

Сила есть, ума не надо

Сборная Руанды по настольному теннису

Давыдов Иван
Ламонов Никита
Михаловский Михаил
Мирсаяпова Элиза
Фрик Виктор

Структура данных

85 субъектов РФ

25 показателей

2019-2022 годы

абсолютные

относительные

Структура данных

абсолютные

демографические

численность постоянного населения в среднем за год, внутрирегиональная миграция, внешняя миграция, межрегиональная миграция

туристические

число коллективных средств размещения (КСР), число номеров в КСР, численность размещенных лиц в КСР, число ночевок в КСР

экономические

средняя цена 1 м² общей площади квартир на рынке жилья

социально-культурные

количество преступлений, число спортивных сооружений

Структура данных

относительные

экономические

доходы населения, объем транспортных услуг, оборот общественного питания, оборот розничной торговли, инвестиции в основной капитал, валовой региональный продукт (ВРП), уровень безработицы, уровень занятости, количество автобусов общего пользования

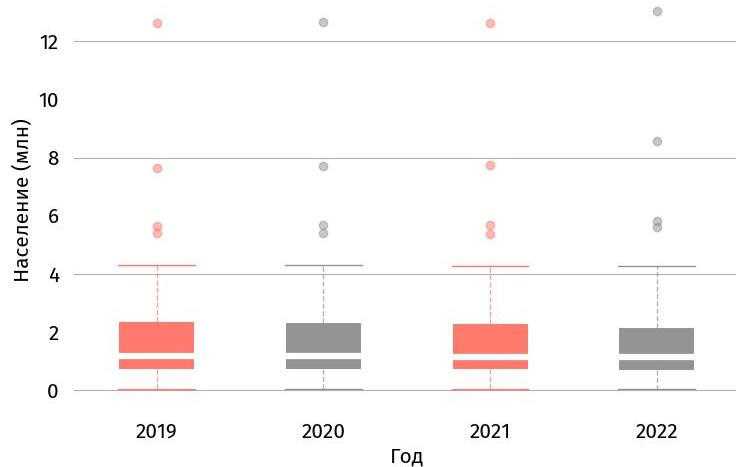
демографические

структура численности постоянного населения на начало года по возрастным группам, доля городского населения в общей численности населения на 1 января

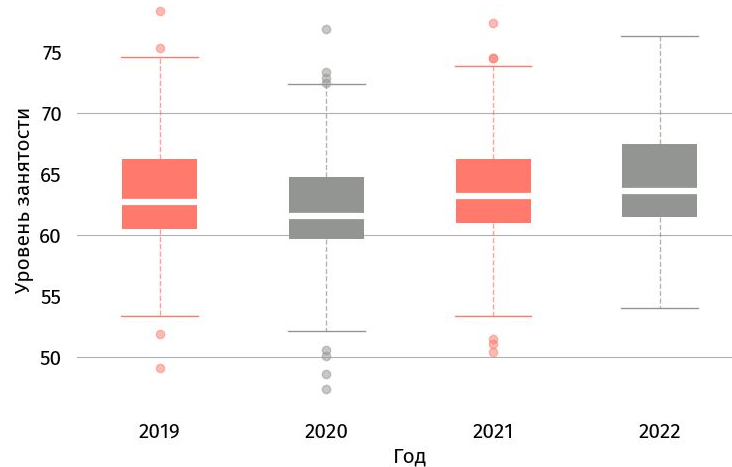
социально-культурные

среднее число зрителей на мероприятиях театров, количество принятых на обучение по программам бакалавриата, специалитета, магистратуры

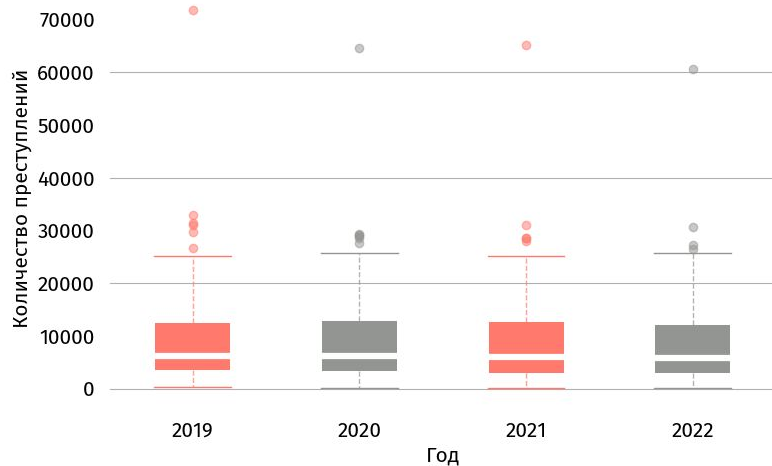
Распределение населения по годам



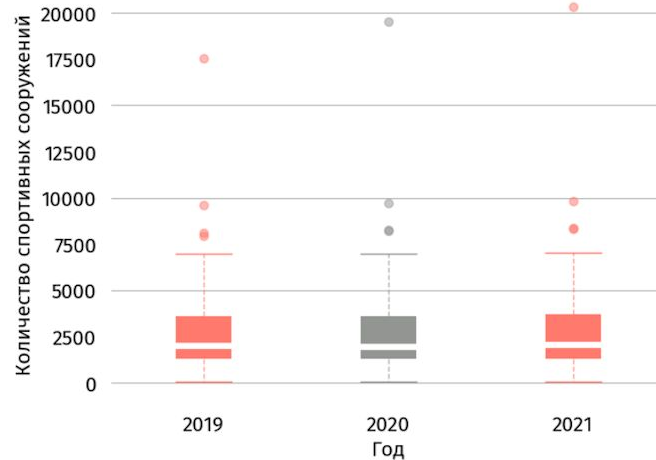
Распределение занятости по годам

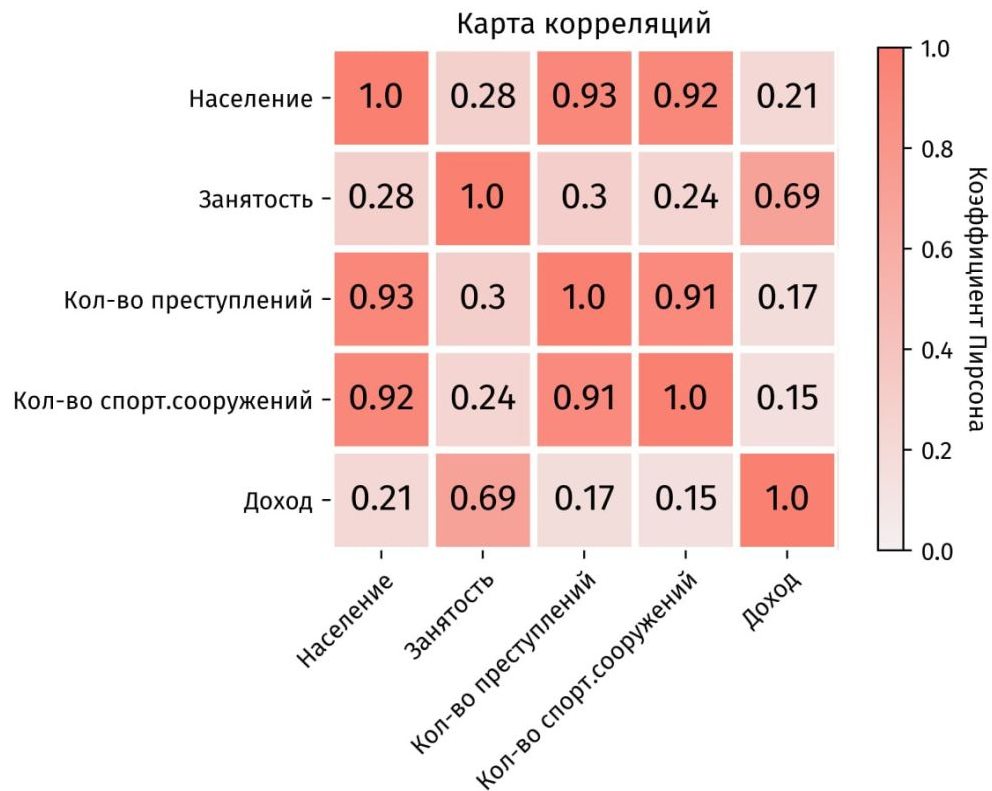


Распределение преступности по годам



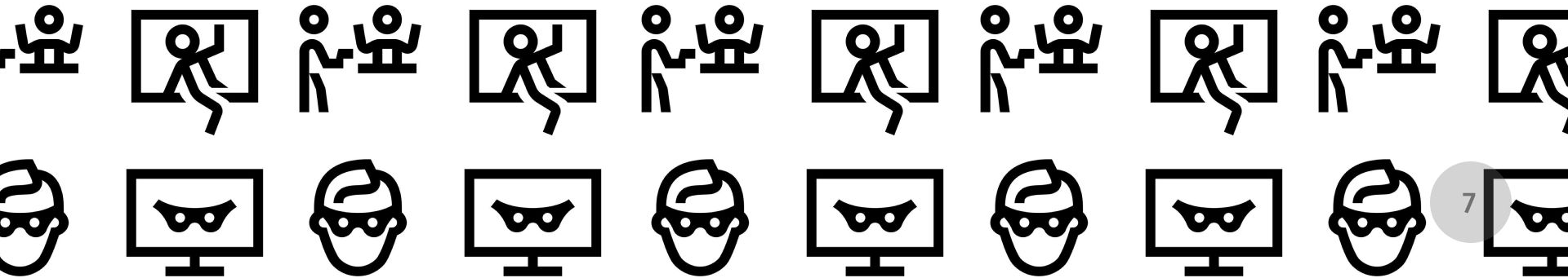
Распределение спортплощадок по годам





Исследовательский вопрос

Как возможность заниматься спортом влияет на преступность?



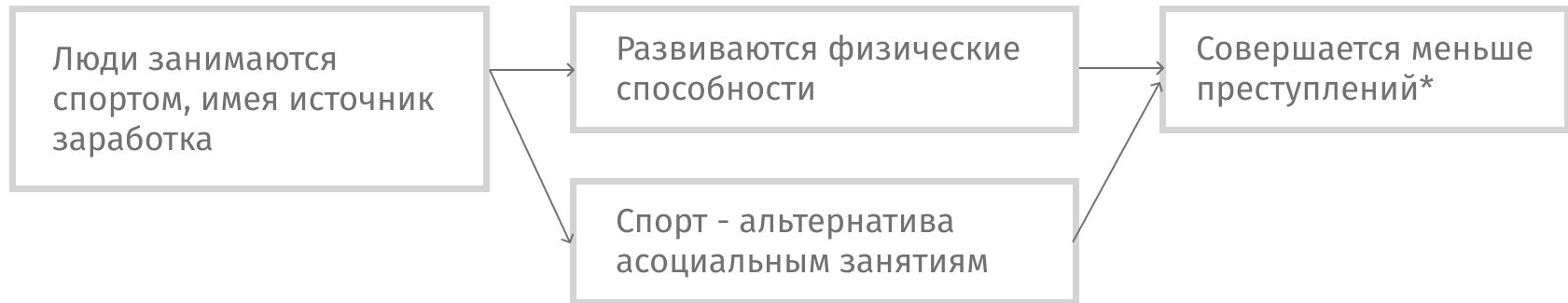
Ввод переменных

новые переменные	в чём измеряется
насыщенность региона спортсооружениями	ед./10 тыс. чел
уровень преступности	ед./10 тыс. чел
доля молодёжи	%
покупательная способность*	индекс

*эта переменная рассчитывается как отношение среднедушевого дохода к прожиточному минимуму - показателю из внешних данных (Росстат)

Первая гипотеза

В регионах с высоким уровнем занятости населения увеличение количества спортивных сооружений снижает преступность



*Спорт и преступность: правомерна ли такая постановка вопроса? - 1996.

*Новикова И. Занятия спортом снижают уровень преступности // Pravda.ru

Вторая гипотеза

В регионах с низким уровнем занятости населения с увеличением количества спортивных сооружений увеличивается преступность

Люди занимаются спортом, не имея источника заработка

Спортивные сооружения становятся местами, где люди сбиваются в банды

Совершается больше преступлений

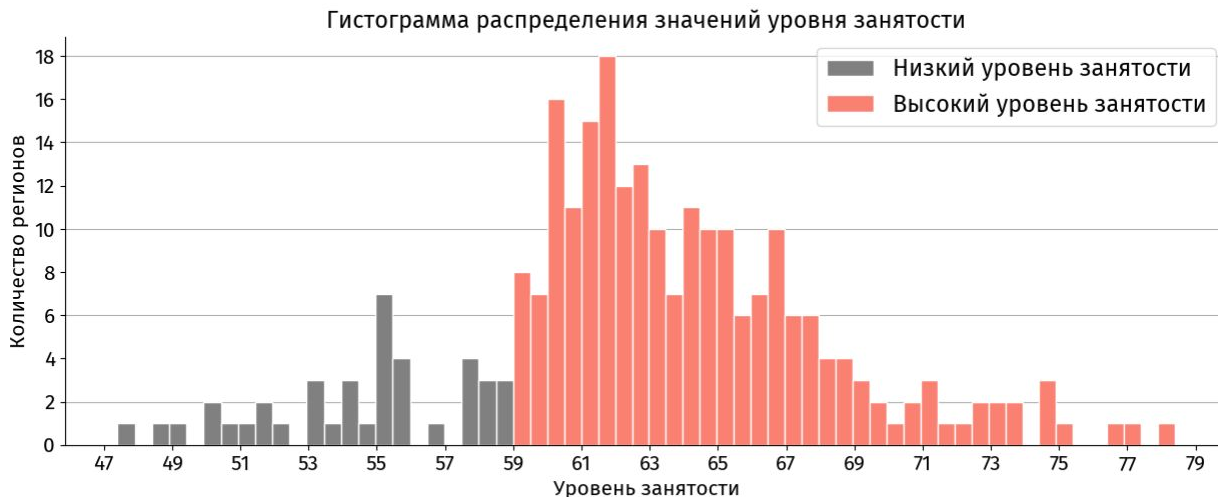
Обработка данных

Пустые значения

удаляем 2022 год (25% от исходных данных) по причине отсутствия данных по спортплощадкам

Выделение подвыборок

делим регионы на две группы: регионы с низким и высоким уровнем занятости



Исследуемые показатели

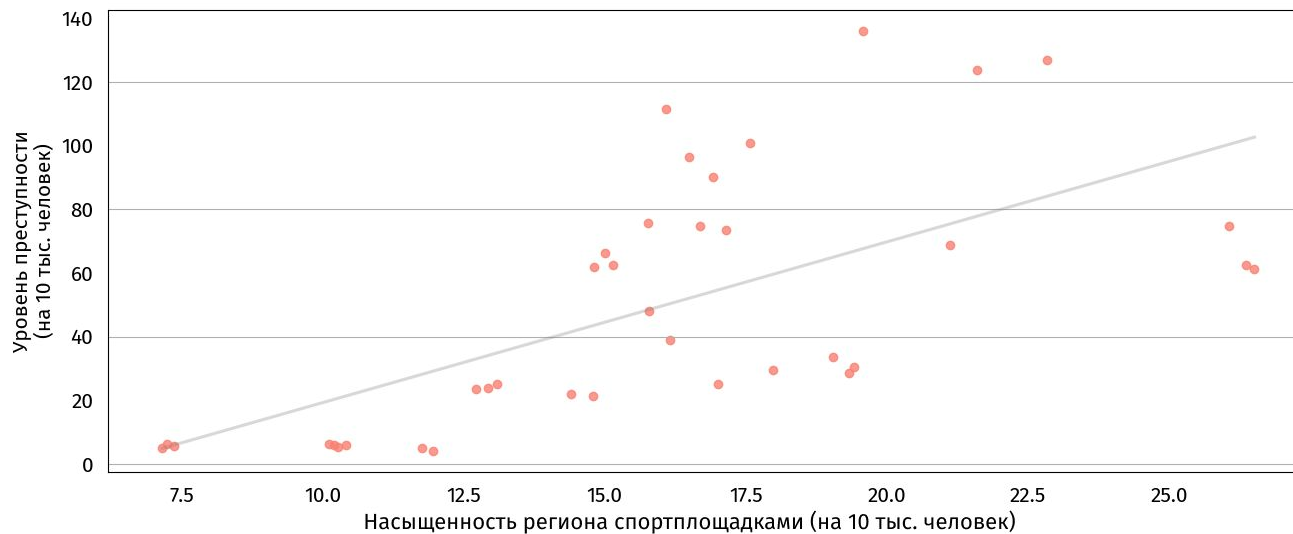
объясняемая переменная	в чём измеряется
уровень преступности*	ед./10 тыс. чел

объясняющие переменные	в чём измеряется
занятость	%
насыщенность региона спортооружениями	ед./10 тыс. чел

*В датасете в качестве преступности рассматриваются количество преступлений, подпадающих под одну из этих статей УК РФ: убийство, умышленное причинение вреда здоровью, изнасилование, кража, разбой, вымогательство, хулиганство

Графики взаимосвязи уровня преступности от насыщенности региона спортплощадками

низкий уровень занятости

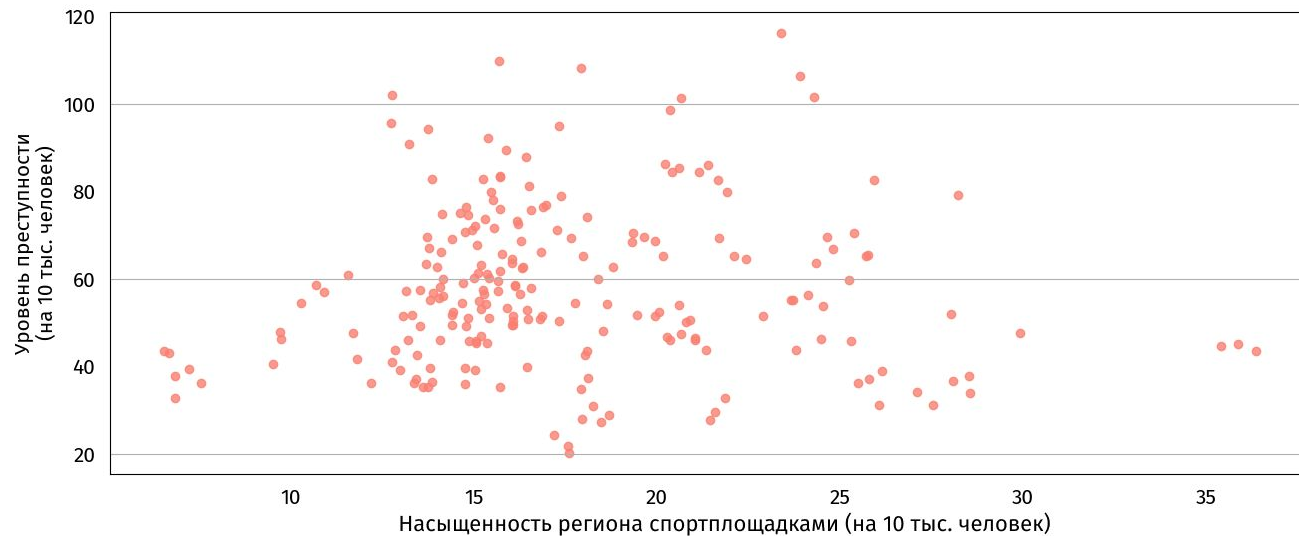


$r = 0,64$

$p\text{-value} = 1,3 \times 10^{-5}$

Графики взаимосвязи уровня преступности от насыщенности региона спортплощадками

средний и высокий уровень занятости



$r = 0,01$
 $p\text{-value} = 0,88$

Нормализация данных

критерии нормализации

насыщенность спортооружениями

покупательная способность

безработица

доля молодёжи

доля городского населения

$$X'_i = \frac{X_i - \bar{X}}{\sigma_x}$$

где X_i - величина, \bar{X} - среднее значение, σ_x - стандартное отклонение

Линейная регрессия

$$\text{crimes} \sim a_0 \times \text{sports} + a_1 \times \text{power} + a_2 \times \text{unem} + a_3 \times \text{youth} + a_4 \times \text{urban} + a_5$$

где sports - насыщенность региона спортооружениями, power - покупательная способность, unem - уровень безработицы, youth - доля молодёжи, urban - доля городского населения, $a_0, a_1, a_2, a_3, a_4, a_5$ - коэффициенты

Линейная регрессия

при низкой занятости

объясняемая переменная	коэффициент детерминации		
количество преступлений	0,8		
объясняющие переменные	коэффициент	p-value	sd
насыщенность спортсооружениями	0,9802	0,001	0,074
покупательная способность	-0,8075	0,001	0,039
безработица	-0,3657	0,063	0,067
доля молодёжи	0,3490	0,257	0,090
доля городского населения	0,2141	0,269	0,074
свободный член	0,1383	0,576	0,076

Линейная регрессия

при высокой занятости

объясняемая переменная	коэффициент детерминации		
количество преступлений	0,85		
объясняющие переменные	коэффициент	p-value	sd
доля молодёжи	0,5963	0,012	0,065
доля городского населения	0,4830	0,008	0,058
покупательная способность	-0,3841	0,001	0,019
насыщенность sportсооружениями	0,1472	0,086	0,023
безработица	-0,0590	0,501	0,021
свободный член	0,0275	0,749	0,032

Результат проверки гипотез

Первая гипотеза **опровергнута**

Вторая гипотеза **подтверждена**

Интерпретация полученных результатов

При высокой занятости количество спортивных сооружений **не влияет на преступность**

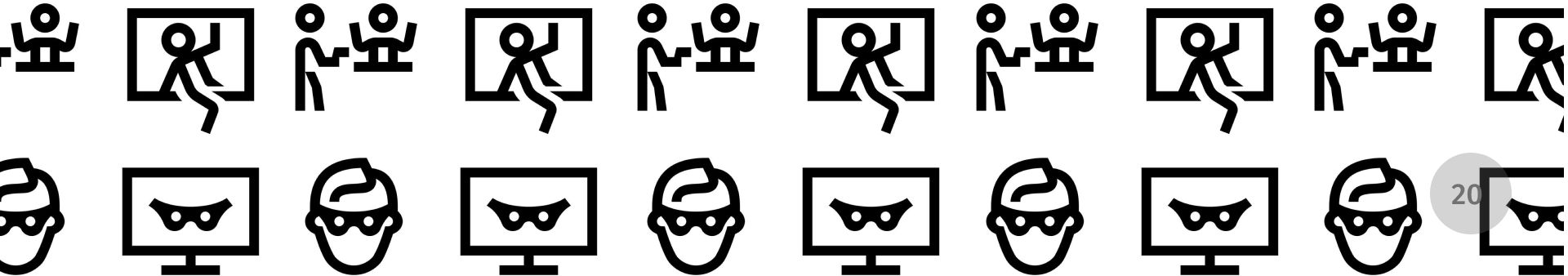
При низкой занятости с увеличением количества спортивных сооружений **увеличивается преступность**

Ограничения

Результаты исследования нельзя обобщить на весь мир

Исследование не даёт возможности проследить динамику изменения показателей за длительный период времени

Нельзя сделать выводы по отдельным социальным группам

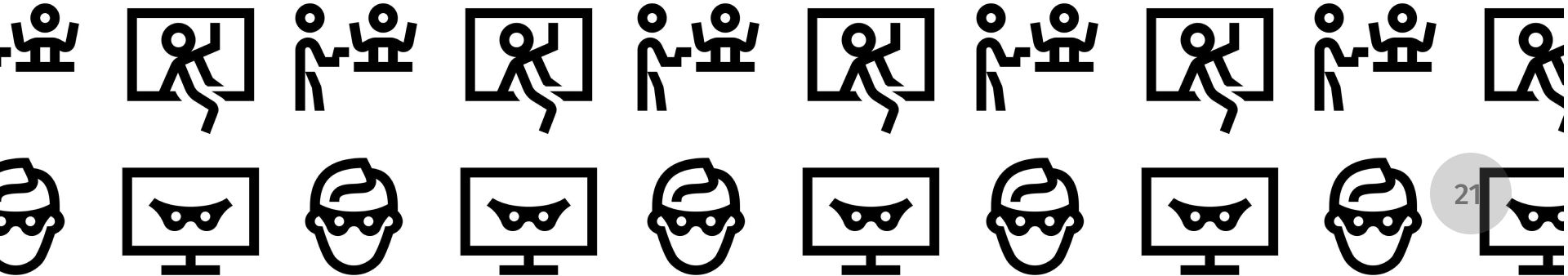


Перспективы

Получить разбивку данных по спортивным сооружениям и преступности

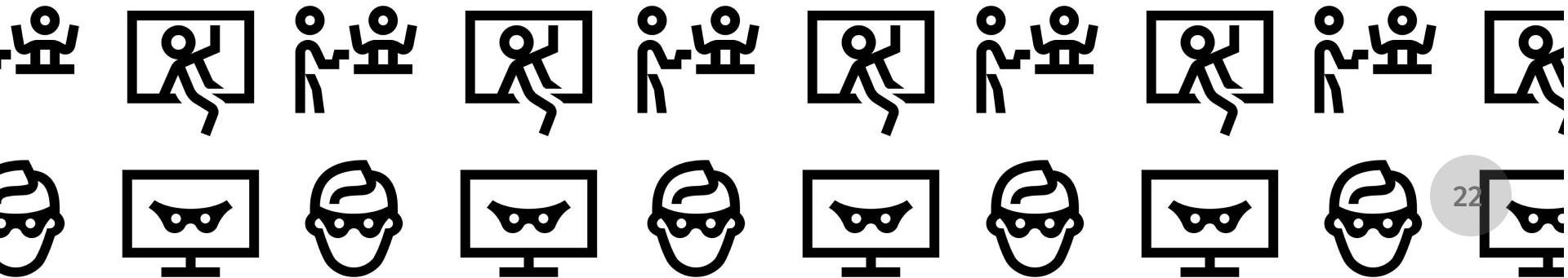
Получить данные по спортивным сооружениям за 2022 год и более ранний период

Получить данные по другим переменным, влияющим на преступность (количество сирот-беспризорников, бездомных и т.п.)



Применение

Наше исследование может быть полезно органам власти для контроля над уровнем преступности и эффективного расходования бюджетных средств



Policy implication

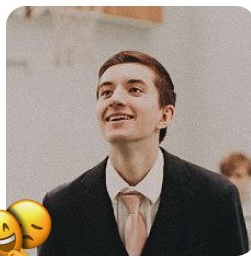
В условиях ограниченности бюджетных средств приоритет строительства новых спортсооружений должен отдаваться регионам с высоким уровнем занятости

В регионах с низким уровнем занятости необходимо перераспределить часть бюджетных средств, выделенных на спорт, в пользу программ, направленных на создание новых рабочих мест (привлечение инвестиций в регион), программ переквалификации работников, а также на реализацию мер социальной поддержки

Команда



Никита Ламонов



Иван Давыдов



Элиза Мирсяева



Михаил Михаловский



Виктор Фрик

Github



Ссылки на другие статьи

Спорт помогает снизить подростковую преступность

<https://nn.aif.ru/society/details/531040>

ПРОПАГАНДА СПОРТА В БОРЬБЕ С ПРЕСТУПНОСТЬЮ МОЛОДЕЖИ

<http://partner.iro63.ru/index.php/stati/2018-god/21-2018-god/78-stati/370-21-20>

Проверка на прочность

при низком уровне занятости

	1	2	3	4	5	6	7	8	9	sd
purchasing_power	-0.855567	-0.855281	-0.798540	-0.785234	-0.756474	-0.825943	-0.792184	-0.743031	-0.817220	0.039430
unemployment	-0.382567	-0.329119	-0.339859	-0.393212	-0.334339	-0.339889	-0.523243	-0.344770	-0.286713	0.067231
sports_facilities_to_population	1.031679	0.906139	1.070109	0.919991	0.935634	1.067998	0.895720	0.944501	1.053323	0.073566
urban_population	0.190693	0.295280	0.234393	0.215524	0.292177	0.135908	0.246805	0.177859	0.066202	0.073690
youth	0.295548	0.333497	0.268704	0.439218	0.279777	0.326966	0.538612	0.375052	0.277629	0.089919
const	0.269565	0.098964	0.144477	0.115708	0.149825	0.157041	0.051647	0.053011	0.245964	0.075823

при высоком уровне занятости

	1	2	3	4	5	6	7	8	9	10	sd
purchasing_power	-0.414867	-0.390643	-0.365574	-0.404441	-0.380582	-0.388549	-0.359522	-0.356980	-0.386000	-0.372991	0.018861
unemployment	-0.062445	-0.077215	-0.062843	-0.079532	-0.058748	-0.088538	-0.037925	-0.055100	-0.032998	-0.023479	0.021180
sports_facilities_to_population	0.116343	0.158920	0.159864	0.140089	0.147590	0.181530	0.121273	0.180065	0.120717	0.149399	0.023420
urban_population	0.458680	0.463371	0.392267	0.501969	0.535348	0.452618	0.577421	0.550246	0.466742	0.426111	0.057896
youth	0.658711	0.625807	0.646227	0.602472	0.537050	0.616251	0.465632	0.518866	0.639893	0.642943	0.065271
const	0.051969	0.038302	0.059873	0.017791	0.030390	0.048289	0.059115	-0.035383	0.010582	-0.012727	0.031888

Линейная регрессия

при низком уровне занятости

OLS Regression Results

Dep. Variable:	crimes_to_population	R-squared:	0.808
Model:	OLS	Adj. R-squared:	0.778
Method:	Least Squares	F-statistic:	22.96
Date:	Wed, 20 Dec 2023	Prob (F-statistic):	9.93e-10
Time:	02:30:46	Log-Likelihood:	-38.103
No. Observations:	38	AIC:	88.21
Df Residuals:	32	BIC:	98.03
Df Model:	5		
Covariance Type:	HC3		

	coef	std err	z	P> z	[0.025	0.975]
purchasing_power	-0.8075	0.187	-4.329	0.000	-1.173	-0.442
unemployment	-0.3657	0.197	-1.858	0.063	-0.752	0.020
sports_facilities_to_population	0.9802	0.243	4.028	0.000	0.503	1.457
urban_population	0.2141	0.194	1.105	0.269	-0.166	0.594
youth	0.3490	0.308	1.133	0.257	-0.255	0.953
const	0.1383	0.248	0.559	0.576	-0.347	0.623

при высоком уровне занятости

OLS Regression Results

Dep. Variable:	crimes_to_population	R-squared:	0.852
Model:	OLS	Adj. R-squared:	0.849
Method:	Least Squares	F-statistic:	92.50
Date:	Tue, 19 Dec 2023	Prob (F-statistic):	3.16e-51
Time:	23:03:50	Log-Likelihood:	-226.59
No. Observations:	217	AIC:	465.2
Df Residuals:	211	BIC:	485.5
Df Model:	5		
Covariance Type:	HC3		

	coef	std err	z	P> z	[0.025	0.975]
purchasing_power	-0.3841	0.095	-4.055	0.000	-0.570	-0.198
unemployment	-0.0590	0.088	-0.672	0.501	-0.231	0.113
sports_facilities_to_population	0.1472	0.086	1.715	0.086	-0.021	0.315
urban_population	0.4830	0.183	2.632	0.008	0.123	0.843
youth	0.5963	0.238	2.506	0.012	0.130	1.063
const	0.0275	0.086	0.319	0.749	-0.141	0.197

