

# МАТЕМАТИЧЕСКИЕ И ИНСТРУМЕНТАЛЬНЫЕ МЕТОДЫ В ЭКОНОМИКЕ

---

УДК 004.8

JEL C88

---

## КЛАСТЕРНЫЙ АНАЛИЗ РЕГИОНОВ РОССИЙСКОЙ ФЕДЕРАЦИИ ПО ПОТРЕБНОСТЯМ В РАБОТНИКАХ НА ОСНОВЕ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ

---

**Быстренина Ирина Евгеньевна**, канд. пед. наук, доц.

**Борин Игорь Михайлович**, студент

Российский экономический университет им. Г.В. Плеханова, Стремянный пер., 36,  
Москва, Россия, 115054; e-mail: iesh@rambler.ru

*Предмет:* в данной работе раскрываются проблемы кластеризации регионов Российской Федерации как необходимого инструмента развития кадрового потенциала страны. В свою очередь, развитие кадрового потенциала страны по различным отраслям экономики должно происходить системно, учитывая тенденцию спроса в кадрах с необходимым набором компетенций и имеющимся кадровым потенциалом. *Цель:* кластеризация регионов согласно потребности организаций в работниках для замещения вакантных рабочих мест по профессиональным группам и субъектам Российской Федерации. *Дизайн исследования:* авторами исследования был реализован алгоритм кластеризации K-средних с использованием библиотеки scikit-learn в Python, который направлен на выбор центроидов с целью минимизации суммы квадратов расстояний внутри кластера. Количество кластеров для набора данных осуществлялся с помощью метода локтя. В ходе проведенного исследования было выявлено четыре кластера регионов Российской Федерации. К первому кластеру в большей степени принадлежат регионы Южного и Северо-Западного федеральных округов, а также города Москва и Санкт-Петербург. Второй кластер, самый количественный по составу, включает многие регионы Приволжского и Центрального федеральных округов. В третий кластер вошли в основном регионы Северо-Кавказского федерального округа. Четвертый кластер включил регионы Центрального, Сибирского, Дальневосточного федеральных округов. *Результаты:* разработанная методика проведения кластерного анализа может способствовать развитию кадрового потенциала Российской Федерации. В частности, она

может стать инструментом в решении вопроса оптимизации подготовки кадров, повышения мобильности трудовых ресурсов, эффективного распределения ресурсов, а также создать условия для устойчивого экономического роста.

**Ключевые слова:** кластерный анализ, машинное обучение, метод локтя, кадровый потенциал, профессиональные группы.

**DOI:** 10.17308/meps/2078-9017/2025/10/8-20

### **Введение**

Одним из основных условий успешного развития страны является формирование отвечающего запросам экономики кадрового потенциала. Наблюдается интерес государства к вопросу нормативного регулирования развития кадрового потенциала Российской Федерации, что находит отражение в принятой еще в 2013 году Стратегии развития системы подготовки рабочих кадров и формирования прикладных квалификаций в Российской Федерации на период до 2020 года. Данная Стратегия была направлена на становление прикладных квалификаций, модернизации образовательных учреждений и сотрудничество с работодателями, что должно способствовать улучшению качества подготовки кадров и их успешной интеграции в экономику. Следующий этап развития кадрового потенциала страны находит отражение в Стратегии развития национальной системы квалификаций до 2030 года, заключающейся в формировании современной, гибкой системы кадрового обеспечения, которая будет соответствовать приоритетным задачам научно-технологического и социально-экономического развития страны. Для реализации целей данной стратегии был запущен в 2025 году национальный проект «Кадры», включающий в себя меры по подготовке и переподготовке кадров в различных отраслях экономики.

В отечественной науке сущность кадрового потенциала изучается в работах ряда ученых [4, 5, 10, 11 и др.]. Так, Тумаева Т.В., давая определение кадрового потенциала, подчеркивает важность образования и профессиональной подготовки, понимая под ним совокупность всех человеческих ресурсов, включая знания, умения и квалификацию, которые могут быть использованы для достижения целей организации или общества [11]. В своих работах В.Н. Батова и В.В. Рассадин рассматривают кадровый потенциал как совокупность физических и духовных качеств человека, которые определяют его трудовые ресурсы [5]. Это включает в себя не только профессиональные навыки, но и личные качества, способствующие эффективной трудовой деятельности.

Отметим также, что развитие кадрового потенциала страны по различным отраслям экономики должно происходить системно, учитывая тенденцию спроса в кадрах с необходимым набором компетенций и имеющимся кадровым потенциалом. Одним из инструментов оценки регионов в кадровом потенциале является кластерный анализ, являющийся одним из мето-

дов машинного обучения, актуальность использования которых продолжает расти, и их применение охватывает все больше сфер [1]. Так, кластерный анализ уже активно используется для решения многих задач: автоматизации процесса управления производством органической продукции [6, 7], анализе влияния климатических условий на налоговый потенциал регионов России [2], анализе регионов России по уровню безработицы в период с 2010 по 2019 год [12], анализе уровней занятости и развития в сельских районах России [13] и др.

Изменения на рынке труда оказывают значимое влияние на экономические процессы страны. И принятие многих решений усложняется существенной региональной разнородностью занятости населения, оплаты труда, миграционных потоков и т.д. И стала актуальна синхронизация знаний, умений и навыков рабочей силы с запросом рынка труда. На практике это позволяет вести подготовку кадров именно тому, что будет востребовано в экономике в ближайшие годы, проводить оптимизацию в данном процессе.

Под кластерным анализом понимают статистический метод, используемый для группировки объектов на основе их подобия, разделяющий набор данных на группы, называемые кластерами, так, чтобы объекты внутри одного кластера были максимально похожи друг на друга, а объекты из разных кластеров максимально различались [8]. В рамках исследования методом кластерного анализа будут выделены группы (кластеры) регионов, дифференцированные по потребностям организаций в работниках для замещения вакантных рабочих мест по профессиональным группам. За виды экономической деятельности взяты следующие сферы: обслуживания и торговли, охраны граждан и собственности; сельского и лесного хозяйств, рыбоводства и рыболовства; промышленности, строительства, транспорта и рабочих родственных занятий; производственных установок и машин, сборки и управления транспортными средствами.

### **Методы и результаты исследования**

Кластеризация, являясь методом анализа данных, используется для группировки объектов на основе их схожести. Различные алгоритмы кластеризации используют различные меры для оценки сходства между объектами. Эти меры подбираются в зависимости от задачи и шкалы измерений. Для числовых данных очень часто применяют евклидово расстояние:

$$D(x, y) = \sqrt{\sum_i (x_i - y_i)^2}$$
, а категориальные атрибуты обычно сравнивают при помощи меры Чекановского-Серпенсена и Жаккара  $(|t_1 \cap t_2| / |t_1 \cup t_2|)$ .

Поиск количества кластеров для выбранных данных осуществляется с помощью метода локтя. Суть данного метода заключается в реализации следующих шагов:

- Определение диапазона значений  $k$ . Обычно это значение варьируется от 1 до 10 или от 1 до 15, в зависимости от объема данных и их сложности.

- Выполнение кластеризации. Для каждого значения  $k$  выполняется алгоритм  $k$ -средних и вычисляется сумма квадратов внутрикластерных расстояний (WCSS). Это значение показывает, насколько хорошо данные сгруппированы внутри кластеров.
- Построение графика. Строится график, где по оси абсцисс указывается количество кластеров  $k$ , а по оси ординат – соответствующее значение WCSS. Это позволит визуально оценить, как изменяется внутрикластерная дисперсия с увеличением числа кластеров.
- Поиск «локтя». На графике определяется точка, где кривая начинает «выпрямляться» или замедлять свое снижение. Эта точка называется «локтем» и указывает на оптимальное количество кластеров. Обычно это значение  $k$  соответствует наименьшему значению WCSS, при котором дальнейшее увеличение числа кластеров не приводит к значительному улучшению.

Применение данного метода позволит выявлять группы регионов со схожими потребностями организаций в работниках для замещения вакантных рабочих мест по профессиональным группам и субъектам Российской Федерации для выработки общих методов и подходов к формированию кадрового потенциала регионов.

В основе метода кластеризации заложен алгоритм  $K$ -средних [3, 9]. Он направлен на выбор центроидов, которые минимизируют суммы квадратов расстояний внутри кластера:

$$\sum_{i=0}^n \min_{\mu_j \in C} (\|x_i - \mu_j\|^2).$$

Анализ был проведен по 82 регионам России, данные по которым имелись. В качестве инструмента был использован язык программирования python. В качестве данных был взят датасет «Распределение потребности организаций в работниках для замещения вакантных рабочих мест по профессиональным группам и субъектам Российской Федерации на 31 октября 2022 г.»<sup>1</sup>. Реализация метода локтя для поиска количества кластеров со схожими потребностями организаций в работниках для замещения вакантных рабочих мест по профессиональным группам и субъектам Российской Федерации представлена ниже (листинг 1).

Листинг 1. Реализация метода локтя

```
inertias = []
for k in range(1, 12):
    kmeans = KMeans(n_clusters=k)
    kmeans.fit(data)
    inertias.append(kmeans.inertia_)
```

В данной реализации создается список `inertias`, который будет исполь-

<sup>1</sup> Федеральная служба государственной статистики [Электронный ресурс]. Режим доступа: <https://rosstat.gov.ru/> (дата обращения: 06.07.2025).

зоваться для хранения значений суммы квадратов расстояний от точек до центроидов кластеров (инерция) для различных количеств кластеров. Далее в цикле происходят следующие действия:

- проходя по значениям  $k$  от 1 до 11, для каждого из значений создается объект KMeans с заданным количеством кластеров;
- метод `fit` обучает модель на данных `data`;
- значение инерции добавляется в список `inertias`.

Для визуализации метода локтя нами используется библиотека `matplotlib` языка программирования Python (листинг 2).

Листинг 2. Реализация метода локтя

```
plt.plot(range(1, 12), inertias, marker='o')
plt.title(«Метод локтя»)
plt.xlabel(«Число кластеров (k)»)
plt.ylabel(«Сумма внутрикластерных расстояний»)
plt.show()
```

Графическая иллюстрация метода локтя представлена на рис. 1. Точка, где кривая начинает «выпрямляться», на графике соответствует  $k$ , равному 4.

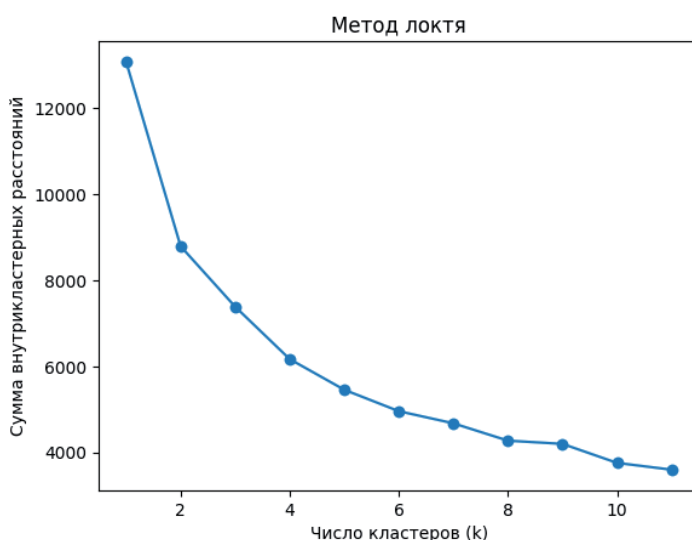


Рис. 1. Реализация метода локтя для поиска количества кластеров со схожими потребностями организаций в работниках для замещения вакантных рабочих мест по профессиональным группам и субъектам Российской Федерации

Для исследования распределения кадрового потенциала страны по регионам России были взяты следующие профессиональные группы:

- руководители;
- специалисты высшего уровня квалификации;

- специалисты среднего уровня квалификации;
- служащие, занятые подготовкой и оформлением документации, учетом и обслуживанием;
- работники сферы обслуживания и торговли, охраны граждан и собственности;
- квалифицированные работники, сельского и лесного хозяйств, рыбоводства и рыболовства;
- квалифицированные рабочие промышленности, строительства, транспорта и рабочие родственных занятий;
- операторы производственных установок и машин, сборщики и водители;
- неквалифицированные рабочие.

Реализация алгоритма кластеризации K-средних с использованием библиотеки `scikit-learn` в Python представлена ниже (листинг 3).

Листинг 3. Реализация алгоритма кластеризации K-средних

```
current_kmeans = KMeans(n_clusters=4)
current_kmeans.fit(data)
labels = current_kmeans.labels_
centroids = current_kmeans.cluster_centers_
```

Массив `labels` содержит метки кластеров для каждой точки данных, указывая, к какому кластеру принадлежит каждая точка.

Состав кластеров представлен в таблице 1. Как видно, распределение регионов по кластерам неравномерно. В 1-й кластер вошли 17 регионов, во 2-й – 32 региона , в 3-й кластер вошли 4 региона, в 4-й – 29 регионов.

Таблица 1

Распределение регионов РФ по кластерам

Номер кластера	Регионы кластера
1	г. Москва, Калининградская область, Новгородская область, Псковская область, г.Санкт-Петербург, Республика Адыгея, Республика Калмыкия, Краснодарский край, Астраханская область, Город федерального значения Севастополь, Кабардино-Балкарская Республика, Республика Северная Осетия – Алания, Ставропольский край, Республика Алтай, Томская область, Республика Бурятия, Камчатский край
2	Белгородская область, Воронежская область, Костромская область, Курская область, Московская область, Тверская область, Тульская область, Ярославская область, Республика Карелия, Архангельская область, Ленинградская область, Мурманская область, Волгоградская область, Карачаево-Черкесская Республика, Республика Башкортостан, Республика Татарстан, Удмуртская Республика, Чувашская Республика, Пермский край, Кировская область, Нижегородская область, Оренбургская область, Самарская область, Саратовская область, Ульяновская область, Курганская область, Свердловская область, Челябинская область, Кемеровская область, Забайкальский край, Приморский край, Хабаровский край
3	Республика Дагестан, Республика Ингушетия, Чеченская Республика, Республика Тыва

Номер кластера	Регионы кластера
4	Брянская область, Владимирская область, Ивановская область, Калужская область, Липецкая область, Орловская область, Рязанская область, Смоленская область, Тамбовская область, Республика Коми, Вологодская область, Республика Крым, Ростовская область, Республика Марий Эл, Республика Мордовия, Пензенская область, Тюменская область, Республика Хакасия, Алтайский край, Красноярский край, Иркутская область, Новосибирская область, Омская область, Республика Саха (Якутия), Амурская область, Магаданская область, Сахалинская область, Еврейская автономная область, Чукотский автономный округ

К первому кластеру принадлежат города Москва и Санкт-Петербург и в большей степени регионы Южного и Северо-Западного федеральных округов. Второй кластер самый количественный по составу. В его состав вошли многие регионы Приволжского и Центрального федеральных округов. В третий кластер вошли в основном регионы Северо-Кавказского федерального округа. Четвертый кластер включил регионы Центрального, Сибирского, Дальневосточного федеральных округов.

В таблице 2 приведены значения различных показателей потребности организаций в работниках для замещения вакантных рабочих мест по профессиональным группам и субъектам Российской Федерации на 31 октября 2022 г. по различным видам экономической деятельности.

Таблица 2

Средние значения ряда показателей (в процентах) по кластерам на 31 октября 2022 г.

Профессиональные группы	Кластер 1	Кластер 2	Кластер 3	Кластер 4
Руководители	4,1	3,7	3,1	4
Специалисты высшего уровня квалификации	27,4	20,9	48,1	23
Специалисты среднего уровня квалификации	14,4	10,3	17,1	12
Служащие, занятые подготовкой и оформлением документации, учетом и обслуживанием	4,1	3,3	4,6	4
Работники сферы обслуживания и торговли, охраны граждан и собственности	14,3	13,3	6,6	9
Квалифицированные работники, сельского и лесного хозяйств, рыбоводства и рыболовства	1,3	1,3	-	2
Квалифицированные рабочие промышленности, строительства, транспорта и рабочие родственных занятий	11,7	22,6	5,7	17
Операторы производственных установок и машин, сборщики и водители	8,9	14,2	6,7	18
Неквалифицированные рабочие	13,8	10,4	8,1	11
Всего	100	100	100	100

Анализ данных распределения регионов по кластерам по профессиональным группам позволяет сделать выводы, что:

– первый кластер характеризуется высокой потребностью в работниках высшего (27,4%) и среднего (14,4%) уровня квалификации, сферы обслуживания и торговли, охраны граждан и собственности (14,3%), а также в неквалифицированных рабочих (13,8%). Данная тенденция обусловлена экономическими, демографическими и социальными факторами. Так, многие из этих регионов, такие как Москва и Санкт-Петербург, являются экономическими центрами, где активно развиваются сферы услуг и торговли, что создает спрос на специалистов данных сфер. В регионах, таких как Краснодарский край и Калининградская область, наблюдается активное строительство и развитие соответствующей инфраструктуры, что требует большого количества рабочих как квалифицированных, так и неквалифицированных. При этом в регионах данного кластера присутствует слабый запрос экономики на следующие профессиональные группы: квалифицированные работники, сельского и лесного хозяйств, рыбоводства и рыболовства (1,3%); служащие, занятые подготовкой и оформлением документации, учетом и обслуживанием, и руководители (по 4,1%);

– второй кластер характеризуется наибольшей потребностью в квалифицированных рабочих промышленности, строительства, транспорта и рабочих родственных занятий (22,6%), специалистах высшего уровня квалификации (20,9%), операторах производственных установок и машин, сборщиках и водителях (14,2%), а также в работниках сферы обслуживания и торговли, охраны граждан и собственности (13,3%). Данная тенденция обусловлена тем, что в последние годы наблюдается активное развитие промышленности, строительства и транспорта в регионах данного кластера. Параллельно с ростом промышленности соответственно увеличивается потребность в работниках сферы обслуживания и торговли. Это связано с ростом населения и увеличением потребительского спроса, что требует большего числа сотрудников в этих секторах экономики. Слабый запрос в регионах второго кластера представлен на квалифицированных работников, сельского и лесного хозяйств, рыбоводства и рыболовства (1,3%), служащих, занятых подготовкой и оформлением документации, учетом и обслуживанием (3,3%);

– регионы третьего кластера испытывают потребность в специалистах высшего (48,1%) и среднего (17,1%) уровня квалификации. Они активно развивают свои экономики, что требует высококвалифицированных специалистов в различных отраслях, включая энергетику, здравоохранение и информационные технологии. А с увеличением населения и развитием социальной инфраструктуры возрастает потребность в специалистах в области образования, медицины и социальных услуг. Это создает дополнительный спрос на квалифицированные кадры, способные обеспечить качественное обслуживание и поддержку. Отметим, что в регионах этого кластера от-



существует запрос на квалифицированных работников, сельского и лесного хозяйств, рыбоводства и рыболовства. А низкий запрос экономики присутствует на профессиональные группы: служащие, занятые подготовкой и оформлением документации, учетом и обслуживанием (4,6%) и руководители (3,1%);

– четвертый кластер характеризуется потребностью в специалистах высшего уровня квалификации (23%), операторах производственных установок и машин, сборщиках и водителях (18%), а также в квалифицированных рабочих промышленности, строительства, транспорта и рабочих родственных занятий (17%). Это обусловлено тем, что в регионах наблюдается активное развитие промышленности, строительства и транспорта, что требует наличия высококвалифицированных специалистов для управления современными технологиями и оборудованием. С переходом на автоматизированные и роботизированные системы производства возрастает потребность в операторах машин и оборудования, которые могут управлять сложными технологическими процессами<sup>2</sup>. В некоторых регионах наблюдается рост населения, что создает дополнительный спрос на рабочие места в сфере услуг, строительства и производства. Миграция рабочей силы также влияет на структуру спроса на квалифицированных специалистов<sup>3</sup>. Несмотря на то, что в этих регионах доля запроса в квалифицированных работниках, сельского и лесного хозяйств, рыбоводства и рыболовства составляет 2%, однако регионы данного кластера являются наиболее остро нуждающимися среди регионов остальных кластеров в данных специалистах.

### **Заключение**

В ходе исследования выполнена кластеризация регионов согласно потребности организаций в работниках для замещения вакантных рабочих мест по профессиональным группам и субъектам Российской Федерации. Для осуществления кластерного анализа применялся метод машинного обучения без учителя k-средних. В исходный набор данных вошли показатели следующих профессиональных групп: руководители; специалисты высшего уровня квалификации; специалисты среднего уровня квалификации; служащие, занятые подготовкой и оформлением документации, учетом и обслуживанием; работники сферы обслуживания и торговли, охраны граждан и собственности; квалифицированные работники, сельского и лесного хозяйств, рыбоводства и рыболовства; квалифицированные рабочие промышленности, строительства, транспорта и рабочие родственных занятий; операторы производственных установок и машин, сборщики и водители; неквалифицированные рабочие. Было установлено, что целесообразно выделить четыре региональных кластера со схожими потребностями организаций

<sup>2</sup> Нехватка специалистов в строительстве: какие профессии в дефиците? [Электронный ресурс]. Доступно: <https://xn--b1agapfwapgc1.xn--p1ai/nehvatka-specialistov-v-stroitelstve-kakie-professii-v-deficite/> (дата обращения: 06.09.2025).

<sup>3</sup> Ситуация на рынке труда [Электронный ресурс]. Доступно: [https://czn-smolensk.admin-smolensk.ru/sityac\\_na\\_rynke/](https://czn-smolensk.admin-smolensk.ru/sityac_na_rynke/) (дата обращения: 06.09.2025).

в работниках для замещения вакантных рабочих мест по профессиональным группам и субъектам Российской Федерации. Методом k-средних получено, что первый кластер включает в большей степени регионы Южного и Северо-Западного федеральных округов, а также такие крупные города, как Москва и Санкт-Петербург. Для него характерна высокая потребность в работниках высшего и среднего уровня квалификации, сферы обслуживания и торговли, охраны граждан и собственности, а также в неквалифицированных рабочих. На территории кластера наблюдается активное строительство и развитие соответствующей инфраструктуры, что требует большого количества рабочих как квалифицированных, так и неквалифицированных. При этом в регионах данного кластера присутствует слабый запрос экономики на следующие профессиональные группы: квалифицированные работники, сельского и лесного хозяйств, рыбоводства и рыболовства; служащие, занятые подготовкой и оформлением документации, учетом и обслуживанием, и руководители.

Второй кластер содержит многие регионы Приволжского и Центрального федеральных округов. Он характеризуется наибольшей потребностью в квалифицированных рабочих промышленности, строительства, транспорта и рабочих родственных занятий, специалистах высшего уровня квалификации, операторах производственных установок и машин, сборщиках и водителях, а также в работниках сферы обслуживания и торговли, охраны граждан и собственности. Данная тенденция обусловлена тем, что в последние годы наблюдается активное развитие промышленности, строительства и транспорта в этих регионах. Параллельно с ростом промышленности увеличивается потребность в работниках сферы обслуживания и торговли. Это связано с ростом населения и увеличением потребительского спроса, что требует большего числа сотрудников в этих секторах экономики.

Третий кластер, наименьший по количеству регионов, состоит в основном из регионов Северо-Кавказского федерального округа. Регионы третьего кластера испытывают потребность в специалистах высшего и среднего уровня квалификации. Они активно развивают свои экономики, что требует высококвалифицированных специалистов в различных отраслях, включая энергетику, здравоохранение и информационные технологии. А с увеличением населения и развитием социальной инфраструктуры возрастает потребность в специалистах в области образования, медицины и социальных услуг. Это создает дополнительный спрос на квалифицированные кадры, способные обеспечить качественное обслуживание и поддержку.

Четвертый кластер включил регионы Центрального, Сибирского, Дальневосточного федеральных округов. Он характеризуется потребностью в специалистах высшего уровня квалификации, операторах производственных установок и машин, сборщиках и водителях, а также в квалифицированных рабочих промышленности, строительства, транспорта и рабочих родственных занятий. Это обусловлено тем, что в регионах наблюдается ак-

тивное развитие промышленности, строительства и транспорта, что требует наличия высококвалифицированных специалистов для управления современными технологиями и оборудованием. В некоторых регионах наблюдается рост населения, что создает дополнительный спрос на рабочие места в сфере услуг, строительства и производства. Несмотря на то, что в этих регионах доля запроса в квалифицированных работниках, сельского и лесного хозяйств, рыбоводства и рыболовства составляет 2%, однако регионы данного кластера являются наиболее остро нуждающимися среди регионов остальных кластеров в данных специалистах.

Таким образом, разработанная методика проведения кластерного анализа может способствовать развитию кадрового потенциала Российской Федерации. В частности, она может стать инструментом в решении вопроса оптимизации подготовки кадров, повышению мобильности трудовых ресурсов, эффективного распределения ресурсов, создать условия для устойчивого экономического роста.

### Список источников

1. Воронин Е.А., Семкин А.Г. Управление продовольственной безопасностью методами машинного обучения в информационном пространстве аграрно-промышленного комплекса // *Надежность и качество сложных систем*, 2023, no. 4, с. 179-188.
2. Герасимова А.Е. Анализ влияния климатических условий на налоговый потенциал регионов России методом кластерного анализа // *Налоги и налогообложение*, 2024, no. 3, с. 97-114.
3. Грас Д. *Data Science. Наука о данных с нуля*: пер. с англ. 2-е изд., перераб. и доп. СПб., БХВ-Петербург, 2021. 416 с.
4. Гуриева Л.К. *Развитие кадрового потенциала малого предприятия*. Уфа, Аэтерна, 2023. 97 с.
5. Дещенко А.Ю. Эволюция теоретических подходов к определению категории «Кадровый потенциал» // *Крымский научный вестник*, 2016, no. 1 (7), с. 108-120.
6. Зарук Н.Ф., Быстренина И.Е., Харитонов А.Е. Автоматизация процесса управления производством органической продукции средствами интеллектуальной системы // *Экономические системы*, 2023, T. 16, no. 3, с. 45-63.
7. Зарук Н.Ф., Мигунов Р.А., Кагирова М.В. [и др.]. *Размещение производства органической продукции растениеводства в условиях формирования зеленой экономики*. Москва, ЭйПиСиПабблишинг, 2023. 188 с.
8. Клименко А.В., Слащев И.С. Кластерный анализ данных // *Вестник науки*, 2019, T. 1, no. 1 (10), с. 159-163. Доступно: <https://www.вестник-науки.рф/article/783> (дата обращения: 11.08.2025).
9. Нельсон Х. *Базовая математика для искусственного интеллекта*: пер. с англ. Астана, АЛИСТ, 2024. 592 с.
10. Преображенский Б.Г. *Государственное управление: развитие кадрового потенциала и трансформация функций*. Воронеж, Научная книга, 2015. 139 с.
11. Тумаева Т.В. Кадровый потенциал государственного и муниципального управления // *Вестник науки*, 2025, T. 3, no. 2 (83), с. 404 – 410. Доступно: <https://www.вестник-науки.рф/article/21519> (дата обращения: 11.08.2025).
12. Blinova T.V., Rusanovskii V.A., Markov V.A. Estimating the Impact of Economic Fluctuations on Unemployment in Russian Regions Based on the Okun Model // *Stud Russ Econ Dev*, 2021, no. 32(1), pp. 103-110.
13. Salimova G., Nigmatullina G., Habirov G., Ableeva A., Gusmanov R. Employment and development levels in rural areas of the Russian Federation // *Regional Sustainability*, 2024, T. 3, no. 5, pp. 100-164.

---

# CLUSTER ANALYSIS OF REGIONS OF THE RUSSIAN FEDERATION BY EMPLOYEE NEEDS BASED ON MACHINE LEARNING METHODS

---

**Bystrenina Irina Evgenevna**, Cand. Sci. (Ped.), Assoc. Prof.

**Borin Igor Mikhailovich**, B. IST

Plekhanov Russian University of Economics, Stremyanny lane, 36, Moscow, Russia, 115054; e-mail: iesh@rambler.ru

*Importance:* this paper reveals the problems of clustering regions of the Russian Federation as a necessary tool for developing the country's human resources potential. In turn, the development of the country's human resources potential in various sectors of the economy should occur systematically, taking into account the trend in demand for personnel with the necessary set of competencies and the existing human resources potential. *Purpose:* clustering of regions according to the needs of organizations for workers to fill vacant jobs by professional groups and subjects of the Russian Federation. *Research design:* the authors of the study implemented a K-means clustering algorithm using the scikit-learn library in Python, which is aimed at selecting centroids in order to minimize the sum of squared distances within a cluster. The number of clusters for the data set was determined using the elbow method. The study identified four clusters of regions of the Russian Federation. The first cluster mostly includes regions of the Southern and Northwestern Federal Districts, as well as the cities of Moscow and St. Petersburg. The second cluster, the most numerous in composition, includes many regions of the Volga and Central Federal Districts. The third cluster mainly includes regions of the North Caucasian Federal District. The fourth cluster includes regions of the Central, Siberian, and Far Eastern Federal Districts. *Results:* the developed methodology for conducting cluster analysis can contribute to the development of human resources potential of the Russian Federation. In particular, it can become a tool in solving the issue of optimizing personnel training, increasing labor mobility, effectively distributing resources, and creating conditions for sustainable economic growth.

**Keywords:** cluster analysis, machine learning, elbow method, human resources, professional groups.

## References

1. Voronin E.A., Semkin A.G. Upravlenie prodovol'stvennoj bezopasnost'yu metodami mashinnogo obucheniya v informacionnom prostranstve agrarno-pro-myshlennogo kompleksa [Food security management using machine learning methods in the information space of the agro-industrial complex]. *Nadezhnost' i*

*kachestvo slozhnykh sistem*, 2023, no. 4, pp. 179-188. (In Russ.)

2. Gerasimova A.E. Analiz vliyaniya klimaticheskikh uslovij na nalogovyj potencial regionov Rossii metodom klasternogo analiza [Analysis of the impact of climatic conditions on the tax potential of Russian regions using cluster analysis]. *Nalogi i nalogooblozhenie*, 2024, no. 3, pp. 97-114. (In Russ.)

3. Gras D. *Data Science. Nauka o dannykh s nulya* [Data Science. Data Science from Scratch]: per. s angl. 2-e izd., pererab. i dop. SPb., BHV-Peterburr, 2021. 416 p. (In Russ.)

4. Gurieva L.K. *Razvitie kadrovogo potenciala malogo predpriyatiya* [Development of human resources potential of small enterprises]. Ufa, Aeterna, 2023. 97 p. (In Russ.)

5. Deshchenko A.Yu. Evolyuciya teoreticheskikh podhodov k opredeleniyu kategorii «Kadrovyy potencial» [Evolution of theoretical approaches to defining the category «Human Resources Potential»]. *Krymskiy nauchnyy vestnik*, 2016, no. 1 (7), pp. 108-120. (In Russ.)

6. Zaruk N.F., Bystrenina I.E., Haritonova A.E. Avtomatizatsiya processa upravleniya proizvodstvom organicheskoy produkcii sredstvami intellektual'noy sistemy [Automation of the process of managing the production of organic products by means of an intelligent system]. *Ekonomicheskie sistemy*, 2023, T. 16, no. 3, pp. 45-63. (In Russ.)

7. Zaruk N.F., Migunov R.A., Kagirowa M.V. [i dr.]. *Razmeshchenie proizvodstva organicheskoy produkcii rasteniyevodstva v*

*usloviyakh formirovaniya zelenoy ekonomiki* [Placement of production of organic plant products in the context of the formation of a green economy]. Moskva, EjPiSiPublishing, 2023. 188 p.

8. Klimenko A.V., Slashchev I.S. [Cluster analysis of data]. *Vestnik nauki*, 2019, T. 1, no. 1 (10), pp. 159-163. (In Russ.). Available at: <https://www.vestnik-nauki.rf/article/783> (accessed: 11.08.2025).

9. Nel'son H. *Bazovaya matematika dlya iskusstvennogo intellekta* [Basic Mathematics for Artificial Intelligence]: per. s angl. Astana, ALIST, 2024. 592 p. (In Russ.)

10. Preobrazhenskij B.G. *Gosudarstvennoe upravlenie: razvitie kadrovogo potenciala i transformatsiya funktsij* [Public Administration: Development of Human Resources and Transformation of Functions]. Voronezh, Nauchnaya kniga, 2015. 139 p. (In Russ.)

11. Tumaeva T.V. [Human resources potential of state and municipal administration]. *Vestnik nauki*, 2025, T. 3, no. 2 (83), pp. 404-410. Available at: <https://www.vestnik-nauki.rf/article/21519> (accessed: 11.08.2025). (In Russ.).

12. Blinova T.V., Rusanovskii V.A., Markov V.A. Estimating the Impact of Economic Fluctuations on Unemployment in Russian Regions Based on the Okun Model. *Stud Russ Econ Dev*, 2021, no. 32(1), pp. 103-110. (In Eng.)

13. Salimova G., Nigmatullina G., Habirov G., Ableeva A., Gusmanov R. Employment and development levels in rural areas of the Russian Federation. *Regional Sustainability*, 2024, T. 3, no. 5, 100164. (In Eng.)