

GEOSTAT 4:

A Business Case for Geospatial Statistics
as a service (GSaaS)

Geostat 4

Title: GEOSTAT 4: A Business Case for Geospatial Statistics as a service (GSaaS)

Project: Eurostat ESSnet grant project GEOSTAT 4

Grant agreement number: 945503 - 2019-FI-GEOSTAT4

Author: GEOSTAT 4

It is permitted to copy and reproduce the content in this report. When quoting, please state the source.

© GEOSTAT 4 and Eurostat 2022

Content

Content.....	2
1. Introduction.....	3
2. The concept of GSaaS.....	4
3. Good examples.....	5
3.1 Geostatistics Portal in Statistics Poland.....	5
3.2 STATatlas – interactive online Atlas in Austria.....	8
3.3 Digital Atlas – BKG Germany.....	11
4. Conclusions.....	13
5. References.....	15

1. Introduction

Many people associate numbers, statistics and hard data with something not very interesting. It makes a quite sensible association, because each of us has come across a situation in which we were completely bored with the way of providing various types of data.

It is for this reason that more and more statistical data is best shown on maps. Due to the regional differentiation of phenomena and the dynamism of changes, the cartographic method has become the optimal form of presentation of research results. To achieve this goal, we must have data with a spatial reference (i.e. it is possible to locate it on a map) and an Internet service thanks to which we will be able to publish (share) it. Moreover, statistical data with precise geographical reference is essential to facilitate comparative analysis and to achieve better statistical geospatial data production.

Web portals presenting statistical data are a very important element supporting the process of making statistical data available to users. Nowadays, the ability to independently view data and select the scope that interests us is a kind of basic requirement put in the first place when we ask about the accessibility and usability of statistical data. The data should be made available in the most readable and intuitive way, while meeting the conditions imposed by statistical confidentiality. In addition, the possibility of their presentation on maps is now also becoming a basic functionality without which users often lose the desire to browse and analyze data. Therefore, when designing the data collection process, it is worth planning the possibility of publishing the results in public web portals.

In addition, by giving a wide audience access to data, you give others the opportunity to create new unique data sets and their presentation, which may turn out to be very valuable.

2. The concept of GSaaS

The main assumption and goal of GSaaS is that it should be a tool that provides users with access to geospatial statistics data in an easy and intuitive way. Ideally it should also enable users to define their own area of interest and to download these data using different services so that they can continue to conduct their research by themselves. The best-known form of such data sharing are various types of web portals and digital atlases.

Portals are used for spatial presentation of largest information resource, enabling the publication of aggregated statistical data in the form of various types of spatial analysis, presented on maps with statistical confidentiality. A portal is a tool for interactive cartographic presentation and the publication of data acquired for example in censuses. It can serve the following functions:

- storing,
- presenting,
- sharing information for a broad group of recipients.

The Portal can function on two levels: for internal (official statistics) and external users, and the scope of presented data is defined through the appropriate roles and authorisations. Internal users could have access to both unit and aggregated data, whereas external users only to aggregated data, published taking into account statistical confidentiality.

Important part of a portal is its interface. It should allow its users quick and easy access to resulting statistical information. Data should be possible to be presented using such cartographical presentation methods as cartograms (choropleth map) and various cartodiagrams. It could be also possible to set one's own parameters for the visualisation of a thematic area for a given cartogram. These include measure, aggregation level (territorial division unit), the number of intervals, etc. Aside from the possibility of using ready-made spatial analyses, internal users could draw up custom thematic maps based on a selected feature of the data model, using dynamic spatial analyses, i.e. linear or distance analyses, or object buffering.

3. Good examples

3.1 Geostatistics Portal in Statistics Poland

In Statistics Poland the Project "Spatial statistical data in the state information system" was a project extending the capabilities of the former Geostatistics Portal. The main objective of that Project, hereinafter referred to as PDS, was to expand the scope and availability of statistical information and methods of geostatistical analyzes using the resources of official statistics, which information and analyzes (through their referentiality) are an important component of the state information system.

This goal was achieved by providing new or expanding services of the existing Geostatistics Portal, taking into account the resulting information developed on an ongoing basis under the Statistical Research Program of Public Statistics, as well as other public statistics projects, including Projects implemented under grant agreements (grants from the European Commission - Eurostat). The new portal presents the data held by public statistics and the results of geostatistical analyses, critical for the functioning of the state, local governments and local communities, which are able to be presented in a convenient graphic (mainly map) form, accelerating decision-making processes. An additional goal was to create solutions supporting the presentation of statistical research results.

The main goal consists of individual specific goals:

- Improving the availability of e-services;
- Providing new functionalities and extending the functionality of existing services;
- Expanding the infrastructure with the elements necessary to provide e-services at a high level of maturity.

The following public e-services are the project products:

- Service of access from computer devices to the resulting statistical information collected on the Portal with the possibility of performing advanced spatial analyzes and to data and metadata of the spatial information infrastructure;
- Service of access from mobile devices to the resulting statistical information collected in the Portal and to its visualization on maps;

- A service that enables the use of exploratory analyzes of spatial data using statistical information provided by the Portal;
- Service enabling the performance of analyzes in the field of geostatistical modeling;
- A service supporting the enrichment of the user's own content with geostatistical information and analyzes provided by the Portal;
- Service supporting the update of selected TERYT (internally maintained - National Official Register of the Territorial Division of the Country) objects together with related georeferencing objects.

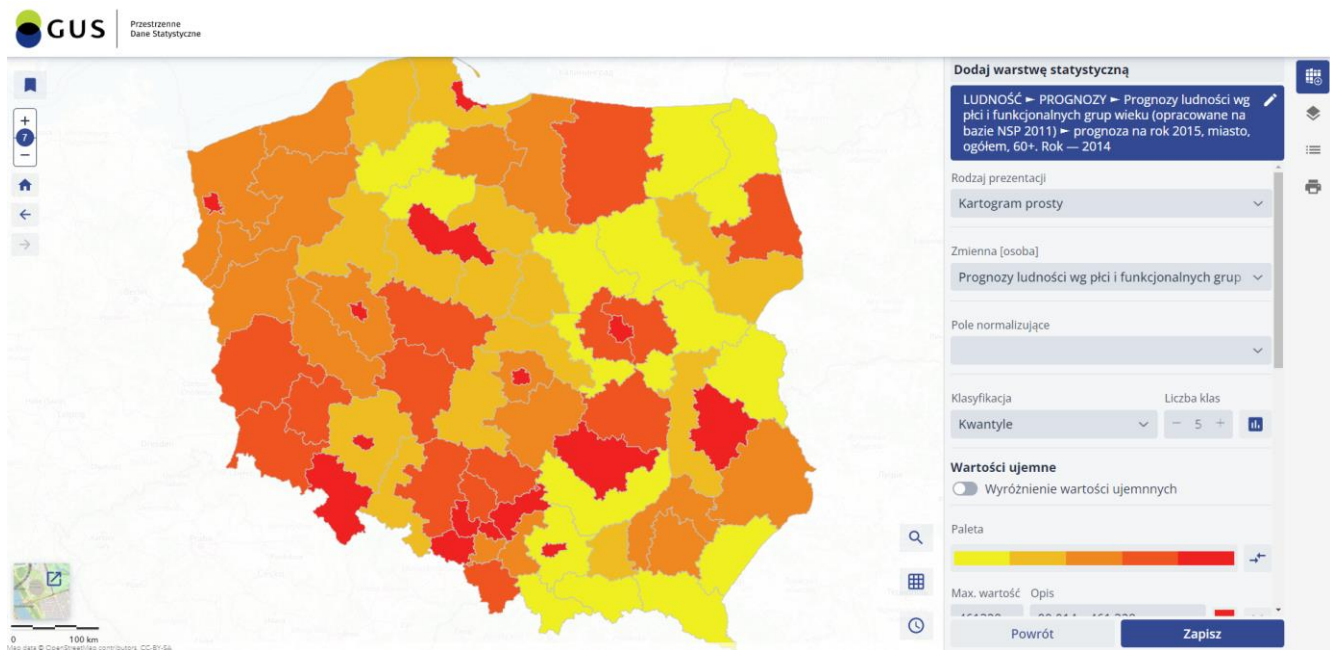


Figure 1 Geostatistics Portal - Statistics Poland

As part of the PDS Project, the functionalities of the previously provided services were expanded and modernised, including preparing statistical analyses in any division of space, e.g. defined in the application by the user, taken from external spatial services (WFS), in a "dynamic" square grid, as well as the possibility of combining statistical data and user's own data, geocoding user objects used for geostatistical analyses.

New services were also developed to enable the use of exploratory spatial data analyzes using statistical information, performing geostatistical modeling analyzes and supporting the enrichment of the user's own content with geostatistical information and analyzes.

These services allow users to perform advanced geostatistical analyzes on data collected in surveys of public statistics. Additionally, the user is able to combine own data with statistical data and on their basis create unique multi-dimensional analyzes.

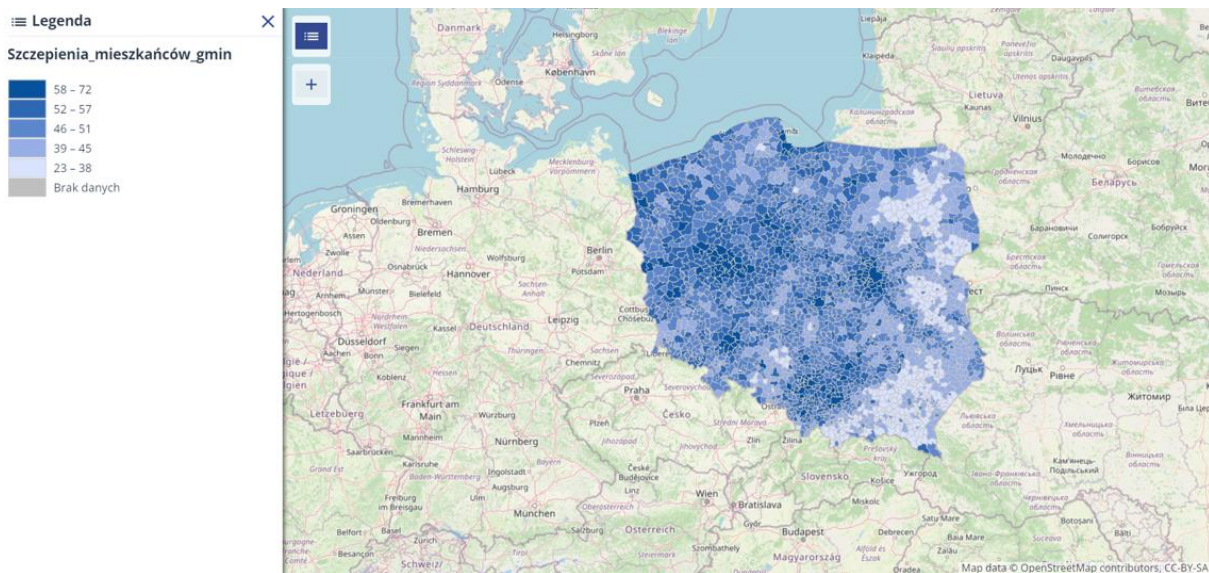


Figure 2 Geostatistics Portal - Statistics Poland

The result of the project is analytical support for portal users in making business decisions related to spatial information. The designed services allow for the practical application of spatial analyzes of a data mining nature, both in commercial activities, as well as those carried out by government and local administration and the scientific sphere. The system's functionalities make it possible to conduct previously unavailable analyzes on spatial data sets and on public statistics resources, allowing to obtain previously unavailable results, and to optimize business processes, significantly enriching the state's information system. In addition, due to the innovative and unique nature of the project's products, they can be used in scientific research to analyse the relationship between phenomena from various research fields.

3.2 STATatlas – interactive online Atlas in Austria

STATatlas (<https://www.statistik.at/atlas/>) is a very intuitive platform for serving data in spatial reference. It is a new cartographic product from Statistics Austria which is a collection for all online maps of Statistics Austria. STATatlas embeds all interactive maps and offers a variety of different statistical topics. It is the central collection for online map products of Statistik Austria and offers a broad overview of various statistical products. A wide variety of topics can be comprehensively explored here at one point. The digital thematic maps are intended to visualize complex topics easily and illustrate regional differences at a glance. The content is constantly being expanded and updated.

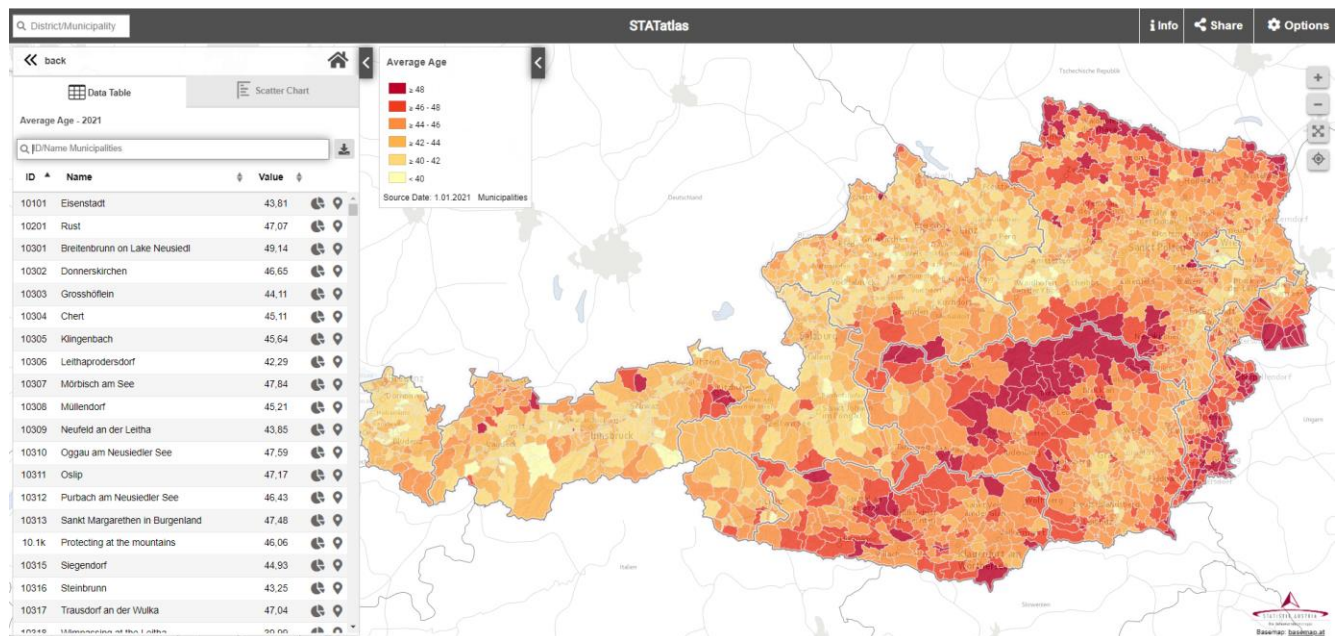


Figure 3 STATatlas data visualization

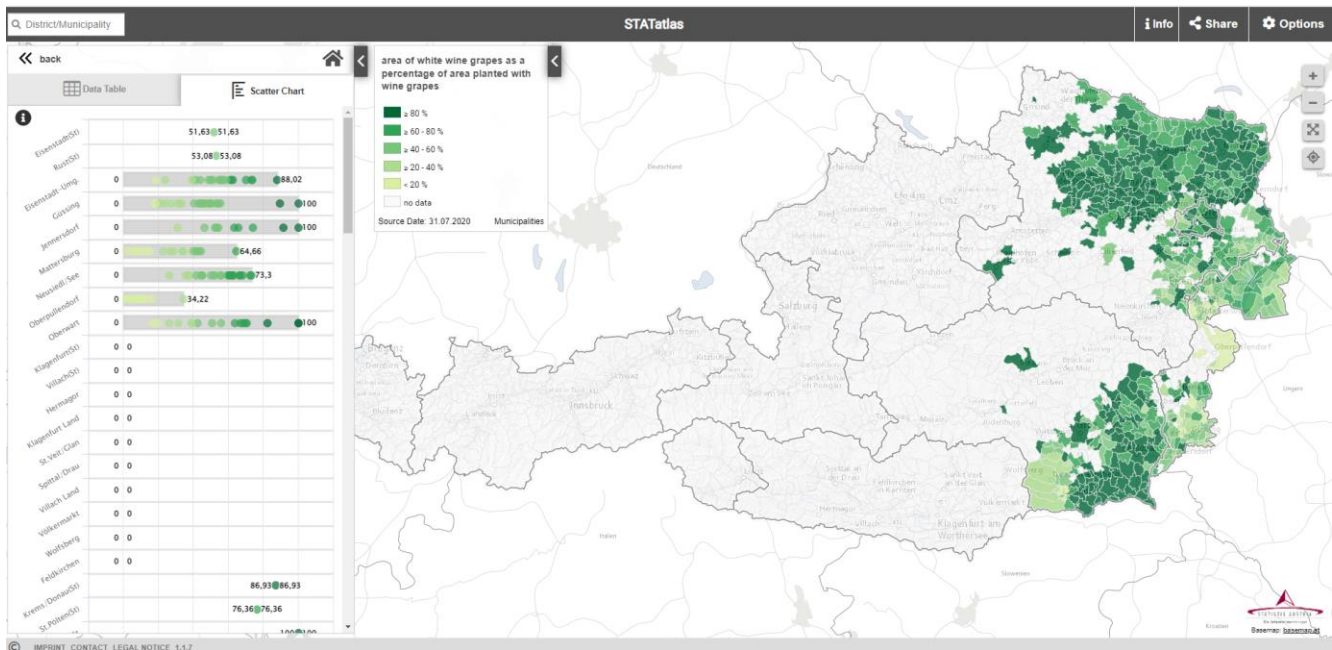


Figure 4 STATAtlas data visualization

Moreover, Statistics Austria provides aggregated data of any customer defined polygons. Georeferenced data is used as basis for these aggregations independently from administrative boundaries. These data can refer to the number of buildings and their characteristics, the number of residents, etc. Data is offered in package, but also single characteristics can be ordered. The regional extent for the requested data can be chosen without restrictions. Of course, the confidentiality criteria have to be met: characteristics relating to the resident population and to buildings are released only if at least 31 persons with main residence (resident population) and 4 buildings are present in the polygon. Details about characteristics can be found at <http://www.statistik.at/reg-datenkatalog/>.

It is possible to obtain for example objects from the Population Stock and the Buildings and Dwellings Register like:

- Resident population (number of persons with main residence)
- Number of persons with secondary residence
- Number of buildings
- Number of residential buildings
- Number of dwellings
- Number of buildings with dwellings
- Number of usage units

State of data: yearly updates with date as of 1 January.

3.3 Digital Atlas – BKG Germany

A very interesting example of the 'atlas' is the BKG atlases presented at the EFGS2021 conference as interactive web applications on various topics with partly up-to-date content (<https://gdz.bkg.bund.de/index.php/default/interaktive-atlanten.html>). The BKG has set itself the task of exploiting the potential of this distributed geoinformation and the technical possibilities of the latest, web-based technologies in order to give this information additional added value. With the interactive atlases, the information published in different places is taken up, merged and made available via a central access. Through the visual representation in a singular map application, synergies between the integrated information should be created and finally novel insights can be gained. Users also have the option of displaying statistics and forecasts, as well as individually adapting the content of the application and exporting specially configured maps for further use outside the Atlas applications.

Information on natural phenomena, such as precipitation, floods, wildfire, drought or heat can be understood and predicted better than ever before. The reasons are based on highly advanced technologies on earth observation in combination with a high density of specific sensors, some of them in real time. This leads to a large number of geospatial datasets provided by different organizations, e.g. via their own web portals. However, that means this high-value data is just information on its own. The BKG wants to enhance the great potential of mostly stand-alone shared geospatial information and aspires to extend its value by combining existing datasets through modern web-based technologies.

With BKGs digital interactive atlases the information is gained, joined and provided via a central access point. By visually representing the data in its respective map applications, synergies between the joined data can be determined in order to gain further insights. Users have several options to display statistics and forecasts as well as to tailor the content to their specific demands. Finally, they can go on to export their individually configured maps for further use.

The wildfire and drought atlases are combining wildfire/drought probability indices, weather information, infrastructure elements and near real-time data while primarily catering to experts. Even though the flood and heatwave atlases focus more on providing information to the general public, they still offer a vast amount of stand-alone and combined data, suitable for experts as well.

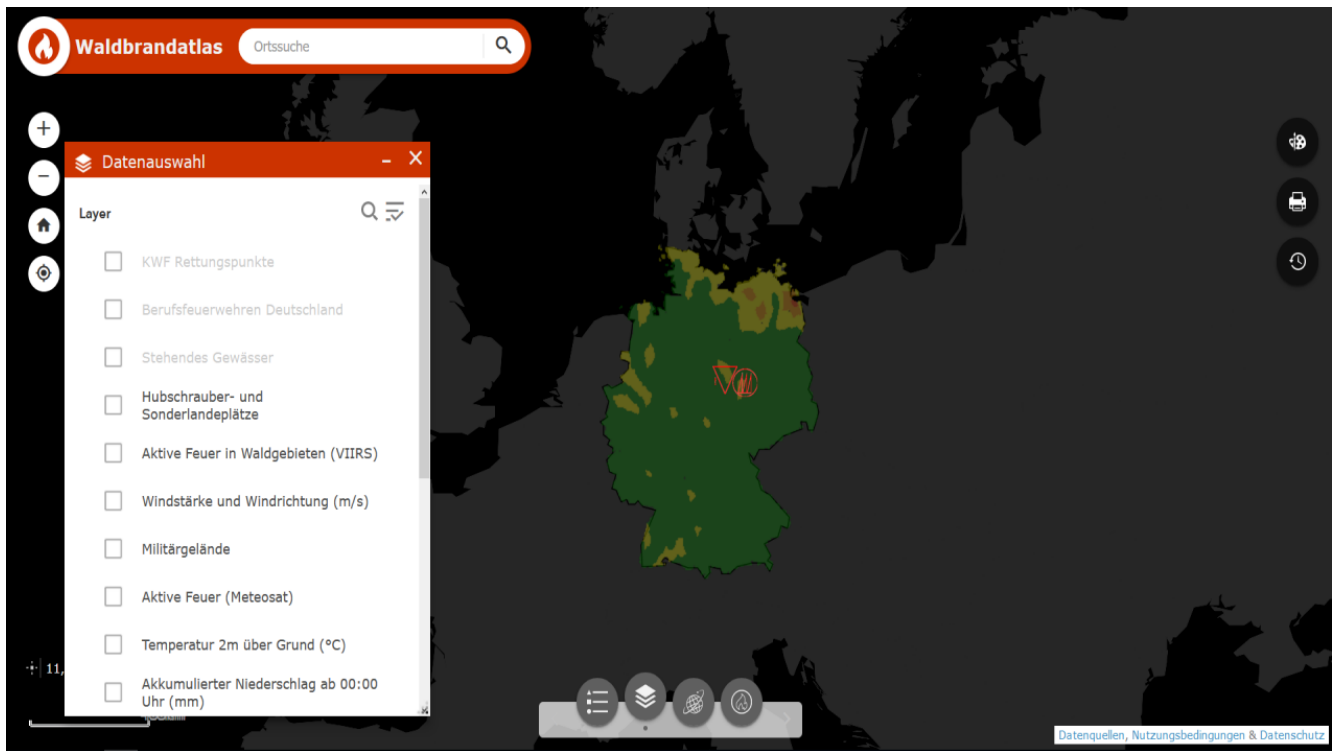


Figure 5 Digital Atlas – BKG Germany

4. Conclusions

Geospatial Statistics as a public service give a lot of possibilities for different kind of users with diverse need. Users without specialist knowledge benefit from easy discovery of interesting maps, professionals can obtain raw data with a few clicks, precise search filters allow data providers to review how their records appear to the public, and last but not least, interactive toolbox with side-by-side comparison, a time slider, drawing and measuring tools, and PDF export lowers the bar for using geospatial data in education.

When designing the process of obtaining data in statistics, it is worth taking into account from the very beginning that classification of the analyses conducted by points with x,y coordinates gives also the possibility to become independent from boundaries changes (in the regional division of the country), usually resulting in changes of census districts and laborious recalculations. This facilitates a comparative analysis of time series, regardless of the changes taking place in this division. An additional advantage is the possibility of the data aggregation both in the structure of the NUTS administrative division and the GRID divisions. It enables the presentation and comparability of data in a wider - pan-European perspective. The closest example of such data prepared in this way is the population data collected and published on the portal run by EUROSTAT (<https://ec.europa.eu/statistical-atlas/viewer/?config=REF-GRID.json>).

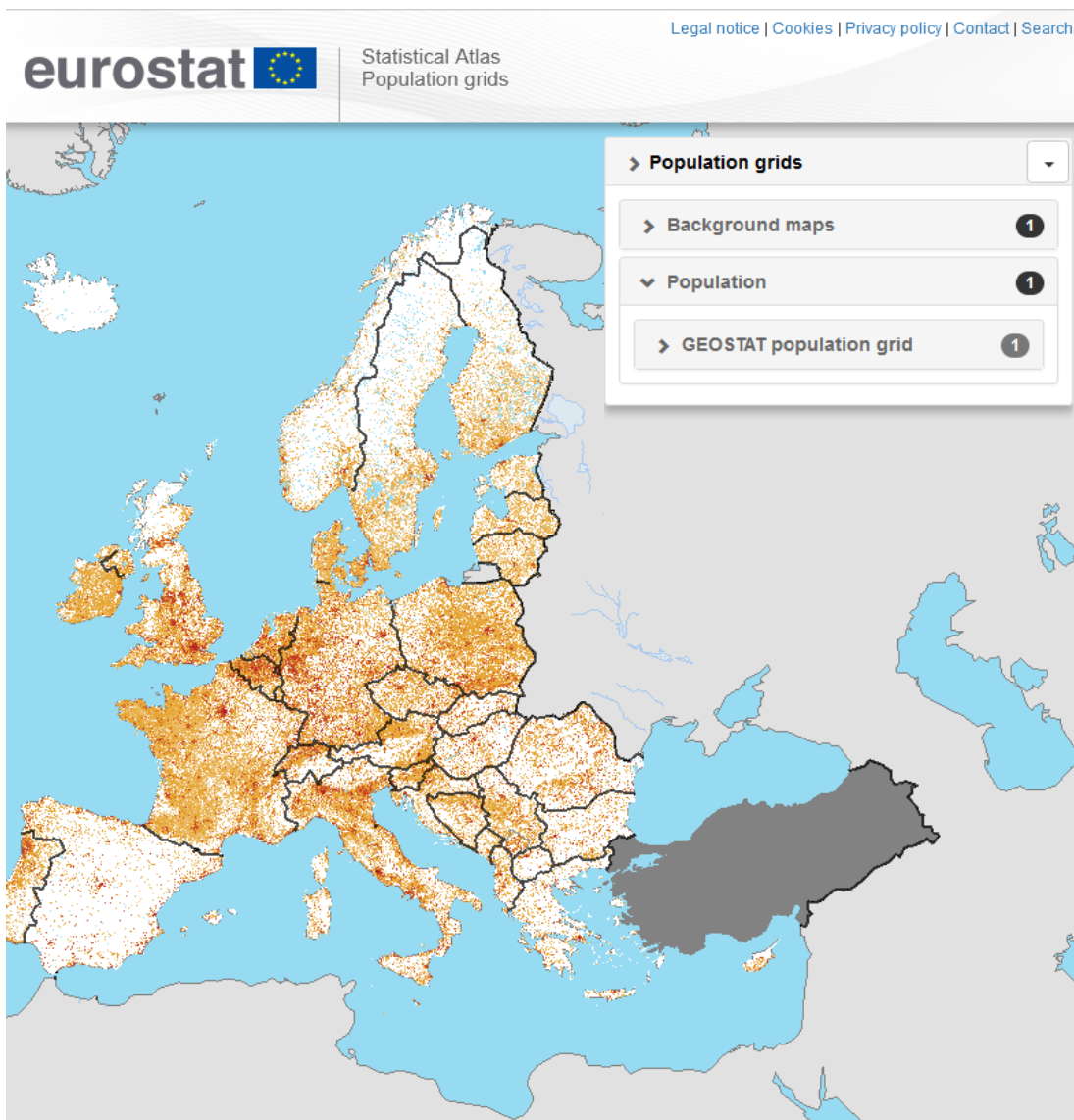


Figure 6 Sttistical Atlas – Eurostat

5. References

Information from Statistics Austria website:

(http://www.statistik.at/web_en/publications_services/customer_defined_polygons/index.html)

'Digital Atlases – the way of combining spatial data for better decision making' Karsten Schmidt - Federal Agency for Cartography and Geodesy – Abstract from EFGS2021:

[Microsoft Word - 08_SCHMIDT_Abstract_DigitalAtlases_EFGS2021.docx](#)

Information from BGK website:

<https://gdz.bkg.bund.de/index.php/default/interaktive-atlanten.html>