

*Назина Софья Леонидовна,  
студентка 4 курса,  
направление 09.03.03 «Прикладная информатика»,  
Белгородский государственный национальный исследовательский  
университет  
Россия, г. Белгород  
e-mail: 1319279@edu.bsu.ru*

*Научный руководитель: Гахова Нина Николаевна,  
доцент,  
Белгородский государственный национальный исследовательский  
университет  
Россия, г. Белгород*

## **АНАЛИЗ ЗАВИСИМОСТИ ЧИСЛА НАСЕЛЕНИЯ ПО ВОЗРАСТНЫМ ГРУППАМ В РОССИИ**

*Аннотация. В статье проведен корреляционно-регрессионный анализ зависимости числа населения по возрастным группам в России с помощью ПО Microsoft Excel. Корреляционно-регрессионный анализ рассмотрен, как способ анализа статистических показателей.*

**Ключевые слова:** корреляционный анализ, регрессионный анализ, Microsoft Excel, населения РФ, возрастная группа, коэффициент корреляции.

*Nazina Sofya Leonidovna,  
4th year student,  
Belgorod State National Research University  
Russia, Belgorod*

*Scientific adviser: Gahova Nina Nikolaevna,  
associate professor,  
Belgorod State National Research University  
Russia, Belgorod*

## **ANALYSIS OF THE DEPENDENCE OF THE POPULATION BY AGE GROUPS IN RUSSIA**

*Abstract: The article presents a correlation and regression analysis of the dependence of the population by age groups in Russia using Microsoft Excel software. Correlation and regression analysis is considered as a way of analyzing statistical indicators.*

**Keywords:** correlation analysis, regression analysis, Microsoft Excel, population of the Russian Federation, age group, correlation coefficient.

Статистика рассматривается как наука о методах изучения массовых явлений. Статистика населения является одним из самых обширных разделом статистики. Перепись населения является главным источником сведений о численности населения в Российской Федерации [1]. Население РФ представляет собой объект всестороннего исследования и является участником производственного процесса и потребителем его результатов. Более того, интерес к статистическому изучению населения не снижается, а возрастает все больше с каждым годом. В исследованиях обращается внимание на различные показатели: численность, изменение численности, пол, возраст и др. Для того, чтобы охарактеризовать различные процессы, происходящие в популяции, надо обратить внимание на различные показатели. Прежде всего, это численность, ее изменение и структура (пол, возраст и т.д.).

В данной работе для проведения анализа была использована статистика распределения населения по возрастным группам в России за 2020г., часть данной статистики приведена на рисунке 1 [2].

Возраст	Общее население	Мужчины	Женщины
0	1 476 347	759 386	716 961
1	1 599 323	824 719	774 604
2	1 685 166	866 044	819 122
3	1 883 092	967 291	915 801
4	1 935 184	994 677	940 507
5	1 937 337	996 077	941 260
6	1 920 005	985 682	934 323
7	1 927 854	990 917	936 937
8	1 823 754	936 961	886 793
9	1 700 467	872 491	827 976
10	1 721 231	882 728	838 503
11	1 712 873	876 491	836 382
12	1 621 157	831 176	789 981

***Рисунок 1 – Статистика***

Выбранные статические данные были перенесены в Excel. После чего проведены 3 вида анализа: регрессионный, корреляционный, корреляционно-регрессионный анализ.

Регрессионный анализ представляет собой методы оценки отношений между переменными. Он используется для оценки степени взаимосвязи между переменными и для моделирования будущей зависимости. По сути, регрессионные методы показывают, как по изменениям независимых переменных можно зафиксировать изменение «зависимой переменной» [3].

Регрессионный анализ со временем становился все более востребованным, чем корреляционный анализ, т.к. является более информативным. На основе данного анализа можно выделять приоритеты, а также прогнозировать, планировать развитие приоритетных направлений или принимать управленческие решения.

Регрессионный анализ рассмотренной статистики был выполнен в Microsoft Excel с помощью инструмента «Анализ данных», который является надстройкой, содержит коллекцию функций и инструментов, расширяющих встроенные аналитические возможности Excel[3]. Результаты регрессионного анализа представлены на рисунке 2.

Вывод итогов								
Регрессионная статистика								
Множественный R	0,679405839							
R-квадрат	0,561592295							
Нормированный R-квадрат	0,456153833							
Стандартная ошибка	21,6076823							
Наблюдения	101							
Дисперсионный анализ								
	df	SS	MS	F	Значимость F			
Регрессия	1	39627,6985	39627,6985	84,87552596	5,70194E-15			
Остаток	99	46222,3015	466,8919344					
Итого	100	85850						
	Коэффициенты	Стандартная ошибка	t-статистика	P-значение	Нижние 95%	Верхние 95%	Нижние 95,0%	Верхние 95,0%
Y-пересечение	185246,4029	28229,27774	6,562208378	2,44921E-09	129233,3915	241259,414	129233,3915	241259,4143
Переменная X 1	-0,723343552	0,036150138	51,99206932	1,10562E-73	1,807790763	1,9512502	1,807790763	1,951250196

**Рисунок 2 – Регрессионный анализ**

В полученных результатах основным показателем является коэффициент детерминации (R-квадрат), показывающий качество модели. В рассматриваемом случае он равен 0,56 (56%), что является хорошим качеством. Стоит отметить, что значение меньше 0,5 является плохим показателем, а выше 0,8 хорошим.

Показатель, расположенный в ячейке на пересечении строки «Y-пересечение» и столбца «Коэффициенты», показывает какое значение будет иметь Y при всех остальных факторах равных нулю. В данном случае показатель равен 185246.

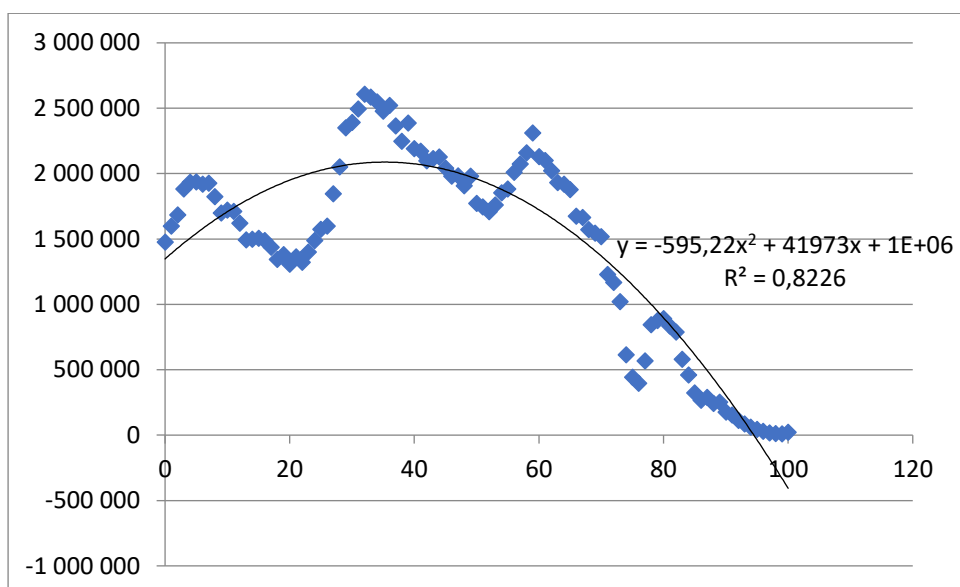
Значение на пересечении линий «Переменная X1» и «Коэффициенты» показывает уровень зависимости Y от X. В нашем случае — это уровень населения РФ от возраста. В рассмотренном случае коэффициент -0,72. Минус указывает на отрицательное влияние. И это действительно так, ведь начиная с 35-36 лет общее количество населения уменьшается.

Корреляционный анализ помогает установить, есть ли между показателями в одной или двух выборках связь. Если связь имеется, то влечет ли увеличение одного параметра повышение (положительная корреляция) либо уменьшение (отрицательная) другого. Корреляционный анализ помогает аналитику определиться, можно ли по величине одного показателя предсказать возможное значение другого.[3]

Коэффициент корреляции варьируется в пределах от +1 до -1. Стоит отметить, что при значении коэффициента 0 линейной зависимости между выборками не существует.[4]

Корреляционный анализ рассмотренной статистики был выполнен в Microsoft Excel с помощью функции КОРРЕЛ. Результаты корреляционного анализа составляют -0,67941. Данный коэффициент показывает, что случайные величины зависимы. В данном случае коэффициент равен -0,67, что близко к -1. Полученная связь является отрицательной, т.е. как при увеличении возраста уменьшается количество населения.

Как правило, на практике рассмотренные выше методики применяются вместе. Для данного анализа необходимо на основе статических данных было построено корреляционное поле с линией тренда и величиной достоверности аппроксимации. [4]. На рисунке 3 представлена полиномиальная линия тренда.



***Рисунок 3 – Корреляционное поле с полиномиальной линией тренда***

Из рисунка 3 можно прийти к выводу, что линия тренда располагается достаточно близко к данным, а коэффициент достоверности равен 0,8226 (82%). Уравнение зависимости линии тренда:  $y = -595,22x^2 - 41973x + 1 \cdot 2,718 + 06$ .

Таким образом, проведя анализ между возрастом и количеством определенной возрастной группы в России, следует, что возраст граждан влияет на численность населения в России в большей степени, нежели иные показатели.

### **Список литературы:**

1. Гончарова Н.П. Демографическая динамика и оценка результативности демографической политики в регионах России // Государственный советник. 2019. № 3 (27). С. 101-108.
2. Распределение населения России. [Электронный ресурс] // Режим доступа: URL: <https://rusind.ru/raspredelenie-naseleniya-rossii.html> (дата обращения 05.10.2021 г.).
3. Корреляционно-регрессионный анализ. [Электронный ресурс] // Режим доступа: URL: <https://exceltable.com/korrelyacino-regressiony> (дата обращения 01.10.2021 г.).

4. Регрессионный анализ. [Электронный ресурс] // Режим доступа:  
URL: [http://www.machinelearning.ru/wiki/index.php?title=Регрессионный\\_анализ](http://www.machinelearning.ru/wiki/index.php?title=Регрессионный_анализ)  
(дата обращения 01.10.2021 г.).