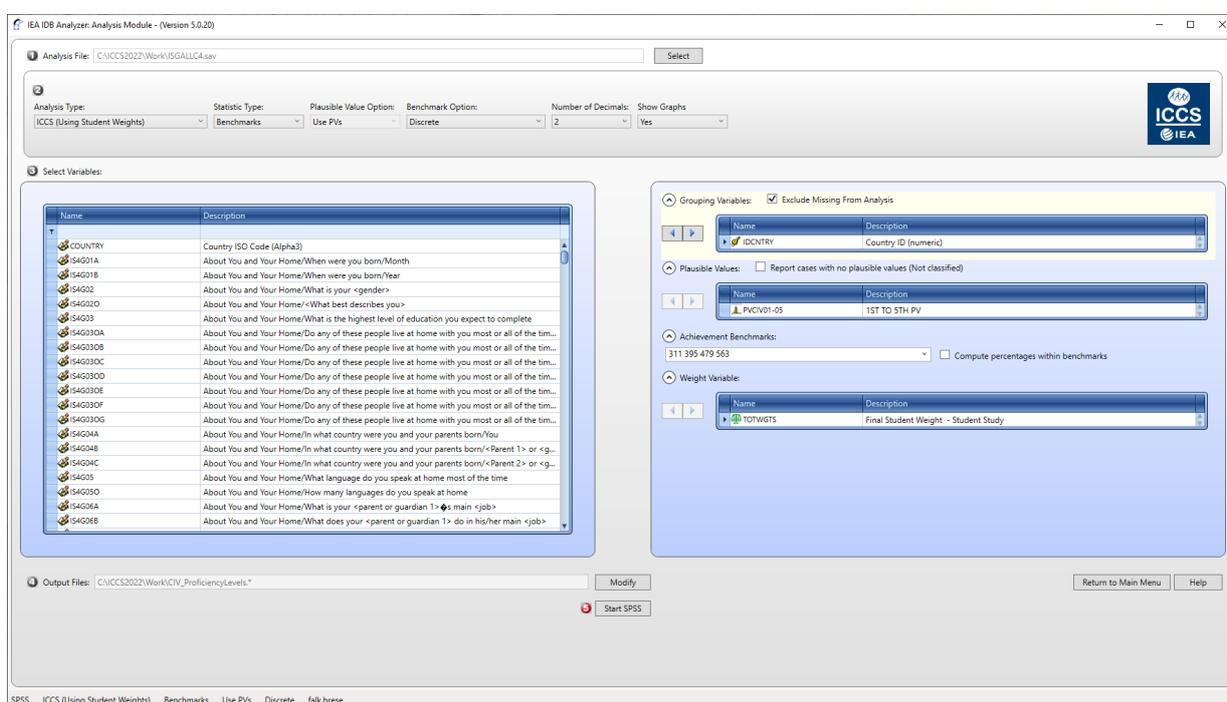
The results of this analysis are presented in Figure 4.18. In Bulgaria 7.94 percent of the target-grade students are below proficiency Level D of 311 score points, with a standard error of 1.00 percent. 22.65 percent of students reached proficiency Level D (from 311 to below 395 score points), with a standard error of 1.57 percent. 27.45 percent of the students reached proficiency level C (from 395 to below 479 score points), with a standard error of

1.41 percent. 24.66 percent of students reached proficiency Level B (from 479 to below 563 score points) with a standard error of 1.43 percent. 17.30 percent of the students scored at proficiency Level A (at or above 563 score points) with standard error of 1.32 percent.

Figure 4.19: Output for example benchmark analysis of levels of civic knowledge

```
Percents by Performance Groups of PVCIV
PAGE    1

                                                           Sum of
                                             N of    Sum of   TOTWGTS              Percent
Country ID (numeric)       Performance Group  Cases   TOTWGTS   (s.e.)    Percent   (s.e.)

Brazil                     1.Below 311          379    187798  16841,89     6,79      ,59
                           2.From 311 to Below 395  1370    626495  29138,02    22,66      ,95
                           3.From 395 to Below 479  1829    847281  36669,19    30,65     1,01
                           4.From 479 to Below 563  1430    668106  34505,02    24,17      ,96
                           5.At or Above 563      877    434592  29935,70    15,72     1,06

Bulgaria                   1.Below 311          214      4448    582,13     7,94     1,00
                           2.From 311 to Below 395   646     12686    969,84    22,65     1,57
                           3.From 395 to Below 479   827     15374    885,18    27,45     1,41
                           4.From 479 to Below 563   811     13812    825,29    24,66     1,43
                           5.At or Above 563      615      9687    714,29    17,30     1,32

Chinese Taipei             1.Below 311           11       537    218,61      ,29      ,12
                           2.From 311 to Below 395    95      5104    817,61     2,75      ,44
                           3.From 395 to Below 479   333     17276   1208,41     9,31      ,65
                           4.From 479 to Below 563   944     47362   2289,96    25,53     1,06
                           5.At or Above 563     2469    115260   2936,51    62,12     1,22

Colombia                   1.Below 311          173     38335   4411,89     5,17      ,60
                           2.From 311 to Below 395   915    175016  13000,37    23,60     1,54
                           3.From 395 to Below 479  1387    240116  13111,13    32,38     1,39
                           4.From 479 to Below 563  1174    197954  12405,64    26,69     1,35
                           5.At or Above 563      544     90133   8961,18    12,15     1,12

Croatia                    1.Below 311            8       126     78,27      ,34      ,21
                           2.From 311 to Below 395   135      1841    225,79     5,04      ,60
                           3.From 395 to Below 479   573      7646    437,13    20,94     1,15
                           4.From 479 to Below 563  1045     13719    511,09    37,57     1,29
                           5.At or Above 563      998     13188    488,56    36,11     1,28

...

Table Average              1.Below 311            .         .        .        2,64      ,11
                           2.From 311 to Below 395    .         .        .       11,76      ,22
                           3.From 395 to Below 479    .         .        .       23,48      ,25
                           4.From 479 to Below 563    .         .        .       30,46      ,27
                           5.At or Above 563        .         .        .       31,65      ,32
```

4.5.5   Computing correlations with context or background variables and civic knowledge scores

In addition to the analyses described above, the IEA IDB Analyzer also is able to compute correlations between context or background variables, and between context or background variables and civic knowledge scores. The example shown here will be a correlation analysis with civic knowledge scores. Correlation analysis between two context or background variables follows the same steps. The only difference is that the correlation between two context or background variables will require adding two variables in the **Analysis Variables** field instead of one.

The examples of correlation between a context or background variable and civic knowledge scores presented here cannot be found in the ICCS 2022 international report. The steps of conducting correlation analysis with IEA IDB Analyzer are described below and Figure 4.19 shows how the completed analysis window looks like. In this example we will calculate the correlation between students' discussion of political and social issues outside of school (S_POLDISC) and the civic knowledge score (represented by the five plausible values PVCIV01-05).

1. Open the **Analysis Module** of the IEA IDB Analyzer.

2. Specify the data file ISGALLC4.SAV as the **Analysis File**.

3. Select **ICCS (Using Student Weights)** as the Analysis Type.

4. Select **Correlations** as the **Statistic Type**. The IDCNTRY (country ID) is selected by default. No other variable needs to be selected for this analysis.

Figure 4.20: IDB analyzer setup for example correlation analysis
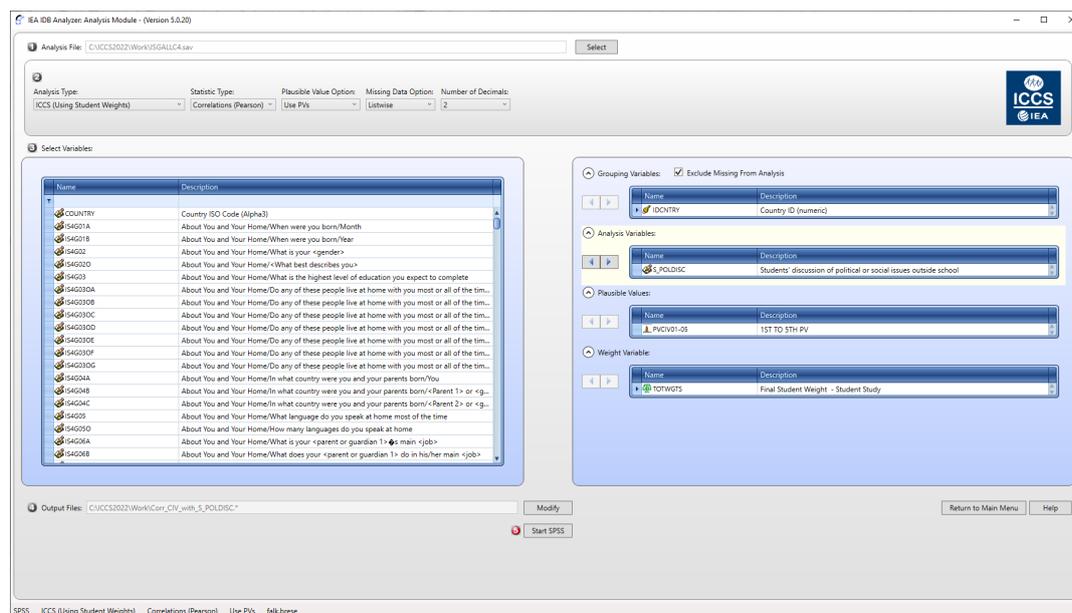
5. Specify the civic knowledge scores to be used for the analysis by clicking the **Plausible Values** field to activate it.

6. Specify the variable POLDISC as the analysis variable by clicking the **Analysis Variables** field to activate it. Select POLDISC from the list of available variables and move it to the **Analysis Variables** field by clicking the **right arrow (▶)** button in this section.

7. Select PVCIV01-05 from the list of available variables and move it to the **Plausible Values** field by clicking the **right arrow (▶)** button in this section.

8. The **Weight Variable** is automatically defined by the software. As this example analysis uses student questionnaire data, TOTWGTS is selected by default.

9. Specify the name and the folder of the output files in the **Output Files** field by clicking the **Define/Modify** button.

Click the **Start SPSS** button to create the SPSS syntax file. The file will open in an SPSS syntax window. The syntax file will be executed by opening the **Run** menu of SPSS and selecting the **All** option. If necessary, the IEA IDB Analyzer will prompt you to confirm overwriting already existing files.

Figure 4.20 shows the output from the correlation analysis conducted by IEA IDB Analyzer. As the SPSS output shows, the correlation between students' discussion of political and social issues outside of school and the civic knowledge score in Chinese Taipei is 0.17 with a standard error of 0.02. To verify whether this correlation is statistically significant, we can compute the t-value by dividing the correlation coefficient by its standard error. Using the exact numbers from the Excel output, we get a t-value of 7.56, which indicates that on a 95 percent confidence level there is a statistically significant correlation in Chinese Taipei between students' discussion of political and social issues outside of school and students' civic knowledge.

Figure 4.21: Output for example correlation analysis

```
Correlation Coefficients
PAGE    1

                                                    Correlation
                                    Correlation        with                        Correlation
                                       with          S_POLDISC     Correlation      with PV_CIV
IDCNTRY                  Variable     S_POLDISC        (s.e.)      with PV_CIV         (s.e.)

Brazil                   S_POLDISC       1,00            ,00           ,14              ,02
                         PV_CIV           ,14            ,02          1,00              ,00

Bulgaria                 S_POLDISC       1,00            ,00           ,07              ,03
                         PV_CIV           ,07            ,03          1,00              ,00

Chinese Taipei           S_POLDISC       1,00            ,00           ,17              ,02
                         PV_CIV           ,17            ,02          1,00              ,00

Colombia                 S_POLDISC       1,00            ,00          -,01              ,03
                         PV_CIV          -,01            ,03          1,00              ,00

Croatia                  S_POLDISC       1,00            ,00           ,14              ,03
                         PV_CIV           ,14            ,03          1,00              ,00

...

Table Average            S_POLDISC       1,00            ,00           ,14              ,01
                         PV_CIV           ,14            ,01          1,00              ,00
```

4.5.6   Calculating percentiles of students' civic knowledge

To calculate percentiles of civic knowledge scores you need to select the **Percentiles** analysis type. This computes the percentiles within the distribution of student civic knowledge scores within specified subgroups of students. This analysis type also computes the appropriate standard errors for those percentiles.

As an example, we will compute the percentiles of student civic knowledge scores and their standard errors within each country, using the weighting variable TOTWGTS, as in Table 3.9 of the ICCS 2022 international report. The data will be read from the data file ISGALLC4.sav and the standard errors will be computed based on replicate weights.

Figure 4.22: Percentiles of student civic knowledge taken from the ICCS 2022 international report (Table 3.9)

Table 3.9: Distributions of civic knowledge

Country	Grade	Average age	Civic knowledge scale scores							Average scale score	HDI
			250	350	450	550	650	750			
Chinese Taipei	8	14.2								583 (2.3) ▲	0.93
Sweden[1]	8	14.8								565 (3.5) ▲	0.95
Poland	8	14.4								554 (2.5) ▲	0.88
Estonia	8	15.0								545 (5.5) ▲	0.89
Croatia[1]	8	14.7								531 (2.6) ▲	0.86
Norway (9)[1]	9	14.9								529 (2.8) ▲	0.96
Italy	8	13.8								523 (3.6) ▲	0.90
Spain	8	14.0								510 (3.3)	0.91
Lithuania	8	14.8								509 (4.0)	0.88
Netherlands†	8	14.1								508 (4.1)	0.94
France	8	13.9								508 (3.3)	0.90
Slovenia	8	13.9								504 (2.3)	0.92
Slovak Republic	8	14.3								501 (3.3) ▼	0.85
Latvia[1]	8	14.8								490 (2.8) ▼	0.86
Malta	8	13.6								490 (7.4) ▼	0.92
Romania	8	15.0								470 (9.1) ▼	0.82
Serbia	8	14.6								464 (3.4) ▼	0.80
Cyprus	8	13.9								459 (2.5) ▼	0.90
Bulgaria	8	14.8								456 (4.6) ▼	0.80
Colombia	8	14.1								452 (3.8) ▼	0.75
ICCS 2022 average		**14.4**	Below D	D	C	B	A			**508 (0.9)**	
					Proficiency Level						
Countries not meeting sample participation requirements											
Brazil	8	14.1								457 (3.3)	0.75
Denmark	8	14.8								556 (3.5)	0.95
German benchmarking participant meeting sample participation requirements											
North Rhine-Westphalia	8	14.3								524 (2.6) ▲	0.94
German benchmarking participant not meeting sample participation requirements											
Schleswig-Holstein	8	14.5								544 (4.4)	0.92

Notes:
() Standard errors appear in parentheses.
(9) Country deviated from international defined population and surveyed adjacent upper grade.
† Nearly met guidelines for sampling participation rates only after replacement schools were included.
[1] National defined population covers 90% to 95% of national target population.
[2] Country surveyed target grade in the first half of the school year.
HDI = Human Development Index.

Percentiles of performance
5th 25th 75th 95th
Mean and confidence interval (±2SE)

▲ Achievement significantly higher than international average
▼ Achievement significantly lower than international average

The steps in the IEA IDB Analyzer required to follow the example are:

1. Open the **Analysis Module** of the IEA IDB Analyzer.

2. Specify the data file ISGALLC4.SAV as the **Analysis File**.

3. Select ICCS (Using Student Weights) as the Analysis Type.

4. Select Percentiles as the **Statistic Type**.

5. The IDCNTRY (country ID) is selected by default as a grouping variable. No other grouping variable needs to be selected for this analysis.

6. Specify the civic knowledge scores to be used for the analysis by clicking the **Plausible Values** field to activate it. Select PVCIV01-05 from the list of available variables and move it to the **Plausible Values** field by clicking the **right arrow (▶)** button in this section.

7. The **Weight Variable** is automatically defined by the software. As this example analysis uses student questionnaire data, TOTWGTS is selected by default.

8. Click on the **Percentiles** radio button and specify the percentile points in the distribution. For our example we will use the 5th, 25th, 75th, and 95th percentiles. These need to be typed in increasing order, separated by spaces.

9. Specify the name and the folder of the output files in the **Output Files** field by clicking the **Define/Modify** button.

10. Click the **Start SPSS** button to create the SPSS syntax file. The file will open in an SPSS syntax window. The syntax file will be executed by opening the **Run** menu of SPSS and selecting the **All** option. If necessary, the IEA IDB Analyzer will prompt you to confirm overwriting already existing files.

Figure 4.22 shows the IDB Analyzer Setup Screen for this analysis, Figure 4.23 shows the SPSS output obtained from SPSS after running the analysis.

Figure 4.23: Analysis Module setup screen for computing percentiles
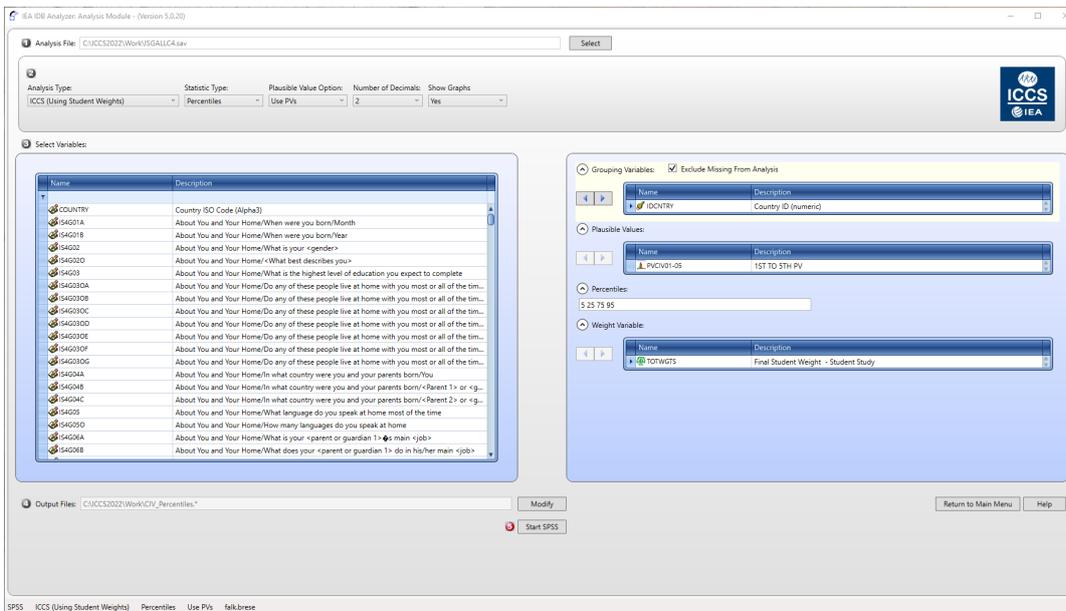
Figure 4.24: SPSS output for percentiles

```
Percentiles for PVCIV by IDCNTRY
PAGE    1

              N of     Sum of
Country ID   Cases    TOTWGTS       p5    p5_se     p25   p25_se     p75   p75_se     p95   p95_se

Brazil        5885   2764273    300,60     3,76  381,75     3,91  526,10     4,40  633,78     7,54
Bulgaria      3113     56007    292,50     7,24  376,61     5,89  534,41     5,23  623,70     6,10
Chinese Taipei 3852   185538    423,78     5,98  528,59     3,65  645,00     2,70  714,25     4,41
Colombia      4193    741554    309,73     5,62  384,96     4,46  517,50     3,94  601,73     5,71
Croatia       2759     36520    391,76     5,53  474,82     4,70  589,96     3,22  665,02     5,10

...

Table Average    .        .    348,13     1,44  443,06     1,18  578,12     1,01  661,73     1,24
```

Notes:

N = number, s.e. = standard error, p5 = 5th percentile, p25 = 25th percentile, p75 = 75th percentile, p95 = 95th percentile. The output shows the first five countries and is for illustrative purposes only.

From the results shown in Figure 4.23, we see that in Bulgaria the score of the 5th percentile of the score distribution is 292.50 points (standard error 7.42), for the 25th – 376.61 (5.89), for the 75th – 534.41 (5.23), and for the 95th – 623.70 (6.10).

4.6 Performing analyses with teacher-level data

As already noted, student and teacher data cannot be merged and analyzed together due to the sampling design of ICCS 2022. The example of an analysis using teacher questionnaire data presented here will investigate the percentage of teachers who report taking part in cultural activities (e.g., theatre, music, cinema) with their target grade classes they teach. The results of such an analysis are presented in Table 6.11 (column five) of the ICCS 2022 international report and are reproduced here in Figure 4.24. We will use the **Percentages Only** analysis type to estimate the percentages of teachers reporting taking part in cultural activities with target classes.

Figure 4.25: Table of example teacher-level analysis taken from the ICCS 2022 international report (Table 6.11)

Table 6.11: Teachers' perceptions of student activities in the community

Country	Activities related to environmental sustainability (e.g., energy and water saving, recycling)			Activities related to human rights			Activities for underprivileged people or groups			Cultural activities (e.g., theatre, music)			Multicultural and intercultural activities within the local community (e.g., promotion and celebration of cultural diversity, food street market)		
Bulgaria†	56	(2.0)	▽	40	(1.7)	▼	39	(2.3)	▽	54	(1.9)	▽	42	(2.1)	△
Chinese Taipei	65	(1.0)	△	44	(1.3)	▽	37	(1.2)	▽	51	(1.4)	▽	36	(1.5)	
Croatia	57	(3.1)		56	(1.7)	△	37	(1.3)	▽	54	(2.4)	▽	29	(2.3)	▽
Italy	64	(1.8)	△	73	(1.4)	▲	48	(1.6)		57	(1.6)		25	(1.8)	▽
Lithuania	72	(1.6)	▲	56	(1.8)	△	53	(1.8)	△	84	(1.2)	▲	76	(1.2)	▲
Malta	42	(3.0)	▼	25	(2.7)	▼	31	(2.1)	▼	32	(4.5)	▼	24	(4.2)	▼
Norway (9)	63	(2.0)		60	(1.8)	△	31	(1.8)	▼	77	(1.6)	▲	17	(1.6)	▼
Poland	66	(1.2)	△	55	(1.5)	△	76	(1.0)	▲	49	(1.5)	▼	30	(1.4)	▽
Romania	63	(2.1)		55	(1.9)	△	48	(1.6)	△	60	(2.1)		40	(2.2)	△
Serbia	56	(1.8)	▽	50	(1.6)		59	(2.7)	▲	51	(2.5)	▽	35	(1.6)	
Slovak Republic	63	(2.3)		54	(2.2)		35	(2.2)	▼	60	(2.3)		30	(1.8)	▽
Slovenia	64	(1.2)	△	49	(1.3)		51	(1.4)	△	81	(1.1)	▲	38	(1.2)	△
Spain	55	(1.8)	▽	47	(1.3)	▽	44	(1.5)		66	(1.6)	△	29	(1.6)	▽
ICCS 2022 average	61	(0.6)		51	(0.5)		45	(0.5)		60	(0.6)		35	(0.6)	

Countries not meeting sample participation requirements															
Brazil	66	(2.1)		70	(2.3)		49	(2.3)		76	(2.6)		62	(2.0)	
Colombia	68	(2.3)		65	(2.4)		40	(2.6)		68	(3.2)		55	(5.2)	
Cyprus	80	(1.2)		59	(1.8)		49	(1.6)		85	(1.3)		37	(2.0)	
Denmark	53	(4.2)		45	(3.6)		26	(4.4)		76	(4.8)		19	(3.2)	
Estonia	72	(1.3)		42	(1.5)		42	(1.7)		87	(1.2)		41	(1.7)	
France	35	(1.6)		37	(1.6)		26	(1.4)		57	(2.0)		12	(1.0)	
Latvia	54	(2.5)		44	(1.5)		31	(1.4)		70	(1.3)		41	(1.2)	
Netherlands	21	(2.8)		17	(1.6)		20	(2.6)		46	(2.1)		12	(1.8)	
Sweden	33	(2.0)		45	(1.7)		20	(1.2)		64	(2.1)		16	(1.4)	
German benchmarking participant not meeting sample participation requirements															
North Rhine-Westphalia	35	(1.8)		27	(1.6)		30	(1.2)		46	(1.7)		18	(1.2)	

Notes:
Because results are rounded to the nearest whole number, some aggregate statistics may appear inconsistent.
() Standard errors appear in parentheses.
(9) Country deviated from international defined population and surveyed adjacent upper grade.
† Met guidelines for sampling paticipation rates only after replacement schools were included.

National ICCS 2022 results are:
▲ More than 10 percentage points above ICCS 2022 average
△ Significantly above ICCS 2022 average
▽ Significantly below ICCS 2022 average
▼ More than 10 percentage points below ICCS 2022 average

Similar to the previous examples, we first proceed to identify the variables relevant to the analysis in the appropriate files and review the documentation for any specific national adaptations to the questions of interest (see Appendix B of this IDB User Guide). Since we are interested in teacher-level data, we need to look in the

teacher questionnaire data files for the variable that contains the information on the teachers' report on taking part in cultural activities of target grade classes (IT4G08D).

The merged data file ITGALLC4.SAV will be used for this example. The variable that identifies the country (IDCNTRY) is selected automatically by the IEA IDB Analyzer Analysis Module, and so are the variables that contain the sampling information and that will be used to generate the replicate weights for the analysis.

The example teacher-level analysis is performed by the analysis module of the IEA IDB Analyzer with the completed analysis window shown in Figure 4.25, using the following steps:

1. Open the **Analysis Module** of the IEA IDB Analyzer.

2. Select the data file ITGALLC4.SAV as the **Analysis File**.

3. Select **ICCS (Using Teacher Weights)** as the Analysis Type since we want to analyze the responses of the teachers.

4. Select **Percentages only** as the **Statistic Type**.

5. Specify the variable IT4G08D as a second grouping variable by clicking the **Grouping Variables** field to activate it. Select IT4G08D from the list of available variables and move it to the **Grouping Variables** field by clicking the **right arrow (▶)** button in this section.

6. The **Weight Variable** is automatically defined by the software. As this example analysis uses only teacher questionnaire data, TOTWGTT is selected by default.

7. Specify the name and the folder of the output files in the **Output Files** field by clicking the **Define/Modify** button.

8. Click the **Start SPSS** button to create the SPSS syntax file. The file will open in an SPSS syntax window. The syntax file will be executed by opening the **Run** menu of SPSS and selecting the **All** option. If necessary, the IEA IDB Analyzer will prompt you to confirm overwriting already existing files.

Figure 4.26: IDB Analyzer setup for example teacher-level analysis
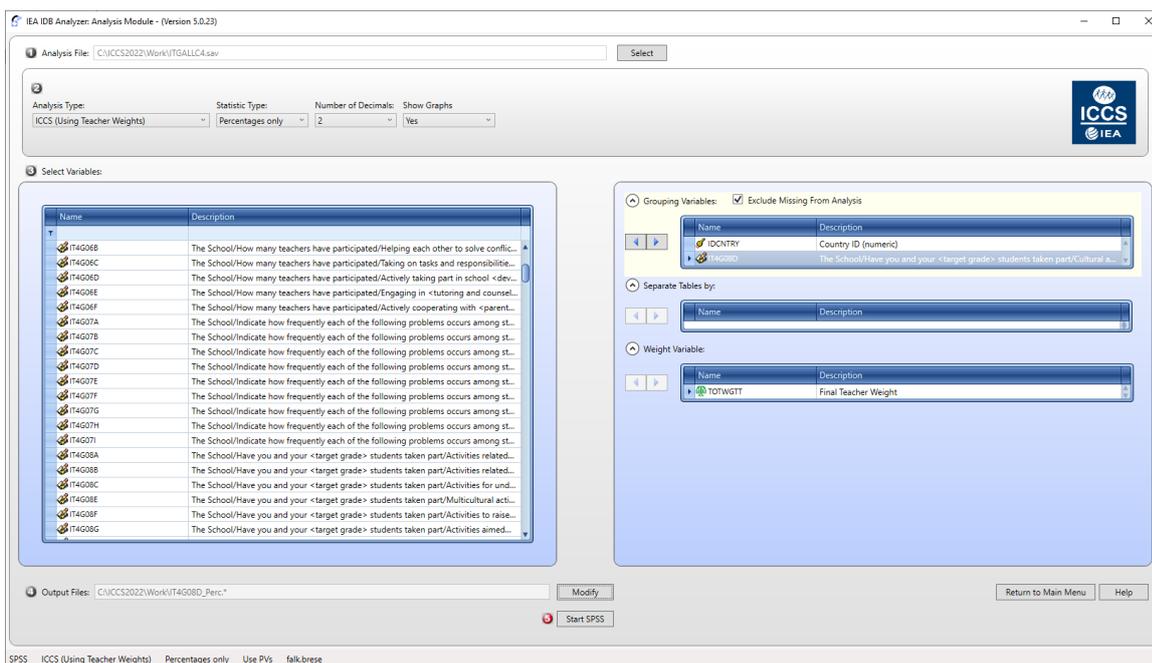
The results of this analysis are presented in Figure 4.26. Each country's results are presented on two lines, one for each value of the IT4G08D variable. In this case the selected variable has two categories (Yes/No). For categorical variables with more than two categories the output will show one line per category for each single country. The results are presented in the same manner as in the previous examples, with countries identified in the first column and the second column describing the categories of IT4G08D. From Figure 4.26 we can see that 54.45 percent of teachers in Bulgaria report they have taken part in cultural activities with any of the target classes they teach, 45.55 percent of the teachers reported they did not, and the standard error in both cases is equal to 1.90 percent.

Figure 4.27: Output for example teacher-level analysis

Percentages by (IDCNTRY IT4G08D)							PAGE 1
Country ID (numeric)	The School/Have you and your <target grade> students taken part/Cultural activities (e.g. theatre, music)	N of Cases	Sum of TOTWGTT	Sum of TOTWGTT (s.e.)	Percent	Percent (s.e.)	nvarstrat
Brazil	Yes	1244	332490	19080,10	75,93	2,61	75
	No	421	105382	12302,59	24,07	2,61	71
Bulgaria	Yes	988	7719	360,28	54,45	1,90	71
	No	784	6457	302,68	45,55	1,90	71
Chinese Taipei	Yes	1139	14275	460,76	50,57	1,40	75
	No	1147	13954	463,81	49,43	1,40	75
Colombia	Yes	797	93285	8493,81	67,65	3,20	62
	No	387	44619	5308,84	32,35	3,20	62
Croatia	Yes	1280	7436	568,80	54,41	2,39	75
	No	978	6230	301,67	45,59	2,39	75
...							
Table Average	Yes	.	.	.	62,32	,55	.
	No	.	.	.	37,68	,55	.

4.7   Performing analyses with school data

When performing analyses with the merged student and school data, the data are analyzed to make statements about the number or percentages of students attending schools with a given characteristic, rather than about the number or percentages of schools with a given characteristic. Our example of an analysis with school data will investigate the percentage of students who attend schools in which civic and citizenship education is taught as a separate subject by teachers in civic education. We will additionally calculate the average civic knowledge within each of the two categories of the variable. The example presented here is not taken from the ICCS 2022 international report.

In the current example we will use the ICGISGALLC4.SAV data file that contains school- and student-level data merged as described earlier. Please note that in merging school- and student-level data only the **Total Student Weight** TOTWGTS (and student **Jackknifing Variables** JKZONES and JKREPS) are included in the merged file, but not the school-level ones. To conduct this analysis, we will use the **Means and Percentages** analysis type in the IEA IDB Analyzer with the **Plausible Value Option** checked. The first step in our analysis is to identify the variables of interest from the appropriate files and review the documentation on specific national adaptations to the questions of interest (see Appendix B). Variable IC4G15A in the school questionnaire data file contains information on whether civic and citizenship is taught as a separate subject in the school.

The variable identifying the country (IDCNTRY) is selected automatically by the IEA IDB Analyzer.

The example analysis with school data is performed by the analysis module of the IEA IDB Analyzer with the completed analysis window shown in Figure 4.27, using the following steps:

1. Open the **Analysis Module** of the IEA IDB Analyzer.

2. Select the merged data file ICGISGALLC4.SAV as the **Analysis File**.

3. Select **ICCS (Using Student Weights)** as the Analysis Type because we want to analyze school questionnaire data as student attributes.

4. Select **Percentages and Means** as the Statistic Type.

5. Select **Use PVs** as the Plausible Value Option.

6. Specify the variable IC4G15A as a second grouping variable by clicking the **Grouping Variables** field to activate it. Select IC4G15A from the list of available variables and move it to the **Grouping Variables** field by clicking the **right arrow (▶)** button in this section.

7. Specify the civic knowledge scores to be used for the analysis by clicking the **Plausible Values** field to activate it. Select PVCIV01-05 from the list of available variables and move it to the Plausible Values field by clicking the **right arrow (▶)** button in this section.

8. The **Weight Variable** is automatically selected by the software; As this example analysis uses student questionnaire data as well as school questionnaire data disaggregated to student level, TOTWGTS is selected by default.

9. Specify the name and the folder of the output files in the Output Files field by clicking the **Define/Modify** button.

10. Click the **Start SPSS** button to create the SPSS syntax file and open it in a SPSS syntax window. The syntax file can be executed by opening the **Run** menu of SPSS and selecting the **All** option. If necessary, the IEA IDB Analyzer will display a prompt to confirm the overwriting of existing files.

Figure 4.28: IDB Analyzer setup for example analysis with school data
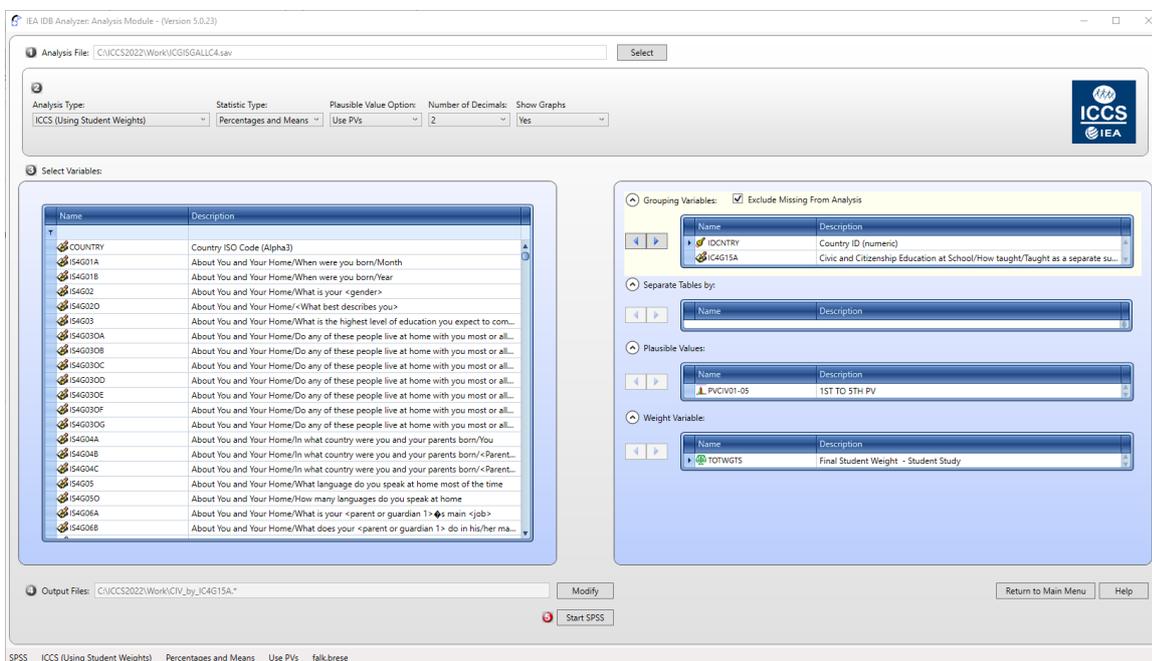
The results of this analysis are presented in Figure 4.28. In this example, each country's results are presented on two lines, one for each value of the IC4G15A variable. The results are presented in the same manner as in the previous examples, with countries identified in the first column and the second column describing the categories of IC4G15A (Yes/No). The third column shows the number of students with data in each category. Columns four and five provide the estimate for the number of students in the student population per category of the grouping variable IC4G15A and its standard error. The sixth and seventh columns represent the percentage of students for each one of the two categories of the IC4G15A selected by the principals and its respective standard error. The eighth and ninth columns represent the mean civic knowledge of the students for which the principals selected "Yes" or "No" and its standard error. Columns ten and eleven provide the standard deviation and its standard error of the mean civic knowledge within each category of IC4G15A. The last but one column contains the percentage of missing cases, and the last column shows the number of variance strata contributing cases to the corresponding group.

Figure 4.28 shows that 19.00 percent of target-grade students in Bulgaria attend schools where civic and citizenship education is taught as a separate subject, and 89.00 percent attend schools where this is not a separate subject. The standard errors of these percentages are 3.71 percent in both cases. We also see that in Bulgaria the estimated mean civic knowledge of students in schools with civic and citizenship education as a separate subject is 451.07 score points (with a standard error of 14.40 score points), whereas the estimated mean civic knowledge of students in schools where citizenship education is not taught as a separate subject is 457.10 score point (standard error of 5.81 score points).

Figure 4.29: Output for example analysis with school data

```
Average for PVCIV by IDCNTRY IC4G15A
PAGE    1

                   Civic and
                   Citizenship
                   Education at
                   School/How
                   taught/Taught as          Sum of
                   a separate    N of   Sum of   TOTWGTS          Percent   PVCIV    PVCIV              Std.Dev.
Country ID         subject       Cases  TOTWGTS  (s.e.)   Percent  (s.e.)   (Mean)   (s.e.)   Std.Dev   (s.e.)   pctmiss  nvarstrat
_____

Brazil             Yes           308    172835   53633,85  9,28    2,84    440,38   21,61    103,14   9,75     ,00      12
                   No            3755   1689503  121074,5  90,72   2,84    450,32   5,66     98,28    2,67     ,00      71

Bulgaria           Yes           557    10091    1992,34   19,00   3,71    451,04   14,40    98,36    5,25     ,00      25
                   No            2401   43019    2396,19   81,00   3,71    457,10   5,81     104,80   2,76     ,00      71

Chinese Taipei     Yes           3092   151344   6840,54   81,57   3,15    582,11   2,94     88,76    1,61     ,00      72
                   No            760    34194    5773,24   18,43   3,15    587,24   6,00     86,04    2,78     ,00      26

Colombia           Yes           571    107967   23928,02  16,90   3,90    454,59   11,90    88,72    5,69     ,00      21
                   No            2975   530869   42143,23  83,10   3,90    453,77   5,14     88,54    1,97     ,00      71

Croatia            Yes           165    1662     630,68    4,64    1,76    523,60   9,35     77,43    6,31     ,00      7
                   No            2543   34164    941,68    95,36   1,76    531,73   2,74     83,10    1,66     ,00      72

...


Table Average      Yes          .       .        .         40,56   ,85     511,76   2,36     95,44    1,02     .        .
                   No           .       .        .         59,44   ,85     505,08   1,58     94,40    ,66      .        .
```

When looking up the output file with the results of the significance test (filename ending "_Sig"), we see that the t-value of the difference in civic knowledge of students in Bulgaria at schools where civic education is taught as a separate subject and students at schools where it is not is 0.36, meaning that this difference in civic knowledge is not statistically significant.

4.8 Trend analyses

When performing trend analyses, it needs to be considered that the process of equating the tests across the cycles of ICCS introduces some additional error into the calculation of any test statistic. Therefore, this additional error, called equating error, needs to be added to the calculation of the standard error of any differences between results from different cycles. Please see also Chapter 13 in the ICCS 2022 technical report (Schulz et al., 2024b), which include more information on the reporting of differences.

In order to estimate the standard error (SE) of a trend statistic, the mean and its standard error needs to be estimated for each cycle. The trend estimate is the difference between the means $\mu 2022$ and $\mu 2016$. To estimate the standard error of the difference between results from ICCS 2022 and ICCS 2016, the two standard errors of the means need to be combined with the equating error:

$$SE(\mu_{ICCS22} - \mu_{ICCS16}) = \sqrt{SE_{(ICCS22)}^{2} + SE_{(ICCS16)}^{2} + EqErr_{2022\_2016}^{2}}$$

For many ICCS 2022 countries, it was also possible to compare results with the first ICCS cycle in 2009. When testing the difference of means between the first ($\mu2009$) and third ICCS survey, the standard error of the difference should include the equating errors between the cycles ICCS 2016 and 2022 as well between ICCS 2009 and 2016:

$$SE(\mu_{ICCS22} - \mu_{ICCS09}) = \sqrt{SE_{(ICCS22)}^{2} + SE_{(ICCS09)}^{2} + EqErr_{2022\_2016}^{2} + EqErr_{2016\_2009}^{2}}$$

Table 4.4 below provides an overview of the equating errors in the reporting metric for test and questionnaire scales (including regional scales) between ICCS 2016 and 2022, as well as between ICCS 2009 and 2016.

Table 4.4: Overview of the equating errors applied for the reporting of ICCS 2022 results

Scale	Equating Error 2022/2016 (in reporting metric)	Equating Error 2016/2009 (in reporting metric)
Civic Knowledge (PVs)	2.649	3.086
S_CITCON	0.240	0.411
S_CITEFF	0.202	0.261
S_CITSOC	0.169	0.428
S_CIVLRN	0.240	-
S_ELECPART	0.211	0.019
S_ETHRGHT	0.318	0.250
S_GENEQL	0.329	0.423
S_IMMRGHT	0.574	0.667
S_INTACT	0.413	-
S_INTRUST	0.155	0.291
S_ILLACT	0.133	-
S_LEGACT	0.168	-
S_OPDISC	0.212	0.340
S_POLDISC	0.127	0.438
S_POLPART	0.061	0.351
S_RELINF	0.115	0.115
S_SCACT	0.173	-
S_STUTREL	0.113	0.454
S_CITCON	0.240	0.411
S_CITEFF	0.202	0.261
S_CITSOC	0.169	0.428

The equating error should not be applied directly when testing the significance of differences between percentages at or above Level B or at or below Level D between ICCS 2022 and ICCS 2016 or ICCS 2009, respectively. Based on the equating errors of 2.649 (between 2016 and 2022) and 3.086 (between 2009 and 2016), for both ICCS 2016 and ICCS 2022 the international project team applied a tailored replication method to estimate

the equating error for percentages at or above a certain level of proficiency (see details in Schulz, 2018, 2024b). This procedure had to be applied for each national sample separately resulting in different estimates for each country or benchmarking entity.

When comparing percentages of students at or above Level B (or for those at or below Level D) between ICCS 2022 and ICCS 2016, within each participating country, using the calculated equating error based on the ICCS 2022 data (EqErr(p_cnt22)), the standard errors for the differences between percentages (at or above Level B, or at or below Level D) were calculated as:

$$SE(\rho_{ICCS22} - \rho_{ICCS16}) = \sqrt{SE_{(\rho_{ICCS22})}{}^2 + SE_{(\rho_{ICCS16})}{}^2 + EqErr_{(\rho\_cnt22)}{}^2}$$

When comparing percentages of students at or above Level B (or for those at or below Level D) between ICCS 2022 and ICCS 2009, for each participating country that participated in all three cycles, using the calculated equating error based on the ICCS 2022 data (EqErr(p_cnt22)) as well as those calculated based on ICCS 2016 data (EqErr(p_cnt16)), the standard errors for the differences between percentages (at or above Level B, or at or below Level D) were derived as:

$$SE(\rho_{ICCS22} - \rho_{ICCS09}) = \sqrt{SE_{(\rho_{ICCS22})}{}^2 + SE_{(\rho_{ICCS09})}{}^2 + EqErr_{(\rho\_cnt22)}{}^2 + EqErr_{(\rho\_cnt16)}{}^2}$$

Table 4.5 presents the country-specific equating errors that were estimated and should be used when comparing percentages at or above Level B or at or above Level D across ICCS cycles.

Table 4.5: Overview of the country-specific equating errors applied for the reporting of percentages across ICCS 2022 results

Country	At or above Level B		At or below Level D	
	2022/2016	2016/2009	2022/2016	2016/2009
Bulgaria	0.81	0.83	0.78	0.56
Chinese Taipei	0.51	0.70	0.17	0.19
Colombia	1.05	1.45	1.25	0.84
Croatia	0.90	1.16	0.35	0.36
Cyprus	0.79		0.90	
Estonia	0.77	1.08	0.31	0.25
Italy	0.83	1.12	0.52	0.40
Latvia	1.39	1.39	0.53	0.73
Lithuania	1.00	1.23	0.51	0.60
Malta	0.97	0.95	0.63	0.62
Netherlands	0.79	1.07	0.44	0.53
Norway	0.88	0.61	0.34	0.34
Poland	0.71		0.23	
Slovak Republic	0.93		0.82	
Slovenia	1.02	1.11	0.52	0.43
Spain	0.93		0.43	
Sweden	0.81	0.62	0.27	0.25
ICCS (pooled sample)	0.90	0.98	0.57	0.52

References

IBM Corp. (2019). IBM SPSS Statistics for Windows, Version 26.0. [Computer Software]. Armonk, NY: IBM Corp

IEA (2024). IDB Analyzer (Version 5.0.23) [Computer Software]. International Association for the Evaluation of Educational Achievement. Available online at http://www.iea.nl/data.

Posit Team (2023). RStudio: Integrated Development Environment for R. Posit Software. [Computer Software]. PBC, Boston, MA. URL http://www.posit.co/

R Core Team (2023) R: A language and environment for statistical computing. [Computer Software]. R Foundation for Statistical Computing, Vienna, Austria.

SAS Institute Inc. (2016). SAS® 9.4. [Computer Software]. Cary, NC: SAS Institute Inc.

Schulz, W., Ainley, J., Fraillon, J., Losito, B., Agrusti, G., Damiani, V., & Friedman, T. (2024a). *Education for Citizenship in Times of Global Challenge. IEA International Civic and Citizenship Education Study 2022 International Report.* International Association for the Evaluation of Educational Achievement (IEA).

Schulz, W., Friedman, T., Losito, B., & Fraillon, J. (Eds.) (2024b). *ICCS 2022 technical report.* International Association for the Evaluation of Educational Achievement (IEA).

The IEA's International Civic and Citizenship Education Study (ICCS) investigates the ways in which young people around the world are prepared to undertake their roles as citizens. ICCS 2022 is the third cycle of the study, which was initiated in 2009 and repeated in 2016. Based on data from 24 countries or benchmarking participants from Europe, Latin America, and Asia, ICCS 2022 studies contexts for and learning outcomes of civic and citizenship education in a wide range of national contexts at the beginning of the third decade of the 21st Century.

The ICCS 2022 user guide describes the content and format of the ICCS 2022 international database. It introduces the use of weighting and variance estimation variables for analyzing the ICCS 2022 data and provides a comprehensive overview of how to work with the IEA's International Database (IDB) Analyzer software. The ICCS 2022 user guide is accompanied by four appendices: the international versions of all questionnaires; an overview of national adaptations to the national versions of the ICCS 2022 international questionnaires; derived variables used in the ICCS 2022 international and regional reporting; and a set of restricted-use civic knowledge items, together with their respective scoring guides, to illustrate the test contents.

Over the past 50 years, the IEA has conducted comparative research studies in a range of domains focusing on educational policies, practices, and outcomes in many countries around the world. Prior to ICCS 2022, the IEA had conducted four international comparative studies of civic and citizenship education: The first IEA Civic Education study in 1971, the second IEA Civic Education Study (CIVED) in 1999, the first cycle of ICCS in 2009 and the second cycle in 2016. ICCS 2022 data will allow education systems to evaluate the strengths of educational policies, from a comparative perspective, and to measure progress in achieving critical social objectives of their educational policy agendas.

IEA