

Социальный потенциал детства и региональная среда: особенности кластеризации российских регионов

В современных условиях детство имеет не только индивидуальную ценность как особый период развития человека, но и все возрастающую общественную ценность, поскольку от социального потенциала детства напрямую зависит социальный потенциал общества – его репродукция, производительные силы, инновационные ресурсы и др. Поэтому для общества и государства важной задачей выступает поддержание особых условий для сохранения и, по возможности, развития потенциала детства. В условиях федеративного типа государственного устройства большую роль в реализации социального потенциала детства играет региональный уровень – уровень субъекта РФ. Под социальным потенциалом детства в регионе мы понимаем "совокупность ресурсов, резервов и возможностей детей как особой социально-демографической общности, проживающей на территории какого-либо региона (субъекта Российской Федерации), реализующихся под воздействием внешних и внутренних факторов, направленных на достижение качественного состояния детства в области здоровья, образования и духовно-нравственного развития, формирование необходимого набора стартовых ресурсов, необходимых для выхода во взрослую жизнь" [1, с. 137].

Интервью с экспертами в области детского образования, здравоохранения, культуры и спорта, а также анализ данных российской региональной статистики привели к выделению двух групп целевых показателей – "здоровье детей" и "образование детей". От третьей группы "духовно-нравственное развитие" решено было отказаться по причине сложности ее фиксации в разрезе статистических данных, в особенности, региональных.

Вообще работа с данными российской статистики была сопряжена с рядом проблем: изменением характеристик собираемых данных, явными искажениями статистических данных в отдельных регионах (искусственное занижение/завышение показателей), отсутствием данных по отдельным регионам. Так, например, в Мониторинге системы образования в 10 субъектах Российской Федерации отсутствовал пункт 2.6 "Результаты аттестации лиц, обучающихся по образовательным программам начального общего, основного общего образования и среднего общего образования" (Архангельская, Сахалинская, Тюменская области, Ненецкий и Чукотский автономные округа, Республика Алтай, Дагестан, Крым, г. Севастополь, Чеченская республика). Если же говорить о замере показателей, выделенных Национальной стратегией действий в интересах детей, то можно отметить "провисание" раздела, касающегося активности самих детей ("Дети – участники реализации Национальной стратегии"). Количественное насыщение шести разделов Национальной стратегии статистическими показателями выглядит следующим образом: I. Семейная политика детствосбережения – 25 показателей; II. Доступность качественного обучения и воспитания – 28 показателей; III. Здравоохранение, дружелюбное к детям, и здоровый образ жизни – 34 показателя; IV. Равные возможности для детей, нуждающихся в особой заботе государства – 8 пока-

© Филипова А.А., Высоцкая А.В., 2018

ФИЛИПОВА Александра Геннадьевна, д-р социол. наук, профессор департамента социальных наук Дальневосточного федерального университета (г. Владивосток). **E-mail:** alexgen77@list.ru

ВЫСОЦКАЯ Алёна Валерьевна, и.о. заведующего кафедрой "Информационные системы" Комсомольского-на-Амуре государственного университета (г. Комсомольск-на-Амуре). **E-mail:** al-w-buaa@rambler.ru

Работа выполнена при финансовой поддержке гранта РФФИ (проект № 16-36-60041)

зателей; V. Создание системы защиты и обеспечения прав и интересов детей и дружественного к ребенку правосудия – 10 показателей; VI. Дети – участники реализации Национальной стратегии – 1 показатель. В последнем случае этот единственный показатель "Доля детей в возрасте 15–17 лет, осуществляющих членство в каких-либо организациях (движениях)" пока не заполнен ни по РФ в целом, ни по российским регионам [6, с. 509].

Подобную диспропорцию И.Е. Калабихина, О.В. Кучмаева объясняют отсутствием в приказе Росстата "Об утверждении системы индикаторов..." для мониторинга Национальной стратегии (2012) раздела, посвященного участию детей [3].

Помимо сайта Росстата и ЕМИСС данные собирались с сайтов органов исполнительной власти субъектов Российской Федерации, аналитического агентства ООО "Автомобильная статистика" и др. [1; 2; 5].

Гипотезой настоящего исследования является предположение о наличии разных региональных сред, оказывающих влияние на социальный потенциал детства, а именно на здоровье и образование детей.

Поэтому наряду с целевыми факторами, характеризующими здоровье и образование детей, были выделены управляющие факторы, связанные с региональными условиями проживания детей. Конечно, обращение к уровню региона в определённой степени нивелирует различия в условиях проживания детей в пределах этого региона, например, в городе и сельской местности, разных типах городов. Однако целью настоящего исследования является выявление именно региональных факторов, влияющих на детство.

В качестве целевых показателей, характеризующих социальный потенциал детства в регионе, взяты у1 – средний балл ЕГЭ по русскому языку, балл; у2 – доля выпускников без аттестата (численность без аттестата к общей численности), %; у3 – доля детей, не обучающихся в образовательных организациях к численности населения младше трудоспособного возраста, %; у4 – охват детей дошкольным образованием, %; у5 – доля детей, обучающихся в школах искусств, музыкальных и пр., к численности населения младше трудоспособного возраста, %; у6 – доля детей в возрасте до 14 лет, систематически занимающихся физической культурой, к численности населения младше трудоспособного возраста, %; у7 – доля выявленных безнадзорных и беспризорных к численности населения младше трудоспособного возраста, %; у8 – доля детей, больных туберкулезом, %; у9 – доля детей, больных злокачественными новообразованиями, %; у10 – число умерших в возрасте до пяти лет на 1000 родившихся живыми, чел.; у11 – число детей-инвалидов (0–17 лет), состоящих под наблюдением в амбулаторно-поликлинических учреждениях, оказывающих медицинскую помощь детям, чел.; у12 – численность детей I–II групп здоровья к численности населения младше трудоспособного возраста, %; у13 – смертность детей от внешних причин смерти в возрасте 0–17 лет на 100 000 чел. соответствующего возраста,

Выбор данных факторов связан, с одной стороны, со смысловой нагрузкой понятия "социальный потенциал детства", а, с другой, – с имеющимися статистическими данными, собираемыми Росстатом.

Данные, собранные на предыдущих этапах работы [8; 9], были взяты за основу при проведении корреляционного анализа. В итоге был сформирован набор из 36 управляющих факторов, включающий группы инфраструктурных факторов (образование, здравоохранение, культура и спорт, транспорт), социально-экономических, территориально-поселенческих, демографических и экологических факторов:

- x1 – средняя наполняемость классов, чел.;
- x2 – отношение частных общеобразовательных организаций ко всем общеобразовательным организациям, %
- x3 – доля школ, имеющих столовую или буфет, %;
- x4 – доля школ, в которых созданы условия для обучения инвалидов, %;
- x5 – среднее число участников клубных формирований в расчете на 1 тыс. человек населения, чел.;
- x6 – процент обучающихся во 2-3 смены, %;
- x7 – доля педагогов высшей категории, %;
- x8 – число ПК образовательных организаций с доступом к интернету на 100 учащихся, шт.;

- x9 – отношение расходов на образование к расходам всего, %;
- x10 – доля автобусов для перевозки детей (выпуск не более 10 лет) к численности лиц младше трудоспособного населения, %;
- x11 – отношение числа детей, отдохнувших в оздоровительных лагерях, к числу обучающихся детей всего, %;
- x12 – пропускная способность спортивных сооружений, чел.;
- x13 – среднее число вакцинированных детей в возрасте 12 мес., чел.;
- x14 – обеспеченность педиатрами (отношение педиатров к численности лиц младше трудоспособного населения), %;
- x15 – отношение заработной платы медицинских работников к средней заработной плате региона, %;
- x16 – расходы на обязательное медицинское страхование, руб/чел.;
- x17 – прирост мигрантов нетрудоспособного возраста, чел.;
- x18 – количество жителей с высшим образованием на 1 тысячу человек, чел.;
- x19 – отношение числа ВУЗов и филиалов к численности населения на 10000 чел.;
- x20 – число профессиональных образовательных организаций, осуществляющих подготовку специалистов среднего звена, шт.;
- x21 – число преступлений в отношении несовершеннолетних к численности лиц младше трудоспособного населения, %;
- x22 – соотношение браков и разводов (на 1000 браков приходится разводов), шт.;
- x23 – число посетителей театров и музеев на 1000 чел, чел.;
- x24 – жилая площадь на 1 человека, м²;
- x25 – уровень занятости населения, %;
- x26 – инновационная активность организаций, %;
- x27 – средняя месячная температура воздуха в июле (фактическая), С°;
- x28 – средняя месячная температура воздуха в январе, С°;
- x29 – плотность населения, чел/км²;
- x30 – расстояния от Москвы до центров регионов, км.;
- x31 – ВРП на душу населения, руб.;
- x32 – отношение среднедушевого дохода населения к прожиточному минимуму в регионе, %
- x33 – плотность автомобильных дорог общего пользования с твердым покрытием, км путей на 1000 км² территории;
- x34 – выбросы загрязняющих веществ в атмосферный воздух, отходящих от стационарных источников, тыс. т/чел.;
- x35 – сброс загрязненных сточных вод в поверхностные водные объекты, млн. м³/чел.;
- x36 – использование свежей воды, млн. м³/чел.

Некоторые из вышеперечисленных показателей нуждаются в дополнительных пояснениях. Под клубными формированиями (x5) понимаются добровольные объединения людей, основанные на общности интересов, запросов и потребностей в занятиях любительским художественным и техническим творчеством, в совместной творческой деятельности, способствующей развитию дарований его участников, освоению и созданию ими культурных ценностей, а также основанное на единстве стремления людей к получению актуальной информации и прикладных знаний в различных областях общественной жизни, культуры, литературы и искусства, науки и техники, к овладению полезными навыками в области культуры быта, здорового образа жизни, организации досуга и отдыха [4]. Число вакцинированных детей рассчитывалось как среднее арифметическое детей, которым сделаны прививки против коклюша, полиомиелита и кори по достижению ими возраста 12 месяцев (x13). Инновационная активность предприятий (x26) определяется через долю организаций, осуществивших технологические, организационные, маркетинговые инновации, в общем числе организаций региона. Использование свежей воды (x36) – это использование забранных из различных источников водных ресурсов (включая морскую воду) для удовлетворения хозяйственных нужд. Сюда не включается оборотное водопотребление, а также повторное использование сточной и коллекторно-дренажной воды [5].

Далее определим наличие линейной связи между выделенными показателями, ее тесноту и статистическую значимость. Поскольку все рассматриваемые факторы имеют метрическую шкалу измерения, то для анализа используем критерий корреляции Пирсона.

На рисунке 1 приведены значения коэффициентов корреляции целевых и управляющих факторами, а на рисунке 2 – корреляции управляющих факторов.

Чем более интенсивным цветом выделены коэффициенты, тем более интенсивная связь между показателями. Принимаем во внимание только значения коэффициента больше 0,3.

Анализ полученных связей целевых и управляющих факторов, в особенности со значениями более 0,5, позволил сделать несколько предположений относительно природы полученных связей. При этом мы понимаем, что "наличие корреляционной зависимости между переменными не всегда означает наличие непосредственной связи этих величин друг с другом: наблюдаемая связь часто существует благодаря другим переменным (не двум рассматриваемым), а изучаемые величины могут быть связаны между собой через латентные (скрытые от исследователя) переменные" [10].

Можно выделить две укрупненные группы корреляций: первая группа включает очевидные корреляции, вторая – неочевидные, назовем их "фоновые" корреляции. Возвращаясь к нашей гипотезе, можно "фоновые" факторы объединить в разные типы региональных сред, отличающиеся разной плотностью заселения территорий, разными социально-экономическими условиями проживания, природно-климатическими условиями, удаленностью от Москвы и др.

Первая группа, демонстрирующая очевидные связи целевых и управляющих факторов, станет основой для последующего регрессионного анализа. Например, обратная корреляция y_3 ("доля детей, не обучающихся в образовательных организациях к численности населения младше трудоспособного возраста") и x_4 ("доля школ, в которых созданы условия для обучения инвалидов") может быть объяснена инклюзией детей-инвалидов в образовательную систему, а отрицательная связь y_2 и x_7 – вероятным снижением доли выпускников образовательных учреждений, не получивших аттестат, при увеличении числа педагогов, имеющих высшую категорию. Некоторые корреляции имеют как очевидный, так и фоновый характер. Так, связь y_{12} и x_3 может быть объяснена как через влияние горячего питания, организованного в школах, на здоровье детей ("доля детей I-II групп здоровья"), так и опосредованно, через уровень социально-экономического развития региона, позволяющий наряду с прочими инфраструктурными достижениями, ориентированными на детей, благоустраивать школы, оснащать их не только столовыми/буфетами, но и бассейнами, тренажерными залами и т.п.

Что же касается второй группы связей, то наше основное предположение относительно их природы связано с "насыщением" региональной среды культурными объектами, учреждениями среднего профессионального и высшего образования, спортивными сооружениями с высокой пропускной способностью, инновационно активными организациями и пр. Это насыщение происходит вследствие концентрации человеческих ресурсов (высокая плотность населения), близости к Москве, более высокого уровня ВРП и более высоких среднедушевых доходов. Подобная концентрация населения приводит к высокой плотности автомобильных дорог, повышению наполняемости школьных классов и, соответственно, увеличению заработной платы педагогов вследствие подушевого финансирования. Эти регионы лучше обеспечены педагогами, а также педагогами высшей категории. Обозначим данный тип региональной среды как "тип А".

Много "фоновых" управляющих факторов фиксируется относительно y_{10} . В данном случае мы высказываем предположение о том, что смертность детей в возрасте до 5 лет выше в тех российских регионах, где хуже общая социально-экономическая ситуация – выше уровень безработицы (x_{25}), больше малообеспеченных жителей (x_{32}), ниже инновационная активность предприятий (x_{26}) и, соответственно, меньше благоустроенных школ, имеющих столовую или буфет (x_3), выше сменность занятий в школах (x_6), а нехватка учреждений дополнительного образования компенсируется менее затратны-

Рис. 2. Коэффициенты корреляции управляющих факторов

	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10	x11	x12	x13	x14	x15	x16	x17	x18	x19	x20	x21	x22	x23	x24	x25	x26	x27	x28	x29	x30	x31	x32	x33	x34	x35	x36			
x1	1																																						
x2	0,5	1																																					
x3	0	0,1	1																																				
x4	0,2	0,4	-0,1	1																																			
x5	-0,1	-0,1	-0,1	-0,1	1																																		
x6	0,1	-0,1	-0,1	-0,1	-0,1	1																																	
x7	0	0	0,4	0,2	-0,1	-0,1	1																																
x8	0,3	0,2	0,2	0,1	-0,1	-0,1	0	1																															
x9	-0,1	-0,1	0,1	-0,1	0,3	0,1	-0,1	0	1																														
x10	-0,1	-0,1	0,1	-0,1	0,3	0,1	-0,1	0	0,2	1																													
x11	-0,1	-0,1	0,1	-0,1	0,4	0,1	-0,1	0	0	0,3	1																												
x12	0,3	0,5	0,2	0,3	-0,1	-0,2	0,1	-0,1	-0,1	-0,1	0,9	1																											
x13	0,5	0,6	0,2	0,3	-1	-0,2	0,1	-0,1	-0,1	-0,1	0,9	0,2	1																										
x14	0,3	0,3	0,2	0	-0,1	0	0,2	0	-0,1	-0,1	0,2	0,2	1	1																									
x15	-0,1	-0,1	0	-0,3	0,1	-0,1	-0,1	0,1	0,3	0,5	-0,1	-0,1	0,5	1	1																								
x16	0,2	0	0,1	-0,1	0,3	0,2	-0,1	0,5	-0,2	-0,1	-0,1	-0,1	0,2	-0,1	1	1																							
x17	0,5	0,7	0,1	0,3	-0,1	-0,2	0,1	-0,1	-0,1	0,6	0,7	0,2	-0,1	0	1	1	1																						
x18	0,3	0,6	0	0,5	-0,1	-0,2	0,3	-0,1	-0,1	0,3	0,4	0,3	-0,3	0,5	1	1	1	1																					
x19	0,1	0,2	0,1	0	0,3	-0,1	0,3	-0,1	-0,1	-0,1	-0,3	-0,4	0,1	0,4	1	1	1	1	1																				
x20	0,3	0,4	0,2	0,1	-1	-0,2	0,1	0,1	-0,1	0,9	0,9	0,1	-0,1	-0,5	0,2	-0,1	1	1	1	1																			
x21	-0,1	-0,1	0,1	-0,2	0,4	0,4	-0,1	0,3	0,8	0,3	-0,1	-0,1	-0,4	0,2	-0,1	-0,1	0,1	1	1	1	1																		
x22	-0,1	-0,4	-0,1	-0,2	-0,1	0,2	0	0,2	-0,2	0,3	-0,1	-0,1	-0,2	0,3	-0,1	-0,1	0,2	0,3	1	1	1	1																	
x23	0,3	0,5	0,1	0,2	-0,1	0,1	0	-0,1	-0,1	0,2	0,3	0,4	-0,1	-0,2	0,2	0,1	0,2	0	-0,1	1	1	1	1																
x24	-0,1	-0,4	-0,1	-1	0,1	0,1	-0,1	0	-0,1	0	-0,1	0	-0,1	0	-0,1	0	-0,2	0	-0,4	0,1	1	1	1	1															
x25	0,3	0,3	0,5	0	-0,1	0,5	-0,1	0	-0,3	0,3	0,5	-0,4	0,3	0,3	0,4	0,3	-0,3	0,4	0,4	1	1	1	1	1															
x26	-0,2	0,3	0,2	-0,1	0,1	0,3	-0,1	-0,1	0,4	0,2	0,3	0	0,1	0,2	0,1	0,2	0,2	-0,3	0,3	0,5	1	1	1	1	1														
x27	0	-0,2	0,2	-0,1	-0,1	-0,1	-0,1	0	-0,1	0,1	0	-0,1	0,1	-0,1	0	-0,1	0,1	0	-0,1	0,1	1	1	1	1	1														
x28	0,1	0,2	-0,2	-0,2	-0,2	-0,1	-1	-0,1	0	0	-1	0,3	0,2	0,1	0	-0,1	-0,2	0,3	-0,1	0,7	1	1	1	1	1														
x29	0,4	0,9	0,1	0,4	-0,1	0,1	0,1	-0,1	0,5	0,5	0,4	-0,1	0,6	0,5	0,1	0,4	-0,1	-0,7	-0,3	0,3	0,7	-0,2	1	1	1	1													
x30	0	-0,1	-0,1	0,3	0,5	-0,1	0	0,4	0,3	-0,1	-0,1	0,2	0,5	-0,1	0	0,1	-0,5	0,2	-0,1	-0,1	-0,1	-0,1	-0,1	-0,1	-0,1	-0,1	-0,1	-0,1	-0,1	-0,1	-0,1	-0,1	-0,1	-0,1	-0,1	-0,1	-0,1		
x31	0,2	0,2	0,1	-0,1	-0,1	0,5	-0,1	-0,1	0,1	0,1	0,4	-0,4	0,2	0,1	0,3	0,1	-0,1	0,3	0	0,4	0,1	-0,1	-0,2	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
x32	0,3	0,4	0,4	0,2	-0,1	0,2	0,5	-0,1	-0,1	0,6	0,6	0,3	-0,1	0,4	0,3	0,2	0,6	-0,1	0,3	0,3	0,6	0,5	-0,2	0,4	-0,4	1	1	1	1	1	1	1	1	1	1	1	1	1	
x33	0,4	0,8	-0,5	-0,1	-0,1	-0,1	-0,1	-0,1	-0,5	0,5	0,3	-0,1	0,6	0,5	0,1	0,3	-0,1	-0,7	-0,2	0,3	0,2	0,4	0,9	-0,1	0,4	1	1	1	1	1	1	1	1	1	1	1	1	1	
x34	0,2	-0,1	0,1	-0,1	-0,2	-0,5	0,1	0,2	-0,1	-0,1	-0,1	-0,6	-0,1	-0,1	-0,1	-0,1	-0,2	0,3	-0,1	-0,3	-0,1	-0,1	-0,2	0,3	0,2	1	1	1	1	1	1	1	1	1	1	1	1	1	
x35	0,5	0,1	0,3	-0,1	-0,1	0,1	0,1	0,1	0,1	0,2	0,2	-0,3	0,1	0,1	0,1	0,2	0,1	0,2	0,2	0,1	0,3	-0,1	-0,1	0	0,2	0,2	-0,3	1	1	1	1	1	1	1	1	1	1	1	
x36	0,1	0,2	-1	0,1	-0,3	-0,1	-0,1	-0,1	-0,1	-0,1	-0,1	-0,1	-0,1	0	-0,1	0	-0,1	0	-0,1	-1	0	-0,1	-0,1	0	0,2	0,3	0,2	-0,1	-0,1	-0,1	-0,1	-0,1	-0,1	-0,1	-0,1	-0,1	-0,1	-0,1	

Источник: составлено авторами

ми клубными формированиями (x5). Такой тип региональной среды обозначим как "тип В".

Ситуация с детской инвалидностью (y11), согласно рисунку 1, также определяется "фоновыми" факторами. Можно предположить, что в регионах, где социально-экономическая ситуация отличается стабильностью, лучше диагностируются детские заболевания, в т.ч. заболевания, приводящие к установлению инвалидности, а также лучше организована работа учреждений медико-социальной экспертизы.

Для проверки нашей гипотезы относительно разных типов региональной среды было решено прибегнуть к методу кластеризации. В качестве входных показателей были выбраны шесть факторов, сигнализирующих, с одной стороны, о "насыщенности" региональной среды в образовательно-культурном и инфраструктурном планах, а с другой, – о возможностях ее развития (инновационного, экономического) – это факторы x18, x23, x24, x25, x26, x33.

Для проведения анализа данных и дальнейшей кластеризации была использована аналитическая платформа Deductor. Система Deductor предназначена для решения широкого спектра задач, связанных с обработкой структурированных и представленных в виде таблиц данных. После импорта исходных данных в систему Deductor была проведена процедура предобработки исходных значений с помощью узла "Качество данных", который предназначен для проведения профайлинга и аудита данных с целью определения степени пригодности полей набора данных для решения задач анализа по объективным критериям: пропуски, выбросы, экстремальные значения [11].

Для кластеризации были поочередно использованы разные алгоритмы формирования кластеров – k-means, g-means, карта Кохонена и др. В ходе изучения результатов формирования кластеров был выбран алгоритм EM-кластеризации, являющийся итерационной процедурой. Были заданы следующие исходные параметры алгоритма: способ разделения исходного множества данных – случайное; фиксированное количество кластеров, рав-

ное последовательно 2, 3, 4 и 5; максимальное количество итераций, равное 300; требуемый уровень точности модели, равный 10-5.

В ходе изучения матриц схожести кластеров, профилей кластеров, данных статистики решено было остановиться на выделении четырех кластеров. Распределение регионов по кластерам получилось следующее: нулевой – 22 региона (29,33 %), первый – 10 (13,33 %), второй – 32 (42,67 %), третий – 11 (14,67 %). В таблице 1 приведена оценка значимости шести входных факторов. Как видно из таблицы, наиболее весомый вклад в формирование всех четырех кластеров у факторов x18 и x24.

При дальнейшей интерпретации данных кластеризации нужно будет учитывать снижение влияния фактора x25 при формировании второго кластера и фактора x23 при формировании нулевого кластера, а также экстремально низкое значение фактора x25 при формировании второго кластера. Вероятно, в дальнейшем нужно будет поработать над разделением самого большого, второго кластера, по факторам, чей вклад оказался существенно ниже – это x25 и x26.

В нулевой кластер включены регионы: Амурская, Астраханская, Иркутская, Кемеровская, Кировская, Курганская, Мурманская, Новосибирская, Омская, Свердловская, Томская области, Еврейский автономный округ, Забайкальский, Красноярский, Пермский, Приморский и Хабаровский края, Республики Бурятия, Карелия, Коми, Саха (Якутия), Хакасия. В первый кластер вошли: Белгородская, Липецкая, Магаданская, Московская, Пензенская области, г. Москва, Санкт-Петербург, Камчатский край, Республика Татарстан, Чувашская Республика. Во второй кластер входят: Алтайский, Краснодарский края, Брянская, Владимирская, Волгоградская, Вологодская, Воронежская, Ивановская, Калининградская, Калужская, Костромская, Курская, Ленинградская, Нижегородская, Новгородская, Оренбургская, Орловская, Псковская, Ростовская, Рязанская, Саратовская, Смоленская, Тамбовская, Тверская, Тульская, Ульяновская, Челябинская, Ярославская области, Республики Башкортостан, Марий Эл, Мордовия, Удмуртия. В третий кластер включены следующие субъекты: Республики Кабардино-Балкария, Карачаево-Черкессия, Адыгее, Ингушетия, Калмыкия, Северная Осетия – Алания, Тыва, а также Самарская область, Ставропольский край, Ханты-Мансийский и Ямало-Ненецкий автономные округа.

На рисунке 3 выделенные кластера обозначены цветом, не расцвеченными остались российские регионы, не вошедшие в выборку настоящего исследования по причинам отсутствия статистических данных по отдельным показателям. Картограмма позволяет сделать некоторые выводы относительно географии субъектов РФ, объединенных в четыре кластера. Регионы первого и второго кластеров преимущественно сосредоточены в европейской части России, в то время как регионы нулевого кластера – в центральной и восточной частях, т.е. в Сибири и на Дальнем Востоке. Наибольшим пространственным

Табл. 1. Оценка значимости факторов при кластеризации (%)

Показатель	Кластер			
	0	1	2	3
x18 (количество с высшим образованием на 1 тысячу человек, чел.)	87,2	99,7	99,1	99,5
x23 (число посетителей театров и музеев на 1000 чел, чел.)	48,5	71,8	82,8	99,8
x24 (жилая площадь на 1 человека, м ²)	96,7	89,5	99,6	99,6
x25 (уровень занятости населения, %)	63,2	99,9	26,3	97
x26 (инновационная активность организаций, %)	83,5	100	1,9	99,8
x33 (плотность автомобильных дорог общего пользования с твердым покрытием, км путей на 1000 км ² территории)	100	99,6	62,8	74,1

Источник: составлено авторами

Табл. 2. Описательные статистики входных факторов кластеризации

Фактор	Показатель	Кластер			
		0	1	2	3
X18	Среднее значение в кластер	29,18	35,57	28,93	34,95
	Среднеквадратическое отклонение	3,25	5,926	2,595	4,26
	Минимальное значение в кластере	22,8	29,6	25	26,5
	Максимальное значение в кластере	33,8	43,45	34,1	42,1
X23	Среднее значение в кластер	702,68	909,966	883,656	372
	Среднеквадратическое отклонение	127,255	534,3368	441,661	158,7388
	Минимальное значение в кластере	477	466	333	131
	Максимальное значение в кластере	940	1775,332	1775,332	710
X24	Среднее значение в кластер	23,477	26,7	26,466	21,905
	Среднеквадратическое отклонение	1,4899	3,876	2,0857	3,6
	Минимальное значение в кластере	20,4	19,1	21,6	16,83
	Максимальное значение в кластере	26,7	33,4	30	27,6
X25	Среднее значение в кластер	63,386	69,18	64,48	60,895
	Среднеквадратическое отклонение	2,7	3,5289	2,4345	7,195
	Минимальное значение в кластере	58,6	64,2	59,2	52,97
	Максимальное значение в кластере	69,3	74,3	68,5	72,6
X26	Среднее значение в кластер	6,66	15,85	8,04469	3,6
	Среднеквадратическое отклонение	2,386	4,11	2,64	2,5469
	Минимальное значение в кластере	2,1	8,5	3,2	0
	Максимальное значение в кластере	12,2	20,39	13,4	9
X33	Среднее значение в кластер	67,4	474,98	275,4375	331,401
	Среднеквадратическое отклонение	49,73	305,946	90,4	294,09
	Минимальное значение в кластере	3,8	4,2	117	3
	Максимальное значение в кластере	175	812,62	559	812,62

Источник: составлено авторами

разбросом отличаются регионы третьего кластера, расположенные как на юге, так и на севере России, в ее западной и центральной частях.

Приступим к характеристике полученных кластеров. В таблице 2 приведены значения среднего и среднего квадратического отклонения по всем факторам для каждого из полученных кластеров.

Значение всех средних шести входных факторов максимальны по выборке из четырех кластеров у кластера 1 и в четырех случаях из шести минимальны – у кластера 3. Изучение данных, представленных в таблице 2, обнаруживает определенные сходства описанного ранее "типа А" региональной среды с региональной средой кластера 1, а "типа В" с кластером 3. Кластеры 0 и 2 занимают промежуточное положение между двумя обозначенными крайними типами региональных сред (А и В), при этом кластер 2 тяготеет к типу А, а кластер 0 – к типу В. Высокие значения среднего квадратического отклонения в первом кластере указывают на больший разброс значений в представленном множестве относительно средней величины множества при явном доминировании гг. Москвы и Санкт-Петербурга, а также республики Татарстан.

Табл. 3. Описательные статистики целевых факторов

Фактор	Показатель	Кластер			
		0	1	2	3
Y1	Минимальное значение в кластере	62,27	61,58	63,4	43,6
	Максимальное значение в кластере	72,87	75,58	73,33	71,14
	Среднее значение в кластер	67,004	67,894	68,258	64,384
	Среднеквадратическое отклонение	2,874	4,024	2,271	8,005
Y2	Минимальное значение в кластере	0,4	0,7	0,02	1,12
	Максимальное значение в кластере	5,03	2	4,79	28,43
	Среднее значение в кластер	2,021	1,38	1,693	6,643
	Среднеквадратическое отклонение	1,306	0,393	1,001	7,659
Y4	Минимальное значение в кластере	57,1	56	55,4	21,3
	Максимальное значение в кластере	86,2	83,5	85,7	74,5
	Среднее значение в кластер	68,941	71,13	71,369	60,055
	Среднеквадратическое отклонение	7,26	7,373	6,857	14,657
Y5	Минимальное значение в кластере	3,399	4,313	4,68	1,217
	Максимальное значение в кластере	8,766	10,581	10,727	8,537
	Среднее значение в кластер	5,429	7,324	7,275	5,122
	Среднеквадратическое отклонение	1,23	2,101	1,482	1,717
Y6	Минимальное значение в кластере	30,654	31,918	40,362	18,339
	Максимальное значение в кластере	106,905	85,53	90,971	76,458
	Среднее значение в кластер	55,988	63,478	64,98	60,302
	Среднеквадратическое отклонение	15,861	18,527	13,109	16,446
Y8	Минимальное значение в кластере	0,012	0,003	0,003	0,005
	Максимальное значение в кластере	0,073	0,088	0,045	0,087
	Среднее значение в кластер	0,031	0,021	0,015	0,028
	Среднеквадратическое отклонение	0,016	0,026	0,009	0,022
Y10	Минимальное значение в кластере	4,5	4,5	5,1	5,1
	Максимальное значение в кластере	17,9	9,9	9,7	14,7
	Среднее значение в кластер	8,118	6,52	7,378	8,864
	Среднеквадратическое отклонение	2,556	1,755	1,126	3,01
Y12	Минимальное значение в кластере	30,584	40,193	35,854	23,512
	Максимальное значение в кластере	47,958	53,246	51,754	50,902
	Среднее значение в кластер	41,413	45,93	43,956	39,686
	Среднеквадратическое отклонение	4,612	3,72	4,425	9,378
Y13	Минимальное значение в кластере	9,3	8,3	10,5	5,7
	Максимальное значение в кластере	33,8	25,2	25,5	72,2
	Среднее значение в кластер	23,186	16,47	18,241	19,273
	Среднеквадратическое отклонение	6,513	5,605	4,001	18,313

Источник: составлено авторами

На следующем этапе анализа проверим влияние "фоновых" региональных факторов, изучив описательные статистики выделенных ранее целевых факторов. Из обозначенных 13-ти факторов остановимся на тех, которые имеют большее число выраженных корреляций с управляющими факторами – это y_1 , y_2 , y_4 , y_5 , y_6 , y_8 , y_{10} , y_{12} и y_{13} .

В таблице 3 приведены описательные статистики целевых факторов в разрезе четырех кластеров. Для y_1 , 4, 5, 6 и 12 – желателен рост, в то время, как для y_2 , 8, 10 и 13 – желательно снижение.

Изучение средних значений целевых факторов по четырем кластерам подтверждает нашу гипотезу о влиянии региональной среды на социальный потенциал детства. Тип среды "А", характерный для российских регионов, входящих в кластеры 1 и 2, приводит к более высоким значениям показателей – средний балл ЕГЭ по русскому языку; доля детей, охваченных ДОУ; доля детей, занимающихся в школах искусств; доля детей, систематически занимающихся физической культурой; доля детей 1-2 групп здоровья. Средние остальных целевых показателей принимают минимальные значения среди средних значений показателей всех четырех кластеров – это доля выпускников без аттестата; детская заболеваемость туберкулезом; смертность до 5 лет; детская смертность от внешних причин.

В кластерах 0 и 3, определяемых как тип среды "В", наблюдается обратная ситуация. Однако здесь необходимо обратить внимание на высокие значения среднего квадратического отклонения в третьем кластере, что указывает на больший разброс значений в представленном множестве со средней величиной множества.

Также большой разброс значений переменных y_4 , y_6 и y_{12} относительно средних величин наблюдается во всех четырех кластерах. Что заставляет нас с осторожностью относиться к связи типа региональной среды с этими целевыми факторами.

Кластерный анализ позволил подтвердить высказанное предположение о связи показателей социального потенциала детства с факторами региональной среды. Условно выделенные типы "А" и "В" региональной среды представляют собой крайние варианты насыщения/ не насыщения региона культурно-образовательными и социально-инфраструктурными ресурсами, по-разному сказывающиеся на показателях детского образования и детского здоровья.

В ходе исследования были выделены две группы корреляций – "очевидные" и "фоновые". На основе "фоновых" корреляций была проведена кластеризация российских регионов. Это поможет в дальнейшем снять проблему интерпретации полученных данных по причине неочевидности обнаруженных связей. Проблема снимается выделением типа региональной среды, связанного с "насыщением" региона культурными объектами, учреждениями среднего профессионального и высшего образования, спортивными сооружениями с высокой пропускной способностью, инновационно активными организациями и др.

Для разных типов региональных сред будут построены свои регрессионные уравнения, описывающие влияние управляющих факторов на целевые показатели социального потенциала детства.

Литература

1. Аналитическое агентство «Автомобильная статистика», [Электронный ресурс]. URL: <https://www.autostat.ru/pages/radar/> (дата обращения: 30.05.2018).
2. Единая межведомственная информационно-статистическая система, [Электронный ресурс]. URL: <https://fedstat.ru> (дата обращения: 16.05.2018).
3. Калабихина И.Е., Кучмаева О.В. Проблемы и перспективы мониторинга участия детей в реализации «Национальной стратегии действий в интересах детей на 2012–2017 гг.» // Журнал исследований социальной политики. Том 14. № 4. С. 507–520.
4. О нормативах штатной численности работников государственных и муниципальных учреждений культурно-досугового типа и библиотек: Приказ Министерства культуры РФ от 1 сентября 2011 г. № 906. Доступ из справочно-информационной системы «ГАРАНТ»

5. Регионы России. Социально-экономические показатели. 2016. Стат. Сб., [Электронный ресурс]. URL: http://www.gks.ru/bgd/regl/B16_14p/Main.htm (дата обращения: 06.06.2018).
6. Федеральная служба государственной статистики. Семья, материнство и детство, [Электронный ресурс]. URL: http://www.gks.ru/wps/wcm/connect/rosstat_main/rosstat/ru/statistics/population/motherhood/ (дата обращения: 01.06.2018).
7. Филипова А.Г., Еськова А.В. Оценка влияния региональных показателей на здоровье детей // Ойкумена. Регионоведческие исследования. 2017. № 4 (43). С. 124–132.
8. Филипова А.Г., Еськова А.В. Социальный потенциал детства в регионе: построение онтологии предметной области // Регионология. 2016. №3. С. 137–146.
9. Филипова А.Г., Еськова А.В. Социальный потенциал региона: опыт использования кластерного анализа // Регионология. 2017. Т. 25. №3 (100). С. 438–455.
10. Харченко М.А. Корреляционный анализ, [Электронный ресурс]. URL: <https://www.twirpx.com/file/481991/> (дата обращения: 03.06.2018).
11. Deductor. Описание платформы, [Электронный ресурс]. URL: <https://basegroup.ru/deductor/description/> (дата обращения: 25.05.2018).

Транслитерация по ГОСТ 7.79-2000 Система Б

1. Analiticheskoe agentstvo «Avtomobil'naya statistika», [Электронный ресурс]. URL: <https://www.autostat.ru/pages/radar/> (дата обращения: 30.05.2018).
2. Edinaya mezhvedomstvennaya informatsionno-statisticheskaya sistema, [Электронный ресурс]. URL: <https://fedstat.ru> (дата обращения: 16.05.2018).
3. Kalabikhina I.E., Kuchmaeva O.V. Problemy i perspektivy monitoringa uchastiya detej v realizatsii «Natsional'noj strategii dejstvij v interesakh detej na 2012–2017 gg.» // Zhurnal issledovaniy sotsial'noj politiki. Tom 14. № 4. S. 507–520.
4. O normativakh shtatnoj chislennosti rabotnikov gosudarstvennykh i munitsipal'nykh uchrezhdenij kul'turno-dosugovogo tipa i bibliotek: Prikaz Ministerstva kul'tury RF ot 1 sentyabrya 2011 g. № 906. Dostup iz spravochno-informatsionnoj sistemy «GARANT»
5. Регионы России. Социально-экономические показатели. 2016. Стат. Сб., [Электронный ресурс]. URL: http://www.gks.ru/bgd/regl/B16_14p/Main.htm (дата обращения: 06.06.2018).
6. Federal'naya sluzhba gosudarstvennoj statistiki. Sem'ya, materinstvo i detstvo, [Электронный ресурс]. URL: http://www.gks.ru/wps/wcm/connect/rosstat_main/rosstat/ru/statistics/population/motherhood/ (дата обращения: 01.06.2018).
7. Filipova A.G., Es'kova A.V. Otsenka vliyaniya regional'nykh pokazatelej na zdorov'e detej // Ojkumena. Regionovedcheskie issledovaniya. 2017. № 4 (43). S. 124–132.
8. Filipova A.G., Es'kova A.V. Sotsial'nyj potentsial detstva v regione: postroenie ontologii predmetnoj oblasti // Regionologiya. 2016. №3. S. 137–146.
9. Filipova A.G., Es'kova A.V. Sotsial'nyj potentsial regiona: opyt ispol'zovaniya klaster'nogo analiza // Regionologiya. 2017. Т. 25. №3 (100). S. 438–455.
10. KHarchenko M.A. Korrelyatsionnyj analiz, [Электронный ресурс]. URL: <https://www.twirpx.com/file/481991/> (дата обращения: 03.06.2018).
11. Deductor. Описание платформы, [Электронный ресурс]. URL: <https://basegroup.ru/deductor/description/> (дата обращения: 25.05.2018).

Филипова А.А., Высоцкая А.В. Социальный потенциал детства и региональная среда: особенности кластеризации российских регионов.

Социальный потенциал детства, понимаемый через показатели детского здоровья и образования, соотносится в статье с разными типами региональной среды. Корреляционный анализ целевых (показатели детства) и управляющих (показатели среды) факторов помогает выйти на оппозицию двух типов региональной среды, условно обозначенных как "тип А" и "тип В". Первый тип насыщен культурно-образовательными и социально-инфраструктурными ресурсами, второй же испытывает потребность в большем насыщении. Для установления связи между показателями детства и типами региональной среды использован инструмент кластеризации – выделены четыре кластера российских регионов. Анализ целевых факторов позволил установить связь между их желательными значениями и характеристиками региональной среды.

Ключевые слова: *российские регионы, детство, корреляционный анализ, кластеризация, региональная среда*

Filipova A.A., Vysotskaya A.V. Social potential of childhood and the regional environment: features of clustering of Russian regions.

The social potential of childhood, understood through the indicators of child health and education, is correlated in the article with different types of regional environment. The correlation analysis of the target (childhood indicators) and managers (environmental factors) helps to overcome the opposition of two types of regional environment conventionally designated as "type A" and "type B". The first type is saturated with cultural, educational and socio-infrastructural resources, while the latter is in need of greater saturation. To establish the connection between the indicators of childhood and the types of regional environment, a clustering tool was used – four clusters of Russian regions were identified. The analysis of the target factors made it possible to establish a connection between their desired values and the characteristics of the regional environment.

Key words: *Russian regions, childhood, correlation analysis, clustering, regional environment*

Для цитирования: Филипова А.А., Высоцкая А.В. Социальный потенциал детства и региональная среда: особенности кластеризации российских регионов // Ойкумена. Регионоведческие исследования. 2018. № 3. С. 108–120. DOI: 10.24866/1998-6785/2018-3/108-120

For citation: Filipova A.A., Vysotskaya A.V. Social potential of childhood and the regional environment: features of clustering of Russian regions. // Ojkumena. Regional researches. 2018. № 3. P. 108–120. DOI: 10.24866/1998-6785/2018-3/108-120

