

УДК 332.012

***АНАЛИЗ ЗАВИСИМОСТИ СРЕДНЕДУШЕВЫХ ДОХОДОВ НАСЕЛЕНИЯ ОТ
СОЦИАЛЬНО-ЭКОНОМИЧЕСКИХ ФАКТОРОВ РЕГИОНОВ РОССИИ С
ПОМОЩЬЮ РЕГРЕССИОННОЙ МОДЕЛИ***

Картавченко А. Е.,

магистр,

НИУ Высшая школа экономики,

Санкт-Петербург, Россия

Аннотация

Данная работа посвящена изучению распределения среднедушевых доходов населения по регионам России, а также поиску и анализу социально-экономических показателей, оказывающих на наибольшее влияние на доходы населения. Повышение среднедушевых доходов граждан России является одной из первоочередных задач, поставленных перед Правительством РФ Президентом РФ, поэтому данная тематика сегодня актуальна. Методы статистики и регрессионного анализа, применяемые в данной работе, позволяют не просто проследить качественные взаимосвязи между показателями, но и получить их количественную оценку.

Ключевые слова: регионы Российской Федерации, среднедушевые доходы населения, математические методы в экономике, статистический анализ данных, регрессионный анализ

ANALYSIS OF THE DEPENDENCE OF THE AVERAGE PERSONAL INCOME

***OF THE POPULATION ON THE SOCIO-ECONOMIC FACTORS OF THE
REGIONS OF RUSSIA USING A REGRESSION MODEL***

Kartavchenko A.E.,

Master

National Research Institute Higher School of Economic,

Saint-Petersburg, Russia

Abstract

This work is devoted to the study of the distribution of average per capita incomes of the population in the regions of Russia, as well as the search and analysis of socio-economic indicators that have the greatest impact on the income of the population. Increasing the average per capita income of Russian citizens is one of the priority tasks set for the Government of the Russian Federation by the President of the Russian Federation, therefore, this topic is relevant today. The methods of statistics and regression analysis used in this work make it possible not only to trace the qualitative relationships between indicators, but also to obtain their quantitative assessment.

Keywords: regions of the Russian Federation, average per capita income of the population, mathematical methods in economics, statistical data analysis, regression analysis

В Российской Федерации традиционно присутствует проблема с неравномерным распределением доходов и неравенством населения. Эта проблема проявляется как внутри одного региона, так и масштабе всей страны на уровне различных регионов. Среднедушевые доходы населения – это один из наиболее важных статистических показателей, характеризующих уровень жизни населения,

поэтому именно повышение этого показателя является задачей, решением которой занимается Правительство РФ. Государство не может повысить этот показатель напрямую, однако может разрабатывать различные государственные программы и инициативы по поддержке населения и модернизации экономики [6]. Однако современная экономика имеет очень сложную структуру, в которой экономические и социальные показатели тесно взаимосвязаны, поэтому государственное воздействие должно использовать и опираться на выявленные взаимосвязи. Современные методы статистики и математики, а также технологии работы с большим объёмом данных позволяют строить математические модели, которые с достаточной степенью точности описывают реальное устройство экономики и обнаруживают скрытые взаимосвязи. В данной работе построена регрессионная модель, описывающая зависимость среднедушевых доходов населения России от социально-экономических факторов определённого региона.

Для первичного анализа в статье использовались 54 социально-экономических показателя для 85 регионов России. Эти данные были собраны Федеральной службой государственной статистики за 2015-2017 годы. Таким образом для каждого региона России имелось по 3 наблюдения. В данной работе среднедушевые доходы населения также будут заменяться в тексте термином – «целевой показатель».

Анализ целевого показателя начнём с изучения распределения по регионам России.



Рис. 1. Распределение среднедушевых доходов населения по регионам РФ

Источник: Составлено автором на основании статистических данных по субъектам РФ за 2015-2017 года.

На рисунке 1 можно видеть, что распределение целевого показателя значительно отличается от нормального распределения. Также видно, что в небольшом числе регионов РФ среднедушевые доходы сильно больше, чем в среднем по стране. Несбалансированность и асимметричность данного распределения подтверждается и расчётами. Среднее значение целевого показателя равно 28328, а медианное – 25313. Такая разница (более 10%) между медианным и средним значением явно говорит нам о неравномерном и смещённом распределении среднедушевых доходов. Разобьём имеющуюся выборку на квартили и рассмотрим примеры регионов, попадающих в различные квартили [5].

Таблица 1. Категоризация регионов по квартилям распределения

Квартиль	Пример регионов	Количество регионов
Q1	Костромская область, Псковская область, Оренбургская область, Кемеровская область, республика Ингушетия	22
Q2	Владимирская область, республика Бурятия, Орловская область, Астраханская область	28
Q3	Белгородская область, Ленинградская область, республика Карелия, Красноярский край	23
Q4	Московская область, г. Москва, Ханты-мансийский автономный округ	12

Источник: Составлено автором на основании статистических данных по субъектам РФ за 2015-2017 года.

В таблице 1 приведено распределение регионов по квартилям целевого показателя. В Q4 находятся регионы с самыми высокими доходами на душу населения – г. Москва, регионы, в которых происходит добыча полезных ископаемых. Регионы квартилей Q2 и Q3 можно отнести к средним по уровню целевого показателя. Доходы в этих регионах в целом соответствуют среднероссийскому уровню. Однако Q2 плавно переходит в Q1, где уже расположились наименее благополучные и наиболее проблемные с точки зрения уровня доходов регионы РФ. Это некоторые области в Центральной России и республики Северного Кавказа.

Для поиска статистической взаимосвязи между целевым показателем и имеющимися в наличии 54 социально-экономическими характеристиками регионов воспользуемся коэффициентом корреляции Пирсона:

$$r_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} \quad (1)$$

Где x_i, y_i – элементы выборок из изучаемой пары показателей.

Данный коэффициент принимает значения от -1 до 1. Значения близкие по модулю к 1 свидетельствуют о сильной взаимосвязи между парой показателей, близкие же к 0 – об отсутствии такой взаимосвязи. Положительные значения коэффициента говорят о наличии положительной взаимосвязи между показателями, то есть в среднем при увеличении одного из пары показателей увеличивается и второй. Противоположная ситуация наблюдается в случае отрицательных значений коэффициента.

Из имеющегося набора социально-экономических показателей выберем те, парный коэффициент корреляции которых с целевым показателем превосходит 0,5 по модулю.

Таблица 2. Величина коэффициентов корреляции Пирсона между показателями и среднедушевыми доходами населения

Показатель	r_{xy}
Величина прожиточного минимума	0,82
Объём платных услуг на душу населения	0,8
Валовой региональный продукт на душу населения	0,78
Инвестиции в основной капитал на душу населения	0,68

ЭЛЕКТРОННЫЙ НАУЧНЫЙ ЖУРНАЛ «ВЕКТОР ЭКОНОМИКИ»

Уровень занятости населения	0,67
Уровень участия в рабочей силе	0,64
Процент населения в трудоспособном возрасте	0,62
Коэффициент демографической нагрузки	-0,61
Процент населения с денежными доходами ниже прожиточного минимума	-0,54
Добыча полезных ископаемых на душу населения	0,53

Источник: Составлено автором на основании статистических данных по субъектам РФ за 2015-2017 года.

В таблице 2 приведены 10 наиболее коррелированных с целевым показателем факторов. Для всех них коэффициент корреляции Пирсона превосходит 0,5 по модулю. Такие значения позволяют говорить о наличии статистически значимой взаимосвязи. Важно отметить, что в таблице 2 присутствуют как положительно, так и отрицательно коррелированные со среднедушевыми доходами показатели. Рассмотрим некоторые показатели совместно с целевым на одном графике, чтобы лучше понять причину найденных взаимосвязей.

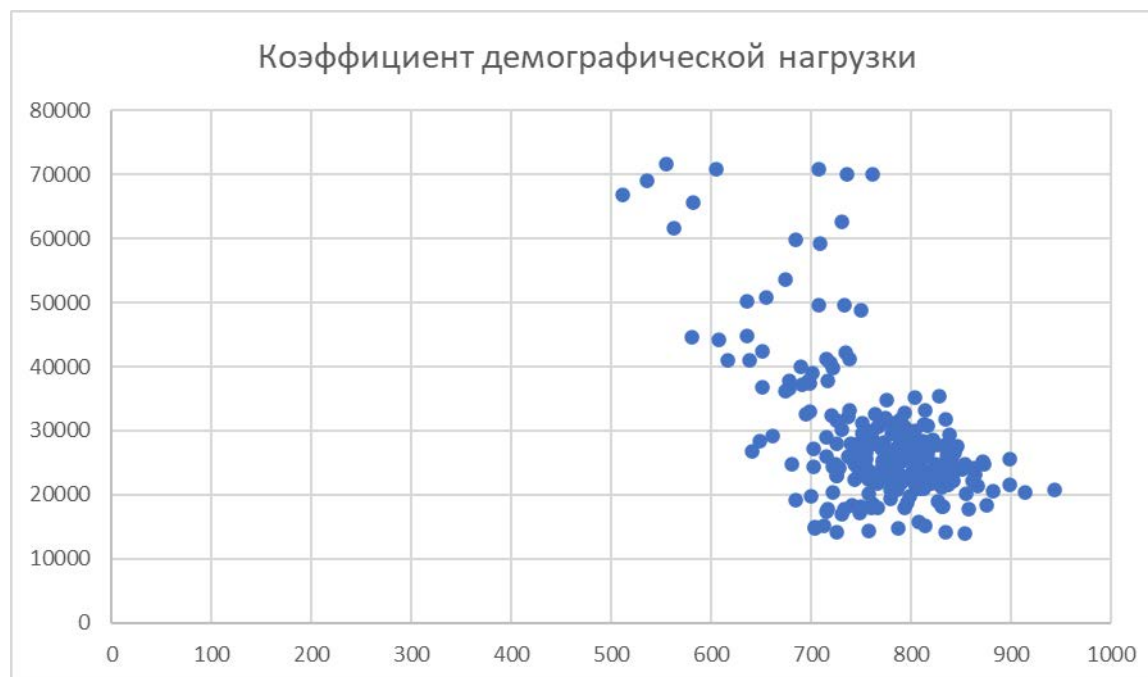


Рис. 2. Зависимость среднедушевых доходов населения от коэффициента демографической нагрузки

Источник: Составлено автором на основании статистических данных по субъектам РФ за 2015-2017 года.

На рисунке 2 изображена зависимость целевого показателя от коэффициента демографической нагрузки. На графике прослеживается негативная взаимосвязь двух показателей. Регионы с наименьшим коэффициентом демографической нагрузки и наибольшим размером среднедушевых доходов – Ханты-Мансийский автономный округ, Чукотский автономный округ и Ямало-Ненецкий автономный округ. Москва имеет также высокий уровень среднедушевых доходов, однако коэффициент демографической нагрузки там выше. Регионом с наибольшим коэффициентом демографической нагрузки является Курганская область. В целом регионы РФ достаточно однородны по коэффициенту демографической нагрузки.

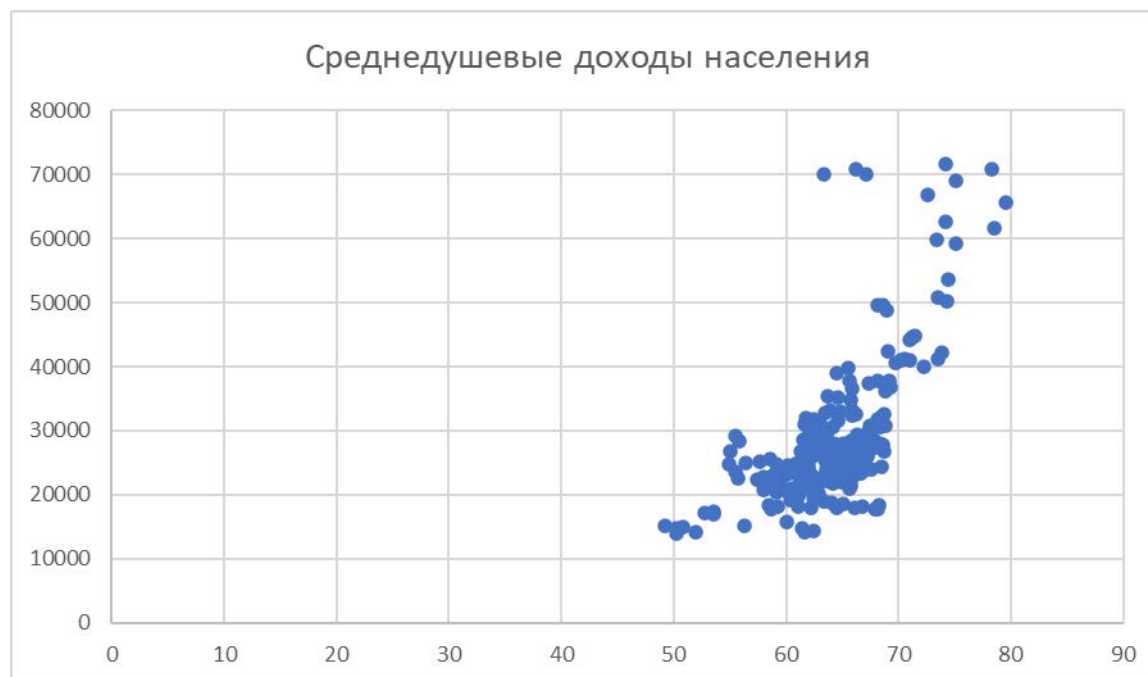


Рис. 3 Зависимость среднедушевых доходов населения от уровня занятости населения

Источник: Составлено автором на основании статистических данных по субъектам РФ за 2015-2017 года.

На рисунке 3 показана зависимость между Среднедушевыми доходами населения и уровнем занятости населения. Обращает внимание на себя группа из 3 точек в диапазоне 60-70% занятости населения и высоким уровнем среднедушевого дохода. Эти точки соответствуют Ненецкому автономному округу за 3 года наблюдения. Это можно объяснить тем, что в данном регионе на средний уровень дохода скорее влияет факт добычи полезных ископаемых, нежели средний уровень занятости населения региона. Традиционно высокий уровень занятости и дохода приходится на Москву, Ямало-Ненецкий автономный округ и Чукотский автономный округ [4]. Наименьшие значения занятости и

дохода приходится на республику Ингушетия, Карачаево-Черкесскую республику и республику Тыва.

Проанализировав часть из найденных зависимостей, можно переходить к построению регрессионной модели [2]. Все показатели из таблицы 2 используются как независимые переменные, в качестве же прогнозируемой переменной выступает среднедушевые доходы населения.

Математическая модель линейной регрессии имеет вид:

$$y = \sum_{i=1}^n w_i x_i + w_0 + \varepsilon \quad (2)$$

Где x_i – независимые переменные модели

y – целевой показатель

w_i, w_0 - определяемые весовые коэффициенты

n – количество независимых переменных в модели

ε – неустраняемая ошибка модели

Для всего имеющегося набора наблюдений модель (2) можно записать в матричной форме:

$$y = Xw + \varepsilon \quad (3)$$

Где $y = (y^1, \dots, y^l)^T$ – вектор целевых показателей

$$X = \begin{pmatrix} 1 & \dots & x_n^1 \\ \dots & \dots & \dots \\ 1 & \dots & x_n^l \end{pmatrix} - \text{матрица независимых переменных}$$

$$w = (w_0, \dots, w_n)^T$$

ε – неустраняемая ошибка

l – количество наблюдений в выборке

Коэффициенты в (2) и (3) определяются согласно методу наименьших квадратов, решая соответствующую задачу оптимизации:

$$\frac{1}{2l} \sum_{i=1}^l (y^i - w^T X_i)^2 = \frac{1}{2l} (y - Xw)^T (y - Xw) \xrightarrow{w} \min \quad (4)$$

Таблица 3. Величина найденных весовых коэффициентов.

Показатель	w_i
Величина прожиточного минимума	105
Объём платных услуг на душу населения	5
Валовой региональный продукт на душу населения	61
Инвестиции в основной капитал на душу населения	75
Уровень занятости населения	185
Уровень участия в рабочей силе	309
Процент населения в трудоспособном возрасте	5753
Коэффициент демографической нагрузки	-116
Процент населения с денежными доходами ниже прожиточного минимума	-496
Добыча полезных ископаемых на душу населения	278

Источник: Рассчитано автором на основании статистических данных по субъектам РФ за 2015-2017 года.

В таблице 3 приведены найденные коэффициенты w_i . Построенная линейная модель (2) допускает количественную интерпретацию найденных взаимосвязей. Например, при увеличении объёма платных услуг на душу населения в регионе на 1 среднедушевые доходы увеличиваются в среднем на 5 единиц. При увеличении коэффициента демографической нагрузки на 1% среднедушевые доходы населения в среднем уменьшаются на 116 единиц. Построенная модель имеет ряд допущений, например, относительно линейности взаимосвязи целевого показателя со вспомогательными, однако она предоставляет возможность приблизительно оценить количественно найденные качественные экономические связи.

Данное исследование выявило основные социально-экономические показатели, оказывающие влияние на среднедушевые доходы населения. Построенная регрессионная модель предоставляет возможность количественно оценивать эффект различных мер государственной поддержки и воздействия. В работе были рассмотрены примеры неравенства регионов России по доходам и другим социально-экономическим показателям. Поэтому одним из приоритетов региональной политики России должно быть сокращение регионального неравенства и поддержка экономики [1; 3].

Библиографический список:

1. Болтов Р.В. Сбалансированность развития территорий / Р.В. Болтов, О.И. Карелин, А.В. Семашко // Евразийский союз ученых. – 2016. – № 4 (25). – С. 109-110.
2. Borkova E.A., Kochkin T.N., Mutareva E.A., Kartavchenko A.E. analysis of the connection between investment indicators and ecology (case of the North-Western macroregion of Russia) В сборнике: International Business Information

- Management Association (IBIMA). Proceedings of the 34th International Business Information Management Association Conference (IBIMA). Editor Khalid S. Soliman. 2019. С. 1224-1230.
3. Дробышевская Л.Н. Оценка структурной сбалансированности инвестиционного развития регионов / Л.Н. Дробышевская, К.И. Шевченко // Региональная экономика: теория и практика. – 2014. – № 8 (335). – С. 34-43.
 4. Ковнир В.Н., Рысина Т.В. Политика доходов и социальное неравенство в РФ // Вестник Московского университета. Серия 21. Управление (государство и общество). 2011. №3. URL: <https://cyberleninka.ru/article/n/politika-dohodov-i-sotsialnoe-neravenstvo-v-rf> (дата обращения: 27.10.2020).
 5. Колосовский Н.Н. Основы экономического районирования / Н.Н. Колосовский. – М.: Госполитиздат, 1958. – 200 с.
 6. Савина Т.Н. Политика формирования доходов населения // Экономический анализ: теория и практика. 2015. №5 (404).

Оригинальность 79%