7COM1079-0901-2024

Team Research and Development Project

Final Report Title: Analysis of Russian Demographic Data

**Group ID: A321**

**Dataset number: DS105 russian\_demography.csv**

**Prepared by:**

Monday Henry chukwu

Mani chandu kommaraneni

Muhammad Naveed

Ravikishore Kommina

University of Hertfordshire

Hatfield, 2024

**Table of Contents**

[1. Introduction 4](#_gjdgxs)

[1.1 Problem statement and research motivation 4](#_30j0zll)

[1.2 The data set 4](#_1fob9te)

[1.3 Research question 4](#_3znysh7)

[1.4 Null hypothesis and alternative hypothesis (H0/H1) 4](#_2et92p0)

[2. Background research 5](#_tyjcwt)

[2.1 Research Papers 5](#_3dy6vkm)

[2.2 Why RQ is of interest 5](#_1t3h5sf)

[3. Visualization 6](#_4d34og8)

[3.1 RQ output of an R script 6](#_2s8eyo1)

[3.2 Understanding the data 8](#_lnxbz9)

[3.3 Useful information for the data understanding 9](#_35nkun2)

[4. Analysis 9](#_1ksv4uv)

[4.1 Statistical test 9](#_44sinio)

[4.2 Null hypothesis is rejected /not rejected based on the p-value 9](#_2jxsxqh)

[5. Evaluation – group’s experience at 7COM1079 10](#_z337ya)

[5.1 What Went Well 10](#_3j2qqm3)

[5.2 Points for improvement 10](#_1y810tw)

[5.3 Group’s time management 10](#_4i7ojhp)

[5.4 Project’s overall judgement 10](#_2xcytpi)

[6. Conclusions 10](#_1ci93xb)

[6.1 Results 10](#_3whwml4)

[6.2 Interpretation of the result 11](#_2bn6wsx)

[6.3 Reasons and/or implications for future work, limitations 11](#_qsh70q)

[7. References 12](#_3as4poj)

[8. Appendices 13](#_1pxezwc)

[A. R code used for analysis and visualization 13](#_49x2ik5)

[B. GitHub log output 13](#_2p2csry)

**List of Figures**

[Figure 1: Birth Rate Distribution 6](#_17dp8vu)

[Figure 2: Net Population Growth 7](#_3rdcrjn)

[Figure 3: Birth Rate vs Net Population 8](#_26in1rg)

# 1. Introduction

## 1.1 Problem statement and research motivation

Russia has demographic threats; it has a low birth rate; it has a growing aging population, and it is demographically uneven. These trends are clearly against the background of economic growth, stability in the labor market, and regional development. While a host of data exist, there is little synthesis across demographic, economic and social differences that can inform policy decisions.

Familiarity with Russian demographic trends are important for identifying the directions of further work to overcome the decline in population and unequal regional development (Arkhangelskiy et al., 2024). This study also seeks to undertake demographic analysis with regards to population characteristics, reveal some of the challenges likely to occur and propose a development solution.

## 1.2 The data set

This research consists of using the dataset named as ‘russian\_demography.csv’, which contains demographic data for different regions around Russia. From the demographic aspects, it includes information for ***Net Population Growth (npg), Year, Region, Birth Rate, Death Rate, Gross Domestic Wealth (gdw), and Urbanisation*** in its column names. Thus, they become very crucial to policy makers and stakeholders for shaping developments about demographic trends as well as factors influencing population change in other regions.

## 1.3 Research question

This research question seeks to examine the relationship between birth rates with the net population growth, with purpose to pinpoint the relationship, regional differences and regional implications on the demographic report for Russia that would help regarding place, life and policy measures to solve the problem of population shrinkage and uneven birth rates and growth across the Russian regions.

***Is there a correlation between birth rate and net population growth in Russian regions?***

## 1.4 Null hypothesis and alternative hypothesis (H0/H1)

**Null Hypothesis (H0)**: There is no significant association between birth rate and net population growth in Russian regions.

**Alternative Hypothesis (H1)**: There is a significant association between birth rate and net population growth in Russian regions.

Birth rate is one of the most important demographic factors which is crucial for affecting the net population growth along with death rate, gross domestic wealth and urbanization in Russian regions. That is why, birth rate also plays an important role in influencing the net population growth for further analysis.

# 2. Background research

## 2.1 Research Papers

Research has examined the Russian demography specifically with regards to regional birth rates and contribution to population growth. According to Kazenin (2021), there was a decline in the first child births, in addition to the third and subsequent child births, across most of the regions in the 2020 birth rate stabilization. This tendency opens up new directions both in childlessness and in the increase in the number of members of a family for Russian conditions. Shubat (2019) used hierarchical cluster analysis together with 30-year dynamics to determine five clusters with similar rates of birth. This study has also disclosed how differently such factors affect fertility rates across these clusters which indeed demands a regional demographic structure.

Pant (2017) also explored Russia’s demographic trend after the de-orbit of the Soviet Union and concluded that it has been growing negatively because of low birth rates and high mortality. This pointed to the high mortality rates as aggravating factors on the generally low birth rates in Russia. SiteSoft (2024) shows that there is still a trend towards a gradual reduction in the birth rate: the number of births in January-October 2023 was 3% lower than in the same period of 2022. Such a trend has upsetting effects on the progress of the population growth, as well as the sustainability of various regions. In combination with each other, these works reveal the multifaceted interaction of birth rates and regionalization of population fluctuations in Russia and calls for the effective anti-demographic policy within the regions.

## 2.2 Why RQ is of interest

As of now, there has been a large amount of demographic research on factors causing population growth but little on Russian regions only. In this research gap, the current study aims to understand the above relationships from the perspective of local conditions. One of the future directions may be to study the effect of other socio-economic factors on population growth and other counties to be compared.

# 3. Visualization

## 3.1 RQ output of an R script

The visualization of scatter plots and histograms with trend bells have shown the distribution of demographic factors and net population growth. This helps in finding patterns and relationships between two variables by giving visual representation of data.

# Plotting the distribution of birth rate with trend bell

ggplot(RussianDemography, aes(x = birth\_rate)) +

geom\_histogram(aes(y = after\_stat(density)), binwidth = 0.1, fill = "green", color = "black") +

geom\_density(color = "red", linewidth = 1) +

labs(title = "Birth Rate Distribution with Trend Bell", x = "Birth Rate", y = "Density") +

theme\_minimal()

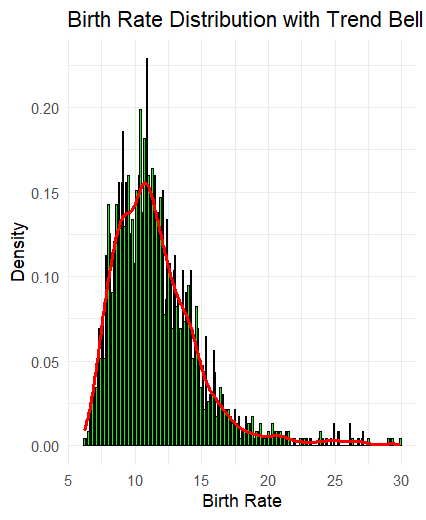


Figure 1: Birth Rate Distribution

This histogram and density plot show that birth rates in Russian regions are distributed the way shown below. The green color depicts the data frequency and the red density curve shows the pattern indicating the right shifted peak where most regions represent moderate births and few higher birth regions.

# Plotting the distribution of Net Population Growth with trend bell

ggplot(RussianDemography, aes(x = npg)) +

geom\_histogram(aes(y = after\_stat(density)), binwidth = 0.1, fill = "purple", color = "black") +

geom\_density(color = "red", linewidth = 1) +

labs(title = "Net Population Growth Distribution with Trend Bell", x = "Net Population Growth", y = "Density") +

theme\_minimal()

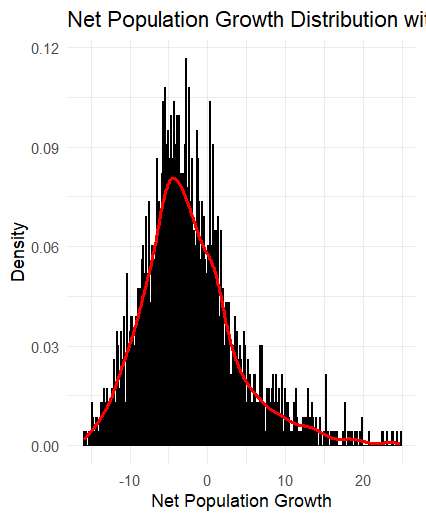


Figure 2: Net Population Growth

The calculated histogram and density plot for net population growth show that distribution of this variable is very close to normal distribution with mean equal to zero, which means that the majority of regions have experienced moderate growth or decline. Purple bars show NPG frequencies and the red density curve enhances the trend as it considers the regional demographics.

# Plotting the relationship between Birth Rate vs. Net Population Growth

ggplot(RussianDemography, aes(x = birth\_rate, y = npg)) +

geom\_point(color = "blue", alpha = 0.6) +

geom\_smooth(method = "lm", color = "red", se = FALSE) +

labs(title = "Scatter Plot of Birth Rate vs Net Population Growth", x = "Birth Rate", y = "Net Population Growth") +

theme\_minimal()

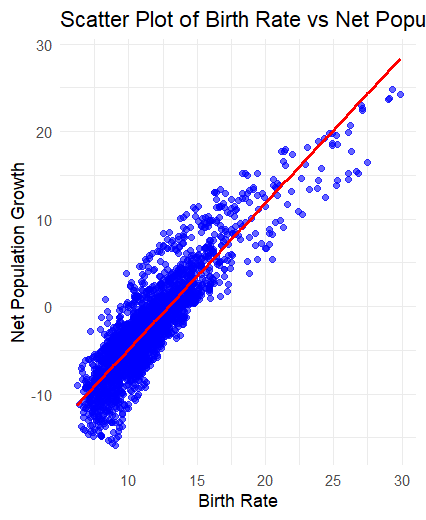


Figure 3: Birth Rate vs Net Population

As it is seen in this scatter plot, there is evidence of a raise and shift pattern in birth rates and net population growth. The geographical areas with commensurate births have a higher population growth rate. The red regression line fit from the linear modeling part proves this hypothesis, this process shows that birth rate has great impacts on population fluctuation.

## 3.2 Understanding the data

**Birth Rate Distribution:**

**Purpose:** To identify the dispersion of birth rates through the Russian regions with a reference to the general trend.

**Insights:** There is right-skewed distribution, which means the majority of regions have average birth rates and few regions have high birth rates; this implies that birth rates may be regional and it may be necessary to make efforts to increase birth rates in low birth rate areas.

**Net Population Growth Distribution:**

**Purpose:** To compare the trending net population growth of the regions of the world.

**Insights:** The fairly balanced bar indicates that there are areas that have population growth and other areas that have population shrinkage which should show that demography is a mix and researchers should try and establish the factors of migration and mortality.

**Scatter Plot of Birth Rate vs. Net Population Growth:**

**Purpose:** In order to examine the correlation between birth rates and net population change.

**Insights**: A positive relationship is observed and proved since the regions with high birth have greater population increment as compared to low birth rate regions supporting the argument that birth rate is a crucial determinant of demographic change in the regions.

## 3.3 Useful information for the data understanding

The plots show that the Russian regions’ total fertility rates are generally average, and only several regions are characterized by higher rates. Net population growth presents movements in several areas as a growth and decline movement in other areas. High positive correlation between birth rates and net population growth underlines the importance of birth rates in the demographic changes.

# 4. Analysis

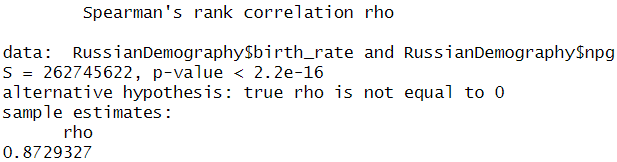
## 4.1 Statistical test

**Test Used:** Spearman’s Correlation Coefficient.

**Choice of Test:** Spearman’s correlation quantifies the form and intensity of a straight relationship between two quantitative variables, namely, birth rate and net population growth (NPG).

**Appropriateness:** This test is proper for the research question because it reveals if and to what extent birth rates affect NPG in Russian regions. Before proceeding with the analysis, a check of the test assumptions is required: the residuals are normally distributed and linear.

**Output:** It shows that NPG has a good positive connection with birth rates based on the perfect coefficient value of 0.8729. The p-value in the present study is less than 0.05 (2.2e-16) thus affirming the significance of the finding at 95% confidence level.



## 4.2 Null hypothesis is rejected /not rejected based on the p-value

As per the given p-value of which is lesser than the significance level of 0.05, the null hypothesis is rejected. The hypothesised null hypothesis in this case states that birth rate does not have a linear causal association with net population growth. Nevertheless, results proved a positive correlation of Birth rate with net population increase in Russian regions affirmed by the value r = 0.8915 which indicates that birth rates correlate highly positively with net population increase. This result supports the alternative hypothesis that birth rate has a significant impact on population trends, and thus the theory should be a concern among the regions in the world.

# 5. Evaluation – group’s experience at 7COM1079

## 5.1 What Went Well

The group co-operated well in tackling assignment requirements which involved working in a team and sharing different abilities in order to make an interpretation of demographic data. This made communication strong, and tasks were well assigned according to specialization hence efficient results were produced. This increased the efficiency of analysis thanks to the use of statistical tools and visualization techniques used. The team members also quickly addressed issues which arose in order to stay on schedule for the project.

## 5.2 Points for improvement

Considering the fact that the project was a success, one could have hoped for pre-definition of roles to be paid more attention and data pre-processing could have been made less of a weighty task. There were problems related to communication process and coordination such as confusion of tasks, work overlapping and eventual delays. Possibly the documentation of the process and discoveries could also increase knowledge export and help in the replication of the study in future.

## 5.3 Group’s time management

The group did adequate progress and was able to meet most due dates set. This check pointed out goals, milestones and any issues arising through the process into the monitoring and evaluation agenda. However, better planning could have avoided these changes at the last minute withholding the progress throughout the project.

## 5.4 Project’s overall judgement

The project was completed successfully achieving the aims with useful information on the population of Russia. However, minor issues noted here did not significantly set a tone of problems in their inter group activities and analytical skills rendered out good results of how the team was capable of handling future research work.

# 6. Conclusions

## 6.1 Results

The analysis gives us t = 94.662 and, the relationship between Russian region birth rates and net population increase is positively significant, with r = 0.8915, and p-value = < 2.2e-16. This supports current population growth theories declaring that birth rates are the direct factors which influence population growth rates in the regions.

## 6.2 Interpretation of the result

The results will contribute to the answer to the research question and show that the birth rate is one of the main factors influencing the net population growth in Russian regions. In relation to the aspects discussed above, this has implications for policymaker’s key for them to debate ways of reversing low birth rate so that more populations can be supported to alleviate or grow for sustainable development.

## 6.3 Reasons and/or implications for future work, limitations

Future work should be designed to assess temporal changes and to consider other factors such as migration and mortality. Limitations include the use of cross sectional data and there is no possibility to control for all related factors. Coordination of these gaps may offer further understanding of demographic concerns and policy solutions.

# 7. References

‌Arkhangelskiy, V., Sivoplyasova, S., Moiseeva, E. and Arkhangelskiy, N. (2024). Exploring Fertility Dynamics and Factors Shaping Russia’s Demographic Prospects. *Changing Societies & Personalities*, [online] 8(2), pp.291–312. doi: https://doi.org/10.15826/csp.

‌Kazenin, K. (2021). *Birth Rate In Russia In 2020: Regional Dynamic*. [online] Ssrn.com. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract\_id=3818100&utm\_source=chatgpt.com [Accessed 3 Jan. 2025].

‌Shubat, O. and Bagirova, A. (2019). *Dynamics Modeling And The Study Of Birth Rate Determinants In Russian Regions*. [online] ECMS 2019 Proceedings edited by Mauro Iacono, Francesco Palmieri, Marco Gribaudo, Massimo Ficco. Available at: https://www.academia.edu/70443236/Dynamics\_Modeling\_And\_The\_Study\_Of\_Birth\_Rate\_Determinants\_In\_Russian\_Regions?utm\_source=chatgpt.com [Accessed 3 Jan. 2025].

SiteSoft (2024). *Federal State Statistics Service — Short-term Economic Indicators*. [online] Rosstat.gov.ru. Available at: https://eng.rosstat.gov.ru/folder/205320 [Accessed 3 Jan. 2025].

‌

‌

‌

# 8. Appendices

## A. R code used for analysis and visualization

## B. GitHub log output