

## Применение многопараметрических методов Data Science для классификации субъектов Российской Федерации по признаку дотационности

**Кузнецова Анна Викторовна<sup>1</sup>**

Канд. биол. наук, ст. науч. сотр. лаборатории математической биофизики

ORCID: 0000-0002-0297-7013, e-mail: azforus@yandex.ru

**Борисова Людмила Робертовна<sup>2</sup>**

Канд. физ.-мат. наук, доц. каф. математики

ORCID: 0000-0002-5757-0341, e-mail: lrborisova@fa.ru

**Хадарцев Валерий Мисостович<sup>3</sup>**

Канд. экон. наук, директор

ORCID: 0009-0006-7330-834X, e-mail: sbm@flotek.ru

<sup>1</sup>Институт биохимической физики имени Н.М. Эмануэля Российской академии наук, 119334, Косыгина ул., 4, г. Москва, Россия

<sup>2</sup>Финансовый университет при Правительстве Российской Федерации, 125993, Ленинградский пр-кт, 49, г. Москва, Россия

<sup>3</sup>Государственное предприятие «ЦНИИТЭИТЯЖМАШ», 123090, Троицкая ул., 17, стр. 1., г. Москва, Россия

### Аннотация

Актуальность исследования обоснована важностью мониторинга и прогнозирования дотационности регионов Российской Федерации (далее – РФ, Россия) с целью выявления основных критериев для классификации субъектов по признаку дотационности. В кратком обзоре литературы рассмотрены математические модели, используемые для моделирования дотационности регионов России. В них в основном фиксировались социально-экономические показатели, которые необходимо учитывать при применении, а также использовались регрессионные модели, но не были представлены математически обоснованные рекомендации для вывода регионов из кластеров дотационности. В работе проведен анализ социально-экономических и демографических показателей регионов РФ с помощью методов, выявляющих закономерности в многопараметрическом датасете. Используются методы традиционного статистического анализа и машинного обучения, в том числе авторские. Определены статистически значимые закономерности, отражающие связь дотационности с такими показателями, как инвестиции в основной капитал, основные фонды, среднедушевые доходы и средний размер назначенных пенсий, уровень безработицы и др. Выполненный логико-статистический анализ убедительно свидетельствует в пользу применения методов машинного обучения (англ. data science) при выявлении статистически значимых связей между различными показателями, характеризующими развитие регионов РФ.

**Ключевые слова:** российские регионы, дотации, бюджет региона, инвестиции, управление бюджетными ресурсами, экономика дотационных регионов, демография, машинное обучение, статистические методы, Data Science

**Для цитирования:** Кузнецова А.В., Борисова Л.Р., Хадарцев В.М. Применение многопараметрических методов Data Science для классификации субъектов Российской Федерации по признаку дотационности//Управление. 2024. Т. 12. № 3. С. 58–73. DOI: 10.26425/2309-3633-2024-12-3-58-73



Received: 01.05.2024

Revised: 19.06.2024

Accepted: 24.06.2024

# Application of multiparametric methods of data science for the classification of Russian subjects on the basis of subsidisation

**Anna V. Kuznetsova<sup>1</sup>**

Cand. Sci. (Biol.), Senior Researcher at the Laboratory of Mathematical Biophysics

ORCID: 0000-0002-0297-7013, e-mail: azforus@yandex.ru

**Lyudmila R. Borisova<sup>2</sup>**

Cand. Sci. (Phys. and Math.), Assoc. Prof. at the Mathematics Department

ORCID: 0000-0002-5757-0341, e-mail: lrborisova@fa.ru

**Valeriy M. Khadartsev<sup>3</sup>**

Cand. Sci. (Econ.), Director

ORCID: 0009-0006-7330-834X, e-mail: sbm@flotek.ru

<sup>1</sup>N.M. Emanuel Institute of Biochemical Physics of the Russian Academy of Sciences, 4, Kosygina ul., Moscow 119334, Russia<sup>2</sup>Financial University under the Government of the Russian Federation, 49, Leningradsky prospekt, Moscow 125993, Russia<sup>3</sup>State enterprise "TSNIITEITYAZHMASH", 17, str. 1, Troitskaya ul., Moscow 123090, Russia

---

## Abstract

---

The relevance of the study is justified by the importance of monitoring and forecasting the subsidisation of the Russian regions in order to identify the main criteria for classifying subjects on the basis of subsidisation. In a brief review of the literature, mathematical models used to model the subsidisation of the Russian regions are considered. They have mainly fixed socio-economic indicators that need to be given attention while applying, and also regression models are used, but mathematically sound recommendations for the withdrawal of regions from clusters of subsidisation are not provided. The paper analyses the socio-economic and demographic indicators of the Russian regions applying methods that identify patterns in a multiparametric dataset. The methods of traditional statistical analysis and machine learning, including the author's ones, are used. Statistically significant patterns have been identified, reflecting the relationship of subsidisation with such indicators as investments in fixed capital, fixed assets, average per capita income and average size of assigned pensions, unemployment rate, etc. The performed logical and statistical analysis strongly supports the use of machine learning (Data Science) methods in identifying statistically significant relationships between various indicators characterising the development of the regions of the Russian Federation.

**Keywords:** Russian regions, subsidies, regional budget, investments, management of budgetary resources, economy of subsidised regions, demography, machine learning, statistical methods, Data Science

**For citation:** Kuznetsova A.V., Borisova L.R., Khadartsev V.M. (2024). Application of multiparametric methods of data science for the classification of Russian subjects on the basis of subsidisation. *Upravlenie / Management (Russia)*, 12 (3), pp. 58–73. DOI: 10.26425/2309-3633-2024-12-3-58-73

---



## Введение / Introduction

Государственные субсидии распространены во многих странах, особенно в сфере защиты окружающей среды, рационального использования природных ресурсов. Распределение дотаций на выравнивание бюджетной обеспеченности субъектов Российской Федерации (далее – РФ, Россия) производится в соответствии с утвержденной Правительством методикой<sup>1</sup>.

Распределение дотаций зависит в том числе от их экономических показателей. Как это соотносится с социально-экономическими и демографическими показателями регионов, отраженными в отчетности Федеральной службы государственной статистики (далее – Росстат), и почему необходимо отдельное исследование их влияния статистическими методами? Актуальными вопросами становятся анализ дотационных регионов, выявление основных критериев, отличающих их от рентабельных регионов, которые пополняют федеральный бюджет страны. Проблема дотационности субъектов РФ стоит достаточно остро. Более 2/3 регионов требуют дотаций из центра (62–65 субъектов).

М.А. Шилов и А.Б. Гусев в своей работе акцентируют внимание на том, что государственная политика по обеспечению однородного развития регионов РФ состоит из двух взаимосвязанных пунктов:

1) фиксирования экономического неравенства между регионами;

2) программы действий по обеспечению экономически однородного развития федеральных округов [Шилов, Гусев, 2009].

Предлагается на первом этапе дифференцировать регионы на две группы. Для этого надо находить отношение душевого валового регионального продукта (далее – ВРП) и максимальное значение данного показателя в стране для региона-лидера. Если значение этого отношения меньше 0,5, то регион предлагается считать проблемным. Заметим, что сами авторы считают такой подход неоднозначным. Проблема дотационности муниципальных образований России была изучена М.Н. Лукьяновой, которая отметила значение учета производственного потенциала для дотаций муниципалитетам [Лукьянова, 2015]. Именно этот показатель в соответствии с ее методикой ранжирования, основанной на использовании экспертных оценок, получил самый

высокий ранг. П.Д. Косинским, А.В. Зубовой, Т.А. Юрзиной также был применен метод использования частных оценок при разработке индикатора риска финансово-бюджетной несостоятельности муниципальных образований на примере анализа бюджета городов Кемерово и Новокузнецка за 2022 г. с целью выяснения причин дотационности муниципальных бюджетов [Косинский, Зубова, Юрзина, 2022]. Авторы этой работы предложили в целях эксперимента увеличить до 40 % поступаемые налоговые доходы в бюджеты городов (вместо 27 % в настоящее время).

А.Б. Симонов, С.А. Баранов, С.П. Сазонов использовали кластерный анализ для разделения всех регионов России на два кластера по бюджетной обеспеченности и показали, что для первого класса бюджетная обеспеченность прямо зависит от основных фондов, а для второго – наоборот [Симонов, Баранов, Сазонов, 2016]. Эту зависимость авторы обосновали как препятствие для развития регионов. Основным вывод, сделанный в работе, состоит в том, что структуру ВРП нужно учитывать при моделировании дотационности. Необходимость при моделировании учитывать размер экономики отметил А.Н. Березняцкий [Березняцкий, 2022]. Он предложил индекс дотационности региона, статистический анализ которого позволяет выделить истинно и технически дотационные регионы; предложил на втором этапе строить коинтеграционную модель для динамики индекса ВРП истинно дотационного региона в зависимости от объема трансфертов, на основе которой строится его вторичный рейтинг.

Самыми дотационными регионами являются субъекты РФ, относящиеся к территориям Дальнего Востока, частично Сибири, Крыма, Северного Кавказа. Они требуют очень высокого уровня финансирования, их бюджеты практически полностью состоят из субсидий [Энеева, Ульбашева, Уянаева, 2010]. Авторы исследования М.К. Алимуратов, А.З. Мидов, С.В. Одинцов отмечают, что в России довольно высока дифференциация регионов по уровню социально-экономического развития [Алимуратов, Мидов, Одинцов, 2021]. Для анализа причин регионального неравенства важно выявить и сгруппировать регионы, наиболее зависимые от федерального центра [Таштамиров, Байсаева, Баташев, 2020].

Регионы-доноры – это Ханты-Мансийский автономный округ (далее – АО), Москва, Ямало-Ненецкий АО, Санкт-Петербург, Московская область. Проблему дотационных регионов в РФ можно решить благодаря комплексным мероприятиям поддержки их развития, включающим в себя следующие меры:

<sup>1</sup> Постановление Правительства Российской Федерации от 22.11.2004 г. № 670 «О распределении дотаций на выравнивание бюджетной обеспеченности субъектов Российской Федерации» (с изменениями и дополнениями). Режим доступа: <https://base.garant.ru/187556/> (дата обращения: 25.04.2024).

1) стимулирование привлечения инвестиций в экономику дотационных регионов через налоговые льготы и другие меры поощрения инвесторов;

2) развитие инфраструктуры в дотационных регионах для повышения их конкурентоспособности, что способствует привлечению новых предприятий и созданию рабочих мест;

3) обучение и повышение квалификации местного населения для улучшения его конкурентоспособности на рынке труда;

4) поддержка развития сельского хозяйства и других отраслей экономики дотационных регионов для снижения зависимости от поставок из других регионов;

5) создание благоприятной инвестиционной среды, прозрачной системы управления и борьба с коррупцией с целью привлечения инвестиций.

Для осуществления стимулирования привлечения инвестиций в экономику дотационных регионов можно использовать следующие меры: налоговые льготы, финансовую поддержку, создание инфраструктуры, партнерство с бизнесом, привлечение иностранных инвесторов. Внедрение таких мер позволяет снизить зависимость дотационных регионов от федерального бюджета, стимулировать их экономическое развитие и улучшение качества жизни населения. Существует также вопрос по перераспределению бюджетных средств в стране, связанный с тем, что предприятия из регионов оформлены в Москве и платят налоги туда же. Тот же Башкортостан – дотационный регион, но он располагает богатыми ресурсами нефти, золота, развитой химической и энергетической промышленностью. Аналогично богатейший Приморский край является дотационным. Как считается эта часть доходов?

Для успешного осуществления мер по выводу регионов из дотационности к самообеспечению и ответа на поставленные вопросы необходимо выявить закономерности, отражающие связь дотационности регионов РФ с социально-экономическими факторами, на которые можно воздействовать. Это возможно с помощью современных подходов и методов Data Science на основе многопараметрического анализа данных. Указанный подход позволяет объединить большие объемы данных из различных источников, таких как экономические показатели, социальные данные, демографические и инфраструктурные характеристики регионов, для более комплексного анализа ситуации.

Все изученные нами литературные источники с математическими моделями не позволяют дать прогноз и рекомендации по выводу региона из дотационного кластера, а представленный в работе подход – позволяет. В этом состоит научная новизна

выполненного исследования. Алгоритмы машинного обучения (далее – МО) и статистические методы помогают выявить взаимосвязи и зависимости между различными параметрами, определить ключевые факторы, влияющие на уровень дотации в каждом конкретном регионе. На основе этих данных можно разработать модели оценки и прогнозирования дотаций, которые будут учитывать не только текущую ситуацию, но и потенциальные изменения и тенденции развития. Благодаря этому подходу возможно повысить эффективность управления бюджетными ресурсами и сделать процесс принятия решений более обоснованным и прозрачным.

Целью работы было с помощью методов МО выявить ключевые факторы, коррелирующие с экономическим развитием региона, и определить, какие из них связаны с увеличением или уменьшением потребности в дотациях. Такой подход позволяет более глубоко и объективно изучить проблему дотационности регионов, выявить скрытые закономерности и оптимизировать процессы принятия решений в этой области.

### **Материалы и методы. Датасет / Materials and methods. Dataset**

В базу данных для исследования вопросов дотационности вошли 46 регионов, разделенных на две группы сравнения. 36 регионов оставили для проведения контроля решающего правила. Количество объектов первой группы – 22 дотационных региона. Это республики: Дагестан, Ингушетия, Кабардино-Балкария, Карачаево-Черкессия, Северная Осетия – Алания, Чеченская, Башкортостан, Тыва, Бурятия, Саха (Якутия), Крым, Чувашская; несколько краев: Ставропольский, Забайкальский, Алтайский, Камчатский; области: Брянская, Ивановская, Ростовская, Кировская, Курганская; Чукотский АО.

Количество объектов второй группы – 24 региона, не нуждающихся в дотациях. Это два города, Москва и Санкт-Петербург; области: Белгородская, Калужская, Липецкая, Московская, Тульская, Ярославская, Вологодская, Ленинградская, Мурманская, Нижегородская, Самарская, Свердловская, Тюменская, Иркутская, Сахалинская; АО: Ненецкий, Ханты-Мансийский, Ямало-Ненецкий; края: Краснодарский, Пермский; Республика Татарстан.

Социально-экономические, демографические, инвестиционные показатели были отобраны на сайте Росстата из баз данных, представленных в ежегодных сборниках «Регионы России. Социально-экономические показатели» за 2023 г., 2022 г., 2021 г. Их исходное число было 45.

## Методы / Methods

Для анализа данных была использована система анализа данных (далее – САД) Data Master Azforus (далее – DMA). Помимо методов традиционной статистики, САД DMA содержит набор методов МО для решения задач классификации – No code [Борисова, Кузнецова, 2022; Кирилук, Кузнецова, Сенько, 2021]. Режим автоклассификации представляет собой набор 9 методов МО, которые могут быть объединены в ансамбли из трех, 5, 7 или 9 наиболее эффективных методов. В данной работе использовали 7 методов из-за малого количества объектов. Нейронные сети неэффективны в подобном случае, поскольку требуют для обучения большое число объектов – десятки и сотни тысяч. Валидация работы методов МО проводилась в режиме скользящего контроля leave-one-out. Метод оптимально достоверных разбиений (далее – ОДР) вычисляет границу между объектами групп сравнения таким образом, чтобы с обеих сторон от границы разбиения преобладали объекты разных групп. Аналогично ставятся границы для двумерного разбиения в пространстве признаков. Значимость найденных закономерностей вычисляется с помощью перестановочного теста, в котором происходит сравнение распознавания исходной обучающей выборки с результатами распознавания на нескольких тысячах случайных таблицах. Малое число верных распознаваний на случайных таблицах говорит о том, что закономерность, найденная на обучающей выборке, достоверная.

Второй авторский метод МО – это метод статистически взвешенных синдромов (далее – СВС), в котором проводится голосование по полученным базовым множествам с дальнейшим отнесением объекта в одну из сравниваемых групп.

Диаграммы рассеяния с результатами одномерных и двумерных разбиений позволяют видеть каждый регион в группе схожих с ним субъектов относительно выставленных границ. Для каждого региона по результатам анализа можно создать план мероприятий по переходу в благоприятную группу рентабельных регионов.

Распознавание субъектов, не вошедших в обучающую выборку, проводится с помощью автономной программы «Прогноз», в которой методом СВС каждый регион из контрольной группы относится алгоритмом к первой или второй группе.

## Результаты автоклассификации / Auto-classification results

Результаты применения методов МО, участвовавших в классификации, при сравнении двух исследуемых групп показали, что самым эффективным из них оказался оригинальный авторский метод СВС. Он показал результат распознавания объектов при скользящем контроле: ROC AUC = 0,943 (англ. ROC – Receiver Operating Characteristic, рабочая характеристика приемника; AUC – area under curve, площадь под кривой ROC). Также хороший результат распознавания получен для методов адаптивный бустинг и деревья решений (табл. 1).

В табл. 1 видно, что чувствительность у большинства методов намного ниже специфичности, что свидетельствует о большем разбросе и разнообразии объектов первой группы – дотационных регионов. Только метод деревьев решений лучше распознал объекты первой группы, чем объекты второй. За счет разных подходов в машинных методах недостатки одних исправляются достоинствами других, и эффективность ансамбля нескольких методов выше, чем каждый из методов в отдельности. Это также увеличивает устойчивость распознавания.

Таблица 1

### Результаты классификации методами МО при сравнении двух групп исследования

Table 1. Results of the classification by machine learning methods when comparing two study groups

Метод	Чувствительность	Специфичность	F-оценка	AUC
Статистически взвешенные синдромы (SWS)	0,773	0,958	0,850	0,943
Адаптивный бустинг (AB)	0,818	1,000	0,900	0,909
Деревья решений (DT)	0,909	0,875	0,889	0,892
Метод опорных векторов (SVM)	0,773	0,917	0,829	0,833
Градиентный бустинг (XGB)	0,636	0,792	0,683	0,821
Метод ближайших соседей (KNN)	0,636	0,833	0,700	0,735
Линейный дискриминантный анализ (LDA)	0,500	0,583	0,512	0,542

Примечание: SWS – statistically weighted syndromes; AB – adaptive boosting; DT – decision trees; SVM – support vector machine; XGB – extreme gradient boosting; KNN – K-nearest neighbors; LNN – linear discriminant analysis

Составлено авторами по материалам исследования / Compiled by the authors on the materials of the study

В табл. 2 приведены результаты распознавания на ансамбле из 7 методов МО. Общий результат распознавания – 84,8 %. Из 22 объектов первой группы правильно были распознаны 17 (77,3 %). 5 регионов были отнесены во 2-ю группу, что свидетельствует о хорошем потенциале у этих методов с точки зрения выхода из дотационной группы. Во второй группе было две ошибки (91,7 %). Эти регионы, наоборот, рискуют попасть из благоприятной группы в дотационную по своим характеристикам.

В табл. 3 можно наблюдать результаты по каждому методу МО и распознавание ансамблем для каждого региона РФ. Обозначения методов приведены в табл. 1. Некоторые регионы (Башкортостан, Якутия, Камчатский край, Ростовская и Ярославская области, Чукотский и Ненецкий АО) алгоритм ошибочно отнес в другую группу, что позволяет для них делать определенные прогнозы.

Таблица 2

**Сводная таблица результата распознавания исследуемых групп регионов для ансамбля из 7 методов МО**

Table 2. Summary table of the recognition result of the studied groups of regions for an ensemble of 7 machine learning methods

Параметр	Количество объектов	Правильно	Ошибочно	Из группы 1	Из группы 2
Первая группа	22 (47,8 %)	17 (77,3 %)	5 (22,7 %)	17 (89,5 %)	5 (18,5 %)
Вторая группа	24 (52,2 %)	22 (91,7 %)	2 (8,3 %)	2 (10,5 %)	22 (81,5 %)
Итого	46 (100,0 %)	39 (84,8 %)	7 (15,2 %)	19 (41,3 %)	27 (58,7 %)

Составлено авторами по материалам исследования / Compiled by the authors on the materials of the study

Таблица 3

**Результат автоклассификации для ансамбля из 7 методов МО**

Table 3. Result of the auto-classification for an ensemble of 7 machine learning methods

Объект	Метод МО							Ансамбль	Верная группа
	AB	LDA	KNN	SVM	DT	SWS	XGB		
Республика Дагестан	1	2	1	1	1	1	2	1	1
Республика Ингушетия	1	2	1	1	1	1	1	1	1
Кабардино-Балкарская Республика	1	2	1	1	1	1	1	1	1
Карачаево-Черкесская Республика	1	1	1	1	1	1	1	1	1
Северная Осетия – Алания	1	1	1	1	1	1	1	1	1
Чеченская Республика	1	1	1	1	1	1	1	1	1
Ставропольский край	1	1	2	2	1	1	1	1	1
Башкортостан	2	1	2	2	1	2	2	2	1
Республика Тыва	1	1	1	1	1	1	1	1	1
Алтайский край	1	2	2	1	1	1	2	1	1
Республика Бурятия	1	2	2	1	1	1	1	1	1
Якутия	2	1	2	2	2	2	2	2	1
Камчатский край	2	2	1	1	2	2	2	2	1
Брянская область	1	2	1	1	1	1	1	1	1
Ивановская область	1	1	1	1	1	1	1	1	1
Республика Крым	1	2	2	2	1	1	1	1	1
Ростовская область	1	1	2	2	1	2	2	2	1
Чувашская Республика	1	2	1	1	1	1	1	1	1
Кировская область	1	2	1	1	1	1	2	1	1
Курганская область	1	1	1	1	1	1	1	1	1
Забайкальский край	1	1	2	1	1	1	1	1	1
Чукотский АО	2	2	1	1	1	2	2	2	1
Белгородская область	2	2	1	2	2	2	2	2	2
Калужская область	2	2	1	2	2	2	2	2	2
Липецкая область	2	1	2	2	2	2	1	2	2

Объект	Метод МО							Ансамбль	Верная группа
	AB	LDA	KNN	SVM	DT	SWS	XGB		
Московская область	2	2	2	2	2	2	2	2	2
Тульская область	2	1	2	2	2	2	2	2	2
Ярославская область	2	1	1	2	2	1	1	1	2
Москва	2	1	2	2	2	2	2	2	2
Ненецкий АО	2	2	1	1	1	2	1	1	2
Вологодская область	2	2	2	2	2	2	2	2	2
Ленинградская область	2	1	2	2	2	2	2	2	2
Мурманская область	2	1	2	2	2	2	2	2	2
Санкт Петербург	2	2	2	2	2	2	2	2	2
Краснодарский край	2	2	2	2	2	2	1	2	2
Республика Татарстан	2	1	2	2	2	2	2	2	2
Пермский край	2	2	2	2	2	2	2	2	2
Нижегородская область	2	2	2	2	2	2	2	2	2
Самарская область	2	1	2	2	2	2	2	2	2
Свердловская область	2	2	2	1	2	2	2	2	2
Тюменская область	2	2	2	2	2	2	2	2	2
Ханты Мансийский АО	2	2	2	2	2	2	2	2	2
Ямало-Ненецкий АО	2	1	2	2	1	2	2	2	2
Красноярский край	2	1	2	2	1	2	2	2	2
Иркутская область	2	2	2	2	2	2	1	2	2
Сахалинская область	2	2	2	2	2	2	2	2	2

Составлено авторами по материалам исследования / Compiled by the authors on the materials of the study

В ошибки при распознавании первой группы попали Башкортостан, Якутия, Камчатский край, Ростовская область и Чукотский АО. Это говорит о том, что по своим социально-экономическим и демографическим показателям эти субъекты больше похожи на группу благополучных регионов-доноров, то есть у них хороший потенциал выйти из перечня дотационных регионов.

Для второй группы оказались ошибочными Ярославская область и Ненецкий АО. Для них, наоборот, прогноз неблагоприятный. По отобранным значимым показателям эти субъекты РФ более близки к группе дотационных регионов.

В табл. 4 можно видеть, что наиболее значимые показатели для сравнения двух групп в данном исследовании – это инвестиции в основной капитал (млн руб. и млн руб. на 1 тыс. граждан), основные фонды (млн руб. и млн руб. на 1 тыс. граждан), а также среднедушевые доходы (руб., 2020 г.), средний размер назначенных пенсий (2019 г.), уровень безработицы (% , 2019 г.). Другие значимые показатели характеризуют демографическую ситуацию, состояние здравоохранения, инфраструктуру регионов.

Таблица 4

#### Результаты теста Манна-Уитни-Уилкоксона (U-test) для сравнения групп

Table 4. Results of the Mann-Whitney-Wilcoxon test (U-test) for comparing groups

Показатель	N1	Mean1	SD1	N2	Mean2	SD2	p<
Смертей на 100 тыс. чел. на 8 сентября 2021 г.	22	94,795	44,915	23	142,782	79,307	0,025
Среднедушевые доходы, руб., 2020 г.	22	29 252,5	15 858,8	24	42 921,08	17 866,1	0,000
Средний размер назначенных пенсий, 2019 г.	22	14 699,9	3 455,88	24	16 562,21	2 687,57	0,000
Младенческая смертность (число детей, умерших до года на 1 тыс. родившихся живыми)	22	5,728	1,896	24	4,288	1,133	0,016
Уровень безработицы, %, 2019 г.	22	9,553	5,893	24	4,084	1,778	0,000

Окончание табл. 4

Показатель	N1	Mean1	SD1	N2	Mean2	SD2	p<
Коэффициент Джини	22	0,366	0,026	24	0,391	0,028	0,015
Площадь жилищ, приходящаяся в среднем на одного жителя, м <sup>2</sup>	22	23,635	4,296	24	27,043	3,791	0,019
Численность населения, тыс. чел.	22	1 436,49	1 163,43	24	2 941,300	2 795,73	0,018
Инвестиции в основной капитал, млн руб.	22	110 794	106 493	24	563 900,3	660 619	0,000
Инвестиции в основной капитал, млн руб. на 1 тыс. граждан	22	100,071	118,504	24	341,624	513,601	0,000
Основные фонды, млн руб. на 1 тыс. граждан	22	1 350,99	983,062	24	5 229,09	6 630,83	0,000
Основные фонды, млн руб.	22	1 782 505,818	1 838 163,241	24	10 905 134,375	14 935 742,394	0,000
Отправлено грузов, млн т (ж/д), 2020 г.	20	22,190	55,844	21	25,276	17,471	0,005
Отправлено пассажиров, тыс. чел. (ж/д), 2020 г.	20	2 415,75	2 243,1	21	35 781,52	88 053,1	0,003
Перевезено автотранспортом, млн т, 2020 г.	21	19,608	23,414	22	33,029	32,725	0,040
Грузооборот, 1 тыс. т, млн 1 тыс. т км, 2020 г.	21	1 457,14	1 689,06	22	4 718,025	3 589,37	0,000
Перевозка пассажиров автобусами, млн чел., 2020 г	21	78,610	75,158	22	179,927	185,28	0,010
Пассажиuroоборот, млн чел. / 1 тыс. км, 2020 г.	21	913,81	935,59	22	1 812,68	1 413,73	0,006
Процент дорог с усовершенствованным покрытием в протяженности автодорог с твердым покрытием общего пользования, 2020 г.	21	55,347	18,070	22	71,267	18,166	0,011
Болезни органов дыхания (на 1 тыс. чел.), 2020 г.	21	327,695	82,115	22	401,582	64,243	0,003
Инфекционные и паразитарные заболевания (на 1 тыс. выявлено впервые)	22	25,036	10,068	24	31,135	8,999	0,037
Новообразования (на 1 тыс. чел. выявлено впервые)	22	9,420	4,568	24	12,913	3,054	0,004

Примечание: ж/д – железнодорожные перевозки

Составлено авторами по материалам исследования / Compiled by the authors on the materials of the study

### ОДР / Optimally reliable partitions

Метод ОДР использует критерий хи-квадрат для постановки границ по каждому показателю и по каждой паре показателей таким образом, чтобы с одной стороны от них преобладали значения одной из сравниваемой групп, а с другой – значения другой группы. Значимость найденной закономерности вычисляется с помощью перестановочного теста путем генерации нескольких тысяч случайных таблиц, в которых целевая функция (номер группы) перемешивается случайным образом. Результат сравнивается с исходной обучающей выборкой. При малом числе случайных таблиц со значением хи-квадрат, близким с таковым в исходной выборке, считается, что найденная закономерность значима для разделения двух исследуемых групп. В случае одномерных разбиений образуется два квадранта, в случае двумерных – четыре, которые нумеруются от левого верхнего по часовой стрелке.

### Одномерные разбиения / One-dimensional partitions

В табл. 5 приведены значимые показатели в порядке возрастания p-value хи-квадрат (убывания

значимости показателя) – правый столбец. Граница, выставленная методом ОДР, и число значений каждой группы (с процентным содержанием) слева от границы – первый квадрант, и число значений справа от границы – второй квадрант.

Как свидетельствует табл. 5, анализ методов ОДР подтвердил результаты статистического анализа (табл. 1), а именно наиболее значимыми показателями стали основные фонды (млн руб. и млн руб. на 1 тыс. граждан), средний размер назначенных пенсий (2019 г.), среднедушевые доходы (руб., 2020 г.), инвестиции в основной капитал (млн руб. и млн руб. на 1 тыс. граждан).

На рис. 1 можно видеть, как слева от границы абсолютно преобладают значения первой группы – квадраты. Справа от границы, кроме регионов второй группы (кружки), попали три региона из первой группы: Якутия, Камчатский и Чукотский края. Значение Республики Башкортостан находится вблизи границы разбиения. Это объясняет, почему этот субъект и Якутия при распознавании попали во вторую – благоприятную – группу. Наглядные диаграммы рассеяния позволяют видеть значения регионов в соседстве с похожими на них регионами.

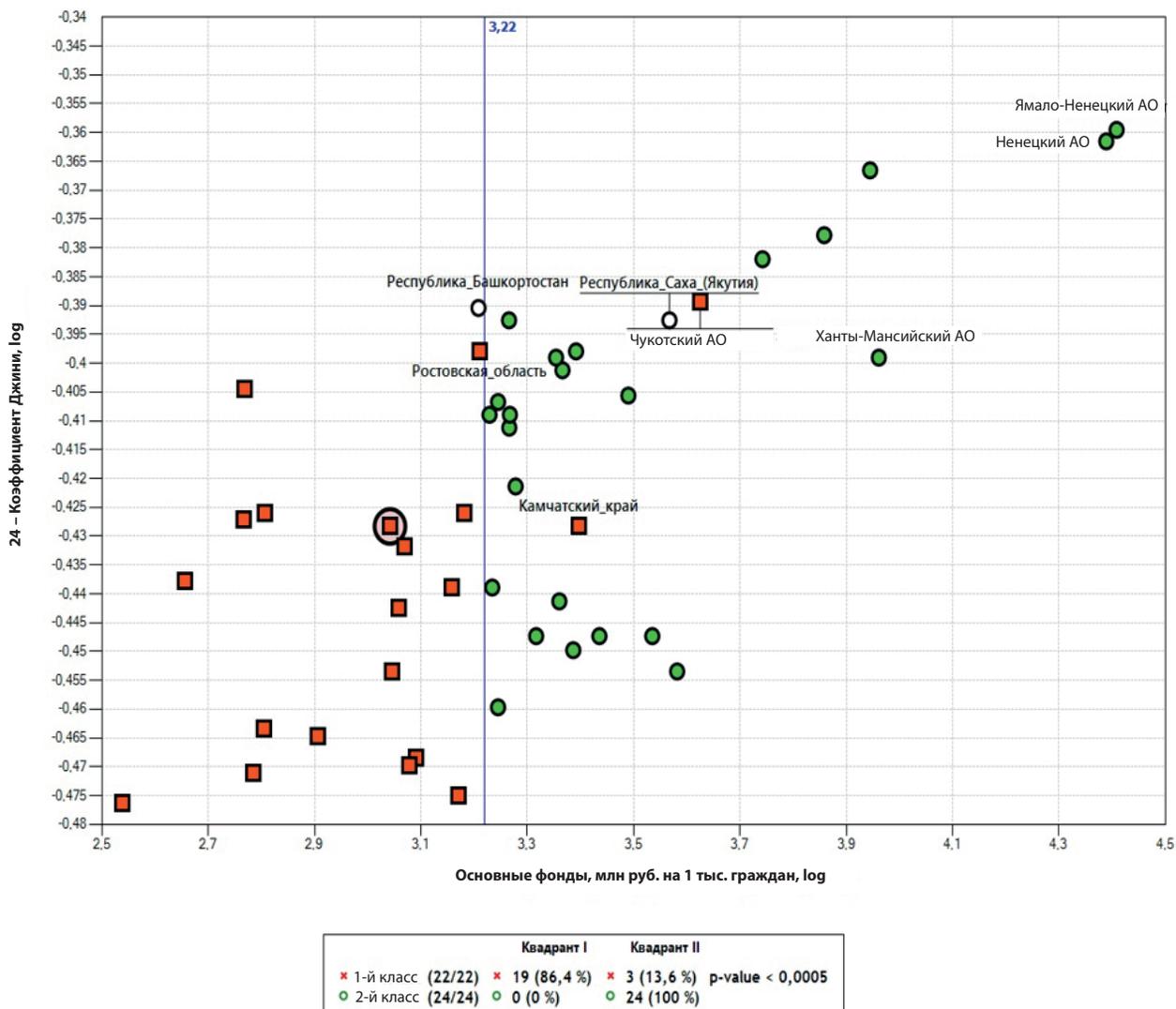
**Одномерные разбиения для сравнения групп по дотациям**  
Table 5. One-dimensional partitions for comparing groups by subsidies

Показатель	Граница	Первый квадрант		Второй квадрант		p-value
		Первая группа	Вторая группа	Первая группа	Вторая группа	
Основные фонды, млн руб. на 1 тыс. граждан	1 660	19 (86,4 %)	0 (0 %)	3 (13,6 %)	24 (100 %)	0,0005
Средний размер назначенных пенсий, 2019 г.	14 234	17 (77,3 %)	1 (4,2 %)	5 (22,7 %)	23 (95,8 %)	0,0005
Основные фонды, млн руб.	1 966 902	17 (77,3 %)	1 (4,2 %)	5 (22,7 %)	23 (95,8 %)	0,0005
Среднедушевые доходы, руб., 2020 г.	28 312	17 (77,3 %)	1 (4,2 %)	5 (22,7 %)	23 (95,8 %)	0,0005
Инвестиции в основной капитал, млн руб. на 1 тыс. граждан	86,12	18 (81,8 %)	2 (8,3 %)	4 (18,2 %)	22 (91,7 %)	0,0005
Инвестиции в основной капитал, млн руб.	88 380	15 (68,2 %)	0 (0 %)	7 (31,8 %)	24 (100 %)	0,0005
Грузооборот 1 тыс. т, млн 1 тыс. т км, 2020 г.	1 427	15 (71,4 %)	3 (13,6 %)	6 (28,6 %)	19 (86,4 %)	0,00133
Новообразования (на 1 тыс. чел. выявлено впервые)	8,5	10 (45,5 %)	0 (0 %)	12 (54,5 %)	24 (100 %)	0,002
Уровень безработицы, %, 2019 г.	6	9 (40,9 %)	22 (91,7 %)	13 (59,1 %)	2 (8,3 %)	0,00233
Отправлено пассажиров, тыс. чел. (ж/д), 2020 г.	5 820	19 (95 %)	10 (47,6 %)	1 (5 %)	11 (52,4 %)	0,005
Процент дорог с усовершенствованным покрытием автодорог с твердым покрытием, 2020 г.	59,67	14 (66,7 %)	4 (18,2 %)	7 (33,3 %)	18 (81,8 %)	0,01
Отправлено грузов, млн т. (ж/д), 2020 г.	1,9	8 (40 %)	0 (0 %)	12 (60 %)	21 (100 %)	0,01433
Младенческая смертность (число детей, умерших до года на 1 тыс. родившихся живыми)	4,35	3 (13,6 %)	14 (58,3 %)	19 (86,4 %)	10 (41,7 %)	0,01433
Болезни органов дыхания (на 1 тыс. чел.), 2020 г.	403,4	19 (90,5 %)	10 (45,5 %)	2 (9,5 %)	12 (54,5 %)	0,016
Перевезено автомобильным транспортом, млн т, 2020 г.	7,392	11 (52,4 %)	2 (9,1 %)	10 (47,6 %)	20 (90,9 %)	0,01867
Пассажирооборот, млн чел. на 1 тыс. км	763	15 (71,4 %)	5 (22,7 %)	6 (28,6 %)	17 (77,3 %)	0,01933
Коэффициент Джини	0,3455	7 (31,8 %)	0 (0 %)	15 (68,2 %)	24 (100 %)	0,024
Заболеваемость (на 1 тыс. впервые выявлено)	640,8	7 (31,8 %)	0 (0 %)	15 (68,2 %)	24 (100 %)	0,03133
Перевозка пассажиров автобусами, млн чел., 2020 г.	81,65	16 (76,2 %)	7 (31,8 %)	5 (23,8 %)	15 (68,2 %)	0,03833
Погибло в происшествиях, человек (на 100 тыс. чел. населения), 2020 г.	10,09	2 (9,5 %)	11 (50 %)	19 (90,5 %)	11 (50 %)	0,03867

Составлено авторами по материалам исследования / Compiled by the authors on the materials of the study

Преобладание групп с двух сторон от границы было следующее: для показателей уровень безработицы (%), младенческая смертность (число детей, умерших до года на 1 тыс. родившихся живыми) и число погибших в происшествиях (на 100 тыс. чел.)

значения первой (дотационной) группы регионов были выше границы, а значения второй – ниже. Для всех остальных значимых показателей первая группа имела значения ниже границы разбиения.



Примечание: данные приведены в логарифмической шкале; ось Y в данном случае дана только для развертки диаграммы и роли не играет; первая группа – квадраты, вторая – кружки

Составлено авторами по материалам исследования / Compiled by the authors on the materials of the study

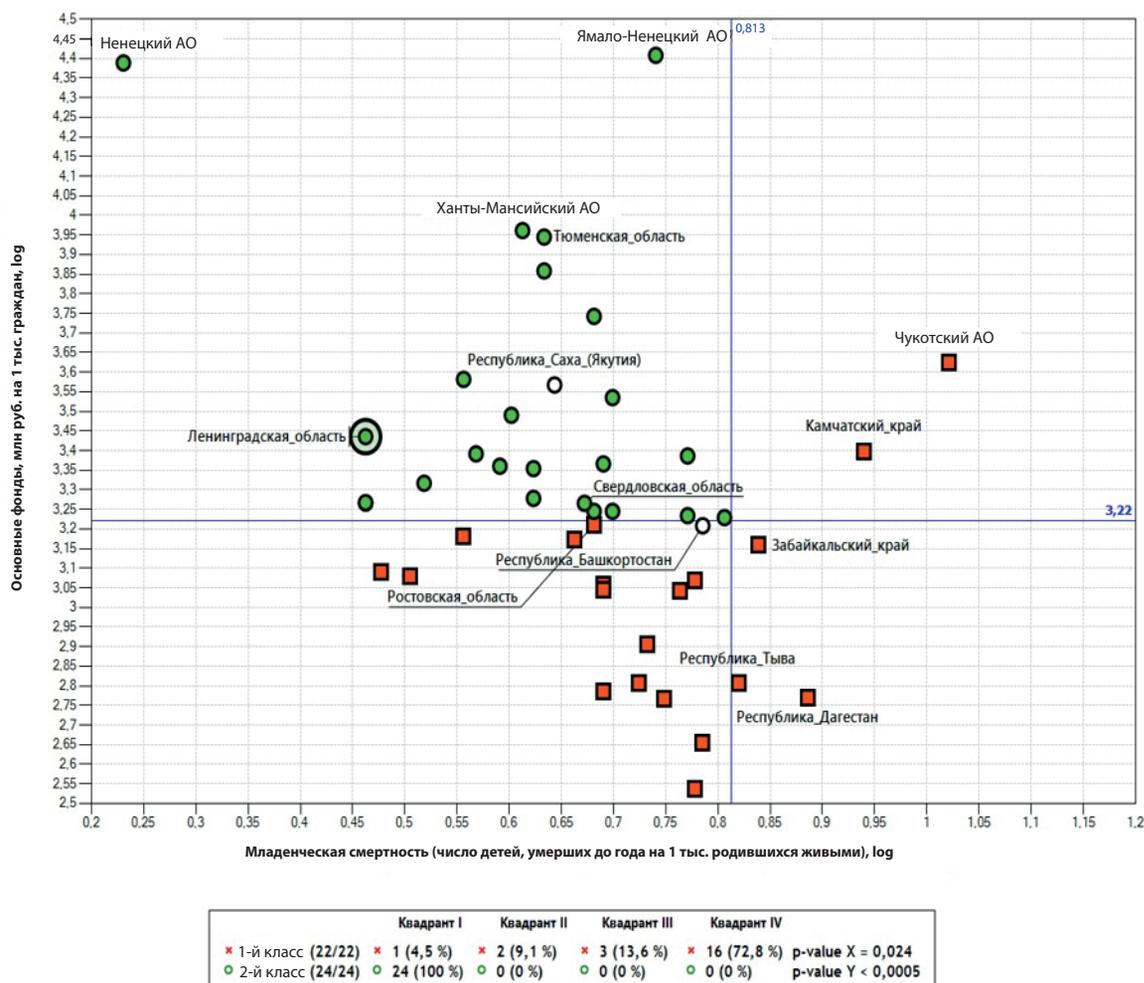
Рис. 1. Одномерные разбиения  
Fig. 1. One-dimensional partitions

### Двумерные разбиения/ Two-dimensional partitions

Когда границы разбиения ставятся для каждого показателя из пары, то в каждом квадранте преобладают значения одной из сравниваемых групп. В этих случаях p-value вычисляется отдельно по оси X и по оси Y. Значимость вычисляется также на перестановочном тесте, и это занимает очень много времени. Пары показателей отбираются так, чтобы оба показателя были значимыми. Таких пар получается очень много. Наглядные диаграммы рассеяния позволяют видеть значения каждого региона.

На рис. 2 в первом квадранте (верхний левый) преобладают значения второй группы. Наблюдения

в этом квадранте имеют низкие значения по младенческой смертности (ниже границы) и высокие значения по основным фондам (на 1 тыс. граждан). Из первой группы там только одно значение региона Республика Саха (Якутия) – незакрашенный кружок. Именно поэтому при распознавании этот регион дает ошибку и относится алгоритмом ко второй благополучной группе. В трех других квадрантах находятся только значения первой группы (квадраты). При голосовании по квадрантам методом СВС в этом случае очень большие баллы будут даны за первую группу.



Примечание: данные приведены в логарифмической шкале; первая группа – квадраты, вторая группа – кружки  
 Составлено авторами по материалам исследования / Compiled by the authors on the materials of the study

Рис. 2. Двумерные разбиения  
 Fig. 2. Two-dimensional partitions

Таблица 6

Двумерные разбиения для сравнения групп по дотациям  
 Table 6. Two-dimensional partitions for comparing groups by subsidies

Показатель	Chi-квадрат	Граница	p-value	Распределение случаев в квадрантах разбиения	
				Квадрант I	Квадрант II
Инвестиции в основной капитал, млн руб.	41,24	55 432	0,0005	2 (9,1 %) / 0 (0 %)	1 (4,5 %) / 24 (100 %)
Основные фонды, млн руб. на 1 тыс. граждан	–	1 660	0,04533	12 (54,5 %) / 0 (0 %)	7 (31,8 %) / 0 (0 %)
Основные фонды, млн руб. на 1 тыс. граждан	41,24	1 660	0,0467	14 (63,6 %) / 0 (0 %)	1 (4,5 %) / 24 (100 %)
Основные фонды, млн руб.	–	823 398	0,0005	2 (9,1 %) / 0 (0 %)	5 (22,7 %) / 0 (0 %)
Младенческая смертность (число детей, умерших до года на 1 тыс. родившихся живыми)	41,24	6,5	0,0005	1 (4,5 %) / 24 (100 %)	2 (9,1 %) / 0 (0 %)
Основные фонды, млн руб. на 1 тыс. граждан	–	1 660	0,024	3 (13,6 %) / 0 (0 %)	16 (72,7 %) / 0 (0 %)
Плотность	40,3	0,95	0,0005	3 (13,6 %) / 2 (8,3 %)	0 (0 %) / 22 (91,7 %)
Основные фонды, млн руб. на 1 тыс. граждан	–	1 660	0,049	19 (86,4 %) / 0 (0 %)	0 (0 %) / 0 (0 %)
Основные фонды, млн руб. на 1 тыс. граждан	39,39	1 660	0,04434	15 (71,4 %) / 0 (0 %)	0 (0 %) / 21 (95,5 %)

Продолжение табл. 6

Показатель	Хи-квадрат	Граница	p-value	Распределение случаев в квадрантах разбиения	
Процент дорог с усовершенствованным покрытием в протяженности автомобильных дорог с твердым покрытием общего пользования, 2020 г.	–	43,55	0,0005	2 (9,5 %) / 1 (4,5 %)	4 (19 %) / 0 (0 %)
Средний размер назначенных пенсий, 2019 г.	37,76	13 782	0,0005	5 (22,7 %) / 0 (0 %)	2 (9,1 %) / 24 (100 %)
Инвестиции в основной капитал, млн руб.	–	88 380	0,00167	10 (45,5 %) / 0 (0 %)	5 (22,7 %) / 0 (0 %)
Удельный вес расходов на оплату ЖКХ, % от общего числа услуг	34,6	42,45	0,0005	1 (4,5 %) / 21 (87,5 %)	3 (13,6 %) / 1 (4,2 %)
Инвестиции в основной капитал, млн руб. на 1 тыс. граждан	–	86,12	0,0067	0 (0 %) / 1 (4,2 %)	18 (81,8 %) / 1 (4,2 %)
Младенческая смертность (число детей, умерших до года на 1 тыс. родившихся живыми)	34,43	5,95	0,0005	2 (9,1 %) / 23 (95,8 %)	6 (27,3 %) / 1 (4,2 %)
Инвестиции в основной капитал, млн руб. на 1 тыс. граждан	–	69,65	0,00233	3 (13,6 %) / 0 (0 %)	11 (50 %) / 0 (0 %)
Средний размер назначенных пенсий, 2019 г.	34,22	13 782	0,0005	4 (18,2 %) / 0 (0 %)	2 (9,1 %) / 23 (95,8 %)
Основные фонды, млн руб.	–	1 819 330	0,00733	10 (45,5 %) / 1 (4,2 %)	6 (27,3 %) / 0 (0 %)
Инвестиции в основной капитал, млн руб.	32,21	88 380	0,031	3 (13,6 %) / 0 (0 %)	2 (9,1 %) / 22 (91,7 %)
Инвестиции в основной капитал, млн руб. на 1 тыс. граждан	–	80,77	0,013	5 (22,7 %) / 2 (8,3 %)	12 (54,5 %) / 0 (0 %)
Среднедушевые доходы, руб., 2020 г.	31,84	31 701	0,001667	1 (4,5 %) / 6 (25 %)	3 (13,6 %) / 15 (62,5 %)
Средний размер назначенных пенсий, 2019 г.	–	14 634	0,0133	0 (0 %) / 3 (12,5 %)	18 (81,8 %) / 0 (0 %)
Инвестиции в основной капитал, млн руб. на 1 тыс. граждан	31,6	69,65	0,025	13 (61,9 %) / 0 (0 %)	2 (9,5 %) / 21 (95,5 %)
Процент дорог с усовершенствованным покрытием в протяженности автомобильных дорог с твердым покрытием общего пользования, 2020 г.	–	43,55	0,0005	5 (23,8 %) / 1 (4,5 %)	1 (4,8 %) / 0 (0 %)
Среднедушевые доходы, руб., 2020 г.	31,56	26 250	0,03067	1 (4,5 %) / 0 (0 %)	4 (18,2 %) / 0 (0 %)
Младенческая смертность (число детей, умерших до года на 1 тыс. родившихся живыми)	–	6,5	0,0005	4 (18,2 %) / 24 (100 %)	13 (59,1 %) / 0 (0 %)
Инвестиции в основной капитал, млн руб. на 1 тыс. граждан	30,47	69,65	0,02467	5 (23,8 %) / 0 (0 %)	1 (4,8 %) / 19 (86,4 %)
Грузооборот 1 тыс т, млн 1 тыс т км, 2020 г.	–	1 427	0,001	6 (28,6 %) / 3 (13,6 %)	9 (42,9 %) / 0 (0 %)
Болезни органов дыхания (на 1 тыс. чел. выявлено впервые)	29,78	457,5	0,0005	3 (13,6 %) / 17 (70,8 %)	2 (9,1 %) / 4 (16,7 %)
Среднедушевые доходы, руб., 2020 г.	–	29 012	0,01233	0 (0 %) / 3 (12,5 %)	17 (77,3 %) / 0 (0 %)
Площадь жилищ, приходящаяся в среднем на одного жителя, м <sup>2</sup>	28,93	28,3	0,0005	3 (13,6 %) / 15 (62,5 %)	0 (0 %) / 1 (4,2 %)
Инвестиции в основной капитал, млн руб.	–	228 237	0,0013	2 (9,1 %) / 8 (33,3 %)	17 (77,3 %) / 0 (0 %)
Площадь жилищ, приходящаяся в среднем на одного жителя, м <sup>2</sup>	28,93	28,3	0,0005	3 (13,6 %) / 15 (62,5 %)	0 (0 %) / 1 (4,2 %)
Основные фонды, млн руб.	–	3 304 900	0,0023	2 (9,1 %) / 8 (33,3 %)	17 (77,3 %) / 0 (0 %)
Средний размер назначенных пенсий, 2019 г.	28,91	14 708	0,02567	18 (81,8 %) / 1 (4,2 %)	3 (13,6 %) / 11 (45,8 %)
Уровень безработицы, %, 2019 г.	–	4	0,0005	0 (0 %) / 8 (33,3 %)	1 (4,5 %) / 4 (16,7 %)
Новообразования (на 1 тыс. чел. выявлено впервые)	28,78	8,95	0,00167	5 (22,7 %) / 0 (0 %)	5 (22,7 %) / 24 (100 %)
Инвестиции в основной капитал, млн руб.	–	71 066	0,01333	7 (31,8 %) / 0 (0 %)	5 (22,7 %) / 0 (0 %)
Инфекционные и паразитарные (на 1 тыс. выявлено впервые)	28,78	18,5	0,00033	5 (22,7 %) / 0 (0 %)	5 (22,7 %) / 24 (100 %)
Инвестиции в основной капитал, млн руб.	–	71 066	0,01167	11 (50 %) / 0 (0 %)	1 (4,5 %) / 0 (0 %)
Инфекционные и паразитарные (на 1 тыс. выявлено впервые)	28,78	18,5	0,00033	4 (18,2 %) / 0 (0 %)	5 (22,7 %) / 24 (100 %)

Показатель	Chi-квadrat	Граница	p-value	Распределение случаев в квадрантах разбиения	
Инвестиции в основной капитал, млн руб. на 1 тыс. граждан	–	65,81	0,028	11 (50 %) / 0 (0 %)	2 (9,1 %) / 0 (0 %)
Новообразования (на 1 тыс. чел. выявлено впервые)	25,1	11,45	0,0005	14 (63,6 %) / 0 (0 %)	3 (13,6 %) / 6 (25 %)
Уровень безработицы, %, 2019 г.	–	4,9	0,00834	5 (22,7 %) / 8 (33,3 %)	0 (0 %) / 10 (41,7 %)
Уровень безработицы, %, 2019 г.	22,64	5,65	0,0467	4 (19 %) / 17 (77,3 %)	13 (61,9 %) / 0 (0 %)
Удельный вес автомобильных дорог с твердым покрытием, %, 2020 г.	–	66,14	0,0005	0 (0 %) / 2 (9,1 %)	4 (19 %) / 3 (13,6 %)
Число смертей на 100 тыс. чел. на 8 сентября 2021 г.	22,18	184,3	0,00567	21 (95,5 %) / 6 (26,1 %)	0 (0 %) / 5 (21,7 %)
Уровень безработицы, %, 2019 г.	–	4	0,0334	0 (0 %) / 2 (8,7 %)	1 (4,5 %) / 10 (43,5 %)
Уровень безработицы, %, 2019 г.	22	7,1	0,049	7 (31,8 %) / 23 (95,8 %)	10 (45,5 %) / 0 (0 %)
Доходы от предпринимательской деятельности, % от общего дохода, 2019 г.	–	2,9	0,00267	1 (4,5 %) / 1 (4,2 %)	4 (18,2 %) / 0 (0 %)
Число смертей на 100 тыс. чел. на 8 сентября 2021 г.	21,68	184,3	0,01	19 (86,4 %) / 4 (17,4 %)	0 (0 %) / 6 (26,1 %)
Младенческая смертность (число детей, умерших до года на 1 тыс. родившихся живыми)	–	4,35	0,0223	0 (0 %) / 1 (4,3 %)	3 (13,6 %) / 12 (52,2 %)
Отправлено грузов, млн т (ж/д), 2020 г.	21,46	22,15	0,0033	0 (0 %) / 7 (33,3 %)	4 (20 %) / 5 (23,8 %)
Перевозка пассажиров автобусами, млн чел., 2020 г.	–	123,2	0,00533	0 (0 %) / 6 (28,6 %)	16 (80 %) / 3 (14,3 %)
Ожидаемая продолжительность жизни, 2019 г.	21,33	72,33	0,00133	6 (27,3 %) / 9 (37,5 %)	13 (59,1 %) / 3 (12,5 %)
Уровень безработицы, %, 2019 г.	–	4,15	0,009	0 (0 %) / 12 (50 %)	3 (13,6 %) / 0 (0 %)
Уровень безработицы, %, 2019 г.	21,26	3,85	0,02533	0 (0 %) / 6 (28,6 %)	8 (40 %) / 12 (57,1 %)
Отправлено грузов, млн т (ж/д), 2020 г.	–	10,25	0,014	12 (60 %) / 0 (0 %)	0 (0 %) / 3 (14,3 %)
Удельный вес расходов на оплату ЖКХ, % от общего числа услуг	20,81	33,4	0,002	2 (10 %) / 0 (0 %)	2 (10 %) / 12 (57,1 %)
Отправлено грузов, млн т (ж/д), 2020 г.	–	21,5	0,008	16 (80 %) / 4 (19 %)	0 (0 %) / 5 (23,8 %)
Отправлено грузов, млн т (ж/д), 2020 г.	20,03	22,15	0,008334	2 (10 %) / 9 (42,9 %)	4 (20 %) / 7 (33,3 %)
Пассажирооборот, млн чел. / 1 тыс. км, 2020 г.	–	805,5	0,03533	0 (0 %) / 4 (19 %)	14 (70 %) / 1 (4,8 %)
Болезни системы кровообращения (на 1 тыс. чел. выявлено впервые)	19,83	34,2	0,00333	11 (52,4 %) / 2 (9,1 %)	2 (9,5 %) / 6 (27,3 %)
Погибло в происшествиях, чел. (на 100 тыс. чел. населения), 2020 г.	–	12,38	0,005667	7 (33,3 %) / 2 (9,1 %)	1 (4,8 %) / 12 (54,5 %)
Болезни нервной системы (на 1 тыс. чел. выявлено впервые)	18,99	20,45	0,00433	2 (9,1 %) / 15 (62,5 %)	4 (18,2 %) / 0 (0 %)
Численность населения, тыс. чел.	–	1 514	0,031	0 (0 %) / 2 (8,3 %)	16 (72,7 %) / 7 (29,2 %)
Новообразования (на 1 тыс. чел. выявлено впервые)	17,92	11,45	0,00967	3 (13,6 %) / 10 (41,7 %)	6 (27,3 %) / 6 (25 %)
Площадь жилищ, приходящаяся в среднем на одного жителя, м <sup>2</sup>	–	25,45	0,016	2 (9,1 %) / 8 (33,3 %)	11 (50 %) / 0 (0 %)
Коэффициент рождаемости (число рождений на одну женщину)	16,91	1,447	0,037	1 (4,5 %) / 0 (0 %)	3 (13,6 %) / 10 (41,7 %)
Коэффициент Джини	–	0,3955	0,03067	15 (68,2 %) / 3 (12,5 %)	3 (13,6 %) / 11 (45,8 %)
Болезни нервной системы (на 1 тыс. чел. выявлено впервые)	13,97	14,7	0,048	4 (19 %) / 10 (45,5 %)	11 (52,4 %) / 5 (22,7 %)
Плотность автомобильных дорог общего пользования с твердым покрытием (км путей на 1 тыс. км <sup>2</sup> территории), 2020 г.	–	131,8	0,047	1 (4,8 %) / 7 (31,8 %)	5 (23,8 %) / 0 (0 %)

Примечание: квадранты представлены как на диаграммах рассеяния: I – верхний левый, II – верхний правый, III – нижний правый, IV – нижний левый; ЖКХ – жилищно-коммунальное хозяйство

Составлено авторами по материалам исследования / Compiled by the authors on the materials of the study

К уже выявленным ранее показателям в данном случае добавились характеристики здоровья населения, социальные условия, состояние дорог, коэффициенты расслоения населения по обеспеченности. Высокая младенческая смертность (число детей, умерших до года на 1 тыс. родившихся живыми) в 5 дотационных регионах: Чукотском АО, Камчатском крае, Дагестане, Забайкальском крае и Тыве сочетается одновременно с низкими значениями основных фондов на 1 тыс. граждан.

Низкие величины для показателя средний размер назначенных пенсий, 2019 г. значимо характеризуют дотационную группу в паре с большими числами показателя уровень безработицы, %, 2019 г. Только два региона (Чукотский АО, Камчатский край) попадают в квадрант с преобладанием значений второй благоприятной группы, что свидетельствует в пользу прогноза перехода их в группу регионов-доноров.

### Чистый контроль / Pure control

Группу регионов, которые в начале исследования не вошли в обучающую выборку, подвергли распознаванию с помощью автономной программы «Прогноз» с «защитым» в нее решающим правилом.

Из 36 субъектов РФ 28 были отнесены в группу дотационных регионов: области Орловская, Рязанская, Смоленская, Тамбовская, Тверская, Новгородская, Псковская, Оренбургская, Пензенская, Саратовская, Ульяновская, Челябинская, Астраханская, Волгоградская, Омская, Томская, Амурская, Магаданская, Кемеровская, Еврейская автономная, Севастополь, республики Марий Эл, Мордовия, Карелия, Адыгея, Калмыкия, Удмуртская, Хакасия.

В благополучную вторую группу были отнесены алгоритмом распознавания Республика Коми, области Архангельская, Архангельская без АО, Калининградская, Тюменская без АО, Новосибирская, Приморский и Хабаровский края. Это говорит о том, что данные субъекты РФ по своему социально-экономическому портрету ближе к бюджетобразующим регионам-донорам.

### Обсуждение / Discussion

Сравнительный анализ методами МО свидетельствует, что наиболее значимыми показателями стали инвестиции в основной капитал (млн руб. и млн руб. на 1 тыс. граждан), основные фонды (млн руб. и млн руб. на 1 тыс. граждан), а также среднедушевые доходы (руб., 2020 г.), средний размер назначенных пенсий (2019 г.), уровень безработицы (%), 2019 г.). Другие значимые показатели характеризуют

демографическую ситуацию, состояние здравоохранения, инфраструктуру регионов.

Направления дальнейших исследований предполагаются следующими: поскольку предложенный математический подход продемонстрировал эффективность в поиске ключевых факторов, можно применить аналогичную технологию для анализа более полной базы данных, содержащей показатели промышленного сектора, предприятия различных сфер деятельности, образовательные учреждения, социальные службы и объекты, привлекательные для инвестиций. Особенный интерес представляют показатели, отражающие основные фонды регионов, поскольку они оказались одними из самых значимых.

Методы ОДР и СВС позволяют создать план перевода объекта из группы с неблагоприятным прогнозом в благоприятную группу, демонстрируя, на какие ключевые показатели надо в первую очередь обращать внимание и каких границ требуется достичь, для того чтобы увеличить вероятность благоприятного прогноза.

Для каждого региона из контрольной группы, распознанного как дотационный, можно дать четкий план мероприятий по переводу его в благоприятную группу в виде набора 20 значимых показателей с границами разбиения, которые нужно достигнуть и перейти. После достижения этой цели вероятность того, что регион сможет отказаться от дотаций из центра, станет вполне реальной. Это выведет практику работы с дотационными регионами на настоящему научный уровень.

### Заключение / Conclusion

Приведем теоретические соображения по поводу помощи слабым регионам. Комплексная программа поддержки развития дотационных регионов включает в себя следующее мероприятие: осуществление стимулирования привлечения инвестиций в их экономику, для чего требуется использовать такие меры, как налоговые льготы, финансовая поддержка, создание инфраструктуры, партнерство с бизнесом, привлечение иностранных инвесторов. В частности, их можно расписать следующим образом:

1) предоставление инвесторам налоговых льгот и освобождение от некоторых налоговых платежей на определенный период времени могут стать мощным стимулом для привлечения инвестиций;

2) государственные гранты, субсидии, льготные кредиты и другие формы финансовой поддержки способны повысить привлекательность дотационных регионов для инвесторов;

3) развитие инфраструктуры, такой как транспортные коммуникации, энергетика, связь и др.,

способствует улучшению инвестиционного климата региона;

4) создание партнерских отношений с частными компаниями и предпринимателями, в том числе через публично-частные партнерства, может привлечь инвестиции в регион;

5) активная маркетинговая кампания за рубежом, участие в международных выставках и форумах, а также создание специальных инвестиционных площадок могут привлечь иностранные инвестиции.

Эти меры в совокупности в состоянии способствовать успешному стимулированию привлечения

инвестиций в экономику дотационных регионов и их развитию.

Для развития инфраструктуры в дотационных регионах и повышения их конкурентоспособности необходимо реализовать прежде всего меры, связанные со строительством и модернизацией транспортной инфраструктуры, а именно развитие дорожной сети, строительство и реконструкция аэропортов, железнодорожной инфраструктуры и портов, что позволит улучшить доступность субъекта и транспортную связь с другими регионами.

## СПИСОК ЛИТЕРАТУРЫ

*Алимурадов М.К., Мидов А.З., Одинцов С.В.* Стратегический анализ бюджетной обеспеченности высокодотационных регионов. *Экономическое возрождение России*. 2021;2(68):113–129. <http://dx.doi.org/10.37930/1990-9780-2021-2-68-113-129>

*Березняцкий А.Н.* Моделирование макроэкономической динамики дотационного региона в России. В кн.: *Применение многомерного статистического анализа в экономике и оценке качества: материалы XII Международной научно-практической конференции имени С.А. Айвазяна, Москва, 21–23 сентября 2022 г. М.: Высшая школа экономики; 2022. С. 42–45.*

*Борисова Л.Р., Кузнецова А.В.* Использование работающего компьютерного тренажера Data Master Azforus для обучения методам машинного обучения. В кн.: *Цифровая трансформация социальных и экономических систем: материалы международной научно-практической конференции, Москва, 28 января 2022 г. М.: Московский университет имени С.Ю. Витте; 2022. С. 264–270.*

*Кирилюк И.Л., Кузнецова А.В., Сенько О.В.* Исследование взаимосвязи производственных функций и социально-экономических показателей российских регионов методом оптимальных разбиений. *Информационные технологии и вычислительные системы*. 2021;1:20–31. <https://doi.org/10.14357/20718632210103>

*Косинский П.Д., Zubova A.V., Yurzina T.A.* Методические подходы к оценке самодостаточности бюджета муниципальных образований. *Региональная экономика и управление*. 2022;3(71).

*Лукьянова М.Н.* Ранжирование причин дотационности муниципальных образований России с применением мозгового штурма и экспертных оценок. *Международный журнал прикладных и фундаментальных исследований*. 2015;10–2:329–331.

*Симонов А.Б., Баранов С.А., Сазонов С.П.* Изучение факторов, влияющих на уровень бюджетной обеспеченности регионов, на основании математических методов. *Экономика и социум*. 2016;6(25):626–633.

*Таштамиров М.Р., Байсаева М.У., Баташев Р.В.* Систематизация факторов и условий высокой дотационности региональных бюджетов. *Фундаментальные исследования*. 2020;11:185–192. <https://doi.org/10.17513/fr.42896>

## REFERENCES

*Alimuradov M.K., Midov A.Z., Odintsov S.V.* Strategic analysis of highly subsidized regions' budgetary sufficiency. *Economic Revival of Russia*. 2021;2(68):113–129. (In Russian). <http://dx.doi.org/10.37930/1990-9780-2021-2-68-113-129>

*Bereznyatsky A.N.* Modelling of the macroeconomic dynamics of the subsidised region in Russia. In: *Application of the multidimensional statistical analysis in economics and quality assessment: Proceedings of the XII International Scientific and Practical Conference named after S.A. Ayvazyan, Moscow, September 21–23; 2022. Moscow: Higher School of Economics. Pp. 42–45. (In Russian).*

*Borisova L.R., Kuznetsova A.V.* Using a working computer simulator Data Master Azforus for teaching machine learning methods. In: *Digital transformation of social and economic systems: Proceedings of the International Scientific and Practical Conference, Moscow, January 28, 2022. Moscow: Moscow Witte University; 2022. Pp. 264–270. (In Russian).*

*Eneeva M.N., Ulbasheva A.R., Uyanaeva H.B.* Factors and causes an endowment of regional economies of the North Caucasus Federal District. *Terra Economicus*. 2010;4–3(8):173–176. (In Russian).

*Kirilyuk I.L., Kuznetsova A.V., Senko O.V.* Investigation of the relationship between production functions and socio-economic indicators of Russian regions by the method of optimal partitioning. *Journal of Information Technologies and Computing Systems*. 2021;1:20–31. (In Russian). <https://doi.org/10.14357/20718632210103>

*Kosinskiy P.D., Zubova A.V., Yurzina T.A.* Methodological approaches to assessing self-sufficiency of municipal budget. *Regional economy and management*. 2022;3(71). (In Russian).

*Lukyanova M.N.* Ranking of the budget deficit reasons of the municipalities in Russia with application of brainstorming and expert estimates. *International Journal of Applied and Fundamental Research*. 2015;10–2:329–331. (In Russian).

*Shilov M.A., Gusev A.B.* Technological modernisation of the regional economy as a factor of balanced development of the Russian territory. *Science Governance and Scientometrics*. 2009;8:143–157. (In Russian).

*Simonov A.B., Baranov S.A., Sazonov S.P.* The study of factors affecting the level of budgetary provision of regions based on mathematical methods. *Economics and socium*. 2016;6(25):626–633. (In Russian).

*Шилов М.А., Гусев А.Б.* Технологическая модернизация экономики регионов как фактор сбалансированного развития территории России. Управление наукой и наукометрия. 2009;8:143–157.

*Энеева М.Н., Ульбашева А.Р., Уянаева Х.Б.* Факторы и причины дотационности региональных экономик СКФО. Terra Economicus. 2010;4–3(8):173–176.

*Tashtamirov M.R., Baysaeva M.U., Batashev R.V.* Systematization of factors and conditions of high subsidization of regional budgets. Fundamental research. 2020;11:185–192. (In Russian). <https://doi.org/10.17513/fr.42896>