

The Office of the Auditor General: Analyzing and visualizing data from Norwegian schools

About the Office of the Auditor General

The Office of the Auditor General (OAG) is a control body under the Norwegian Parliament, the Storting. Our main mission is to examine how the central government does its job by auditing the state's spending, productivity and goal achievement. We do this through investigations of the central government accounts, examining whether money allocated by parliament is used efficiently and controlling state companies' follow-up of laws and regulations.

In 2017, the Office of the Auditor General established a specialized section for data science, the Office of the Auditor General's *Innovation Lab*. The unit currently consists of seven people with backgrounds in both social and natural science. An important part of our work is to develop applications and tools for collecting, distributing and visualizing data, as well as making use of new analytical methods and technology.

Problem statement

The Norwegian school system is divided into primary, lower secondary and higher secondary education. The first two are mandatory and generally known together as *Grunnskole* (GSK). All students also have a legal right to higher secondary education, *Videregående Opplæring* (VGO), where they can choose between general studies (*studieforberedelse*) or vocational studies (*yrkesfag*).

In strict terms the responsibility for primary and lower secondary education belongs to the municipalities, while the responsibility for higher secondary education rests at the county level. However, the broader goals of the educational system are national and set by parliament.

Indeed, the quality of the school system is a central topic of political debate in Norway. This includes questions regarding the number of teachers per students, significant differences in student performance between schools and more efficient use of resources. We at the OAG would therefore like to know more about differences in quality, efficiency and productivity in Norwegian schools.

As a first step, you will use an R package developed by OAG to explore data from the [National School Registry](#), and combine this with information from the [Central Coordinating Register for Legal Entities](#) as well as geographical information from [The Norwegian Mapping Authority](#).

Furthermore, we will provide you with two datasets containing data we put together from the [Norwegian Directorate for Education and Training](#). These include information on educational inputs (teachers, salaries, expenditure), outputs (number of students) and results (grades, dropouts, absence from school). There is one set of tables for GSK schools and another set of tables for VGO schools.

It will be your job to combine, analyze and visualize the data. Communication of your findings will be central to the quality of your work.

Assignment

1. Install and explore the [noAPI](#) package. Put together a dataset of all the currently active VGO schools in Vestland County that includes information on the number of employees and geographic coordinates. Create an interactive map of the schools.
2. Familiarize yourself with the provided datasets for GSK and VGO schools and the [education statistics](#) provided by The Norwegian Directorate for Education and Training. Combine the datasets so that it is easier to make comparisons.
3. Create an interactive report (e.g. using [Quarto](#)) or a simple dashboard (e.g. with [R Shiny](#)) to visualize and analyze the data. Possible topics to explore include are:
 - Qualitative differences between schools and changes over time.
 - Aggregated comparisons of higher-level decision-making units, i.e. municipalities and counties.
 - Comparisons of inputs (e.g. expenditure, teachers), outputs (e.g. number of students) and results (e.g. grades).
 - Geographic visualizations
 - Handling of missing data
4. Create an unsupervised model that clusters schools based on the performance and resource metrics provided. The purpose is to identify groups of comparable schools where similar political measures might be taken. You should preferably explore different clustering methods.

Resources

Wikipedia:

https://en.wikipedia.org/wiki/Education_in_Norway

The Norwegian Directorate for Education and Training

<https://www.udir.no/tall-og-forskning/statistikk/>

Statistics Norway:

<https://www.ssb.no/utdanning/grunnskoler>

<https://www.ssb.no/utdanning/videregaende-utdanning>

<https://www.ssb.no/utdanning/artikler-og-publikasjoner/er-det-forskjeller-i-skolers-og-kommuners-bidrag-til-elevenes-laering-i-grunnskolen>

R packages

<https://riksrevisjonen.github.io/noAPI/>

<https://r-spatial.github.io/sf/>

<https://r-tmap.github.io/tmap/>

<https://rstudio.github.io/bslib/>

<https://github.com/amices/mice>

Contact information

Aleksander Eilertsen Valberg

ale@riksrevisjonen.no