

## **Documentation of statistics for Highest Education Attained 2024**

## 1 Introduction

The purpose of the statistics on educational attainment is to give an overall statistical description of the educational level of the population at any given time. The primary data source to these statistics is the Student Register with data from 1970 onwards. In addition, the Qualification Register is used. Since the Student Register is the primary source for information, the Attainment Register gives nearly complete coverage from 1970 onwards. There is, however information before this time coming from The Qualification Register.

## 2 Statistical presentation

The Attainment Register gathers information about the highest completed education for each single person based on the information in the Student Register and the Qualification Register. It is a longitudinal register based on an assessment of each person's education "career" and shows how the qualification career develops over time. The register is formed by interpreting the qualification career (skills in chronological order) in order to determine a change in the skill level. Once a year a status register is also produced with the population and information about education the 30. September the current year.

### 2.1 Data description

Variables in the register of the highest educational attainment:

- Person number (PERSON\_ID)
- Starting data for the highest level of education (HF\_VFRA)
- Ending data for the highest level of education (HF\_VTIL) .
- Institution (INSTNR)
- Highest attained education (HF\_AUDD)
- Sources for the highest attained education (HF\_KILDE)

For further definitions of the variables see the documentation in TIMES.

The register is available in three other versions with a similar content.

1. Highest attained youth education
2. Highest attained professional education
3. Latest completed education

### 2.2 Classification system

In connection with publications of the highest level of education Statistic Denmark classification system, "DISCED" is used. A description of it can be found at: [DISCED](#)

### 2.3 Sector coverage

The population consist of people who are registered with a qualification in The Student Register or in The Qualification Register.

## **2.4 Statistical concepts and definitions**

Highest education completed: The highest education that a person has completed at any given point in time. The highest completed education is defined in relation to the main groups in the Statistics Denmark's educational nomenclature (ISCED) and follows the hierarchy herein. In the statistics, the highest education is assessed yearly per 30th of September.

## **2.5 Statistical unit**

The statistical entity is completed education qualifications.

## **2.6 Statistical population**

The population consist of people who are registered with a qualification in Statistic Denmark's Student Register or in Statistic Denmark's Qualification register. The collection of information about PhD students covers the period from 1. January till 31. December. Since 2012 it has not been possible to have data about PhD student ready together with data from the other education area because the publication of The Student Register has been advanced. When The Attainment Register is updated, the information concerning PhD is therefore a year behind.

## **2.7 Reference area**

Denmark

## **2.8 Time coverage**

Since the student register is the primary source for the highest attained education, The Attainment Register covers fully from the beginning of 1970 which is where the information in the student register starts, although there are many information before 1970 from other sources.

## **2.9 Base period**

Not relevant for this register.

## **2.10 Unit of measure**

The unit of measure is number of people.

## **2.11 Reference period**

The reference time is 30. September the current year in both the longitudinal register and the status register.

## **2.12 Frequency of dissemination**

There is an annual update of both the longitudinal register and the status register.

## 2.13 Legal acts and other agreements

Law concerning Statistics Denmark § 6

## 2.14 Cost and burden

The registers are based on registers and therefore there is no respondent burden

## 2.15 Comment

Further information can be found on the topic site: [Befolkningens uddannelsesstatus](#).

The main contact person is Eva Lotti Hansen, phone.: (+45) 39 17 30 86, email: ebr@dst.dk

## 3 Statistical processing

The Attainment Register is a longitudinal register based on an assessment of each person's education career in The Student Register and The Qualification Register. It shows how the qualification career develops over time, and it is updated once a year. The status register is produced on the basis of the longitudinal register and contains information about the population and their highest completed education per September 30 of the given year.

### 3.1 Source data

The primary source to produce the Attainment Register is the Student Register. The Student Register is supplemented with additional data from Statistics Denmark's Qualification Register that collects data from multiple sources of different nature.

When the Attainment Register is produced, the source of the information is taken into account and categorized according to how reliable the information is assumed to be, and given a priority (1 to 4). Information coming from the source of highest priority will always be used in the attainment register, even if there is information about a higher qualification coming from a source of lower priority.

Sources with priority 1:

- Examinations achieved through adult education (HF\_KILDE = 1) Qualification from the adult education system are updated annually.
- IDA Engineering association (HF\_KILDE = 5) Statistics Denmark received in 2005 the members of the engineering association.
- Ministry of Health (HF\_KILDE = 8) Since 1980, The Authorisation Register from Sundhedsdatastyrelsen have been used as a source to The Attainment Register.
- Danish Agency for International Education (HF\_KILDE = 11) Since 2007 Statistics Denmark have received information about education that has been assessed by the Agency for international Education.
- Educational Information from Greenland (HF\_KILDE = 12) Once a year, Statistics Denmark receive information about people who have completed an education and received grants from Greenland under their education. The information from Greenland is not ready for the last

year when the register of highest completed education is formed and the information from this source is therefore missing in the last year.

- Maritime Agency (HF\_KILDE = 13) In 2010 Statistic Denmark received information from the Maritime Agency's administrative system about completed maritime education.
- Ph.d. (HF\_KILDE = 14) Statistics Denmark's Ph.D. register contains information about the admission course to the Ph.D. education. The information from Greenland is not ready for the past year when the register of highest completed education is formed and the information from this source is therefore missing in the past year.
- The Student Register (HF\_KILDE = 15)

Sources with priority 2:

- Population & Housing Census 1970 (HF\_KILDE = 2). Population & Housing Census 1970 consist of self reported education information and the quality is considered to be less good. The date of completion on the educations from the Population & Housing Census 1970 is all set to 1-1-1970.
- Immigrants education (IVU) (HF\_KILDE = 3, 17) Self reported education information for people who have immigrated to Denmark before 2007 have source = 3, from 2007. The completion date is set to the date of first immigration to Denmark (added 1-1-1600 if date is missing).

Sources with priority 3: - The admission education (HF\_KILDE = 19) People that Statistics Denmark don't have a educational information about that have immigrated to Denmark and subsequently have started on a education in the student register, receive the admission education for this education, as their highest completed education.

Sources with priority 4: - Immigrants education (IVU) (HF\_KILDE = 9, 18) Information on immigrants' education is imputed when it hasn't been possible to obtain educational information in another way. Imputed values is information that replace unstated values based on experience from similar people. The imputed data is useful in overall statistical statements, but cannot be considered as valid information on individuals' educational attainment. The imputed educational information from immigrants is coded with broad educational codes, ie codes that simply indicate a certain level of education

The completion date is set to the date of first immigration to Denmark (added 1-1-1600 if date is missing).

With the inclusion of new sources, the highest level of completed education is also influenced back in time, depending on how far back the new data covers. This is important when compared to older versions of the attainment register (PSD\_HFUDD) and causes data breaches in the status register (PSD\_Udd\_STATUS), which are not changed in previous years' versions.

When the register of highest completed education is updated, there are two sources that are not ready with data for the last year, Ph.D. (HF\_KILDE = 14) and Greenland (HF\_KILDE = 12). Information from these two sources is therefore not included in the last year of the register.

### 3.2 Frequency of data collection

The register is based on information that are gathered once a year.

### 3.3 Data collection

The register is based on information from Statistic Denmark's two registers, The Student Register and The Qualification Register

### 3.4 Data validation

Data editing is minimal since the register is based on information from other registries, that already has been true a comprehensive error control.

There is a check for: - The development of the main groups for the highest attained education over time. - The development of the use of sources over time.

- if there is a record that is temporarily followed by a new record with lower priority.
- if there are two records with the same priority, where the temporal first record has a higher level than the next record.
- if there are two records with the same priority and the same main group level, where the temporal first record has a higher PRIA than the next record.
- if there are two records with the same AUDD, where the temporary first record has the student register as source and the next not have the student register as source.

### 3.5 Data compilation

Based on The Student Register and The Qualification Register, an assessment is made of each person's educational career in order to determine the person's highest completed education.

Higher education is chosen according to the following principle: 1. Highest priority 2. If there are several AUDD here, highest main group, with main groups 20 and 35 being equated 3. If there are several AUDD here, highest PRIA 4. If there are several AUDD here, the recently completed. 5. If there are 2 records with the same AUDD, the register is only updated if no HF\_KILDE = 15 already exists.

When the register of highest completed education is produced, only completed qualifying courses are included with KOMP = 1 (qualifying general education), 3 (vocational qualifying education) or 7 (secondary education in higher education). The exception to this is education / grade level at the primary school level, where completed grades from grade 6 and up are all included regardless of the value of the variable KOMP and AUDD = 0001, which means that you have never gone to school.

Similarly, a dataset with the highest completed general education (main areas 10, 20 and 35) and a dataset with the highest completed professional qualifying education (main areas 30, 40, 50, 60, 70 and 80) is produced.

Finally, a data set with the latest competency-giving education that a person has completed at a given time is produced. Latest completed education is chosen according to the following principle:

1. Recent completed education.
2. If two completed education have the same time of departure, the first priority sources are selected
3. If two completed education have the same departure time and the same source priority, the highest DISCED level is selected.
4. If two completed education have the same departure time, same source priority and same DISCED level, the highest PRIA is selected.

### **3.6 Adjustment**

There is no correction of the data beyond what has already been described under data validation and data processing.

## **4 Relevance**

There is a great variety of users. The information is generally used in connection with describing the population or sections hereof. The register is used in connection with status reports for other statistical fields. Data reports are thus submitted for (mainly on the population's highest level of education completed) a wide number of integration registers operated by Statistics Denmark. Furthermore, the register is frequently used in connection with external service activities order by Danish ministries, municipalities, research institutions, professional organization, private enterprises, private individuals and, not least, requests made by the news media.

### **4.1 User Needs**

There is a great variety of users. The information is generally used in connection with describing the population or sections hereof. The register is used in connection with status reports for other statistical fields. Data reports are thus submitted for (mainly on the population's highest level of education completed) a wide number of integration registers operated by Statistics Denmark. Furthermore, the register is frequently used in connection with external service activities order by Danish ministries, municipalities, research institutions, professional organisations, private enterprises, private individuals and, not least, requests made by the news media.

### **4.2 User Satisfaction**

In working with the education statistics the users are involved as much as possible. Some users can help secure that the data from the schools have a high quality. Other users can provide inspiration for new statistics and analysis. Finally, we try to get inspiration for the development of methods and models that can contribute to the continued development and improvement of statistics through users.

### **4.3 Data completeness rate**

When the register of the educational attainment level is produced, Denmark deviates from the OECD rules on two points. 1. It is the education levels in the Danish classification of education DISCED and not ISCED that is used when the individual qualifications are being assessed in relation to each other. 2. Educational attainment is classified on the basis of the characteristics of the education program at the time the register is produced and not at the time the program is completed.

When the register of highest completed education is updated, there are two sources that are not ready with data for the last year, Ph.D. (HF\_KILDE = 14) and Greenland (HF\_KILDE = 12). Information from these two sources is therefore not included in the last year of the register.

## **5 Accuracy and reliability**

The Accuracy and reliability vary depending on the source of information. More than 80 pct. of the information comes from administrative sources, such as student systems of educational institutions, which are highly reliable. These sources have priority one when the registry is created and will be used if there is information from one of these sources. Other sources are not so closely linked to the education programs and will often be less reliable. Examples of these sources are the surveys of immigrants' education and the population and housing census in 1970, based on self-reported education. In addition, information is imputed for persons who do not respond in the study of immigrants' education. The imputed data is useful in overall statistical statements, but cannot be considered as valid information on individuals' educational attainment.

In connection with the annual reports from the education institutions there is information which also relate to previous years. These delayed notifications concern particularly the last year.

### **5.1 Overall accuracy**

In connection with the annual reports from the education institutions there is information which also relate to previous years. These delayed notifications concern particularly the last year. See section 5.08.

### **5.2 Sampling error**

Not relevant for these statistics.

### **5.3 Non-sampling error**

More than 80 pct. of the information comes from administrative sources, such as student systems of educational institutions, which are highly reliable. Other sources are not so closely linked to the education programs and will often be less reliable. Examples of these sources are the surveys of immigrants' education and the population and housing census in 1970, based on self-reported education.

Over the years, several studies have been conducted on immigrants' education at arrival in Denmark. The information about immigrant consist of both self-reported and imputed values. The imputed values can not be regarded as valid information about the educational level of individuals as they result from a statistical model based on other people's self-reported education and characteristics, including country of origin. Because the statistical model used for imputation is uncertain, the imputed information is more uncertain than the self-reported education information. A description of the studies of immigrants' education at arrival can be seen here: The paper [Indvandrenerne medbragte uddannelse \(2024\)](#) describes in detail the various studies that have been carried out over the years to provide information on the education brought by the immigrants.

### **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**

In connection with the production of the register, a number of sources are used which have respectively priority 1, 2, 3 and 4.

Sources with priority 1 are information from administrative sources, eg the student register, where the quality is assessed to be high. These sources account for more than 80 per cent. of the educational information. Sources with priority 2 are self-answered educational information from the Immigration Survey and the Folke & Housing Census, where the quality is considered to be less good. The source of priority 3 is educational information based on an ongoing education in the student register. People with unknown highest completed education that have immigrated to Denmark and started on an education in the student register get the education that is required to start on the education, as their highest completed education. For most programs, admission requirements are clear but there are some artistic programs where it is more uncertain. The source of priority 4 is imputed educational information for people for whom it has not been possible to obtain information in connection with the study of immigrants' education. Imputed values is information that replace unstated values based on experience from similar people. . The imputed data is useful in overall statistical statements, but cannot be considered as valid information on individuals' educational attainment.

The table below shows the use of sources

## **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**

In connection with the annual reports of new educational information, some information goes back in time, especially the previous year. Figures for the past year must therefore be considered provisional.

## **6 Timeliness and punctuality**

The statistics are published around 6 months after the end of the reference time. The statistics are usually published without delay in relation to the scheduled date.

## **6.1 Timeliness and time lag - final results**

The statistics are published once a year. In connection with next year's publication, the figures may change. The statistics are primarily based on longitudinal register, where changes can be made on an ongoing basis. Partly because some information comes late, and partly because the statistics will change if data errors are detected.

## **6.2 Punctuality**

The statistics are usually published without delay in relation to the scheduled date.

## **7 Comparability**

The longitudinal register is produced once a year and the entire period is thus calculated in the same way. Based on the longitudinal register, a status register is produced with the population per September 30 that year. In the event of significant changes in the way the longitudinal register is produced, the status registers for all years will be reproduced. It happens that an education changes level from one year to the next. Typically, this will not cause a reproduction of all the status registers and therefore affect comparability over time. Labor force survey provide information too Eurostat about the educational attainment level and this is these figures that are used for international comparison of the attainment level.

### **7.1 Comparability - geographical**

Labor force survey provide information too Eurostat about the attainment level and this is these figures that are used for international comparison of the educational attainment level.

### **7.2 Comparability over time**

The longitudinal register is produced once a year and the entire period is thus calculated in the same way. Based on the longitudinal register, a status register is produced with the population per September 30 that year. In the event of significant changes in the way the longitudinal register is produced, the status registers for all years will be reproduced. It happens that an education changes level from one year to the next. Typically, this will not cause a reproduction of all the status registers and therefore affect comparability over time

### **7.3 Coherence - cross domain**

There is no other Danish statistics on this subject.

### **7.4 Coherence - internal**

Highest completed education is produced on the basis of the student register and the qualification register. The information in the student register comes from the educational institutions administrative system and is considered to be of a very high quality. The qualification register contains information from a variety of sources of different quality. See section 3.08.

## **8 Accessibility and clarity**

Statistics are published once a year in "News from Statistics Denmark" . At the same time data are released in the Statbank and on the homepage: [Homepage](#) Information also appears in the annual publications Statistical 10-Year Review and the Statistical Yearbook.

### **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 Calendar can be accessed on our English website: [Release Calendar](#).

### **8.4 News release**

New releases of "News from Statistics Denmark" can be found on the homepage: [Homepage](#)

### **8.5 Publications**

Statistical 10-Year Review: [Statistical 10-Year Review](#)

### **8.6 On-line database**

Data are released in the Statbank under the subject Education and Knowledge, Population by Statute of Education, Educational Attainment: [Statbank](#)

### **8.7 Micro-data access**

The Attainment Register is a longitudinal register and is stored in Statistic Denmark's Database PSD. Once a year ( since 2005) a status register is produced with the population and information about education the 30. September the current year. These status register are also stored in Statistic Denmark's Database PSD.

### **8.8 Other**

Not relevant for this register.

### **8.9 Confidentiality - policy**

In connection with publication from the register Statistics Denmark data privacy policy is followed, see: [Data privacy policy](#)

### **8.10 Confidentiality - data treatment**

In connection with publication from the register Statistics Denmark's data privacy policy is followed, see: [Data privacy policy](#)

### **8.11 Documentation on methodology**

The paper [Indvandrenes medbragte uddannelse \(2024\)](#) describes in detail the various studies that have been carried out over the years to provide information on the education brought by the immigrants.

### **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**

The administrative placement of these statistics is in the division of Population and Education, Social Statistics. The contact person is Alexander Pfeiffer Cappelen, tel.: + 45 2159 9906, and e-mail: [APF@dst.dk](mailto:APF@dst.dk).