Innovative digital services for working with big data

Sergey Sukhinin¹, Olga Nikulina¹, Konstantin Petrov¹*

¹Don State Technical University, 344000, Gagarin Square 1, Rostov-on-Don, Russia

Abstract. Digital services contain big data, including about the population of countries around the world, its dynamics, demographic processes and population parameters. They are aimed at collecting, processing and storing this data, providing it to consumers for research, analytical, management, and educational activities. The value of digital services is that they capture a large amount of data, integrate them with each other and visualize them. The purpose of this article is to consider a new digital service containing big data about the population of the world's countries and their age and gender structure, visualizing demographic data in population pyramids reflecting the ratio of sexes by age groups. The advantages of this service are associated with a simple, easy-to-learn interface, modularity and interactivity in the use of big data presented in it. The new digital service provides both broad analytical and research opportunities for scientists who use its big data, as well as for schoolchildren and students who use its materials in the educational process.

1 Introduction

Demography, as the science of population, operates with large amounts of information reflecting the demographic behavior of people and socio-demographic indicators. Demographic data are very dynamic – they constantly change over time, by country and region. At the same time, demographic data for past periods are important, since they find their application in retrospective studies. They are used to compare with the current demographic situation, to identify prevailing trends in the development of demographic forecasts and to create demographic models [1].

The accumulation of large amounts of information in demographics makes it necessary to process, organize and store them. At the same time, big data has to be made available to the consumer quickly and presented in various forms – tables, graphics or maps [2, 3]. This creates large databases with demographic content, which are stored on the servers of official statistical services or specialized scientific institutions, are required when conducting socio-demographic research and for publications.

In the formation and work with big data, varieties of digital technologies are used. Their purpose is not only to provide open access to the necessary sources of information, but also

* Corresponding author: pks81@bk.ru
the possibilities of electronic processing of large amounts of information, calculating synthetic indicators based on them and visualizing data [4]. Due to the rapid obsolescence of information, digital statistical services are constantly changing, improving and having new functions and research opportunities. This makes the problem of considering new digital services used in various sciences, including demography, relevant.

2 Research materials and methods

The widespread introduction of electronic technologies in demographic research has been taking place in recent years because of the use of modern computer technology. Many scientists have considered methods and features of using digital databases in demographic research. F. Billari and E. Zagheni consider digital services in demography as an innovative tool that makes it possible to exchange knowledge with other fields of science in the big data environment [5]. R.M. Aliguliyev, R.M. Aliguliyev, F.F. Yusifov, I.Y. Alekperova define digital services in demography as a software system for maintaining and analyzing demographic data [6]. I. Weber and B. State made an important contribution to the consideration of digital sources of knowledge and technologies for demographic research [7]. They showed new opportunities for using information technologies in the study of the population and potential for the development of demography in the digital age. Numerous works of scientists are devoted to the use of digital electronic services in certain thematic areas of demography:

• in the study of population reproduction – M. Fire, Y. Elovici [8], F. Billari, F. D’Amuri, J. Marcucci [9];
• in compilation of population registers – K. Prins [10];
• consideration of population migration – Blumenstock J.E. [11], R. Careja, P. Bevelander [12], A Zyrianov and T Subbotina [13];

At the same time, the issues of using digital services in the study of demographic structures of the population have not yet been considered. The problem of using digital services when studying the composition of the population is the subject of our study.

The purpose of this article is to characterize an innovative digital service that considers the gender and age structure of the population.

The objectives of this study were:

• identification of the essence and specifics of digital services when working with big demographic data;
• analysis of the interface of a digital service with big data on the gender and age structure of the population;
• determination of information and analytical possibilities of using this digital service in demographic research;
• development of a methodology for working with this new digital demographic service.

3 Results

One of the aspects of studying the population of countries and regions of the world is the characteristic of the gender and age structure of population. The study of gender and age structures is carried out with population pyramids. These pyramids are a graphical representation of the distribution of men and women by age group. Such pyramids are depicted using two axes: age groups of the population are plotted vertically in the form of intervals; horizontally in the form of linear diagrams show the number or proportion of men
and women in each age group (Fig. 1). The pyramids make it possible to trace the relationship of the gender and age structure of the population with the reproduction of population, labor resources, as well as to identify demographic changes under the influence of various factors and in connection with historical events (wars, mass migrations, state demographic management).

Fig. 1. The generalized form of population pyramid [15].

However, the construction of such pyramids in demographic research is a very time-consuming process for a scientist, requiring a large amount of quantitative information, which he does not always have. Geographic information systems can help create population pyramids, but their programs are not available to all scientists [16]. This problem can be solved in demographic research through the use of population pyramids from Internet resources and information technologies for working with them. Such resources were developed in the 2000s and quickly entered not only the practice of scientific research but also the education system [17]. Thus, Internet sites with ready-made population pyramids were used in demography when studying the age and gender structure of population, in geography when considering the population of individual countries. Students with the help of such services were taught methods of analyzing pyramids in lessons of geography [18].

The most widely used population pyramids were presented on the website of the U.S. Bureau Censuses – www.census.gov. On the official website of this state statistical organization of United States, you could see the population pyramids of all countries of the world. A similar Internet resource demonstrating population pyramids was the website of the State Committee on Statistics of Russia (Rosstat) – www.gks.ru. On this site, having selected the tab "Russia in numbers" and in it by activating the link "Visualization of the dynamics of demographic processes in Russia", you could see an animated image of a population pyramid showing structural changes in the population of Russia from 1979 to 2030 (forecasts) with a lag one year in time. In recent years, in connection with the reform of these sites, resources on the population pyramids on them have become unavailable. Thus, it became necessary to create a new digital service containing big data on the age and sex structure of the population, and their visualization in the form of population pyramids.

In 2017, the Belgian programmer and scientist Martin De Wulf developed and presented on the Internet his website PopulationPyramid.net, where you can find pyramids of countries of the world integrated with demographic data on population dynamics. As M. De Wulf points out in the annotation to his website, the first version of this digital demographic service was developed by him in 2011 in order to facilitate the activities of schoolchildren when using population pyramids in educational activities. Therefore, it happened - for many years, this site was used in many schools, which greatly facilitated the efforts of schoolchildren, including when doing homework. The author received letters of feedback and comments from teachers who also began to use this site in their courses. All this
prompted M. De Wulf to modify the site. He focused not only on collecting and on processing big data on demographics, but also on visualizing them for users. He completely overhauled the backend code (the backend of the website, which now uses the Django framework) and the visualizations of the population pyramids themselves. Improved support for foreign languages – the service is now available in 11 languages. The author has strengthened social media sharing and openness to search engines (aka SEO: Search Engine Optimization); improved the performance of the site, the speed of displaying images and loading the data on which they are built. As a result, the new digital service has become more informative, visual, easy to understand, but at the same time has not lost its importance in educational activities and has expanded its capabilities for use in scientific research of a country-specific nature. The population pyramids presented by this service make it possible not only to see the ratio of men and women by age groups, but also to download data for calculating various demographic coefficients. Based on these calculations, researchers can make interstate comparisons of the gender and age structure of population.

The interface of the new digital demographic service is easy to use. The site comes off as a population pyramid for the entire population of the world. The "bookmark" located at the top of the page allows you to select the country of interest to the scientist to display its population pyramid on the screen. Next to the image of the population pyramid is a graph of the dynamics of the country's population. If you move the cursor over this graph, then two intersecting lines appear (as coordinate axes), they allow you to select the year for which the researcher wants to see the image of the population pyramid (Fig. 2). The user can obtain an image and a corresponding array of data on the sex and age structure and the population of the country for a fairly long time period – from 1950 to 2100. Thus, the developer of this service, in addition to real data on the gender and age structure of the population and the number of residents, also presents a demographic forecast. The author of this service managed to integrate big data on the population of the country, its composition by gender and age. Big data in this digital service is successfully visualized using a population pyramid and a graph of population dynamics. This service allows you to concretize the graphic image with digital data, to make a comparison between the selected age groups of the population. Next to each diagrammatic figure of the population pyramid, the share (%) of this age group in the total population is presented, which provides a basis for a detailed analysis of the ratio of men and women in age groups.

Fig. 2. Interface of service PopulationPyramid.net [19].
You can get complete data on the population of individual age groups, which are the basis for constructing the population pyramid, using the Excel CSV option. After activating this hyperlink, this service uploads an Excel file to the user with the corresponding statistical data on the ratio of men and women in the age groups of the population of a given country. In addition, this digital service allows you to download the image of the population pyramid itself using the Download function, as well as see the information resources on the basis of which big data was formed. Activating the Sources function for this purpose, the user will see a list of sites that served as the basis for creation population pyramids.

As additional information elements, this digital service presents for the user hyperlinks to information resources that are similar in content to the considered population pyramids. These are the "Number of migrants in a given country" and "Population density for a given country", which open, respectively, a diagram of the distribution of emigrants and immigrants by gender, as well as a map of population density. These information resources further enhance the cognitive function of this digital service.

4 Discussion

The presented new digital big data service on demography is aimed primarily at graphical visualization of data on the age structure of the population of a particular country. The main advantages of this service are information content and functionality. Based on the statistical data contained in this service, you can calculate the indicators of interest to the researcher and perform other calculation and analytical operations. In addition, this service uses a forecasting method, since it allows you to determine the future population size and gender ratio by age groups for a sufficiently long period.

The author-developer himself, considering the essence and values of the presented service, pointed out two of its features - modularity and interactivity. Modularity, according to M. De Wulf, is expressed in the possibility of using not all the information and analytical capabilities of the service as a whole, but its individual components for solving specific research or applied problems. For example, if the user is only interested in the population size and its dynamics, then he operates on the graph of population change, choosing the period necessary for the analysis. At the same time, the use of retrospective data makes it possible to single out the stages of the country's demographic development, and the use of prospective data is predictive in nature and can be used for the purposes of strategic planning and the development of a demographic policy program for the future. The modularity of this service is also important for the educational process, since the materials of the service provide an opportunity to choose those elements that are necessary for a particular course for each specific scientific research [20].

Interactivity, as the second advantage of this digital service, is to provide the user with opportunities not only to get the demographic data necessary for him about the population of the country in the so-called "ready-made form", but also to use them for various analytical operations and actions. The presented population pyramids push the researcher or student to think about why the pyramid looks exactly like this, what are the reasons for its shape, silhouette, structure, and disproportions. For example, for countries with natural growth, the population pyramid has a wide base and a narrow top, which indicates a constantly expanding population in each subsequent generation, compared to the previous one (Fig. 3). Swedish demographer and statistician A.-G. Sundberg called this age structure of the country's population progressive. If there has been a population decline in the country for many years, during which the number of inhabitants is decreasing, then each subsequent generation of people will be smaller than the previous one. This indicates a
regressive type of the age structure of the population, and in this case, the population pyramid has a wide top and a gradually narrowing base (Fig. 3).

![Population Pyramids](https://example.com/population-pyramids.png)

**Fig. 3.** Population pyramids for countries with different types of age structure [21, 22].

In addition, according to the silhouette of the population pyramid, it is possible to determine the stages of demographic transition at which the country under consideration is located. American demographer Frank W. Notestein developed the concept of demographic transition, according to which the country in its demographic development goes through 4 stages. The change in the stages of demographic transition is associated with changes in the types of reproduction, the historical transition from having many children to having few children. Each of the four stages of demographic transition corresponds to its own form of the population pyramid (Fig. 4).

![Demographic Transition Stages](https://example.com/demographic-transition-stages.png)

**Fig. 4.** Population pyramids for stages of the demographic transition [23].

The interactivity of the service, according to the author who developed it, allows students to explore the data and independently think over the facts, observed features, identify patterns. The animation of the population pyramid, combined with the change in the population graph, is of great importance for attracting interest in demographic data and involvement in population research. For example, it is more clear and interesting to compare the age and gender compositions and population of Afghanistan and Germany not using digital data, but using pyramids and graphs of the population dynamics of these countries (Fig. 3).
The presented service also facilitates the procedure for forecasting demographic data, since it eliminates the need to carry out the necessary calculations by the method of moving ages, to extrapolate the current trend in population dynamics for the future [24]. For the researcher, the entire forecasting procedure, based on the available big data, will be compiled by the service program itself and will give the result in the form of a population pyramid and a population graph for the requested year of the future period, will indicate the corresponding calculation results. Because such forecasting was based on an array of big data, its results can be considered quite reliable and objective.

5 Conclusion

The new digital service PopulationPyramid.net, developed by the Belgian teacher and programmer M. De Wulf, solves the important problem associated with the lack of sites with the most complete, reliable information with big data on demographics. This service contains large data on the population of the country and their gender and age composition; it allows you to present them not only in the form of digital arrays, but also clearly visualize them in the form of population pyramids and graphs of population dynamics.

This digital service has an important cognitive value. It provides the scientist with a large array of information for conducting socio-demographic research, allows for data analysis and cross-country comparisons, and graphic visualization facilitates the research process. An important feature of this digital service is its integrating function. This service combines demographic data on the dynamics of the population with information about the gender and age composition of population. This makes it possible to calculate not only synthetic demographic coefficients (coefficients of demographic load, population aging coefficient), but also to establish the type of age structure of the population and to identify its relationship with reproduction and population dynamics.

The described digital service can also find its application in the educational process. It allows you to diversify the course of the educational process, clearly demonstrating the educational material to students on the example of specific countries. Schoolchildren have the opportunity to apply in the classroom graphically new and relevant data that fill the existing gap in the educational system in the illustrative material on the study of the population. With the help of this service, it is convenient to explain the issues of the demographic development of the country, since its materials reflect the population of different countries, show their problems and specifics [25].

We have also used this service in our own research practice. Therefore, we can present a number of proposals for expanding its functionality. It would be possible to improve this service by making it a combination of the population pyramids of one country for different years in one image. For example, to demonstrate at the same time the population pyramids for the two extreme years (1950 and 2100) in one combined drawing. In this case, the pyramid for the base year (1950) should be shown in full, in the standard form, and the pyramid for the last year (2100) could be drawn with a line of a different color in the form of a silhouette. The divergence of the forms of the two pyramids in this case would indicate a change in the gender and age structure of the population. Such a visual presentation of the dynamics of the population pyramids would allow us to clearly trace demographic changes and link them with socio-economic and other processes and phenomena.

Summarizing our research, we can state that the author of this service has done a lot of work to create it. The service contains up-to-date demographic data, allows you to generate them in graphical or tabular form, receive them as separate files and use them for research and educational activities.
References


