

Выявление зависимости рождаемости в России от различных факторов с помощью языка R

Радионов Сергей Владимирович

Приамурский государственный университет им. Шолом-Алейхема

Студент

Баженов Руслан Иванович

Приамурский государственный университет им. Шолом-Алейхема

к.п.н., доцент, зав. кафедрой информационных систем, математики и правовой информатики

Аннотация

В данной статье выявлены зависимости рождаемости в России от различных факторов. Представлены графики, отображающие разные зависимости. Благодаря полученным результатам можно эффективно влиять на фактор рождаемости в России. Решение осуществлено с помощью языка R и среды RStudio.

Ключевые слова: R, RStudio, рождаемость, Россия, дети

Revealing the dependence of the birth rate in Russia on various factors using the language R

Radionov Sergey Vladimirovich

Sholom-Aleichem Priamursky State University

Student

Bazhenov Ruslan Ivanovich

Sholom-Aleichem Priamursky State University

Associate professor of Department of Information Systems, Mathematics and Law Informatics

Annotation

In this article, the dependence of the birth rate in Russia on various factors is revealed. There are graphs showing different dependencies. Thanks to the results obtained, it is possible to effectively influence the fertility factor in Russia. The solution is implemented using the language R and the RStudio environment.

Keywords: R, RStudio, fertility, Russia, child.

В мире существует множество явлений, статистические данные которых собираются на протяжении многих лет. С помощью интеллектуальных методов можно проанализировать эти данные и выявить закономерности, которые помогут спрогнозировать будущие явления и положительно повлиять

на них. В зависимости какие данные мы собираемся анализировать, то нам необходимо выбрать для этого наиболее эффективный интеллектуальный метод, который даст правдивый прогноз самих данных на будущий период и покажет в них зависимость.

Целью исследования является выявление зависимости рождаемости в России с использованием интеллектуального метода линейной регрессии в системе R.

В 2014 году М.А. Абызов, А.В. Коротаев изучали рост рождаемости как фактор подталкивания экономического роста в наиболее экономически развитых странах [1]. Ю.А. Григорьев и С.Н. Добродеев исследовали влияние конфессионального фактора на рождаемость в различных регионах мира в 2008 году [2]. В 2010 году О.В. Кучмаева и О.Л. Петрякова изучали дифференциацию показателей рождаемости в России и роль демографических и социальных факторов [3], а А.Р. Рахимова и Е.В. Кабашова исследовали влияние факторов на изменение рождаемости в 2011 году [4]. Также можно выделить и другие исследования [5-7]. Не менее значимо имитационное моделирование и в англоязычном сегменте [8-9].

Для анализа и выявления зависимостей данных рождаемости в России используем интеллектуальный метод под названием «Метод линейной регрессии». Данный метод позволяет построить регрессионную модель, в которой мы можем увидеть зависимость этих данных за каждый год и как они менялись, а также построить саму линию регрессии, которая покажет, что с данными будет происходить дальше, то есть какой будет вероятный их прогноз и в каком промежутке.

Обработка данных будет осуществляться с помощью языка программирования R, использующийся для статистического анализа, а также в нем есть поддержка построения различных графиков. Программное обеспечение, на котором понадобится использовать данный метод называется «RStudio», сама программа обрела популярность, благодаря ее понятному интуитивному интерфейсу и в ней гораздо удобнее работать, так как она имеет ряд инструментов, которые понадобятся для статистического анализа. Для языка программирования R создано большое количество библиотек, упрощающих работу в сфере анализа данных.

Данные для анализа взяты с различных сайтов, содержащих статистические наблюдения.

С помощью программы «RStudio» используем метод линейной регрессии на данные рабочих мест в Лондоне и Великобритании. Для начала выявим чему равен R^2 используя функцию «lm» рабочих мест в первой группе.

```

Call:
lm(formula = child ~ mother.capital, data = a)

Residuals:
    Min       1Q   Median       3Q      Max
-0.07058 -0.02643  0.01074  0.02277  0.05133

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  9.257e-01  6.090e-02  15.20 3.47e-07 ***
mother.capital 1.893e-06  1.722e-07  10.99 4.17e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.03929 on 8 degrees of freedom
Multiple R-squared:  0.9379,    Adjusted R-squared:  0.9301
F-statistic: 120.8 on 1 and 8 DF,  p-value: 4.171e-06

```

Рис.1 Результат использования функции «lm»

На рисунке 1 представлен результат использования функции «lm» для данных количества детей на женщину и материнского капитала. $R^2 = 0,9301$, что указывает на хорошую связь между этими данными, так как $R^2 > 0,8$. Уравнение регрессии для данной зависимости: $Y=0,926 + 0,00000189x$

Зависимость количества детей на женщину от других данных:

1. ВВП России $R^2 = 0,7534$, что указывает на среднюю связь. Уравнение регрессии: $Y = -0,589 + 0,0034x$.

2. Ипотечная ставка $R^2 = 0,3484$, что показывает очень слабую зависимость.

3. Количество покупок новых машин $R^2 = -0,1243$, наблюдается отсутствие связи.

Для более подробного рассмотрения зависимости с методом линейной регрессии необходимо построить графики. Начнем с материнского капитала.

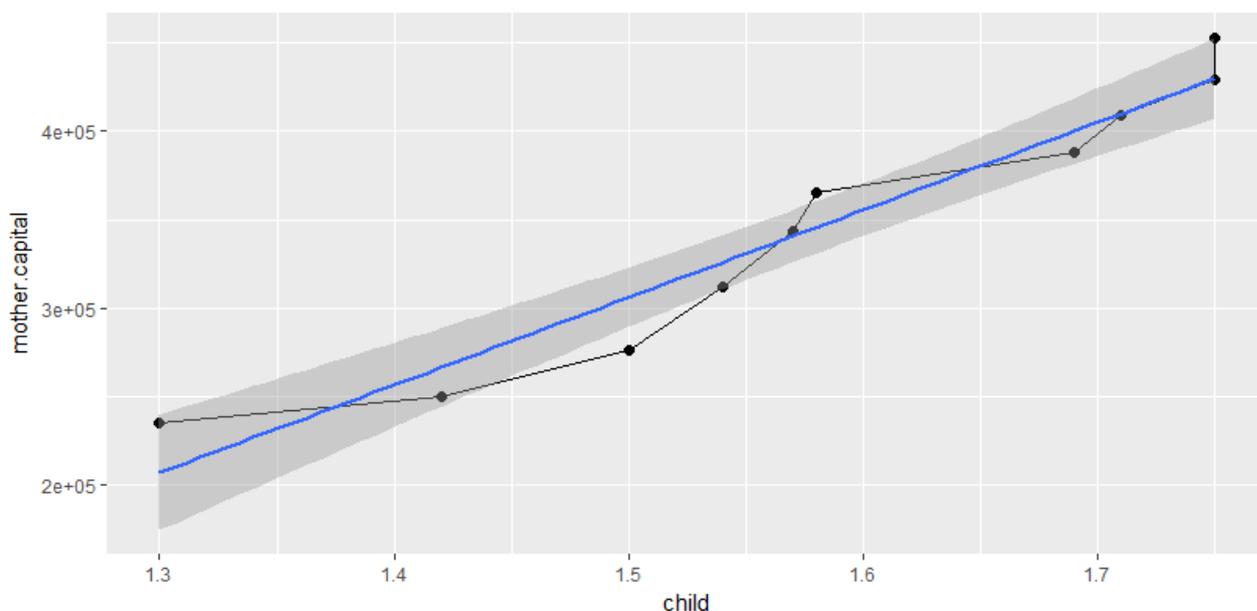


Рис.2 График зависимости рождаемости от материнского капитала

На рисунке 2 изображен график, в котором отображена зависимость между средним количеством детей на женщину и материнским капиталом. Точки графика являются значениями на начало каждого года с 2006 по 2015. На графике видно, что точки не сильно отклонены от линии тренда и постепенно растут, что показывает линейная регрессия, изображенная линией тренда, которая стремится вверх.

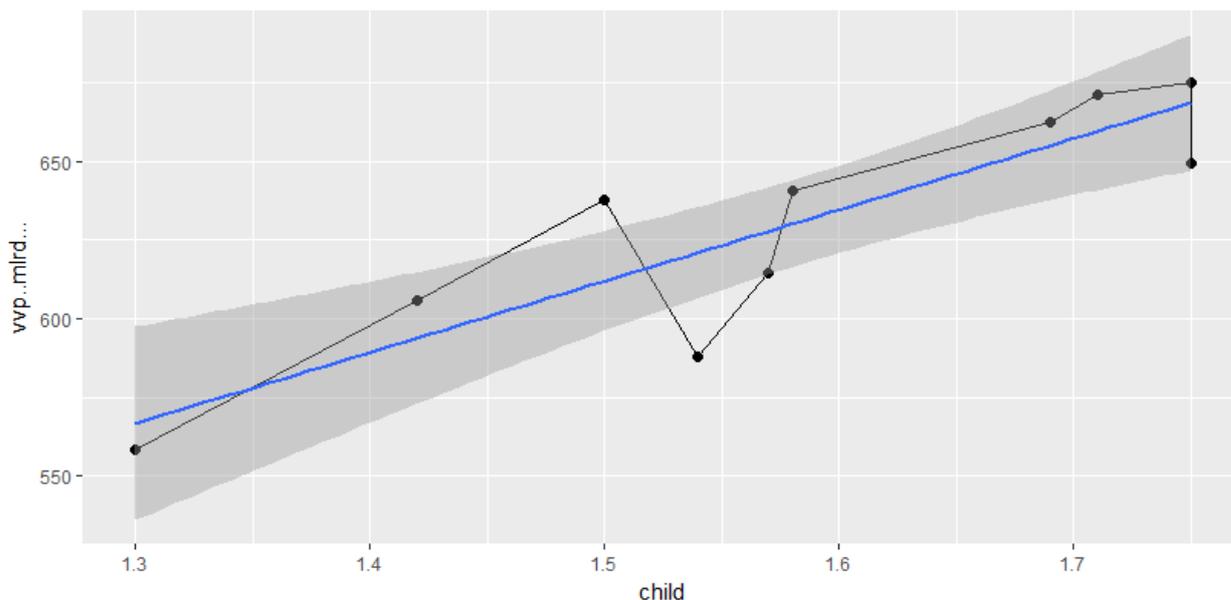


Рис.3 График зависимости рождаемости от ВВП России

На графике из рисунка 3 показана зависимость между средним количеством детей на женщину и ВВП России. На рисунке также наблюдается незначительное отклонение от линии тренда.

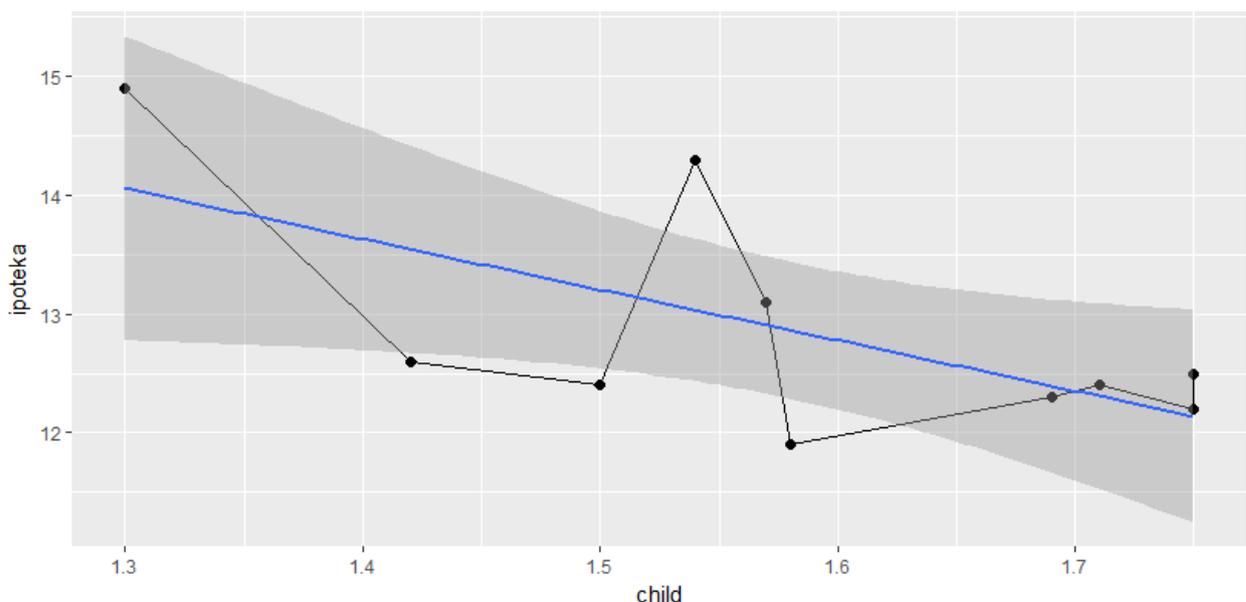


Рис.4 График зависимости рождаемости от ипотечной ставки

Далее идет график зависимости рождаемости от ипотечной ставки (Рис.4). На нем отклонение точек от линии тренда уже значительно, что указывает на слабую связь анализируемых данных.

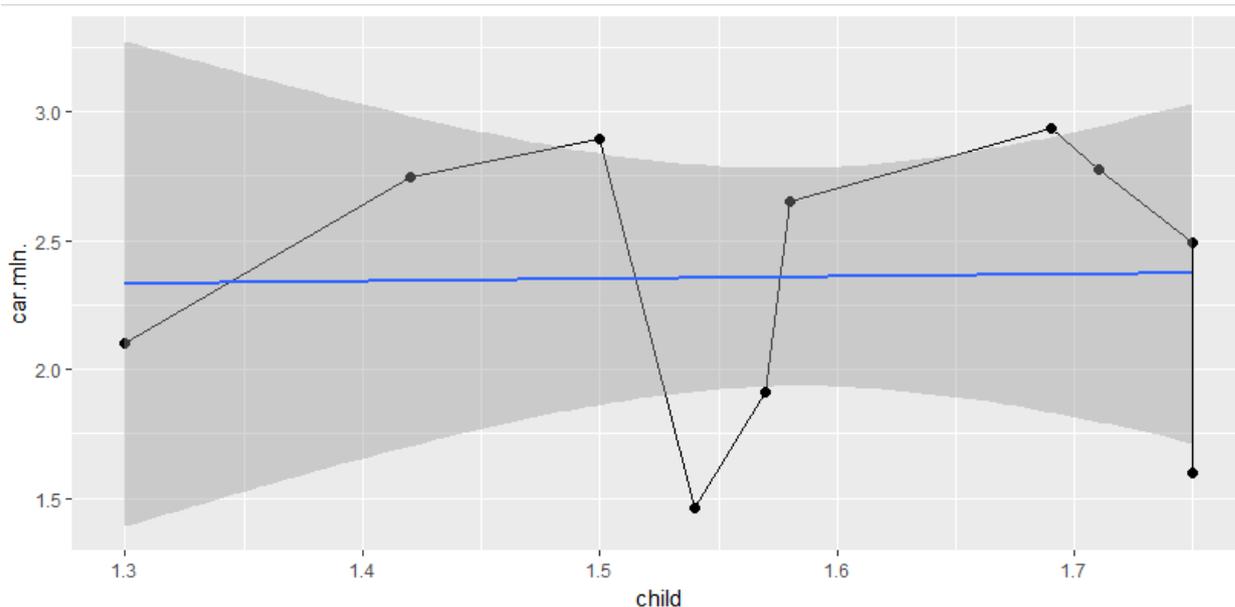


Рис.5 График зависимости рождаемости от покупок новых машин

Посмотрев на график зависимости рождаемости от покупок новых машин, можно визуально определить отсутствие связи между рассматриваемыми параметрами, т.к. линия тренда не возрастает и не убывает.

Можно подвести итог исследования. Рассмотрев зависимость количества детей на женщину в России за период 2006-2015гг. можно сделать вывод, что положительное влияние на рост этого параметра оказало возрастание материнского капитала, в меньшей мере рост ВВП. Влияние ипотечной ставки оказалось довольно низко. Таким образом, чтобы влиять на рождаемость, наиболее эффективно будет изменять материнский капитал.

Библиографический список

1. Абызов М.А., Коротаев А.В. О росте рождаемости как факторе подталкивания экономического роста в наиболее экономически развитых странах // В сборнике: Мировая динамика: закономерности, тенденции, перспективы Акаев А.А., Коротаев А.В., Малков С.Ю. Российская академия наук, Институт востоковедения, МГУ им. М. В. Ломоносова; отв. ред. А. А. Акаев, А. В. Коротаев, С. Ю. Малков. Москва, 2014. С. 274-290.
2. Григорьев Ю.А., Добродеев С.Н. Влияние конфессионального фактора на рождаемость в различных регионах мира // В сборнике: Проблемы общественного здоровья и экологии человека: новые закономерности Материалы XLIII научно-практической конференции с международным участием "Гигиена, организация здравоохранения и профпатология" и IV Межрегиональной научно-практической конференции с международным

- участием "Медико-социальное обслуживание и реабилитация пожилых людей и инвалидов". Российская академия медицинских наук, Сибирское отделение, Кузбасский научный центр, ГУ НИИ комплексных проблем гигиены и профессиональных заболеваний. Кемерово, 2008. С. 36-40.
3. Кучмаева О.В., Петрякова О.Л. Дифференциация показателей рождаемости в России: роль демографических и социальных факторов // Вопросы статистики. 2010. № 10. С. 21-25.
 4. Рахимова А.Р., Кабашова Е.В. Исследование влияния факторов на изменение рождаемости // В сборнике: Актуальные вопросы экономико-статистического исследования и информационных технологий сборник научных статей: посвящается 40-летию создания кафедры "Статистики и информационных систем в экономике". МСХ РФ, Башкирский государственный аграрный университет. Уфа, 2011. С. 329-330.
 5. Орешников В.В. Прогнозирование изменения уровня рождаемости под влиянием социально-экономических факторов // В сборнике: Инновационные технологии управления социально-экономическим развитием регионов России Материалы V Всероссийской научно-практической конференции с международным участием: в 3-х частях. Ответственный редактор Гайнанов Д.А. 2012. С. 264-270.
 6. Шишкина М.А. Измерение факторов детерминации современных тенденций рождаемости в северных регионах России // Корпоративное управление и инновационное развитие экономики Севера: Вестник Научно-исследовательского центра корпоративного права, управления и венчурного инвестирования Сыктывкарского государственного университета. 2017. № 4. С. 118-125.
 7. Журавлева Т.Л., Гаврилова Я.А. Анализ факторов рождаемости в России: что говорят данные РМЭЗ НИУ ВШЭ? // Экономический журнал Высшей школы экономики. 2017. Т. 21. № 1. С. 145-187.
 8. Olesen A. B. et al. Atopic dermatitis and birth factors: historical follow up by record linkage // *Bmj.* – 1997. – Т. 314. – №. 7086. – С. 1003.
 9. Coren S., Porac C. Birth factors and laterality: Effects of birth order, parental age, and birth stress on four indices of lateral preference // *Behavior genetics.* – 1980. – Т. 10. – №. 2. – С. 123-138.