Tatyana N. Belova[1]
Academy of Law and Administration of the Federal Penitentiary Service of Russia
Russian Federation, 119991,
Ryazan, 1 Sennaya st.
http://www.apu.fsin.su/

Dmitriy Mendeleev's Forecasts of the Population of Russia and the Realities of its Present Development

Abstract. In his last work "The Sacred Thoughts" an outstanding Russian scientist and encyclopedist Dmitri Ivanovich Mendeleev has explored demographic processes, acting as an economist and sociologist. The author of this article examines methodical approaches and tools, repeats and continues calculations of Mendeleev using modern data and information technologies. Comparing Mendeleev's forecasts of the size and age structure of population of Russia with the modern statistical data, the author discusses the reasons for the "failure" of forecasts. The arsenal of modern demographic tools can be supplemented with Mendeleev's idea on using the formula (law) of the vertical parabola in studying the age structure (the dependence between the size of the group and age).

Keywords: Mendeleev; size of population; fertility; mortality; population growth; age structure of the population.

JEL CODES: J11, J13

Brilliant discoveries and inventions in the field of natural sciences (the periodic law and the system of elements, the elasticity of gases, smokeless powder, etc.) are just one side of the multifaceted scientific and social activities of the outstanding Russian scientist Dmitri Ivanovich Mendeleev. Throughout his life he actively participated in the economic life of the country; he not only thought, compared, advised, but even gave his own example of how to act. Thus, in the Klin district he bought "...about 400 acres of land, the main mass of which was occupied by forests and meadows" [Mendeleev, 1995: 17], organized there an exemplary agricultural production and in 6-7 years proved that even on scarce land near Moscow it is possible not only to provide its population with bread, but also to trade surpluses. He also directly participated in the formation of the oil industry, factory and plant arrangement in the Urals, foreign trade relations and the introduction of the Single Customs Tariff of 1891, reforms in secondary and higher education, and development of aeronautics. Mendeleev travelled all

[1] Tatyana Nikolaevna Belova, Doctor of Economics, Professor of the Academy of Law and Administration of the Federal Penitentiary Service of Russia. E-mail: belova_t_n@mail.ru

over the country and the entire world, observed, advised, argued with politicians and ministers, gave recommendations to Nicholas II — he was an "annoying" scientist [Sorokin, 2010].

At the end of his life path, in 1903-1906, Mendeleev turns to the pressing social and economic issues of Russia and writes "The Sacred Thoughts" and their sequel "To the Knowledge of Russia". Unlike his natural science works, widely reprinted in the USSR, these works were treated by Soviet academic community as errors of the great scientist, were not published and hushed up. For example, in the series of biographies of "Lives of Outstanding People" the volume devoted to Mendeleev gives so little attention to these works that it is in fact impossible to understand what the scientist was actually writing about when exploring the social and economic issues of Russia [Pisarzhevsky, 1951; Smirnov, 1974]. The full edition of "The Sacred Thoughts" was published only in 1995, for the first time since 1905.

"The Sacred Thoughts" were long nurtured by Mendeleev, then he hurried to express his thoughts, it sounds like a testament to descendants. He is afraid to "sin with silence ..", when "the boiling thoughts tear out" [Mendeleev, 1995: 3]. The book contains more than a dozen sections devoted to the most important aspects of socio-economic and political life: industry and agriculture, secondary and higher education, foreign trade, the Russian-Japanese war. The second chapter of "The Sacred Thoughts" contains so many demographic aspects that would suffice for several modern dissertations. These are: total population growth, birth rate, mortality, age structure and population density, average age, differentiation of demographic indicators by country. The information base of the research is strikingly large for that time — these are all countries in which population censuses were conducted: North American United States, Germany, England, Austria, Hungary, Romania, Bulgaria, Serbia, The Netherlands, Belgium, Switzerland, Finland and France. In Russia, the first census of the population according to international rules was implemented later than in most developed European countries — only in 1897 — and by the time the book was written, only a part of the data was statistically processed.

For a natural scientist who extracts every digit with hard labour during experiments, this huge flow of statistical information has become a treasure indeed. Speaking of the benefits of not so much quantitative as qualitative considerations Mendeleev says this: "... in all parts of science, i.e. in all the search for truth, [it is necessary] if possible, to find numerical, measurable attributes, properties, and relations in order to be guided by them to find quantitative laws called empirical or experimental" [Mendeleev, 1995: 33]. Only in this way one can strengthen confidence "... in the existence of unshakable divine laws, the logical cause of which is often not known at all or is only hypothetically assumed, without any certainty in the truth of the assumption" [Mendeleev, 1995: 33].

While writing the book, Dmitri Ivanovich Mendeleev worked in the Chamber of Weights and Measures, and some employees, as well as his younger chil-

dren Masha and Vasya were involved in processing data that required not only careful calculations, but also good vision. How at that time was work with huge series of statistical data performed? In her memoirs O. E. Ozarovskaya presents Mendeleev's advice on the choice of a calculation device: "Here Vasily Dmitrievich [son — *author's remark*] says that one could learn faster using an Odhner Arithmometer, but it is noisy, maybe gets on one's nerves, while the French typewriter is softer, but it's harder to learn on it[1]."

Touching personal evaluations are combined in "The Sacred Thoughts" with strict scientific logic and mathematical calculations. In the implementation of population projections, its age structure, middle age and other demographic indicators, Mendeleev relies on extensive statistical data, develops appropriate methodology and a thorough mathematical apparatus. It is very interesting for a modern researcher, first, to make similar calculations using modern computer technologies; secondly, to refine and update computational procedures; thirdly, to evaluate the prognostic value of Mendeleev's methodology for historical demography[2] and, finally, to find out the factors underlying the "failure" of Mendeleev's forecasts for our time. Such tasks are posed by the author of this article.

Mendeleev's forecast: in 2000 the population of Russia will be 594 million people.

In his reflections on the size of population Mendeleev relies upon the principal postulate: growth of the population size is necessary for "the good of mankind in general and for the good of individual nations" [Mendeleev, 1995: 35]. He resolutely rejects the theory of Malthus, according to which, with the increase in population, poverty increases and, consequently, there is a restraint connected with the limitedness of natural resources, describing it as "Malthusian nonsense". Indeed, more than 70 years have passed (after the book was written in 1903), and "the benefits in human life have not diminished", hunger, diseases, wars have not increased, but have clearly decreased. According to Mendeleev's calculations, the area of the earth suitable for agricultural production is quite sufficient for the population and for its multiplication, since lots more of "the earth's surface is empty, and the distribution of population is far from even" [Mendeleev, 1995: 36]. In the course of time, there will be progress "in the extraction of food and all means for life" and the land will give much more yield than a hundred years ago, "if real knowledge, industry, trade, communication routes, desire for peace,

[1] O. E. Ozarovskaya. Memoires about Mendeleev. http://www.telenir.net/nauchnaja_literatura_prochee/poznanie_rossii_zavetnye_mysli_sbornik/p12.php

[2] For example, for the assessment of the age structure of the population of the Russian Empire in the late XVIII-first half of the XIX century, if only the total population size is known.

etc. will not weaken and will continue to develop in the same way as recently" [Mendeleev, 1995: 37].

Forecasting method. The forecasting method used by Mendeleev and his predecessors, John Graunt (1662) and later Thomas Robert Malthus (1798), is based on the exponential function (1):

$$S_{n+t} = S_n \left(1 + \frac{K_{\text{общ. пр.}}}{1000}\right)^t, \quad (1)$$

where t is the forecast period;
S_{n+t} is population at the end of the period n+t;
S_n is initial population, at the moment n;
$K_{\text{общ. пр.}}$ — annual population growth rate.

Graunt made a forecast of the population of England and came to the conclusion that it doubles every 280 years. In his work "An Essay on the Principle of Population" Malthus estimates the period of doubling at 25 years. Later, the methods of demographic forecasting became more complicated[1] and improved, but the results of long-term forecasting were invariably far from reality. In fact, only one indicator determines the success of the forecast — total population growth. Since the observation period was too short — it is necessary to wait for the next census, says Mendeleev, to better take into account the change in the coefficient of total growth at different time intervals.

Using the formula (1), Mendeleev makes a famous forecast: in 2000 the population of the Russian Empire will be 594.3 million people. To a modern Russian this forecast seems fantastic, but here it is necessary to consider the following initial assumptions, which Mendeleev operated on [Dyachkov, 2010]. First, territory and, accordingly, the population of these territories and lands, and secondly, the average annual population growth. Below we will discuss these two main components of the forecast.

Initial population. At the turn of the XIX-XX centuries the Russian Empire consisted of 19 districts and lands, including Finland. From the XIV century, Russia had been occupying lands — Siberia, Turkestan, Estland, Finland, Poland, etc. — and had become a major world power. To understand the scale of losses, we compare contemporary realities with the data of the 1897 census.

In the book "To the Knowledge of Russia" Mendeleev cites the 1897 Census data in the context of the districts, counties and regions on various grounds. We have boldly divided this table into two parts, leaving the original data un-

[1] The complication of the mathematical formula (1) was the proposal for the population forecast to use the logistic function (sigmoid), the graph of which resembles the letter S. According to the properties of this function, population growth rates first increase, and then, after passing the inflection point of the function, begin to slow down.

changed. In the left part of the table are the districts and lands of the Russian Empire, which are now part of the Russian Federation. The right side includes the administrative units of the Russian Empire, which, in the course of various historical events, became independent states — "lost" lands. Therefore, to verify the forecast, one must take the population number only in the left part of Table 1. We have received a final figure on the number of inhabitants of both sexes on the territory of the modern Russian Federation of 66,385 thousand people.

Table 1. Data of the 1897 Census of the Russian Empire on the number of people and the territory of administrative units

Administrative units	Number of inhabitants of both sexes, thousands.	Total territory, thousand square versts	Administrative units	Number of inhabitants of both sexes, thousands.	Total territory, thousand square versts
Parts of the Russian Federation (2017)			***"Lost" Districts and Lands***		
Petersburg Region units	**4601**	**181**	**Livonian Region units**	**2386**	**83**
Novgorod Governorate	1367	104	Estland Governorate	413	18
St. Petersburg Governorate.	2112	39	Lifland Governorate	1299	41
Pskov Governorate	1122	38	Courland Governorate.	674	24
Moscow Region units	**9793**	**232**	**Polish Region units**	**9404**	**112**
Tver Governorate	1769	57	Suvalki Governorate	583	11
Smolensk Governorate	1525	49	Łomża Governorate	580	9
Moscow Governorate	2431	29	Płock Governorate	554	8
Vladimir Governorate	1516	43	Warsaw Governorate	1932	15
Kaluga Governorate	1133	27	Siedlce Governorate	772	13
Tula Governorate	1419	27	Radom Governorate	815	11
Central Russia Region units	**12892**	**270**	Kalisz Governorate	841	10
Ryazan Governorate	1802	37	Petrokovskaya Governorate	1404	11
Oryol Governorate	2034	41	Lublin Governorate	1161	15

End of table 1

Administrative units	Number of inhabitants of both sexes, thousands.	Total territory, thousand square versts	Administrative units	Number of inhabitants of both sexes, thousands.	Total territory, thousand square versts
Tambov Governorate	2684	59	Kielce Governorate	762	9
Penza Governorate	1470	34	**Maloross Region units**	**17134**	**283**
Voronezh Governorate	2531	58	Podolian Governorate	3018	37
Kursk Governorate	2371	41	Volhynian Governorate	2989	63
Perm Region units	**10467**	**983**	Kievan Governorate	3559	45
Vyatka Governorate	3031	135	Poltava Governorate	2778	44
Perm Governorate	2994	290	Chernigov Governorate	2298	46
Ufa Governorate	2197	107	Kharkov Governorate	2492	48
Orenburg Governorate	1600	167	**The Lithuanian-Belarusian Region units**	**10063**	**267**
The Ural Oblast	645	284	Kovno Governorate	1545	35
Upper Volga Region units	**6214**	**206**	Vilno Governorate	1591	37
Kazan Governorate	2171	56	Vitebsk Governorate	1489	39
Nizhny Novgorod Governorate	1585	45	Mogilev Governorate	1687	42
Kostroma Governorate	1387	74	Minsk Governorate	2148	80
Yaroslavl Governorate	1071	31	Grodno Governorate	1603	34
Lower Volga Region units	**7689**	**444**	**Southern Russia Region units**	**8231**	**216**
Astrakhan Governorate	1004	190	Bessarabia Governorate	1935	40
Saratov Governorate	2406	74	Kherson Governorate	2734	62

End of table 1

Administrative units	Number of inhabitants of both sexes, thousands.	Total territory, thousand square versts	Administrative units	Number of inhabitants of both sexes, thousands.	Total territory, thousand square versts
Samara Governorate	2751	137	Taurida Governorate	1448	58
Simbirsk Governorate	1528	43	Yekaterinoslav Governorate	2114	56
Caucasus Region units	**6918**	**371**			
Black Sea Oblast	57	6	**Transcaucasia units**	**4935**	**184**
Kuban Oblast	1919	81	Baku Governorate	827	34
Terskaya Oblast	934	61	Kars Oblast	291	17
Stavropol Governorate	873	53	Elisabethpol Governorate	878	39
Dagestan Oblast	571	26	Erivan Governorate	830	23
Don Army Oblast	2564	144	Tiflis Governorate	1051	39
East-Siberian Region units	**1827**	**6769**	Kutais Governorate	1058	32
Irkutsk Governorate	514	638			
The Transbaikal Oblast	672	539	**Transcaspian Region units**	**4292**	**1167**
Yakutsk Oblast	270	3469	Transcaspian Oblast	382	532
Amur Region	120	397	Sir-Darya Oblast	1478	453
Primorskaya Oblast	223	1659	Samarkand Oblast	860	61
Sakhalin Island	28	67	Fergana Region	1572	121
West-Siberian Region units	**3931**	**4198**			
Tobolsk Governorate	1433	1219	**Kyrgyz Region units**	**2809**	**1691**
Tomsk Governorate	1928	745	Akmola Oblast	683	498
Yeniseysk Governorate	570	2234	Semipalatinsk Oblast	685	445
Northern Russia Region units	**2053**	**1211**	Semirechenskaya Oblast	988	348
Arkhangelsk Governorate	347	743	Torgay Region	453	400

End of table 1

Administrative units	Number of inhabitants of both sexes, thousands.	Total territory, thousand square versts	Administrative units	Number of inhabitants of both sexes, thousands.	Total territory, thousand square versts
Vologda Governorate	1342	353			
Olonets Governorate	364	115	**Finnish Region**	**2600**	**311**
Total	**66385**	**14865**	**Total**	**61854**	**4314**
All Russia: 128 239 thousand people Total territory: 19,179 thousand square versts Per capita, on average: 15.6 dessiatina of entire land					

Source: [Mendeleev, 1906: 16-21]. The author's layout of the table.

When we divided the census data into two parts, there were often doubts if the administrative unit was or was not a part of modern Russia. For example, the region of the Don Army with its tragic history: the number of inhabitants was 2,564 thousand people, the entire land is 144 thousand square versts. After the establishment of Soviet power in its territory, part of the counties became part of the Ukrainian SSR, and part became part of the RSFSR. The administrative-territorial division of the former region of the Don Army, including the Lugansk and Donetsk regions, has been repainted many times and to date part belongs to the territory of the Ukraine, and the other part — to the Russian Federation (Rostov Region). Another example is the Russian lands of the South Siberian, or the Kyrgyz region. Included in it: the Urals, Turgai, Semirechenskaya, Semipalatinsk and Akmola regions with a total population of 3,454 thousand people and the territory of 6,769 thousand square versts. Now the structure of the Russian Federation includes only the Ural (Sverdlovsk) region and part of the Semirechesk (Dzhetysu) region — the Kara-Kirghiz Autonomous District, which is part of the RSFSR. The Semipalatinsk, Turgai and Akmola regions are now part of Kazakhstan.

Taking into account the redistribution of territorial borders for such a long historical period, the population of the Russian Empire within the modern territory of the Russian Federation was 65,978 thousand people, according to updated data of modern statistics. [Population of Russia... 1998].

Population growth and forecasts. The second parameter for calculating the forecast of the population of Russia is the average annual population growth, the value of which, in Mendeleev's opinion, was not less than 1.5% [Mendeleev, 1906: 11]. Therefore, according to the formula (1) in 1900, the population of the Russian Empire would be 142.3 million people, and in 2000 — 594.3 million people. [Mendeleev, 1906: p. 12], (Table 2).

By territories, the total growth fluctuated significantly. Thus, for the European part of Russia in 1897, the birth rate was 49.5 per thousand, the death rate was 31.4 and, consequently, the natural increase was 18.1 per thousand. Such rate of overall growth was a long-term trend. The migratory component Mendeleev did not take into account, in view of its very small contribution to the overall increase. In Russian families, thus, there was the highest increase, "the most vigorous population".

Table 2. Population according to the forecast of Mendeleev for 2000 and actual population growth

Territory (country, region)	Data of 1897		Population forecast for 2000	Actual data, 2000	
	Population, thousand pers.	Population growth rate, ‰		Population, thousand pers.	Average population growth rate over the period
Russian Empire	128239	15	594323	-	-
Russian Empire (within the borders of the Russian Federation)	65978	15	305775	143667	7.63
Central Russia Region	12892	18	80973	8522	-4.01
Petersburg Region	4601	15	21323	7887	5.24
Moscow Region	9793	18	61509	23667	8.60
Poland	9404	15	43583	38559	13.79
Finland	2600	15	12050	5168.6	6.69
Entire world	1600000	10	4458836	6126622	13.12

Sources: Census of the population of the Russian Empire of 1897; Rosstat data, URL: http://cbsd.gks.ru/. 2017 (reference date: 18.03.2018). Author's calculations (average population growth)

Mendeleev has compared the demographic indicators with those of other countries. In Russia at the time, population growth was higher than in other parts of Europe. In the Netherlands, Germany and Norway it was about 1.3%, in England, Sweden, Italy — from 0.8 to 1.2%, in Spain, Switzerland — less than 1%. In France, at that time, there was a zero increase in population, in spite of quite a sufficient amount of land resources. On this occasion, Mendeleev expressed himself very ingeniously: "It seems to me that the reason for this must first of all be sought in that the modern French miserliness, in which a man and a woman, wishing to save up some money every month,

take a lot of care not to increase family expenses because of birth of children" [Mendeleev, 1995: 40].

Mendeleev made a prediction of the world population, based on considerations of natural growth and putting aside migration. He believed that "... the annual preponderance of fertility against mortality in different countries is different, and for the inhabitants of the whole world it can on average be accepted today at 1%" [Mendeleev, 1995: 40]. Precise information about the population of all the countries of the planet was not available at the time, and Mendeleev believed that the total population of the Earth was at least 1.6 billion. "If we assume that the increase will continue to be close to 1%, that is, the number of inhabitants of the Earth will double in approximately 60-70 years, then in 100 years by the year 2000, there will be over 4 billion inhabitants on the Earth" [Mendeleev, 1995: 42], more precisely 4 458.8 million people. (Table 2). Mendeleev's 100-year forecast for the world's population, unlike Russia, lags behind reality. In 2000, the world population was over 6 billion people, therefore the real average annual population growth by calculation is equal to 1.31%. Such a tightness (population density), Mendeleev believed, is the reason that "...all advanced countries with a dense population, even a small Belgium, are now concerned about the acquisition of colonies. Here too England and Germany are ahead of all other nations, and Russia occupied its neighbouring empty lands far-sightedly and in advance" [Mendeleev, 1995: 43].

Let's continue our calculations (Table 2). Taking into account the current borders, the population of Russia could have made up 305.8 million people in 2000, but the average increase for over a hundred years was only 0.76%. On the European part of Russia the example of the Central Russia Region is very typical (Ryazan, Orlov, Tambov, Penza, Voronezh and Kursk Governorates): the average increase over the period was negative — 0.4%. At the same time, the increase in the population of Poland, then part of the Russian Empire, is much higher and almost equal to the forecast — 1.4%.

Thus, Mendeleev's forecast of the population of Russia did not come true, and the discrepancies with real data are very significant. However, if we attribute the time of calculations to 1956, when population growth was 1.7%, then there will be no error. In addition, the example of the population forecast of other countries shows the correctness of the approach of Mendeleev. Could an outstanding scientist who had been observing the progress with his own eyes had known what awaited Russia in the future? He wrote: "The reason for the changes that have occurred in the world cannot be attributed to anything other than the spread among all of mankind of what is called humanity, or humaneness, what is contained in the concepts of modern realists about the possibility of avoiding wars, what compels caring for children more than before, and what is contained in the broad concept of freedom of labor" [Mendeleev, 1995: 43].

Age structure: the idea of a "sliding" vertical parabola. From population projections based exclusively on population growth rate, Mendeleev moved on to "the distribution of inhabitants by age and sex". At the same time, he believed that the structure by sex is not of great interest, since "everywhere in the world the number of men and women is close to one another." At the same time, "... distribution by age... is of enormous importance in all social relations, since all of them are determined by the labour of people, while children and old people cannot take part in it" [Mendeleev, 1995: 44]. Mendeleev, apparently, was the first to undertake a study of this kind: "As far as I know, no one has ever taken up the question of the normal law of distribution of the number of inhabitants by age, and if I decide to take up such a difficult new question, it is only for the reason that I believe in the law of large numbers..." [Mendeleev, 1995: 50].

Indeed, the material for the study was data from population censuses of over 25 countries in the world in dynamics. The determination of the regularity of the change in n (age) as a function of y (the proportion of population aged n), comparison of the obtained law (formula) for countries with different levels of economic development was the task set by Mendeleev in this part of the work.

Germany and the North American United States. The hypothesis was to show the proximity of the distribution by age for countries where "the degree of education and wealth of the people are close to one another" [Mendeleev, 1995: 45]. Germany belongs to such countries (a population of 49,425 thousand people) and N.-A. U. States (a population of 34,270 million people) according to the data of 1890, and in the N.-A. U. States only the number of the White births is taken into account (Table 3). The second and third columns of the table show the age structure of these countries for five-year periods, which is extremely similar, despite the significant differences in geographical location, natural resources, state structure, level of development of industries and trade, etc. between these countries. Therefore, it is quite justified to calculate the average for these countries (column 4), with which further calculations are made for levelling.

Table 3. Distribution of the population of Germany and N.-A. U. States by age groups (according to the data of 1890), actual and forecast values

	Proportion of population in age n,%			Discrete n	y based on 1%	y by the formula, %		y by the regression equation (3)
Age	Germany	N.-A. U. States	Average			N=105	N=110	
1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
0-5	13.01	13.28	13.14	3	2.62	2.74	2.62	2.63
5-10	11.19	12.93	12.06	8	2.38	2.47	2.38	2.40
10-15	10.95	11.57	11.26	13	2.24	2.23	2.15	2.18

End of table 3

Age	Proportion of population in age *n*,%			Discrete *n*	*y* based on 1%	*y* by the formula, %		*y* by the regression equation (3)
	Germany	N.-A. U. States	Average			N=105	N=110	
1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
15-20	9.72	10.37	10.04	18	2.00	1.99	1.93	1.97
20-25	8.61	9.29	8.95	23	1.78	1.77	1.73	1.77
25-30	7.58	7.43	7.50	28	1.50	1.56	1.54	1.58
30-35	6.85	6.94	6.89	33	1.38	1.36	1.36	1.40
35-40	5.91	5.99	5.95	38	1.19	1.18	1.18	1.22
40-45	5.44	5.02	5.23	43	1.05	1.01	1.03	1.06
45-50	4.94	4.29	4.62	48	0.94	0.85	0.88	0.91
50-55	4.33	3.66	4.00	53	0.81	0.71	0.77	0.76
55-60	3.50	2.73	3.11	58	0.64	0.58	0.62	0.63
60-65	2.88	2.32	2.60	63	0.53	0.46	0.51	0.51
65-70	2.32	1.69	2.00	68	0.41	0.36	0.40	0.39
70-75	1.56	1.22	1.39	73	0.28	0.27	0.31	0.29
75-80	0.80	0.71	0.76	78	0.15	0.19	0.23	0.19
80-85	0.31	0.37	0.34	83	0.07	0.13	0.16	0.11
85-90	0.09	0.14	0.11	88	0.02	0.08	0.11	0.03
90-N	0.02	0.05	0.03	93	0.01	0.05	0.10	-0.03
Total						$\sigma = 0{,}25$	$\sigma = 0{,}33$	$\sigma = 0.14$

Source: [Mendeleev, 1995: 48]. The last column of the table is based on the author's calculations.

Let's build a correlation field and add a trend line using spreadsheets (Figure 1). Mendeleev's assumption about the dependence of *y* on *n* in the form of a vertical parabola is quite logical[1]:

$$y = A + Bn + Cn^2, \quad (2)$$

where A, B and C are parameters of the regression equation,

[1] Here and further in Mendeleev's calculations the original symbols are retained. In the authors's own calculations using regression analysis (3), modern generally accepted symbols are used.

y is the dependent variable, the share of the population of the *n*-th group in the total population;

n is an independent variable, the age of the population of the n-th group.

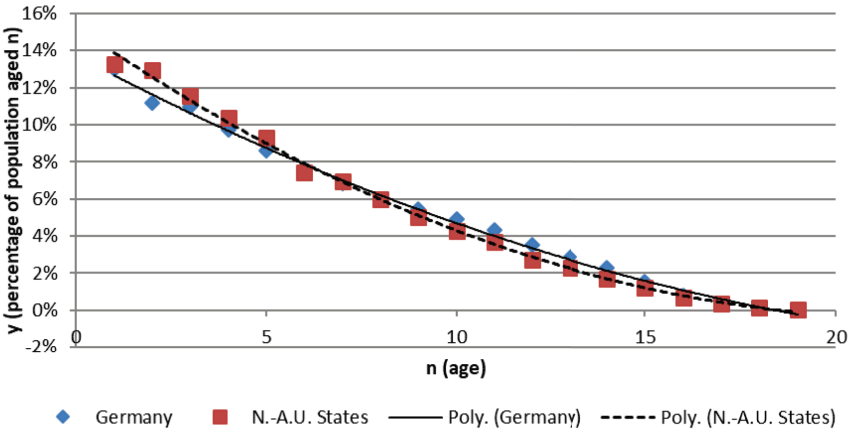

Fig. 1. Dependence of *y* on *n* (Germany, 1890, N.-A. U. States, 1890)
Data source: [Mendeleev, 1995: 48]. Author's calculations.

Mendeleev doubted that the formula (2) "... is final and completely accurate," but asserted "...that it very closely satisfies reality and deviates from it only by minor magnitudes..." [Mendeleev, 1995: p. 51]. Further, Mendeleev described a rather complicated and unusual for modern statistical analysis alignment procedure, to which we will return later. But earlier we will calculate the parameters of the regression equation (2) and all other characteristics of the model, using the standard MS Excel spreadsheet environment. We obtained the following regression equation for the *y* average (4 column)

$$\hat{y}_i = 2{,}775 - 0{,}048 n_i + 0{,}00019 n_i^2. \quad (3)$$

The model (3) provides an ideal description of the investigated dependence (R=0.99; it is adequate by the F-criterion, all parameters are significant by t-statistics, etc.). In column (9) of Table 2, the values of *y* are calculated using the model (3). But with regret, we note that in the last level of the series for the oldest age group of 93 (90 and older), according to our model, a negative value of *y*= -0.03% is obtained, which is essentially impossible. This is a significant qualitative drawback of the model (3).

Mendeleev's approach to the formula (or, as he calls it, "the law") (2) is somewhat different. First, to calculate parameters A, B and C, he assumed the least squares method. Then he leaned toward the "well-known trick of my [Mendeleev's] late friend P. L. Chebyshev, which I [Mendeleev] developed in the study

of the oscillation of weights[1]." A decisive role in the choice of the method was played by "two considerations that simplify the matter". First of all, this is the limitation of n by a certain limit N (the maximum age of survival of a person) and by the fact that the sum of all y is 100%. "Therefore, for $n=N$, *the* value of y can be assumed to be zero" [Mendeleev, 1995: 52]. The first derivative with respect to n is equal to $B + 2\,CN$. Then we obtain $B = -2CN$ and equal to zero. Substituting this expression in the equation (2), we obtain:

$$A = CN^2. \tag{4}$$

Then, using the values of A, B and C found this way in the equation (2), Mendeleev obtained a formula for y:

$$y = CN^2 - 2CNn + Cn^2, \text{ or } y = C(N-n)^2 \tag{5}$$

Further, Mendeleev found the sum of the right and left parts of the expression (5). Since $\sum_{n=0}^{N} y = 100$, then

$$\sum_{n=0}^{N} C(N-n)^2 = C\sum_{n=0}^{N}(N-n)^2 = 100 \tag{6}.$$

But $\sum_{n=0}^{N}(N-n)^2 = 1^2 + 2^2 + \ldots (N-1)^2 = \frac{N(N-1)(2N-1)}{1\times 2\times 3}$, therefore

$$C = \frac{600}{2N^3 - 2N^2 + N} \tag{7}$$

The required dependence of y on n takes the following form [Mendeleev, 1995, p. 53]:

$$y = \frac{600(N-n)^2}{N(N-1)(2N-1)} \tag{8}$$

Using the formula (8), Mendeleev calculated the aligned values of y for N= 105 and N= 110 (in Table 2, these are columns 7 and 8, respectively). In contrast to second-order polynomials (Fig. 1), "fading" in older ages to the negative region, Mendeleev's description of the dependence based on the formula (8) looks like a vertical parabola, "moving" on the abscissa axis as N increases. This is clearly seen in Fig. 2, which shows the kinks at the points N=105 and N=110.

---

[1] This means the Chebyshev polynomial for the expansion of functions on a given interval. O. E. Ozarovskaya, who helped Mendeleev in the calculations, recalls his instructions: "Chebyshev's method is necessary. Few possess it. Besides me, maybe five people in Russia. So, if you master it, you become a valuable person."
http://www.telenir.net/nauchnaja_literatura_prochee/poznanie_rossii_zavetnye_mysli_sbornik/p12.php

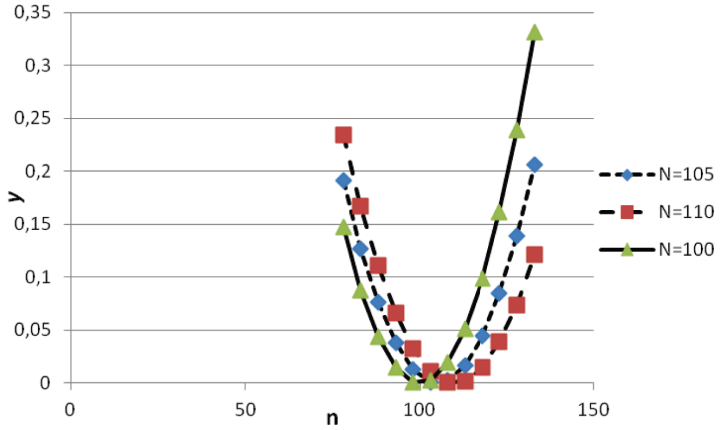

Fig. 2. Dependence of *y* on *n* for different values of the marginal age *N* in the foreseeable age range (78-133) years.

For the initial data for Germany and N.-A. U. States Mendeleev calculated the average value of *N*, and it turned out lying in the interval from N=105 to N =110 years. To reduce the complexity of calculations, he calculated the differences (y-$y_{calculated}$), separately positive $+\Delta$ and negative $-\Delta$. The indicator of the standard deviation σ, similar in meaning, was calculated by us for the aligned data in columns 7 and 8 in the final line. Mendeleev accompanied the selection of N for the smallest deviation from the empirical data with interesting comments: "I am convinced that this ultimate age of N varies with nations and centuries, and I even have reason to believe that it will subsequently grow with the development of education, ... that is, I consider the age of Methuselah[1] not as a single exception, but one should expect the norm in future, not mourn it somewhere behind" [Mendeleev, 1995: 51]. Mendeleev hoped that in the future, "the physical causes of ageing and the means to combat these causes" will be found. The increase in the proportion of the population of older ages should benefit the state system, the scientist believed. He explained it as follows: "... with the increase in the percentage of vigorous old people, mankind will have to improve, because such old people, wise with the experience of life, will have a beneficial effect on young people, no matter how conceited they are. ...The dry formula for the distribution of population over the ages and the indication that there is already an onset of an increase in the number of old people among the most educated people convinces me of the feasibility of such a "professorial dream"[Mendeleev, 1995: 52].

[1] Methuselah is one of the forefathers of mankind, famous for his longevity. In the Bible (Genesis 5: 21-27) it is said that he lived 969 years.

Age structure on the example of Russia. When Mendeleev was working on his book, the results of the first census of the Russian Empire in 1897 were not yet finalized. Thus, detailed data on the distribution by age, were known only for 17 provinces and regions, the island of Sakhalin and the two capitals [Mendeleev, 1995: 62]. It is for this reason of lack and incompleteness of information that Mendeleev investigated the dependence of y and n on the example of Germany and N.-A. U. States. Now that we know the detailed data not only of the 1897 census, but also of the modern (for 2017) demographic situation, it is appropriate to tackle "historical reconstruction", taking the formula (8) as a basis and the idea of a "sliding" vertical parabola. The results of the calculations are presented in Table 4.

Table 4. The distribution of the population of the Russian Empire (in present-day boundaries) by age groups, actual (1897) and projected values

Population groups by age n	Population		y based on 1%	y by the formula (8),% with		y by the regression equation (9)
	thousand persons	in % of the total		N=100	N=105	
1	2	3	4	5	6	7
0-5	9923	15.04	3.01	2.87	2.74	2.75
5-10	7715	11.69	2.34	2.58	2.47	2.49
10-15	7303	11.07	2.21	2.31	2.23	2.24
15-20	6532	9.90	1.98	2.05	1.99	2.00
20-25	5419	8.21	1.64	1.81	1.77	1.78
25-30	4973	7.54	1.51	1.58	1.56	1.57
30-35	4190	6.35	1.27	1.37	1.36	1.37
35-40	4078	6.18	1.24	1.17	1.18	1.19
40-45	3525	5.34	1.07	0.99	1.01	1.02
45-50	2943	4.46	0.89	0.82	0.85	0.86
50-55	2573	3.90	0.78	0.67	0.71	0.71
55-60	1967	2.98	0.60	0.54	0.58	0.58
60-65	1902	2.88	0.58	0.42	0.46	0.46
65-70	1145	1.74	0.35	0.31	0.36	0.35
70-75	948	1.44	0.29	0.22	0.27	0.26
75-80	506	0.77	0.15	0.15	0.19	0.18
80-85	236	0.36	0.07	0.09	0.13	0.11
85-90	81	0.12	0.02	0.04	0.08	0.06
90-N	19	0.03	0.01	0.01	0.04	0.02
	65 978	100%	$\sigma = 0{,}44$	$\sigma = 0{,}39$	$\sigma = 0{,}39$	$\sigma = 0{,}39$

Source: [Population of Russia..., 1998] Author's calculations.

Comparison of the presented data with the age structure of Germany and N.-A.U. States shows significant differences, related both to a higher birth rate in the Russian Empire, and mortality. Let's pay attention to the first two age groups: up to 5 years and from 5 to 10 years. In Germany and N.-A.U. States children under the age of 5 averaged 13.14% (Table 3), children from 5 to 10 years - 12.06%. In the Russian Empire in 1897 - 15.04% and 11.69% respectively. The gap between these groups is explained by the high mortality rate, the highest in infancy, and slightly less up to 5 years. Mendeleev considered high infant mortality to be a lack of education: "... young organisms, especially in the first years of life among the unenlightened and poor people, are dying out in large numbers not only from lacking both medical care and deprivation, but mainly from the underdevelopment of mothers on whom lies the natural duty to care for children of small age, if the fathers are obliged to raise funds for the entire family" [Mendeleev, 1905: 45]. In the older age groups, on the contrary, the proportion of the population of the elderly is higher in Germany and the N.-A.U. States and lower in the Russian Empire. It should be noted that with the similarity of the middle part of age structures, differences begin already with the able-bodied age of 45-50 years.

The alignment of the age structure was done according to the method proposed by Mendeleev - this is the formula (8). The results of the calculations are given in columns (5) and (6) of Table 4. Selection of N was carried out by the criterion of the smallest standard deviation. For $N = 100$, this figure was 0.44, for N=105 slightly less than 0.39. We note that the value of the limiting value of N, as expected, is lower for Russian data. This means that the age of Russian long-livers is less, on average, by 5 years.

The results of the equalization by regression equation (9) are given in column (7) of Table 4. Let us compare its parameters and characteristics with the regression equation (3) by average empirical data of Germany and the N.-A.U. States.

$$\hat{y}_i = 2{,}91 - 0{,}055 n_i + 0{,}00029 n_i^2. \quad (9)$$

The general form of the second-order polynomial did not change (Figure 3), but the quadratic effect became somewhat less. All parameters of the regression equation (9) are significant, the model is adequate and quite objectively describes empirical data for the Russian Empire.

Let's turn to modern statistical data on the distribution of the population of Russia by age (Figure 3). Such a dependence of y on n, or rather the absence of any dependence and order could hardly have been predicted by Mendeleev. Chaotic zigzags at first glance seem to be errors in statistical data. Note that in all three sets presented earlier, the size of each subsequent age group is less than the previous one. Rare cases of discrepancy can be explained either by increased migration, or inaccuracy in the collection of information. In the case of modern

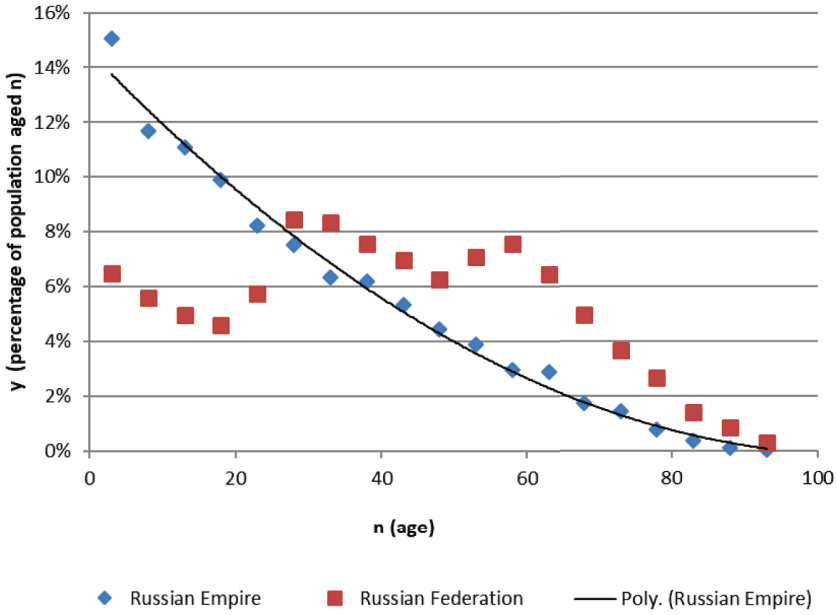

Fig. 3. Dependence of *y* on *n*: The Russian Empire, 1897 and the Russian Federation, 2016.
Sources: census of the population of the Russian Empire (1897) and data of Rosstat [Digital resource]. URL: http://cbsd.gks.ru/. 2017 (reference date: 15.10.2017).

Russia, which has a long negative overall population growth, sharp fluctuations have a pronounced socio-economic connotation. Thus, the age group of 15-19 years has the smallest number of 6,731 thousand people, the subsequent age group of 20-24 years (1992-1997) significantly exceeds it — 8,445 million people. The largest number in groups of young ages is: 25-29 years — 12 412 million people, and 30-34 years — 12 219 million people. The reason for the significant excess of the number of the last two cohorts is primarily the measures of demographic policy of the 1980s.

Further discussions on the contemporary age structure of Russia call for addressing fertility issues, the family and family values crisis, the role of women, income levels and many other acute problems. The dependence of fertility on material factors is now being questioned [Arkhangelsky, 2017; Rybakovsky, 2014], and the ageing of the population becomes a factor of increasing dependency rate [Belova, 2005]. A huge layer of literary sources is devoted to the discussion of these issues. We, on the other hand, will return to Mendeleev and try to summarize our reasoning.

Conclusion The limited journal space did not enable continuing the analysis of Mendeleev's study of the age structure, in particular, its applications. Based

on the alignment of the age structure according to the formula (8), as Mendeleev showed, it is possible to calculate forecasted (predicted) values of the following important demographic indicators: middle age (M), fertility and mortality. Between M and N, Mendeleev believed, there is a link, so that by age structure one can find M, and N corresponds to a certain average age. The deduced formula of the age structure using a sliding vertical parabola can replenish the arsenal of demographic prediction methods. It should be noted that it correctly describes evolutionary, slowly changing processes and for this reason is not suitable for forecasting the Russian demographic situation.

Why was the forecast for the population of Russia in 2000 wrong? Mendeleev considered the main factor in the overall growth of the Russian population to be materialistic — the amount of land suitable for economic management per capita, the development of agricultural and manufacturing industry, transport, trade, etc. He was sure that fertility, mankind's desire for reproduction, is laid in "... the nature of people, like all organisms in general, and the welfare of mankind cannot be spoken of without relying on information on the population" [Mendeleev, 1995: 44]. But the nature of people apparently changed and even then was "...in clear contradiction with the socialists, communists and all sorts of other intriguers..." [Mendeleev, 1906: 14]. Mendeleev expressed the significance of the influence of the state and social system on the birth rate of the population in these words: "For me, the highest or most important and humane goal of any "policy" is expressed more clearly, simpler and more tangibly in the development of conditions for human reproduction" [Mendeleev, 1906: 14]. Modern demographic realities as a criterion for the effectiveness of "policies" are rather dismal, but inspire hope for the best...

Reference List

1. Arkhangelsky V. N., Zinkina Y. V., Korotaev A. V., Shulgin S. G. 2017. Current trends in the birth rate in Russia and the impact of government support measures. *Socis.* 3: 43-50. (in Russian)
2. Belova T. N. 2005. What links the monetization of benefits and the birth rate of the population? *Statistics issues.* 6: 31-36. (in Russian)
3. Dyachkov V. L., Kanishchev V. V. 2011. Forecast for the growth of the population of Russia in the XX century. *Ineternum.* 1. [Digital resource]. URL: https://cyberleninka.ru/article/v/prognoz-rosta-naseleniya-rossii-na-xx-v-i-sostoyavshayasya-realnost-vzglyad-snizu-iz-tambovskoy-krestyanskoy-sredy (in Russian)
4. Mendeleev D. I. 1906. To the Knowledge of Russia. With the map of Russia. The 3rd edition, again amended and supplemented. St. Petersburg. A. S. Suvorin edition. (in Russian)
5. Mendeleev D. I. 1995. The Sacred Thoughts: Full edition (for the first time since 1905) M. Mysl. (in Russian)
6. Ozarovskaya O. E. Memoires about Mendeleev. http://www.telenir.net/nauchnaja_literatura_prochee/poznanie_rossii_zavetnye_mysli_sbornik/p12.php

7. Population of Russia for 100 years (1897-1997). 1998. Statistical Digest. Moscow. Goskomstat of Russia. (in Russian)
8. Pisarzhevsky O. N. 1951. Dmitri Ivanovich Mendeleev. Moscow. Molodaya Gvardiya. (in Russian)
9. Rybakovsky O. L., Tayunova O. A. 2014. Factors of the dynamics of the birth rate of the population of Russia in the early XXI century. *Sociological research*. 9: 19-24. (in Russian)
10. Sorokin A. I. 2010. D. I. Mendeleev on the issues of socio-economic development of Russia at the turn of the XIX-XX centuries. *Bulletin of St. Petersburg State University*. Series 5. Issue 1. [Digital resource]. URL: https://cyberleninka.ru/article/n/d-i-mendeleev-o-problemah-sotsialno-ekonomicheskogo-razvitiya-rossii-na-rubezhe-xix-xx-vv (in Russian)
11. Smirnov G. V. 1974. Mendeleev. Moscow. Molodaya Gvardiya. (in Russian)