**Documentation of statistics for
Teacher-student register for primary school 2022**

1 Introduction

The Matched Educational Data for primary and lower secondary education level (MED-Elementary) serves the purpose of enriching the elementary school statistics with information about the individual lessons. The data registry has been developed for research purposes and provides the possibility to study who receives teaching together and who executes the teaching. So far, the register covers the school year of 2020/2021 and 2021/2022 .

2 Statistical presentation

The database, MED-Elementary, contains information about students, groups, teachers and their activities in public elementary schools and other schools at the level of primary to lower secondary education in Denmark.

The registry covers students from 0th grade to 10th grade, the groups said students are attached to, the employees who execute the teaching and pedagogical support, as well as the activities and lessons, that the students and employees participate in.

In its current form, the registry covers from the school year of 2020/21 and are updated yearly with the addition of another year.

2.1 Data description

MED-Elementary is a unique database on global scale, that creates possibilities for research in teaching at the level of the individual lesson in Danish public elementary schools (primary and lower secondary).

The database contains information about elementary school students, their association to institutions, groups, and grades with corresponding start- and end dates. It also contains information about employees, their association to institutions, and their main role at the institution with corresponding start- and end dates. In addition, there is information about the groups, the type of group, and the creation- and closing date of the group. Last but not least, there is information about the activities in the schools in the form of a description of the activity, the start- and end times, and associations to groups and employees. The students are arranged in institutions and groups, where the groups can be the main class, teams for electoral subjects and similar. All students have exactly one main group pr. institution they are associated to, and that will typically be their main class.

The registry is organized in four fundamental tables - student, employee, group, lesson - and three relational tables that connect the four units to each other. The students are coupled together in groups. The groups and teachers are both coupled to the lessons. That means that teachers can be coupled to students via the lessons and groups.

2.2 Classification system

Education activities in the elementary school are divided according to grade and follow DISCED, which is Statistics Denmark's classification of educations. DISCED is the Danish version of the international classification system ISCED.

2.3 Sector coverage

Not relevant for this statistics.

2.4 Statistical concepts and definitions

Primary school: Teaching for children of the compulsory age

2.5 Statistical unit

The statistic units are activities, groups, and persons, where persons include students, teachers, and other employees.

2.6 Statistical population

The population consists partly of students, that have been enrolled in an elementary school (0th to 10th grade) i Denmark and partly of the personnel, who are attached to the schools.

2.7 Reference area

Denmark

2.8 Time coverage

The school year 2020/21-. There is only information about students and groups from Unilogin from September 1st 2020.

2.9 Base period

Not relevant for this statistics.

2.10 Unit of measure

The statistical units are persons, activities/lessons, and groups.

2.11 Reference period

The reference time is the school year from 1 August to 31 July.

2.12 Frequency of dissemination

Yearly.

2.13 Legal acts and other agreements

The Act on Statistics Denmark §6.

2.14 Cost and burden

The burden on respondents is minimal, since the information is typed into the systems for other purposes.

2.15 Comment

Further information can be obtained by inquiry to Statistics Denmark.

3 Statistical processing

The data for this registry is collected weekly from the schools' digital login system, Unilogin, and the communication platform, AULA, as automatic system-to-system reporting. The weekly updates are compiled into chronological tracks, which undergo a broad validation process followed by and enrichment of the data, also on a weekly basis. Once a year data from the two sources are integrated and combined with data from the Student Registry.

3.1 Source data

The data sources for MED-Elementary are the two administrative systems, AULA and Unilogin, which are used in Danish public schools, and also Statistics Denmark's Student Registry. Unilogin is a digital ID, which all students and teachers in the elementary schools are assigned. It can follow a student all the way from daycare through high school. Data from Unilogin contains information about the persons associations to an institution, grade, and groups. AULA is a digital communication platform, which provides information about scheduled activities, and which groups and employees are associated to the activities. The Student Registry, which is an existing registry over all students in the ordinary educational system, is used to consolidate the student population in MED. It is required that all student records are consistent with the Student Registry on institution, period and grade level.

3.2 Frequency of data collection

The statistics are based on data collected weekly from AULA and Unilogin.

3.3 Data collection

Data is collected via system-to-system report from AULA and Unilogin.

3.4 Data validation

When data is received by Statistics Denmark, they go through an overall quality assurance process, where improbable fluctuations in the number of reported institutions, groups, students, employees and activities are searched for. Various types of potential errors are counted and reported, such as invalid CPR numbers, overlapping student courses, students without a group and employees without a user ID. Finally, calculations are made of how many of the mandatory activities that are associated with teachers and students, how many activities lie outside the normal school day, and the distribution of the length of the activities. The statements are used to monitor data quality, so that the schools or system suppliers can be contacted in the event of unexplained fluctuations. The ongoing monitoring of the data flow also contributes to improving and testing the automated data processing, which is described in section 3.05.

3.5 Data compilation

Information from the two data sources is integrated with the Student Register once a year and then published.

In AULA the activities are only described by a free text field, which may or may not refer to an ordinary school subject. At Statistics Denmark, the reported activity descriptions are translated via a text recognition algorithm into one of 28 mandatory subjects, which are described in the Proclamation of the Law on Basic Schools. If the activity cannot be translated into one of the 28 subject, it is labeled "Other subject".

The weekly obtained data contain information about the student population, the teacher population, the groups and their schedules as they appear on the date when the data is retrieved. The data is then added to all previous weeks' data, such that a chronology is built, where duplicate information is deleted, changes to the existing information are corrected, and tracks are merged or terminated.

Once a year, the information from Unilogin and AULA are integrated and subjected to some logical rules for their mutual relations. For example, a teacher cannot be listed as teaching a lesson, if the teacher is not associated with the institution on the date of the lesson. The student population is then replaced with the primary school students from the Student Register, which is then supplemented with information and group relations from Unilogin to the extent that there is consistency between the two sources in terms of institution, grade level, start and end dates.

3.6 Adjustment

No corrections are made to data other than what is described under Data Validation and Data Processing

4 Relevance

MED-Elementary has been developed for research purposes and creates opportunities for research in teaching at the level of the individual lessons in primary and lower secondary schools. This applies, for example, to the importance of the composition and length of the school day. In the development of MED-Elementary, external stakeholders and future users have been continuously involved.

4.1 User Needs

The users of MED-Elementary are expected to be primarily researchers, analysts and public authorities, with an interest in elementary school education. In connection with, for example, the Primary School Grade Registry or the Transition Registry, MED-Elementary can be used to analyze which factors have an influence on students' grades, further education, dropouts, etc. Such analyses can, among other things, form the basis for policy development in the field of education and administrative decisions at schools.

4.2 User Satisfaction

In the development of MED-Elementary, external stakeholders and future users have been continuously involved. Through meetings with the project's research team, the potential users' interests and preferences in relation to quality, coverage and display of data have helped to shape the final register. Several of the potential users of the registry have also had access to temporary versions of the registry in order to test the usability of the registry in relation to the size of the data sets and the computational performance.

4.3 Data completeness rate

Not relevant for this statistics.

5 Accuracy and reliability

The information in the database comes from the institutions' administrative systems and good agreement has been found with existing registers based on other data sources.

By translating the activity designations into a mandatory subject, the number of hours at each individual school is compared to the minimum requirement or the indicative number of hours according to the executive order on elementary schools. This comparison gives an impression of how well each institution register their activities. It is estimated that in 84 pct. of primary schools, students have on average between 85 pct. and 115 pct. of the recommended number of hours/minimum hours in the executive order in the school year 2021-2022.

5.1 Overall accuracy

There is good agreement between the educational information about the students from Unilogin and the student in the Student Register, which undergoes an extensive quality assurance in connection with the annual data collections. Therefore, the accuracy of information about the students is assessed to be good. Likewise, the precision of information about the groups and teachers is judged to be in reasonable agreement with independent registers to the extent that they can be compared.

The activities cannot be directly verified against other sources. There are 2 conditions in particular which mean that there is uncertainty linked to the activities. The activities are described in a free text field, which Statistics Denmark attempts to translate into one of the mandatory school subjects and, if this is not possible, into the category "other subject" and errors may occur in this translation. Mandatory school subjects can be mistakenly translated to "other subject" because the text cannot be recognized and they can also be translated into an incorrect mandatory subject. Furthermore, not all schools register all their activities in AULA. Some schools are missing most activities, in other schools only some groups lack activities or some special activities are missing.

5.2 Sampling error

Not relevant for this statistics.

5.3 Non-sampling error

The degree to which activity data in MED-Elementary covers the activities that the students actually had, depends to a large extent on the schools' registration practices. Via AULA, the most recently updated schedule data is collected on a weekly basis. The AULA system is used administratively at the schools and contains the activities that students, parents and teachers see in their timetable. It therefore gives the best possible picture of the actual school day, but the completeness depends on the individual recording practices of the school.

A lack of coverage of students can also be due to a lack of reports from a supplier.

All students in the Danish school system have a Unilogin, but it is primarily the public elementary schools that use the AULA system. Therefore, it is primarily public schools that are represented with information about lessons in MED-Elementary.

5.4 Quality management

Statistics Denmark follows the recommendations on organisation and management of quality given in the Code of Practice for European Statistics (CoP) and the implementation guidelines given in the Quality Assurance Framework of the European Statistical System (QAF). A Working Group on Quality and a central quality assurance function have been established to continuously carry through control of products and processes.

5.5 Quality assurance

Statistics Denmark follows the principles in the Code of Practice for European Statistics (CoP) and uses the Quality Assurance Framework of the European Statistical System (QAF) for the implementation of the principles. This involves continuous decentralized and central control of products and processes based on documentation following international standards. The central quality assurance function reports to the Working Group on Quality. Reports include suggestions for improvement that are assessed, decided and subsequently implemented.

5.6 Quality assessment

The quality of data has been assessed through a series of analyses and comparisons with independent data sets.

The student population is based on Statistics Denmark's Student Registry, to which relationships are formed with groups for the students found in Unilogin. When comparing the students in the two data sources, a match percentage of grade level and institution number of between 99.2 pct. and 97.4 pct. has been found for grades 0-9 in the school year 2020-2021.

The number of employees has been assessed by comparing Unilogin and AULA data with the register of Employment Statistic for Employees (BFL). The study shows that 98 pct. of the employees who are associated with activities are employed during their entire career at the institution. If you look at which work functions the employees in question are classified under in BFL, you can see a similar distribution of roles for employees in MED-Elementary. In addition, there is reasonable agreement between the number of teachers associated with activities in MED-Elementary and STIL's inventory of the number of teachers in public basic schools (Competence data).

The groups in MED-Elementary are assessed by comparing the students' main groups with the students' class designation in the Basic School Registry. The study shows that the vast majority of the groups have 100 pct. agreement among associated students in the two registries,, and the second largest part of the groups has 91-99 pct. agreement, corresponding to one or two students who deviate per group. Less than 2 pct. of the groups have less than 90 pct. agreement.

The three analyses are documented in internal reports, but an overall summary of the results can be found here: [Notat om kvalitet 2020/2021]
(https://www.dst.dk/ext/46793106127/0/uddannelse/Notat-om-kvalitet-- pdf)

Activity data cannot be verified by any independent data source. By translating the activity designations into a mandatory subject, the number of hours at each individual school is compared to the minimum requirement or the indicative number of hours in the primary school executive order. This comparison gives an impression of how good the registration practice is at each school. It is estimated that in 84 pct. of primary schools, students have on average between 85 pct. and 115 pct. of the recommended number of hours/minimum hours in the executive order in the schoolyear 2021-2022.

For a detailed description of the statements for the individual schools, see the documentation: Note on coverage

5.7 Data revision - policy

Statistics Denmark revises published figures in accordance with the Revision Policy for Statistics Denmark. The common procedures and principles of the Revision Policy are for some statistics supplemented by a specific revision practice.

5.8 Data revision practice

There are no ongoing revisions of the annual versions of MED-Elementary.

6 Timeliness and punctuality

It is expected that the register will be published in April for the previous school year without significant delays.

6.1 Timeliness and time lag - final results

It is expected that the register will be published once a year in April. The published data relates exclusively to the previous school year and the information is final upon publication.

6.2 Punctuality

The data base is still under development, but the expectation is that it will be published on a fixed date each year without delays.

7 Comparability

The MED-Elementary is currently available for the school year 2020/2021- .

7.1 Comparability - geographical

There are no similar statistics in other countries.

7.2 Comparability over time

The MED-Elementary is available for the school year 2020/2021- .

7.3 Coherence - cross domain

The student population in MED-Elementary is obtained from Statistic Denmark's Student Registry, but there will be discrepancies with previous years' student populations, as revisions in the Student Registry are not incorporated into EMD-Elementary.

7.4 Coherence - internal

In the data processing, consistency is ensured between students, groups, activities and teachers through the formation of three relationship tables. Relationships are only formed between two entities in cases where there is temporal consistency.

8 Accessibility and clarity

Researchers and other analysts from authorized research institutions can gain access to the micro data through Statistics Denmark's Research Scheme, and the ministries will be able to gain access through the ministerial scheme.

8.1 Release calendar

The publication date appears in the release calendar. The date is confirmed in the weeks before.

8.3 User access

Statistics are always published at 8:00 a.m. at the day announced in the release calendar. No one outside of Statistics Denmark can access the statistics before they are published.

8.2 Release calendar access

The Release Calender can be accessed on our English website: Release Calender.

8.4 News release

Statistics Denmark do not produce publications based on the MED-Elementary.

8.5 Publications

No publications are published about MED-Elementary.

.

8.6 On-line database

There are no tables in the Stat Bank.

8.7 Micro-data access

Researchers and other analysts from authorized research institutions can gain access to the micro-data through Statistics Denmark's Research Scheme, and the ministries will be able to gain access through the ministerial scheme.

8.8 Other

LER-Grundskole is not available in other ways.

8.9 Confidentiality - policy

See Statistics Denmark's data privacy policy is followed: Dataprivacy

8.10 Confidentiality - data treatment

No tables are published in the Statbank.

8.11 Documentation on methodology

Report on quality Report on coverage

8.12 Quality documentation

Results from the quality evaluation of products and selected processes are available in detail for each statistics and in summary reports for the Working Group on Quality.

9 Contact

Administratively, the statistics are located in the Population and Education office. The person in charge of statistics is Eva Lotti Hansen, tel.: 3917 3086, e-mail: ebr@dst.dk

9.1 Contact organisation

Statistics Denmark

9.2 Contact organisation unit

Population and Education, Education

9.3 Contact name

Eva Lotti Hansen

9.4 Contact person function

Responsible for the statistics

9.5 Contact mail address

Sejrøgade 11, 2100 Copenhagen

9.6 Contact email address

ebr@dst.dk

9.7 Contact phone number

3917 3086

9.8 Contact fax number

N/A