

НАЦИОНАЛЬНЫЙ ЦЕНТР
КОГНИТИВНЫХ РАЗРАБОТОК

НЦКР

ДУМАТЬ ЧТОБЫ ДЕЙСТВОВАТЬ

Влияние факторов городской среды на частоты развития хронических заболеваний в Санкт-Петербурге

Кадышева Анна Леоднидовна
annakamysheva7@gmail.com

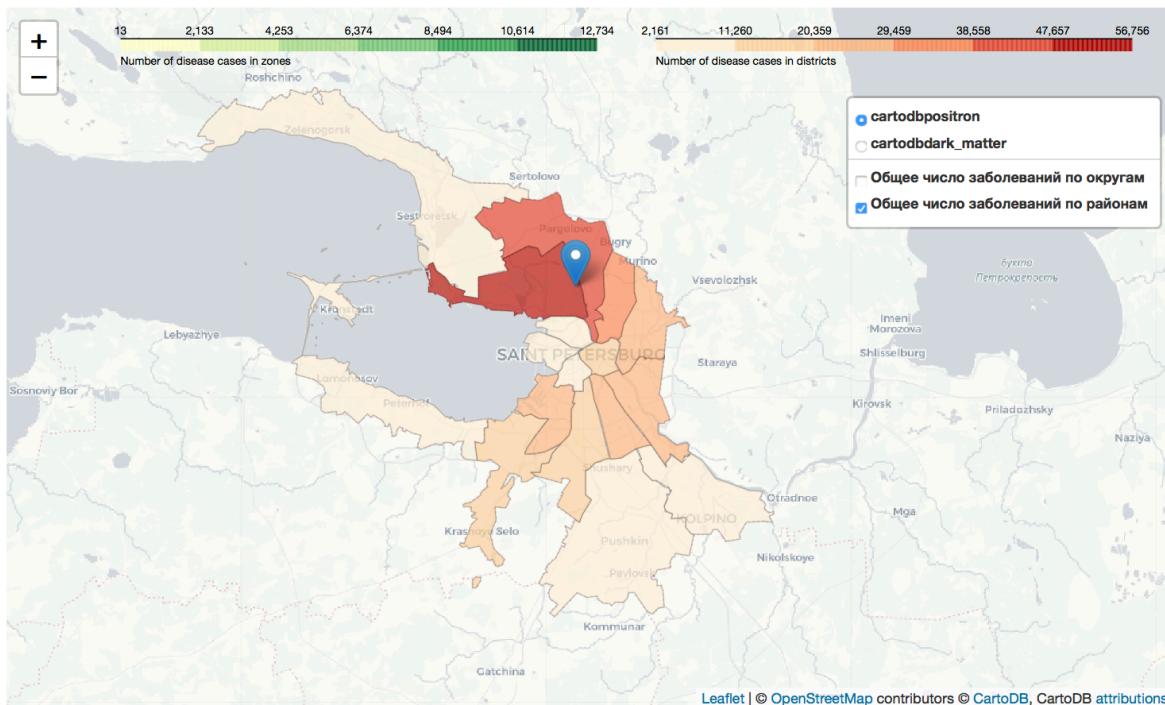
Ментор: Балахонцева Марина Андреевна

Цель исследования - выявление различных социальных факторов и факторов городской среды, влияющих на частоты развития хронических заболеваний и построение интерактивной карты города, отражающих эти закономерности.

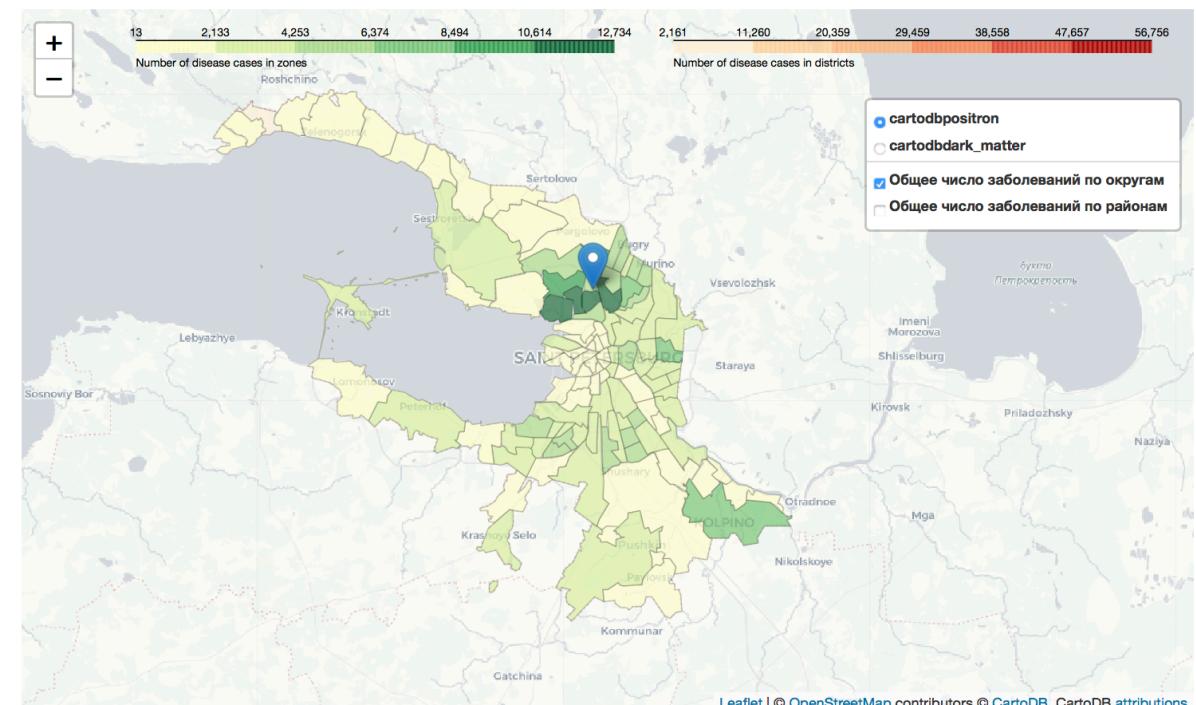
- Из медицинских записей на естественном языке выделить адреса проживания пациентов, а также их различные характеристики (пол, возраст, семейное положение и др.)
- Определить пациентов, проживающих в Санкт-Петербурге, и сопоставить адреса их проживания 18 районам города
- Рассчитать частоты возникновения 35 заболеваний в районах Санкт-Петербурга
- Отранжировать районы по частоте возникновения заболеваний
- Составить набор характеристик для районов по различным социологическим факторам и факторам городской среды
- Определить влияние этих факторов на частоты развития хронических заболеваний
- Составить интерактивную карту города, отображающую выявленные закономерности

Основной вклад в распределение частот заболеваний по районам/округам вносить фактор близости района/округа к центру Алмазова (пациенты из удаленных районов обращаются за медицинской помощью в другие учреждения из-за чего мы упускаем эту часть выборки из расчетов)

Окраска районов по суммарному количеству заболеваний



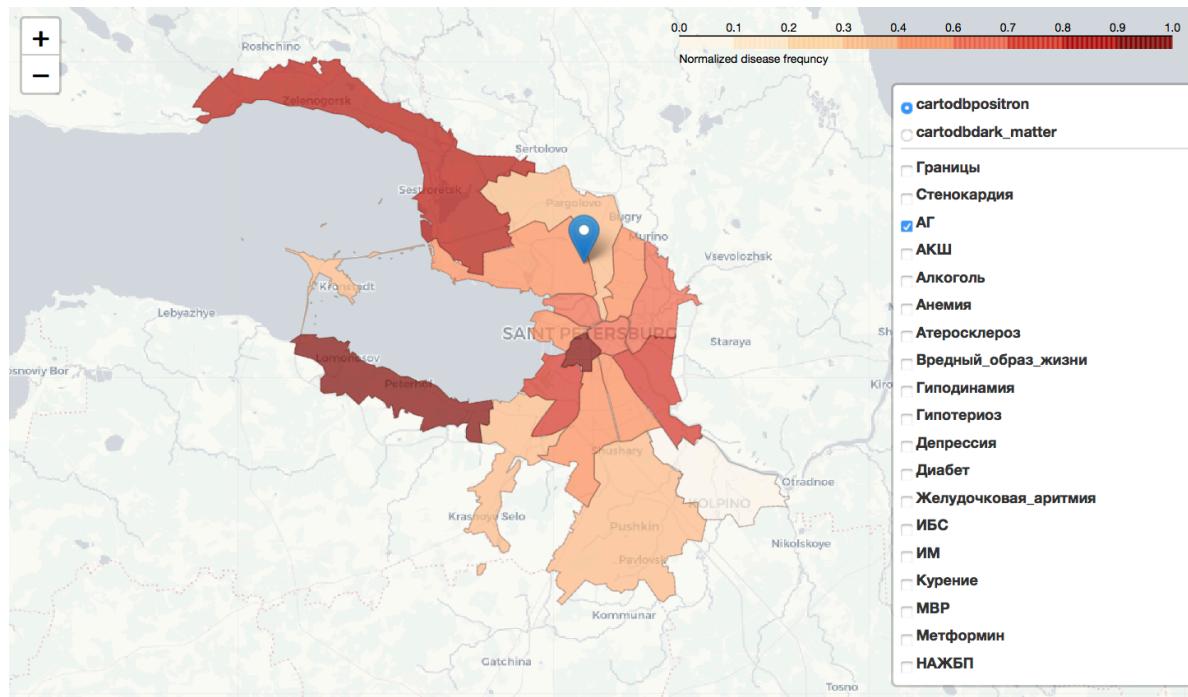
Окраска округов по суммарному количеству заболеваний



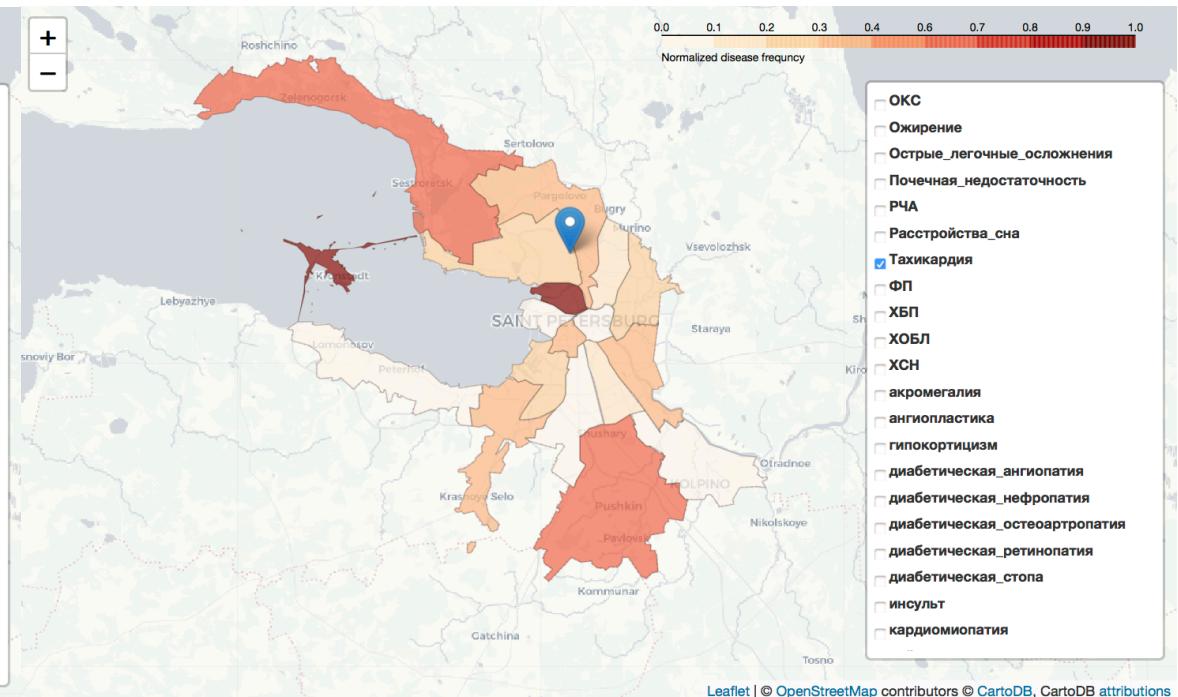
Решение:

Рассматривать не абсолютные частоты хронических заболеваний по районам (доля пациентов центра Алмазова от общего числа жителей в районе), а долю пациентов с конкретной хроникой среди всех пациентов, обратившихся за высокотехнологичной медицинской помощью в центр Алмазова

Нормализованные частоты АГ среди пациентов центра Алмазова

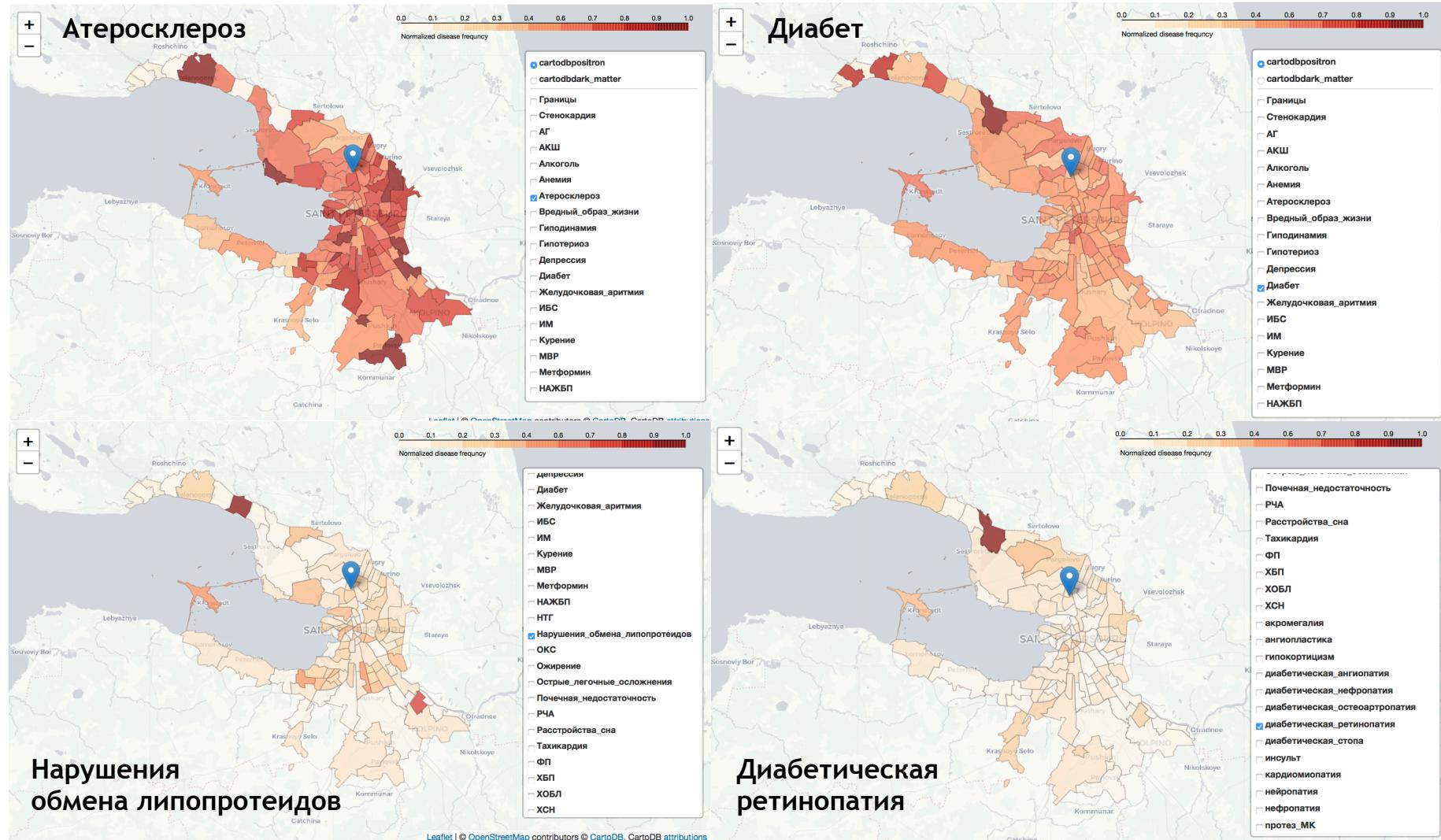


Нормализованные частоты Тахикардии среди пациентов центра Алмазова



- Использовались данные о заболеваемости 15214 равнообследованных пациентов центра Алмазова, проживающих в Санкт-Петербурге
- Частота заболевания в конкретном районе/округе рассчитывалась как число обращений по данному заболеванию в центр Алмазова среди пациентов, проживающих в данном районе/округе, деленное на общее число обращений в центр Алмазова по району/округу

Нормализованные частоты заболеваний среди пациентов центра Алмазова



Частоты распределяются независимо от близости к центру Алмазова!

Тест на влияние района/округа на распределение частот заболеваний

- Для проверки гипотезы о влиянии фактора района/округа использовался критерий согласия Пирсона (хи-квадрат)
- Округ гораздо реже влияет на распределение частот заболеваний

Факторы, для которых зона играет роль(8):

АКШ, ОКС, РЧА, диабетическая_ангиопатия, диабетическая_нефропатия, диабетическая_остеоартропатия, диабетическая_ретинопатия

Факторы, для которых район играет роль(12):

АКШ, Атеросклероз, ИМ, Курение, НТГ, ОКС, РЧА, Расстройства_сна, диабетическая_нефроатия, диабетическая_остеоартропатия, диабетическая_ретинопатия, инсульт, нейропатия, нефропатия

Результаты для разбиения по МО

	Statistics	p-value	p-value < 0.05
Вредные привычки	100.909994	0.815313	False
Заболевания почек	118.429659	0.815313	False
Легочные заболевания	100.652522	0.815313	False
Сердечно-сосудистые заболевания	9.532415	1.000000	False
Сердечные операции	104.800382	0.815313	False
Эндокринные заболевания	105.682901	0.815313	False

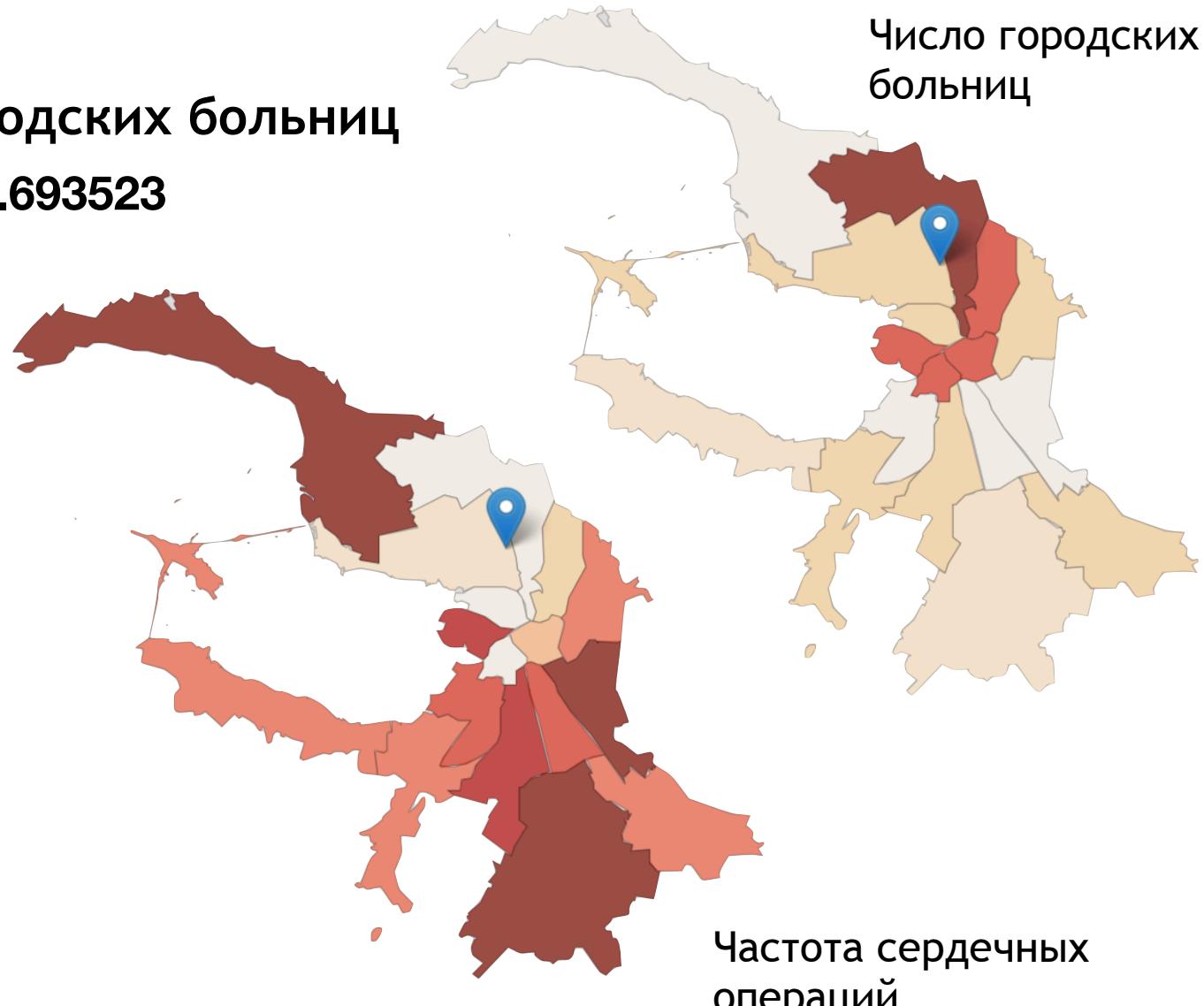
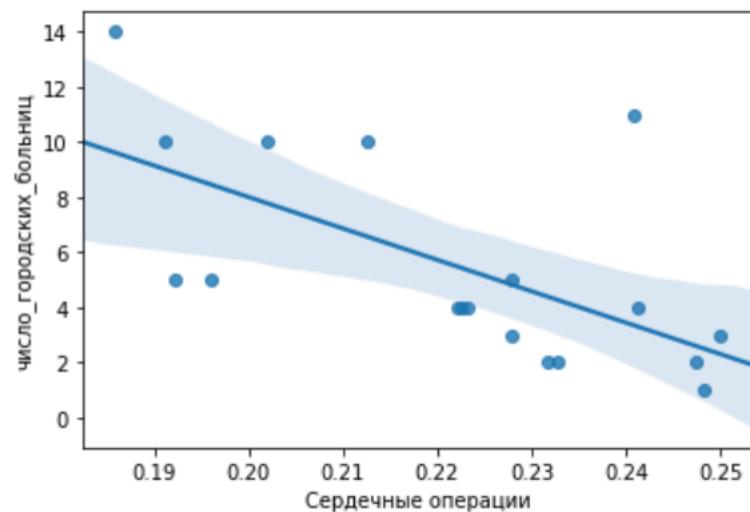
Результаты для разбиения по районам

	Statistics	p-value	p-value < 0.05
Вредные привычки	36.170391	8.704482e-03	True
Заболевания почек	74.673935	1.998803e-08	True
Легочные заболевания	34.511298	1.081127e-02	True
Сердечно-сосудистые заболевания	4.093560	9.993958e-01	False
Сердечные операции	25.348068	1.046226e-01	False
Эндокринные заболевания	47.237189	3.366542e-04	True

Сердечные операции - число городских больниц

Коэффициент корреляции Спиромана: **-0.693523**

В данные по сердечным операциям
вносит сильный вклад наличие
более близких больниц



Легочные заболевания - рейтинг удобства инфраструктуры

Коэффициент корреляции Спиромана: **-0.670242**

Легочные заболевания - доля района очень удобная для жизни

Коэффициент корреляции Спиромана:

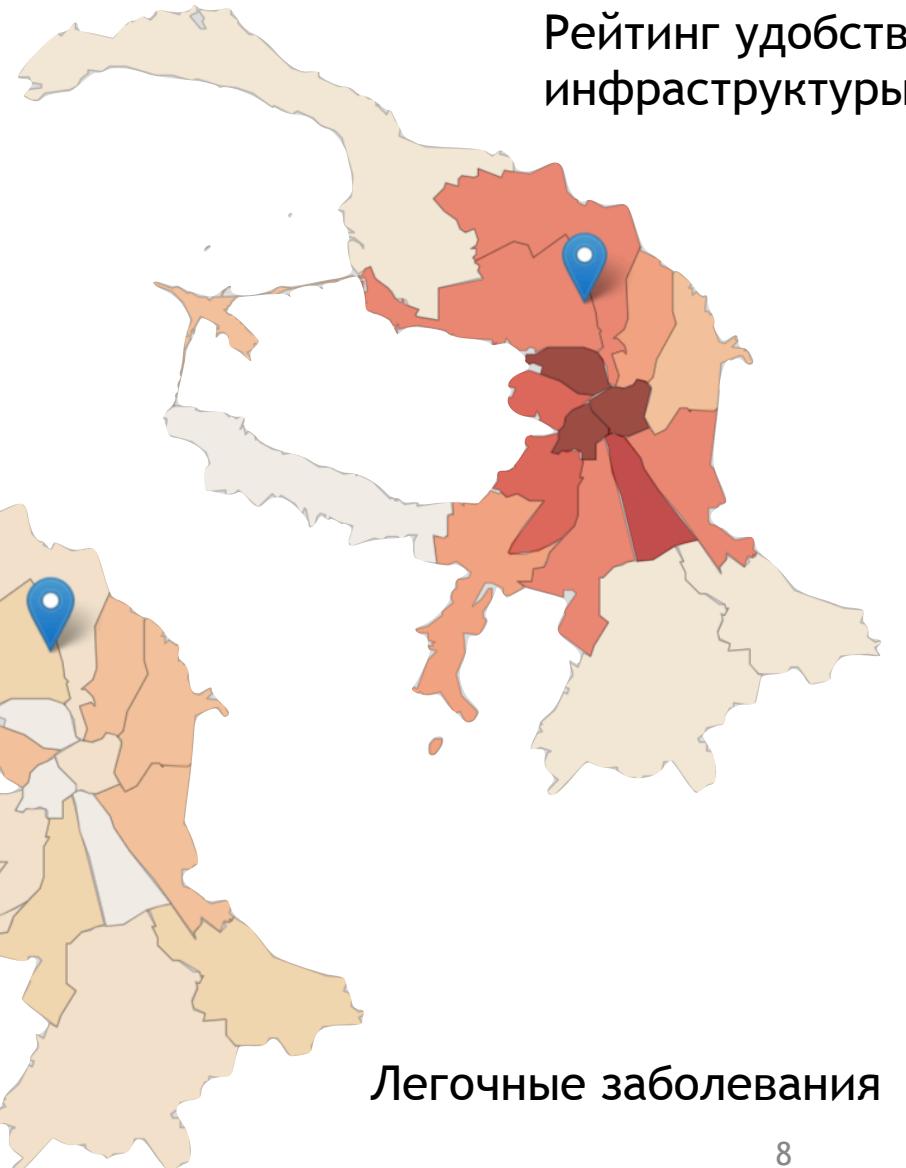
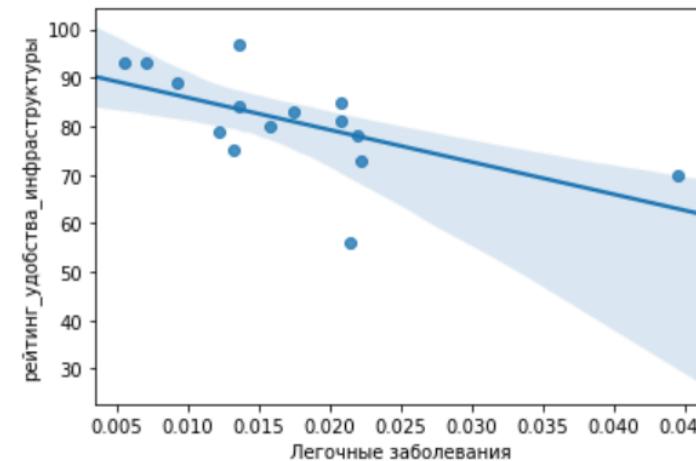
-0.631966

Легочные заболевания - доля района с недостаточным развитием инфраструктуры

Коэффициент корреляции Спиромана:

0.630084

Легочные
заболевания
связаны с
развитием
инфраструктуры
района



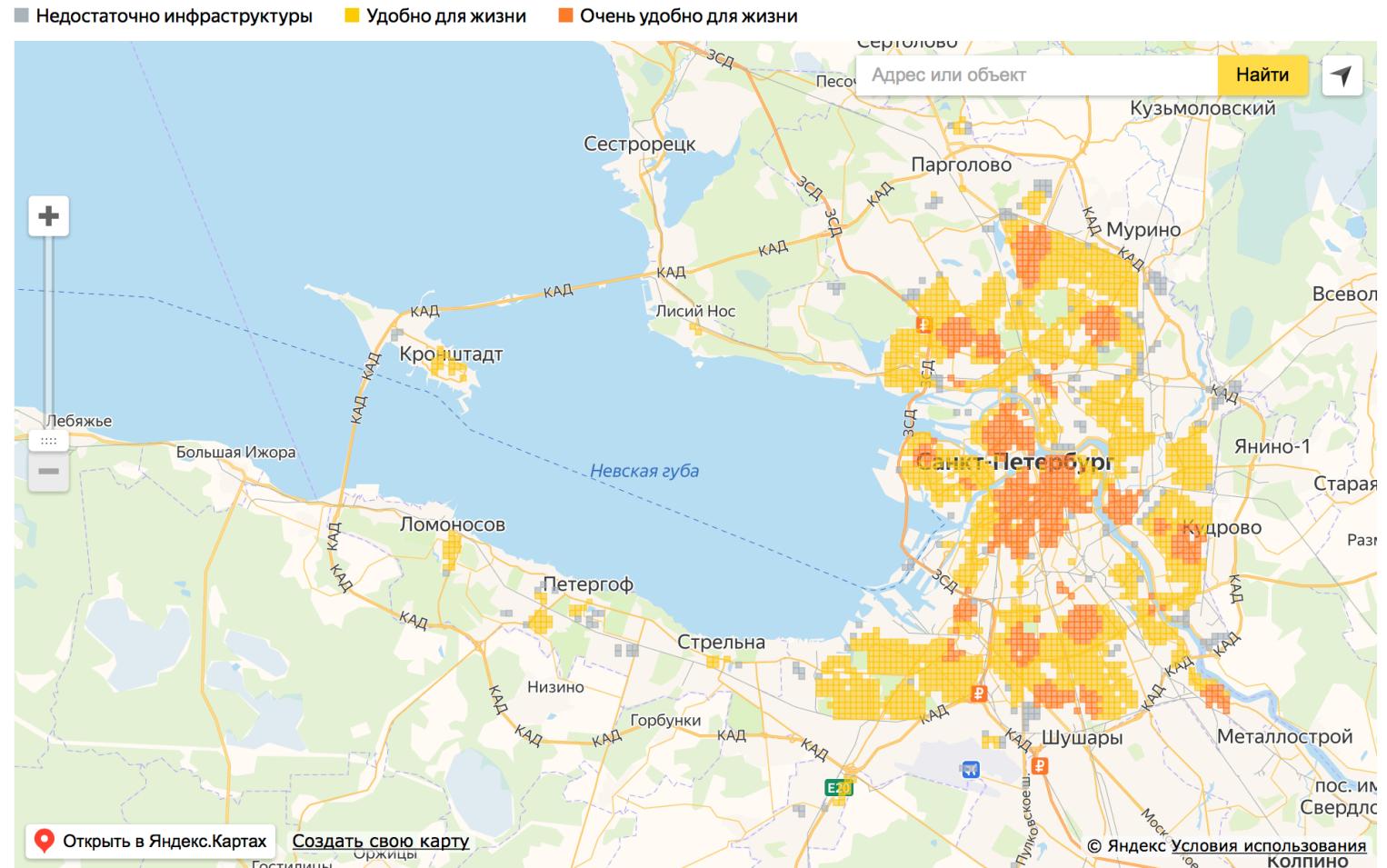
Удобный квадрат - в течение 10 минут можно дойти до:

- хотя бы двух магазинов продуктов,
- детского сада,
- общеобразовательной школы (или гимназии, или лицея),
- аптеки,

и не более чем за 15 минут можно дойти до: салона красоты, магазина товаров для дома, поликлиники, детской поликлиники, ветеринарной клиники, фитнес-клуба, хотя бы одной из следующих организаций: ателье, почта, банк.

Очень удобный квадрат - просто удобный и не более чем за 15 минут можно дойти из центра квадрата до:

- бара,
- кафе (или ресторана, или столовой, или кофейни),
- кинотеатра,
- станции метро.



<https://yandex.ru/company/researches/2017/spb/districts>

Сердечно-сосудистые заболевания - площадь загрязнения воздуха древесной пылью в процентах

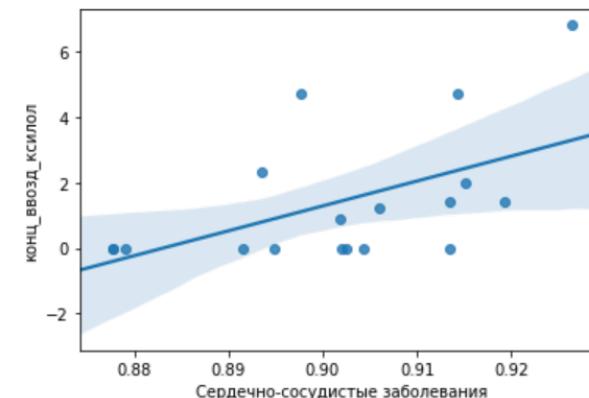
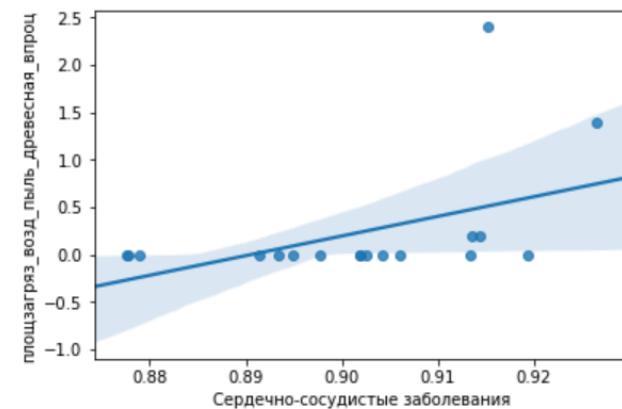
Коэффициент корреляции Спирамана: **0.657718**

Сердечно-сосудистые заболевания - концентрация в воздухе неорганической пыли

Коэффициент корреляции Спирамана: **0.656925**

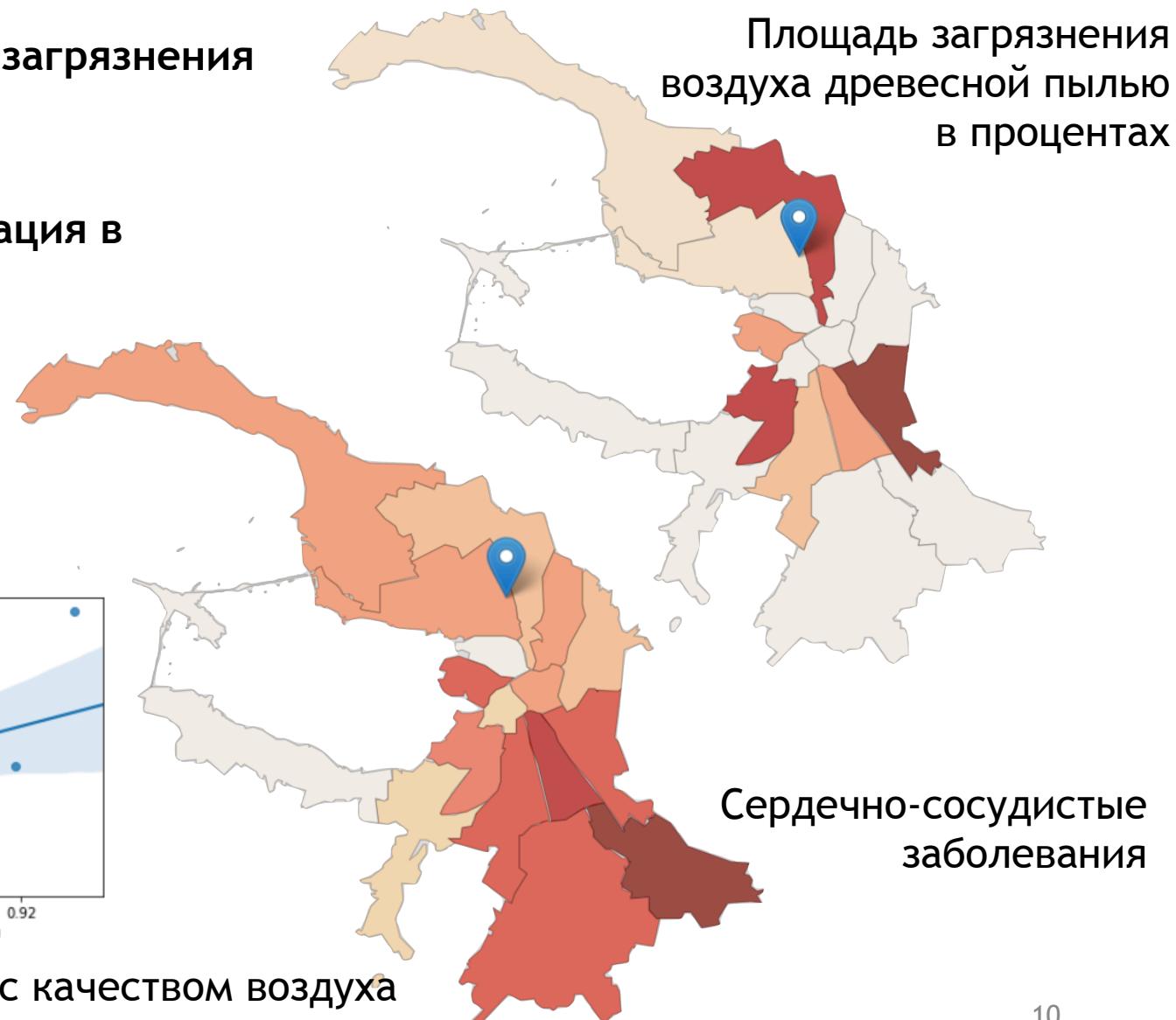
Сердечно-сосудистые заболевания - концентрация в воздухе ксилола

Коэффициент корреляции Спирамана: **0.548596**



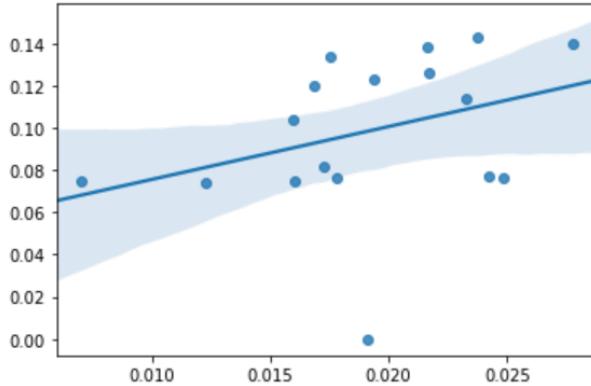
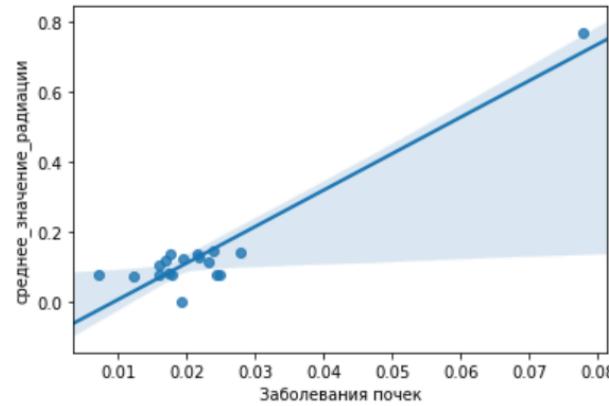
Сердечно-сосудистые заболевания коррелируют с качеством воздуха

<http://www.infoeco.ru/assets/files/godeco/admiralteisky.pdf>



Заболевания почек - среднее значение радиации

Коэффициент корреляции Спиромана: **0.608589**

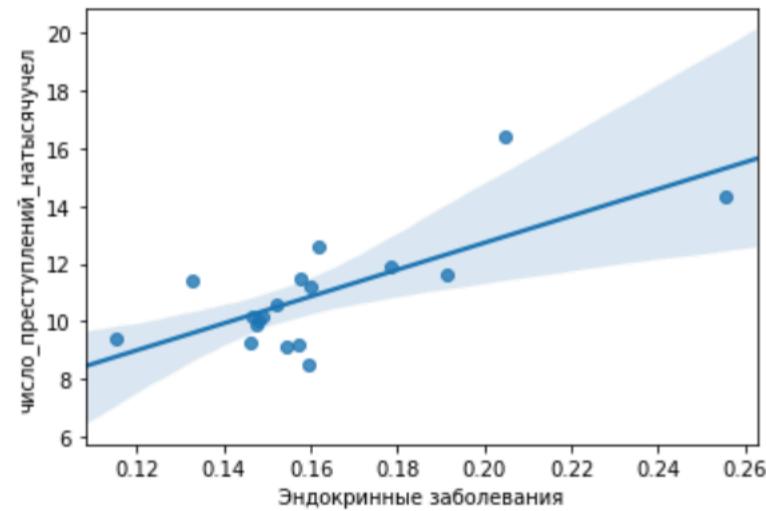


Если убрать выброс (Кронштадт)
Коэффициент корреляции - 0.4895



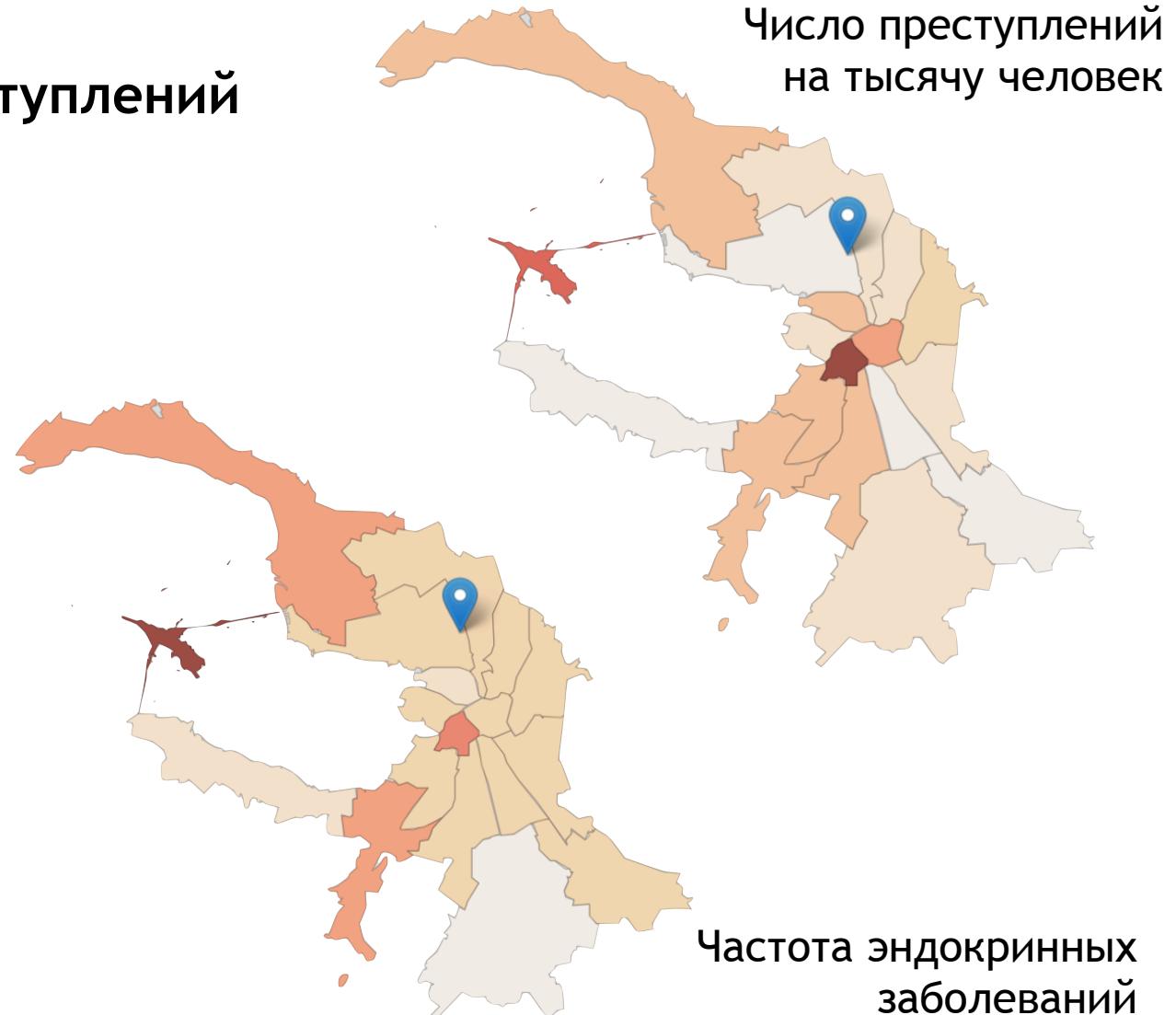
Эндокринные заболевания - число преступлений на тысячу человек

Коэффициент корреляции Спиромана: **0.605059**



Проблема объединения по группам - несбалансированность частот заболеваний.

Основной вклад в частоты эндокринных заболеваний вносит диабет и ожирение

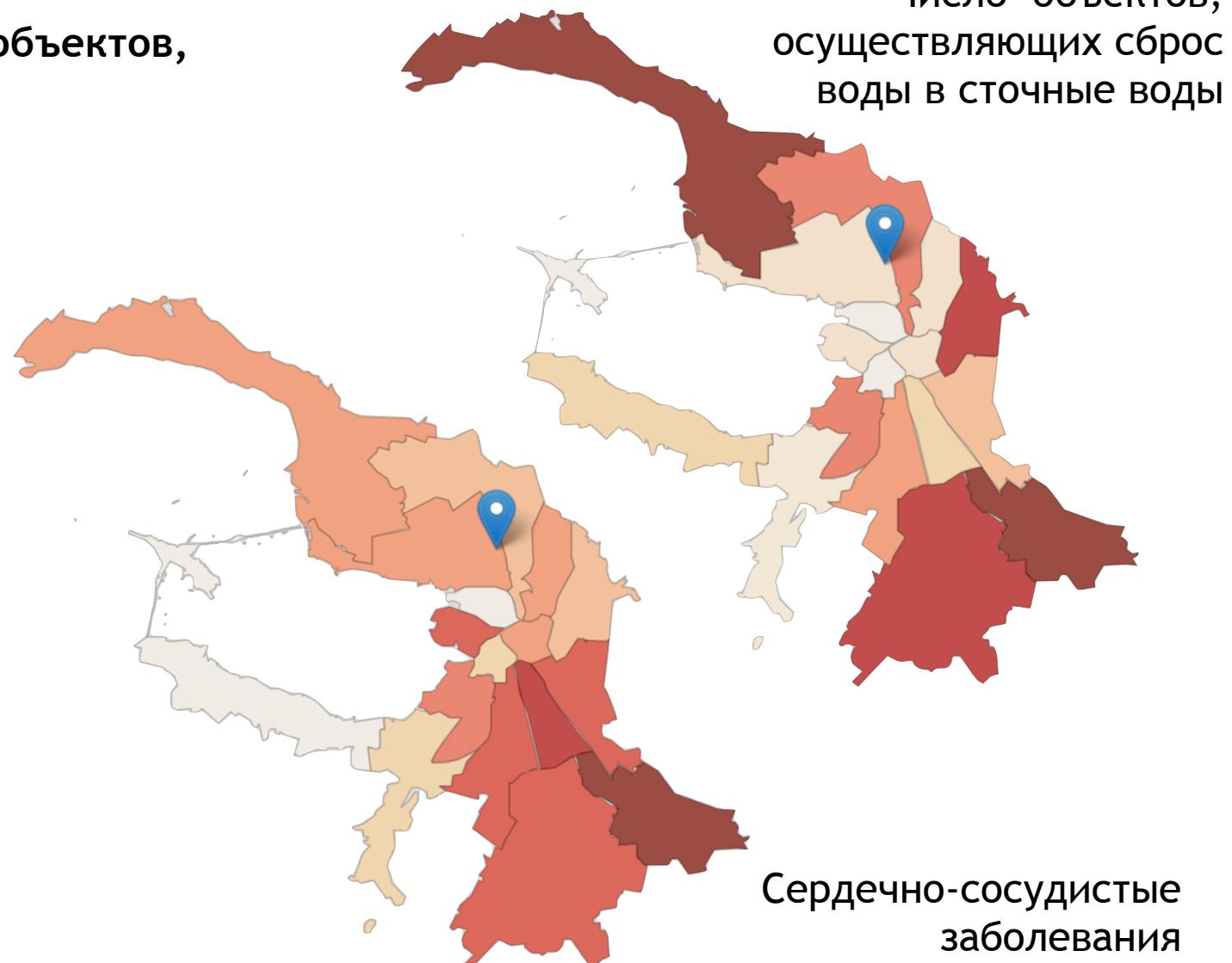
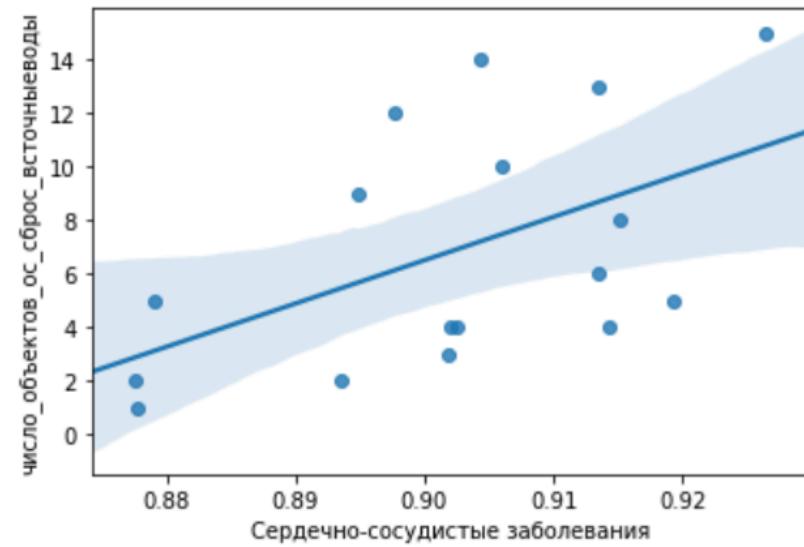


Сердечно-сосудистые заболевания - число объектов, осуществляющих сброс в сточные воды

Коэффициент корреляции Спиромана: **0.568270**

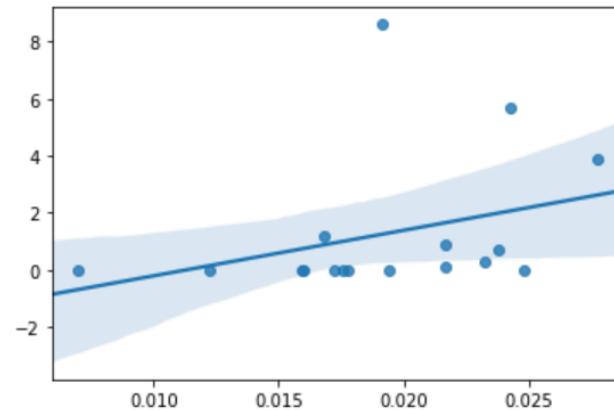
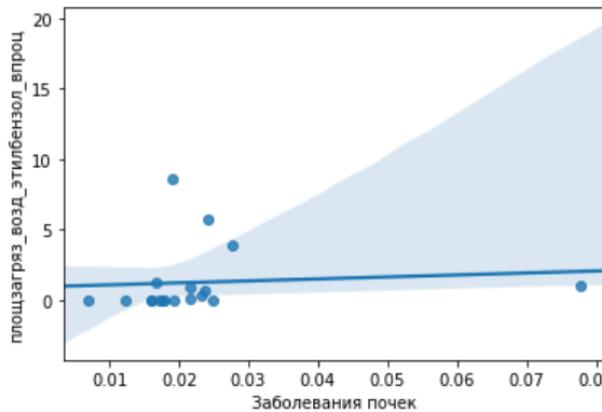
Сердечные операции - число объектов, осуществляющих сброс в сточные воды

Коэффициент корреляции Спиромана: **0.563349**

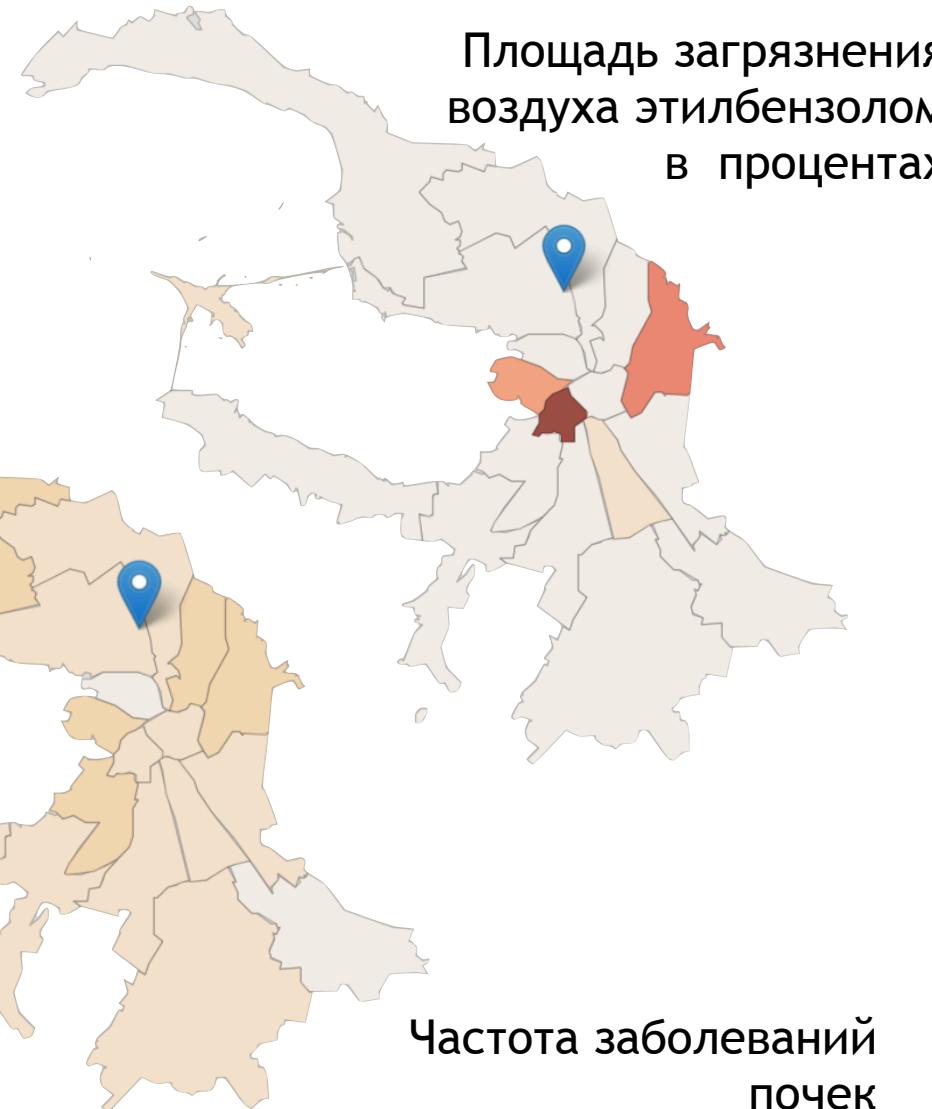


Заболевания почек - площадь загрязнения воздуха этилбензолом в процентах

Коэффициент корреляции Спирамана: **0.552360**

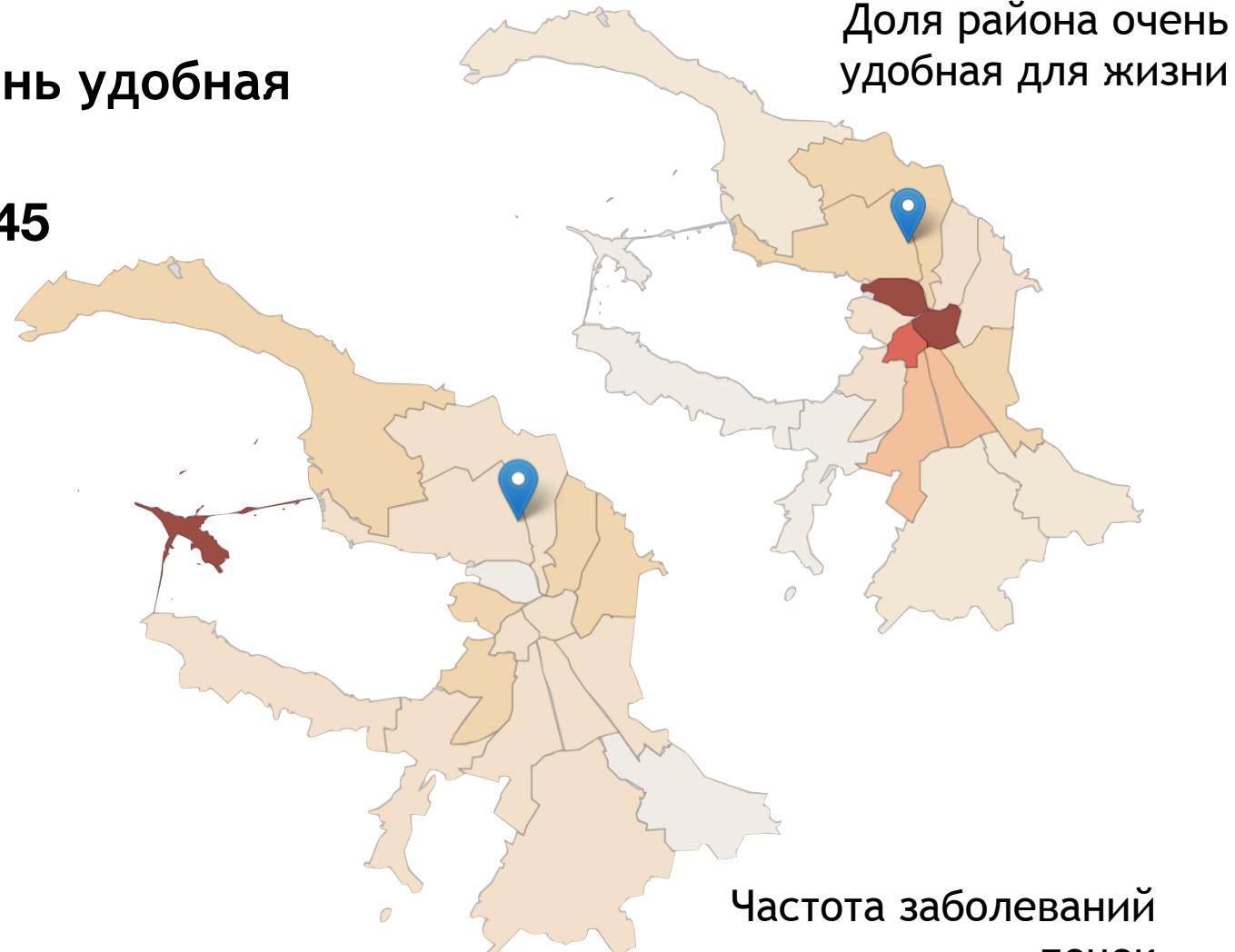
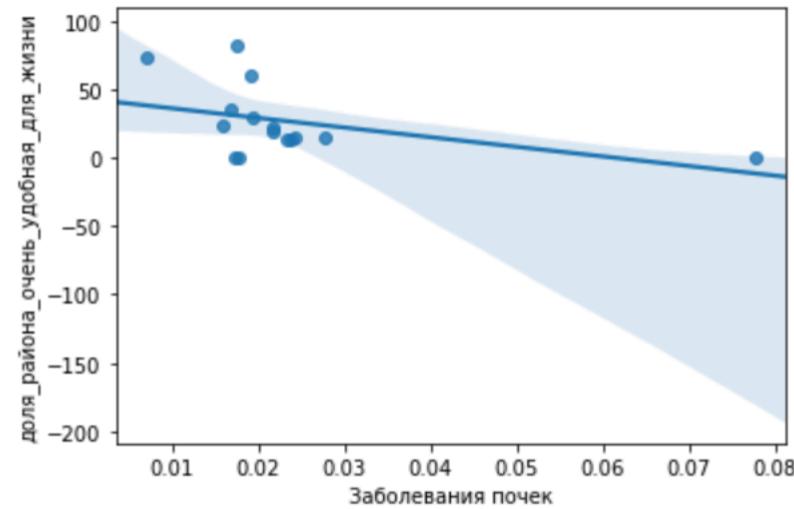


Если убрать выброс (Кронштадт)
Коэффициент корреляции - 0.53077



Заболевания почек - доля района очень удобная для жизни

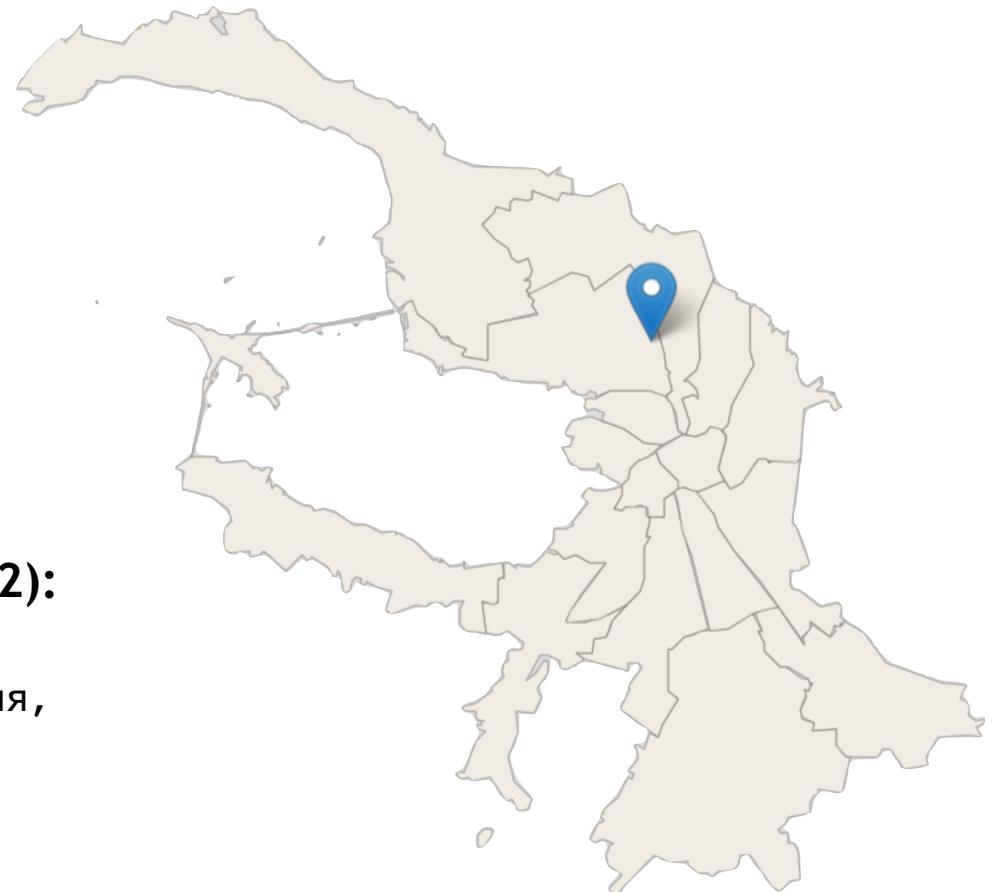
Коэффициент корреляции Спиромана: **-0.524245**



Поиск взаимосвязей между
факторами городской
среды и конкретными
заболеваниями/операциями

Факторы, для которых район играет роль(12):

АКШ, Атеросклероз, ИМ, Курение, НТГ, ОКС,
РЧА, Расстройства_сна, диабетическая_нефроатия,
диабетическая_остеоартропатия,
диабетическая_ретинопатия, инсульт,
нейропатия, нефропатия



Атеросклероз - площадь загрязнения воздуха диоксидом азота в процентах

Площадь загрязнения оксида

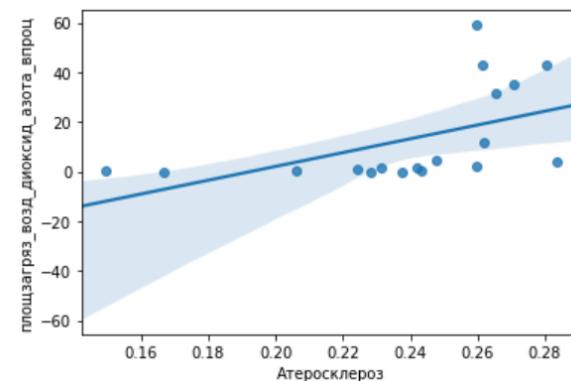
Адмиралтейский район	59.0
Василеостровский район	42.7
Выборгский район	1.3
Калининский район	2.1
Кировский район	31.3
Колпинский район	4.8
Красногвардейский район	3.7
Красносельский район	1.3
Кронштадтский район	0.0
Курортный район	0.0
Московский район	12.0
Невский район	42.7
Петроградский район	0.1
Петродворцовый район	0.2
Приморский район	0.1
Пушкинский район	0.8
Фрунзенский район	35.2
Центральный район	0.0

Коэффициент корреляции Спирмана:

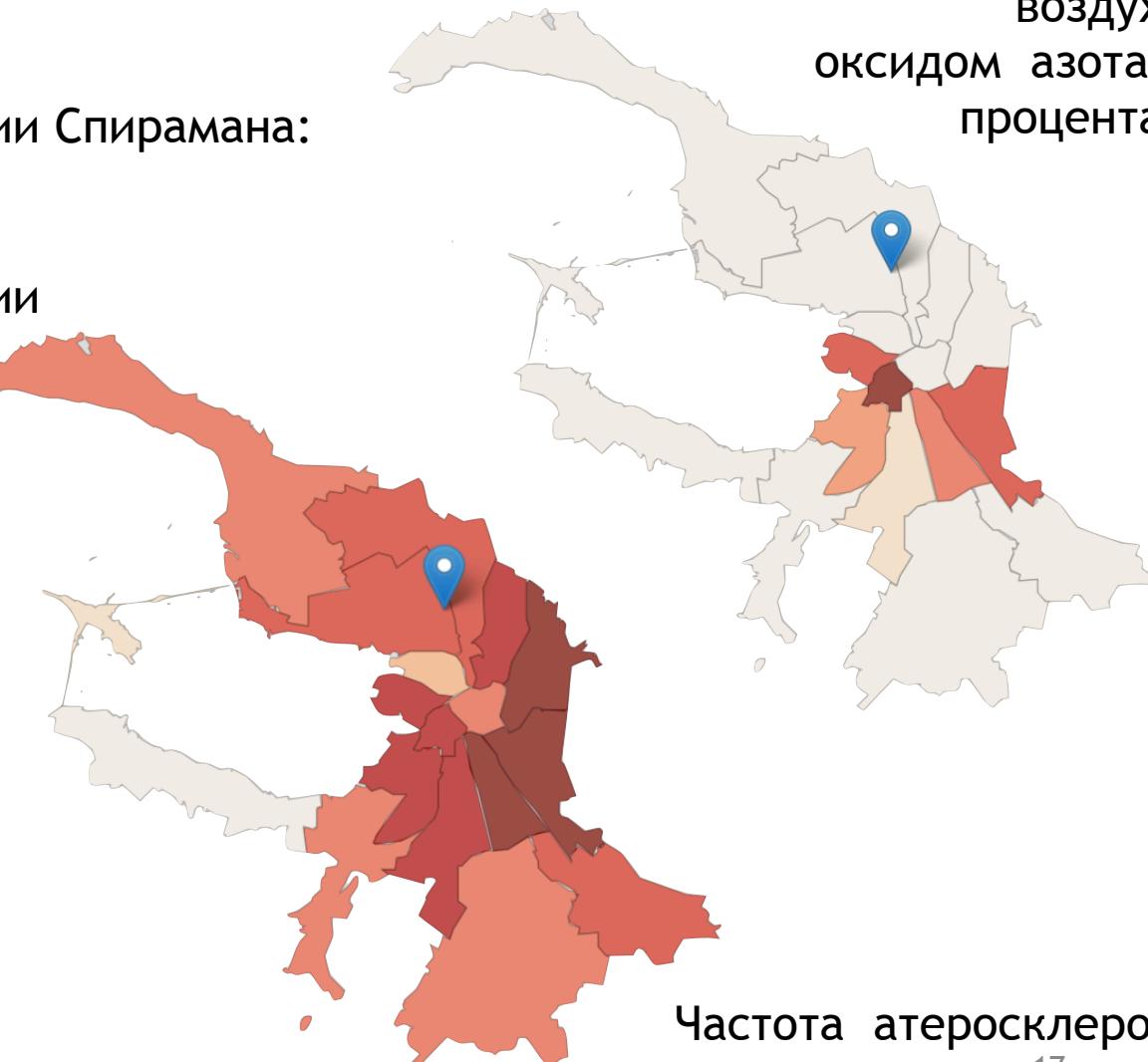
0.784813

Коэффициент корреляции
Пирсона:

0.5164



Площадь загрязнения
воздуха
оксидом азота в
процентах



Частота атеросклероза

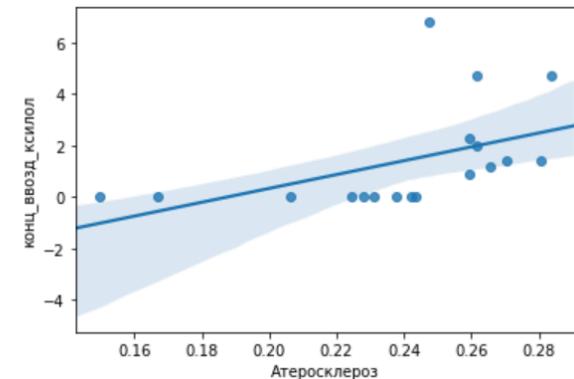
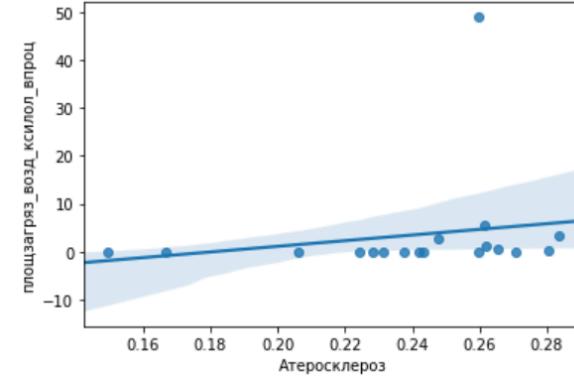
Атеросклероз - площадь загрязнения воздуха ксилолом в процентах

Площадь загрязнения

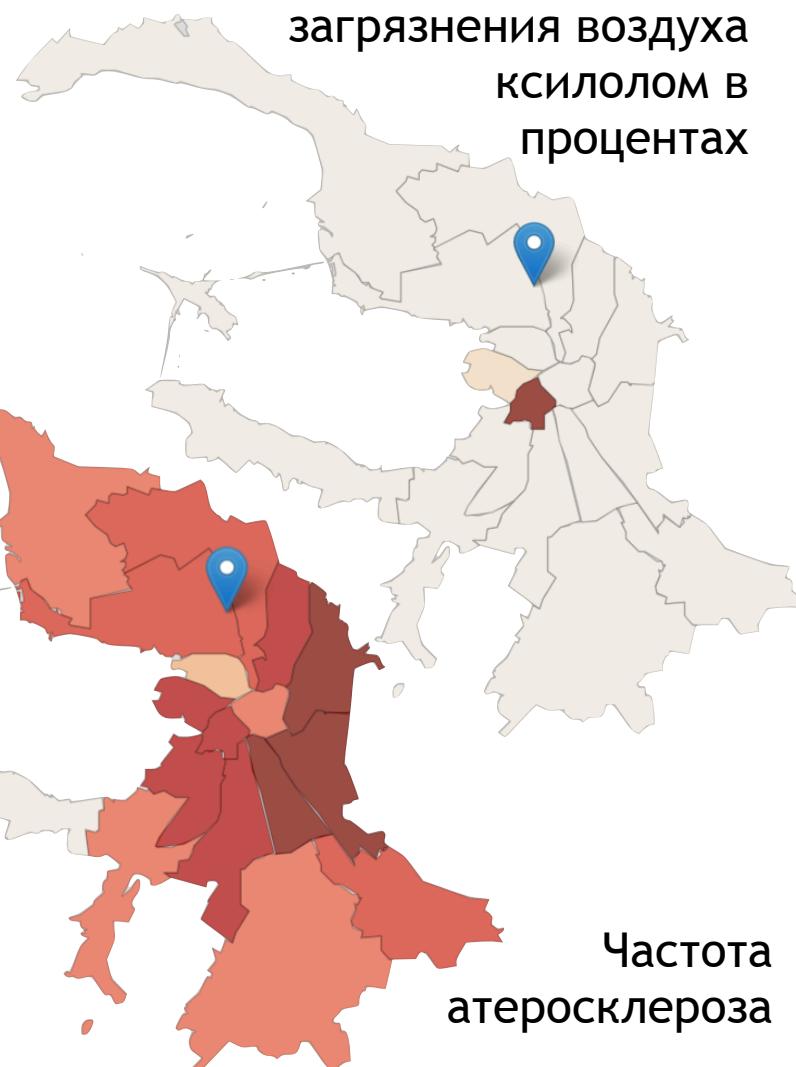
Адмиралтейский район	48.9
Василеостровский район	5.6
Выборгский район	0.0
Калининский район	0.0
Кировский район	0.6
Колпинский район	2.7
Красногвардейский район	3.4
Красносельский район	0.0
Кронштадтский район	0.0
Курортный район	0.0
Московский район	1.1
Невский район	0.3
Петроградский район	0.0
Петродворцовый район	0.0
Приморский район	0.0
Пушкинский район	0.0
Фрунзенский район	0.1
Центральный район	0.0

Коэффициент корреляции Спирмана:

0.746613



Концентрация в
воздухе ксилола
Коэффициент
корреляции
Спирмана
-0.7848

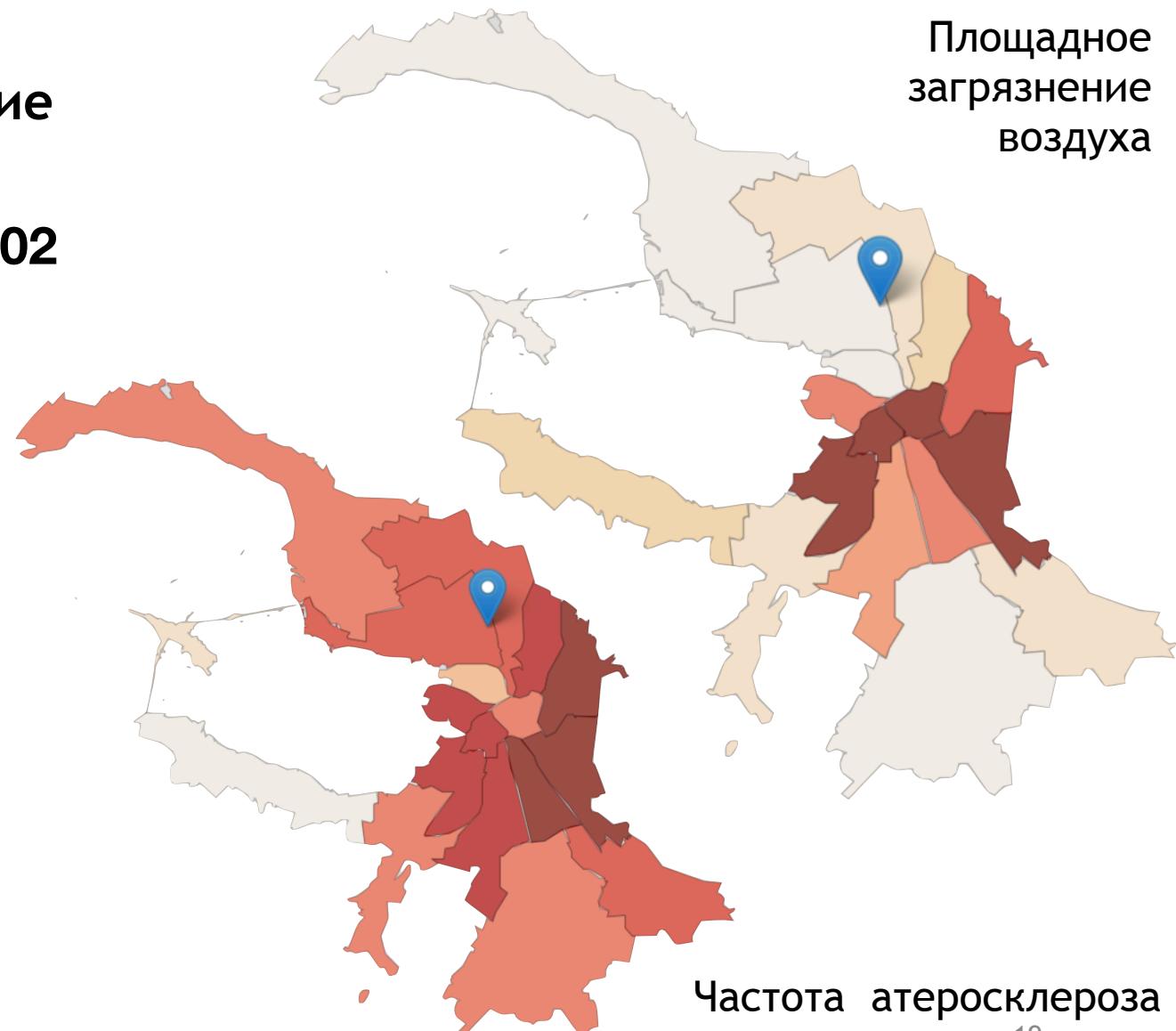
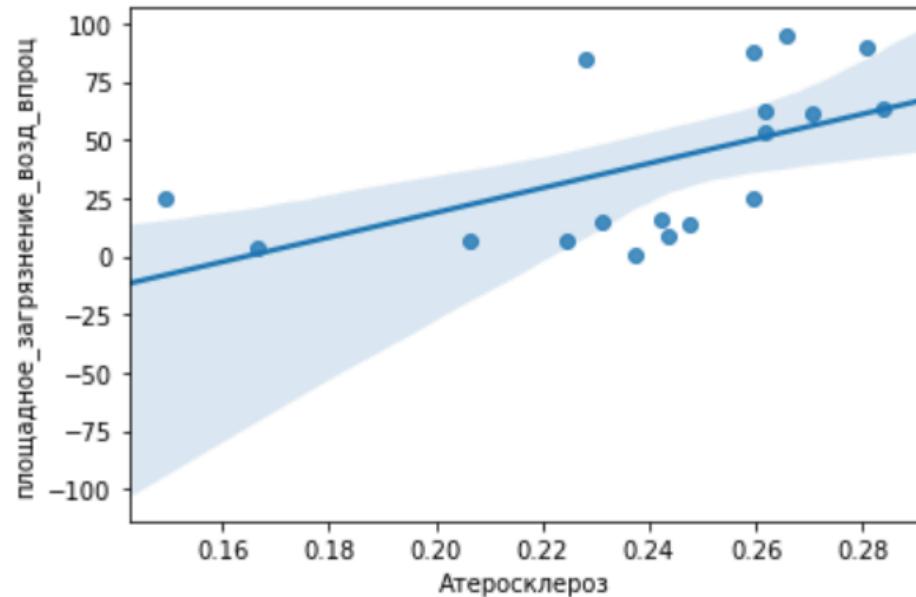


Площадь
загрязнения воздуха
ксилолом в
процентах

Частота
атеросклероза

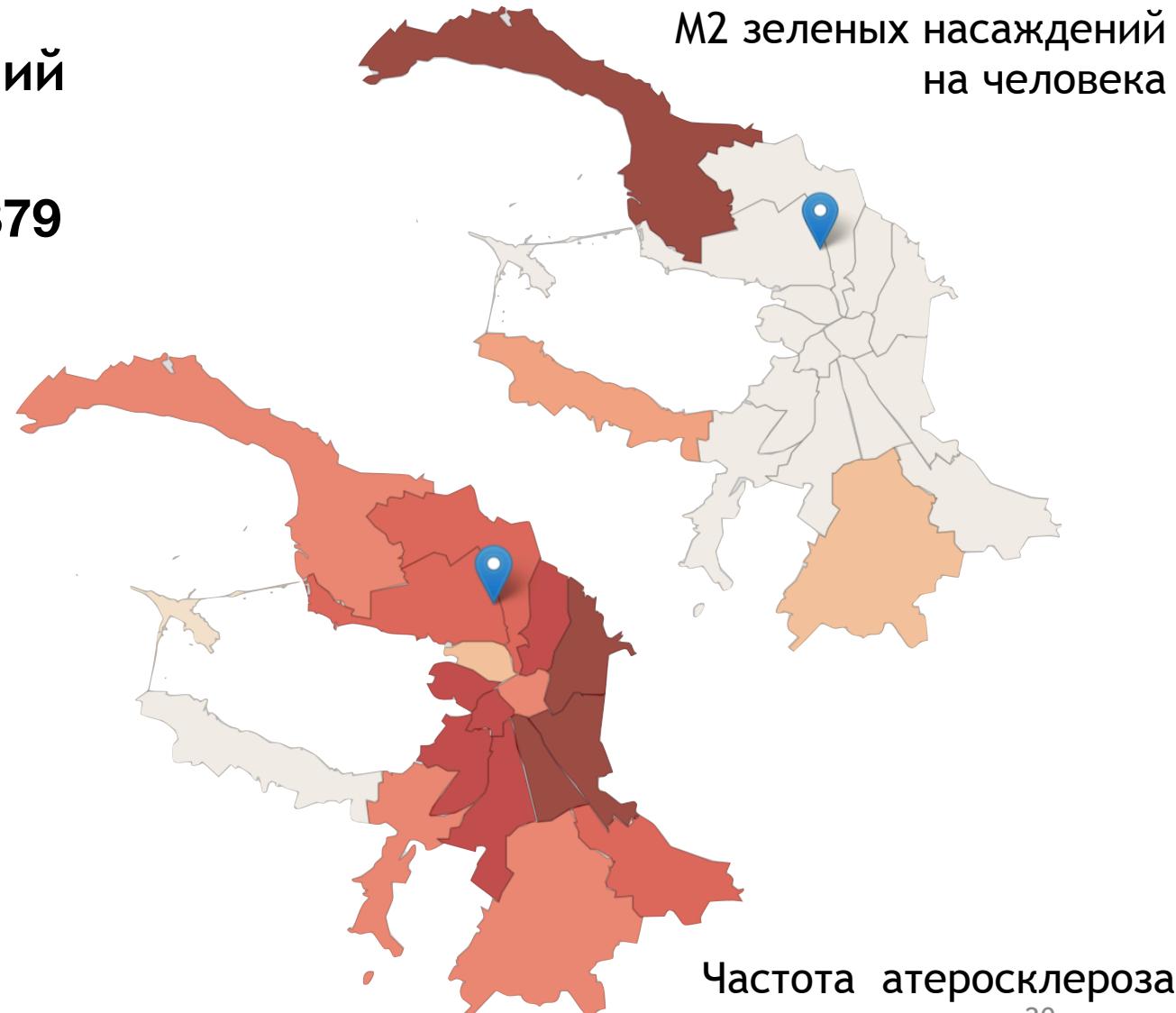
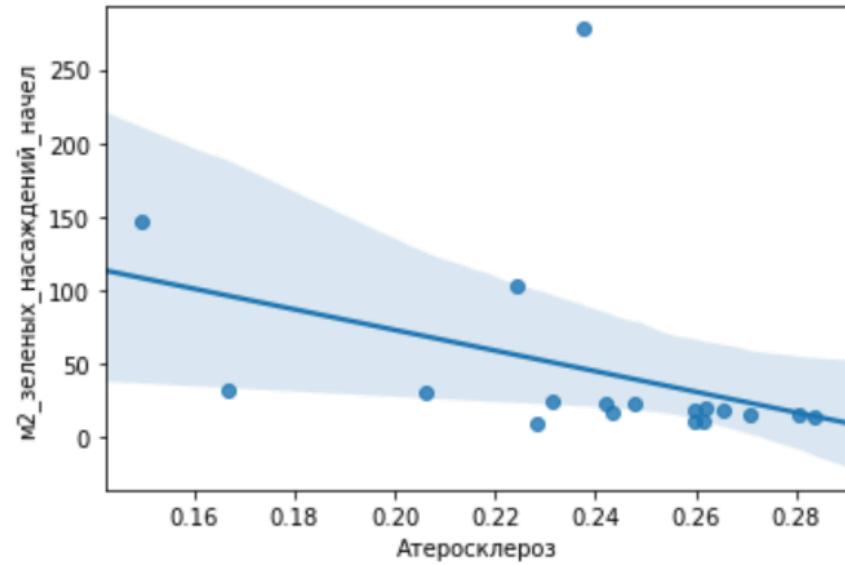
Атеросклероз - площадное загрязнение воздуха в процентах

Коэффициент корреляции Спирамана: **0.690402**



Атеросклероз - м² зеленых насаждений на человека

Коэффициент корреляции Спиромана: **-0.657379**



Атеросклероз - площадь загрязнения воздуха этилбензолов в процентах

Коэффициент корреляции Спирамана: **0.598665**

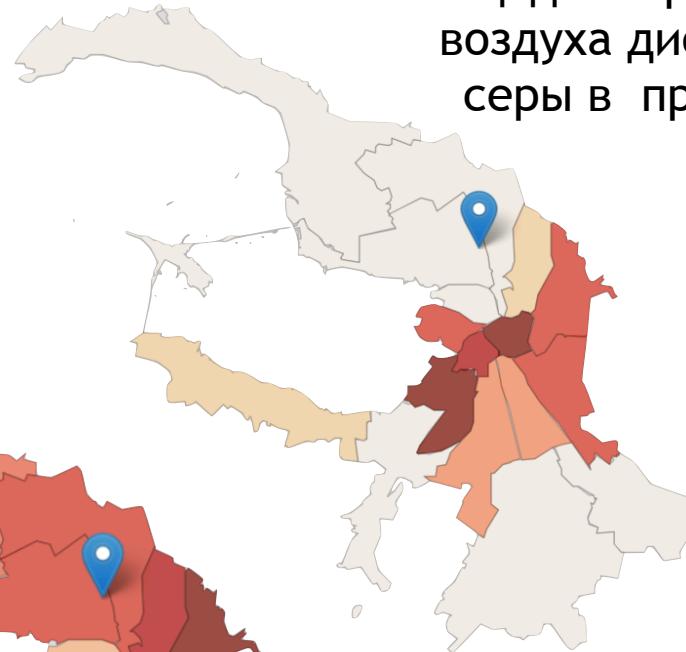
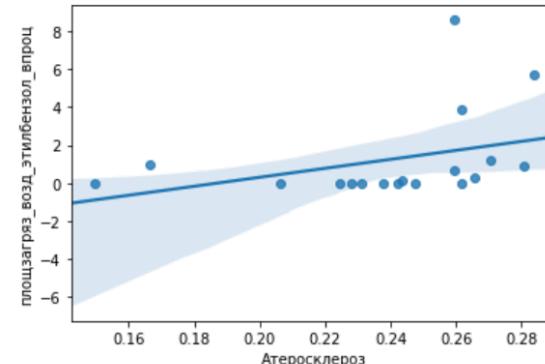
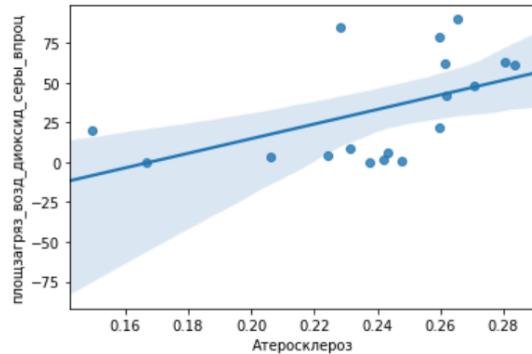
Атеросклероз - площадь загрязнения воздуха диоксидом серы в процентах

Коэффициент корреляции Спирамана: **0.586474**

Атеросклероз - площадь загрязнения воздуха абразивной пылью в процентах

Коэффициент корреляции Спирамана: **0.571776**

+ множество других корреляций с загрязняющими веществами



Площадь загрязнения воздуха диоксидом серы в процентах

Частота атеросклероза

Диабетическая нефропатия - число преступлений на тысячу человек

Коэффициент корреляции Спирамана: **0.752710**

Нейропатия - число преступлений на тысячу человек

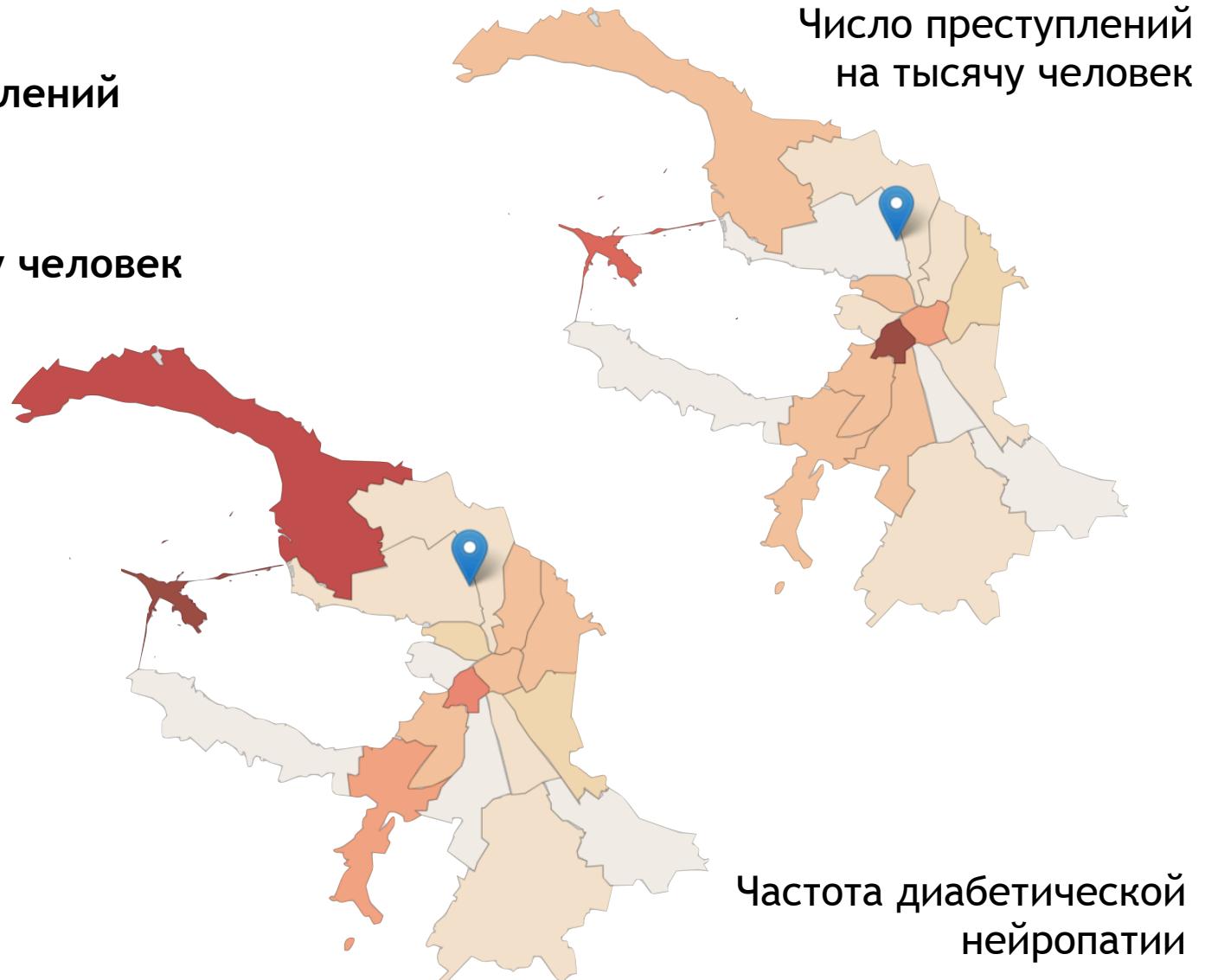
Коэффициент корреляции Спирамана: **0.660816**

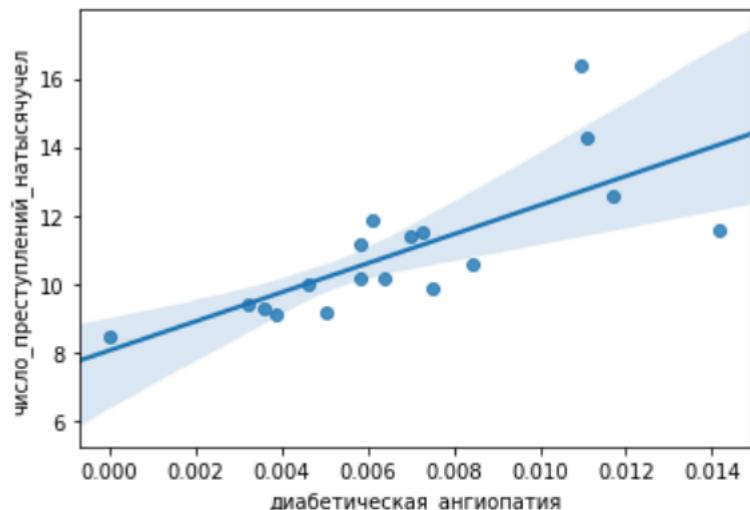
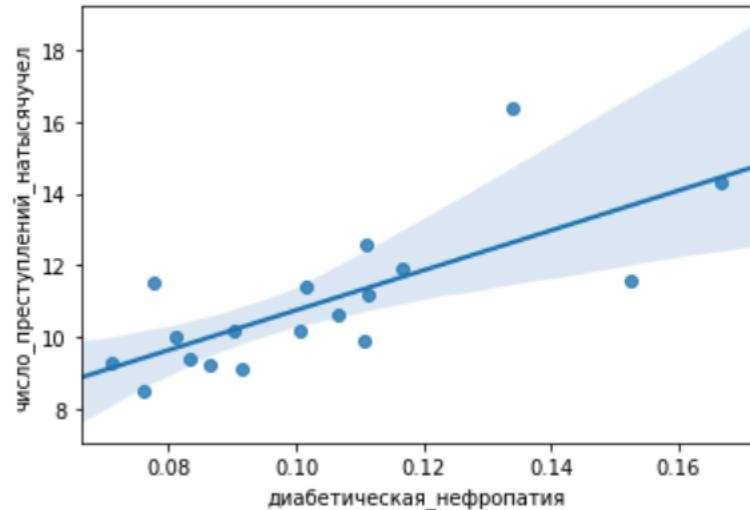
*Диабет -
число преступлений на тысячу человек*

Коэффициент корреляции Спирамана: **0.660816**

*Диабетическая ангиопатия -
число преступлений на тысячу человек*

Коэффициент корреляции Спирамана: **0.818792**





Количество преступлений влияет на частоты диабета и сопутствующие осложнения

- Диабет связан с физической активностью. В районах с большей преступностью люди меньше гуляют (*Fear of walking outdoors: a multilevel ecologic analysis of crime and disorder, Fear of Moving Outdoors and Development of Outdoor Walking Difficulty in Older People*)
- В районах с большей преступностью повышен уровень стресса у жителей, который также является фактором риска диабета (*Neighborhood Disorder, Psychophysiological Distress, and Health*)
- В районах с высоким уровнем преступности выше уровни потребления алкоголя и количество травм, что является причинами нейропатий и факторами риска диабета (*Spatial dynamics of alcohol availability, neighborhood structure and violent crime*)

РЧА - доля района с недостаточным развитием инфраструктуры

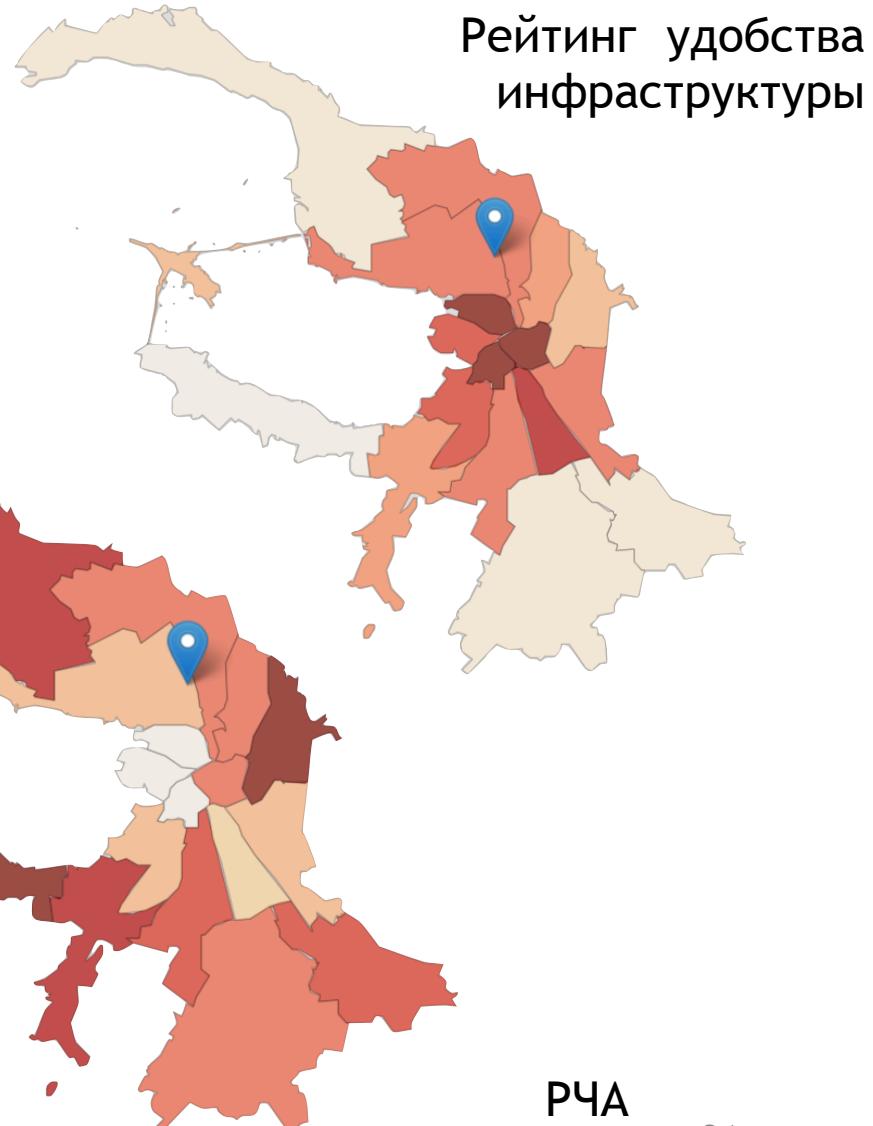
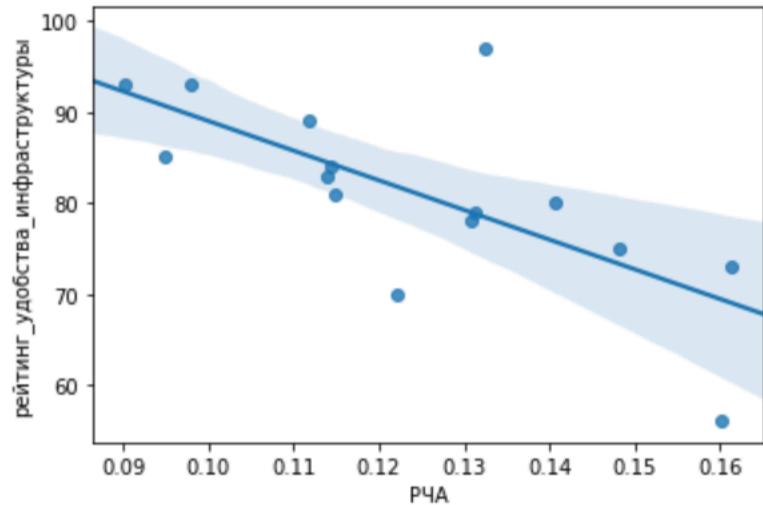
Коэффициент корреляции Спиромана: **0.732697**

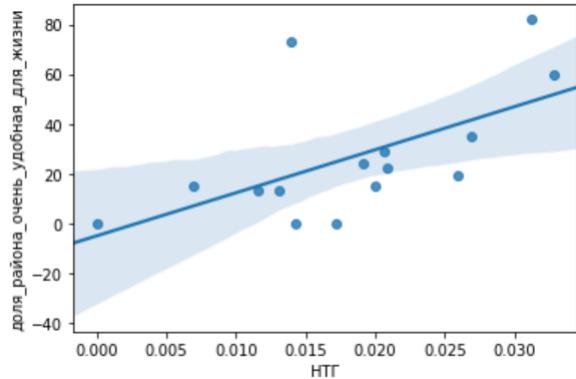
РЧА - рейтинг удобства инфраструктуры

Коэффициент корреляции Спиромана: **-0.693477**

РЧА - расстояние до центра

Коэффициент корреляции Спиромана: **0.683824**



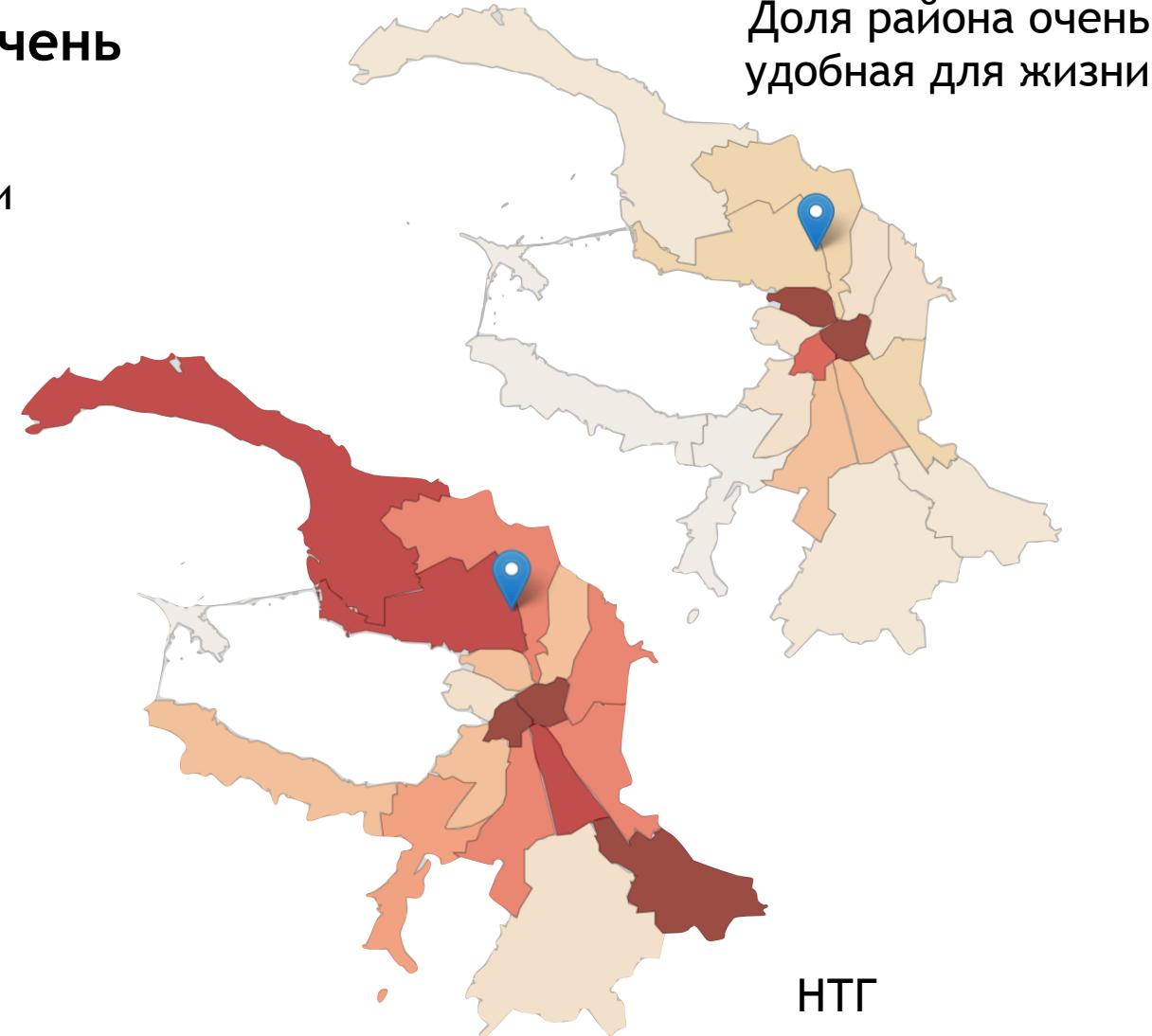


НТГ - доля района очень удобная для жизни

Коэффициент корреляции Спиромана:

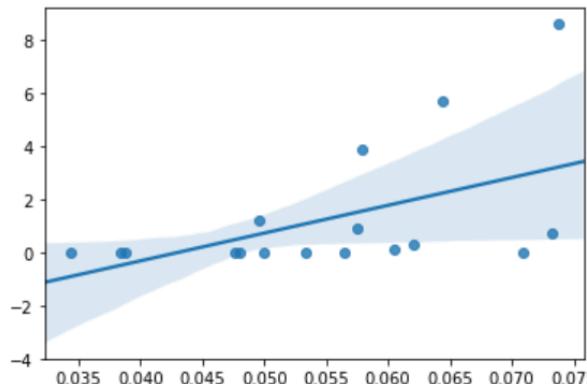
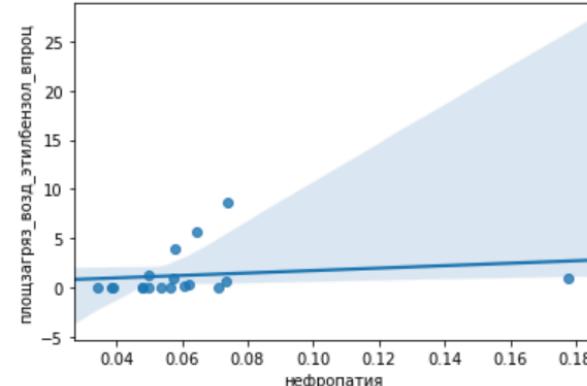
0.671464

- При частом питании вне дома изменяется диета (Eating out of home and its association with dietary intake: a systematic review of the evidence.) и питание вне дома связано с риском возникновения диабета (Consumption of Meals Prepared at Home and Risk of Type 2 Diabetes: An Analysis of Two Prospective Cohort Studies)
- Потребление алкоголя также связывают с риском развития диабета и преддиабета (Alcohol consumption and risk of pre-diabetes and type 2 diabetes development in a Swedish population)

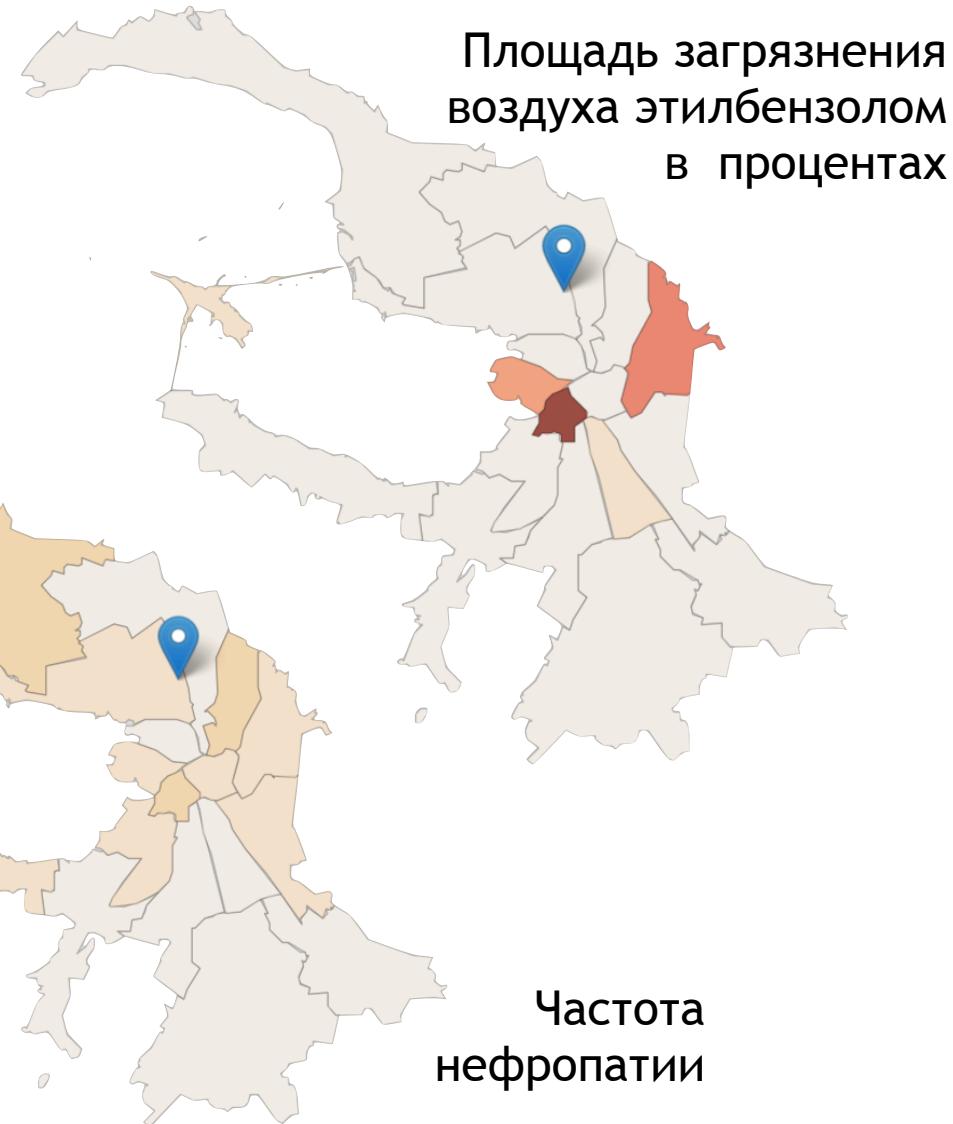


нефропатия - площадь загрязнения воздуха этилбензолом в процентах

Коэффициент корреляции Спирмана: **0.644971**



Без Кронштадта коэффициент корреляции Пирсона - 0.5019, коэффициент корреляции Спирмана = 0.6369



Площадь загрязнения воздуха этилбензолом в процентах

Частота нефропатии

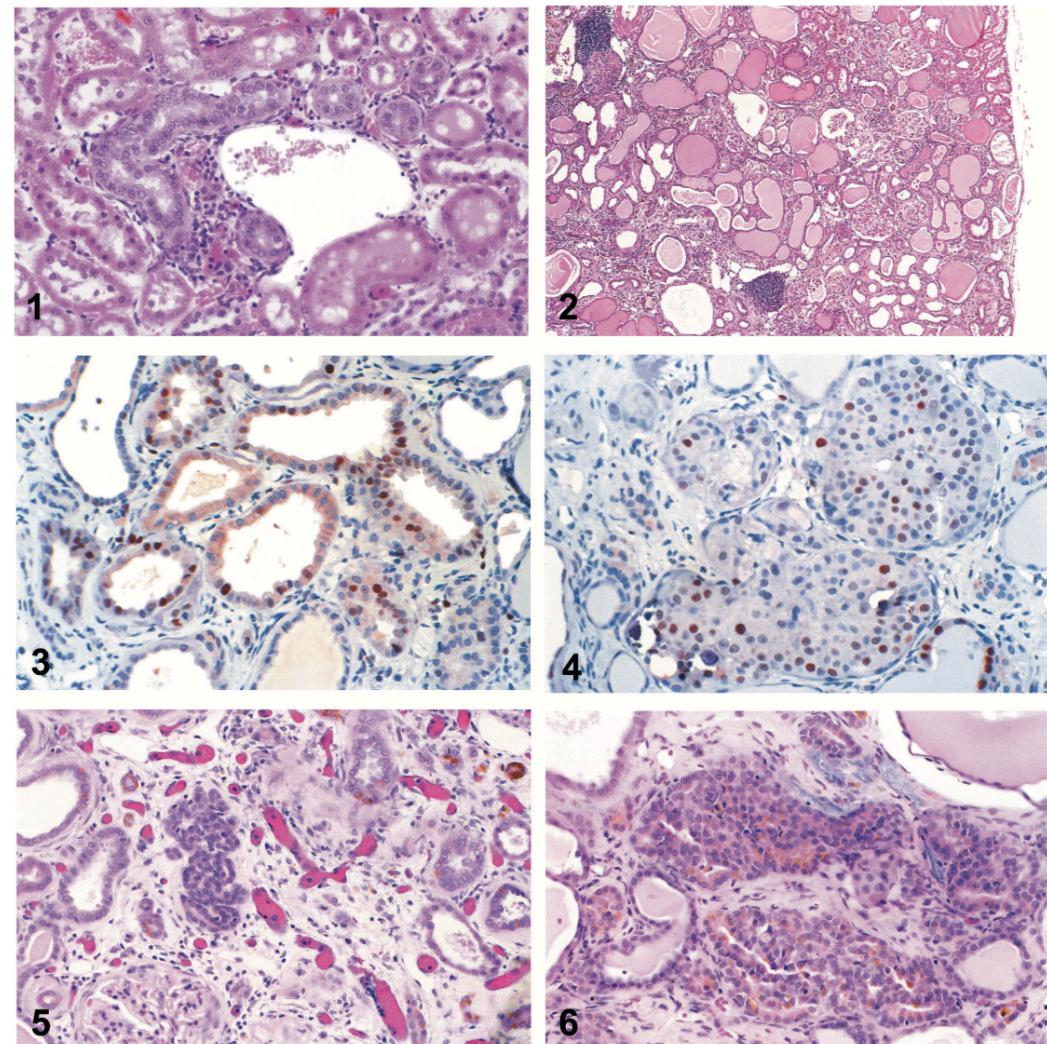
Этилбензол связывают с Хронической прогрессирующей нефропатией и почечным канцерогенезом у крыс. Это заболевание видоспецифично и не имеет аналога у человека, поэтому считается, что для человека такая ассоциация не состоятельна, хотя напрямую этого никто не смотрел.

Table 3. Summary of key neoplastic and non-neoplastic effects of ethylbenzene in rats and mice (NTP 1999, Chan et al. 1998).

Species	Sex	Concentration (ppm)	Kidney		Testes		Lung		Liver	
			Chronic progressive nephropathy (mean severity score)	Adenoma/carcinoma*	Hyperplasia	Adenoma	Hyperplasia/Metaplasia	Alveolar/bronchiolar adenoma/carcinoma	Eosinophilic Foci	Adenoma/carcinoma
Rat	M	0	47/50 (2.3)	3/50	14/50	36/50	2/50	3/50	5/50	0/50
		75	43/50 (2.4)	5/50	19/50	33/50	2/50	1/50	11/50	3/50
		250	47/50 (2.3)	8/50	12/50	40/50	1/50	0/50	4/50	0/50
		750	48/50 (3.5)	21/50	8/50	44/50	2/50	1/50	9/50	0/50
	F	0	38/50 (1.3)	0/50	NA		1/50	1/50	2/50	0/50
		75	42/50 (1.6)	0/50	NA		5/50	1/50	3/50	0/50
		250	43/50 (1.7)	1/50	NA		2/50	1/50	8/50	0/50
		750	46/49 (2.3)	8/49	NA		5/49	0/49	5/49	0/49
Mouse	M	0	34/50	0/50	0/50	1/49	1/50	7/50	6/50	29/50
		75	38/50	0/50	1/50	0/50	6/50	10/50	8/50	25/50
		250	40/50	1/50	0/50	1/50	4/50	15/50	8/50	30/50
		750	36/50	0/50	0/50	1/50	10/50	19/50	12/50	28/50
	F	0	13/50	0/50	NA		0/50	4/50	5/50	13/50
		75	7/50	0/50	NA		1/50	6/50	7/50	12/50
		250	9/50	0/50	NA		3/50	5/49	6/50	15/50
		750	21/50	0/50	NA		1/50	8/50	22/50	25/50

*Standard and extended evaluations combined; bolded tumor incidence values indicated a statistically significant increase over untreated control animals.

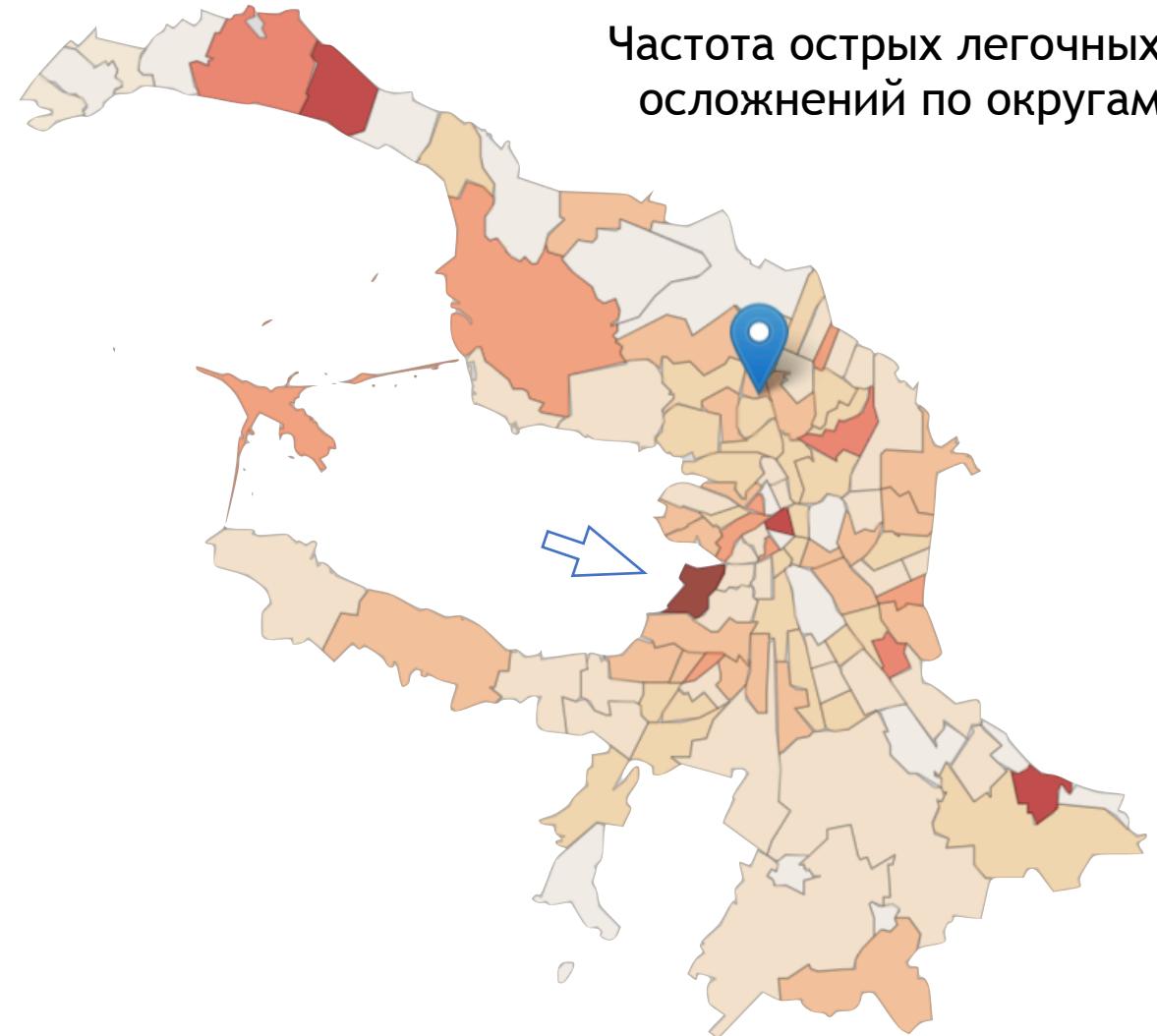
NA not applicable.



Самая высокая частота острых легочных осложнений в округе Морские ворота, на территории которого располагается Кировский завод



Частота острых легочных осложнений по округам



Выводы:

- 1) Если рассчитывать частоты заболеваний среди людей, обратившихся за высокотехнологичной медицинской помощью в центр Алмазова, удается избавиться от влияния близости центра Алмазова на распределение частот
- 2) Фактор округа практически не влияет на распределение частот, поэтому было принято решение использовать разбиение на районы и смотреть только на заболевания, для которых фактор района играет роль
- 3) Выявленные взаимосвязи суммированы в таблице на следующем слайде
- 4) Обнаружено несколько интересных фактов, которые требуют более детального изучения (связь нефропатии и загрязнения воздуха этилбензолом, связь НТГ и инфраструктуры района, влияние Кировского завода на острые легочные осложнения)

	correlation_coefficient	p_value
Сердечные операции - число_городских_больниц	-0.693523	0.001414
Легочные заболевания - рейтинг_удобства_инфраструктуры	-0.670242	0.006253
Сердечно-сосудистые заболевания - площзагряз_возд_пыль_древесная_впроц	0.657718	0.004109
Сердечно-сосудистые заболевания - конц_ввозд_пыль_неорг	0.656925	0.003058
Заболевания почек - среднее_значение_радиации	0.608589	0.009528
Эндокринные заболевания - число_преступлений_натысячучел	0.605059	0.007803
Сердечные операции - число_объектов_ос_сброс_всточныеводы	0.568270	0.017313
Сердечно-сосудистые заболевания - число_объектов_ос_сброс_всточныеводы	0.563349	0.018530
Заболевания почек - площзагряз_возд_этилбензол_впроц	0.552360	0.017453
Сердечно-сосудистые заболевания - конц_ввозд_ксилол	0.548596	0.018398
Сердечные операции - конц_ввозд_пыль_древесная	0.541277	0.020350
Заболевания почек - доля_района_очень_удобная_для_жизни	-0.524245	0.044847
Эндокринные заболевания - коэффициенты_нагрузок_натрудоспособное	0.494324	0.037044

		correlation_coefficient	p_value
	Атеросклероз - площзагряз_возд_диоксид_азота_впроц	0.790269	0.000095
	Атеросклероз - конц_ввозд_ксилол	0.784813	0.000115
	диабетическая_нефропатия - число_преступлений_натысячучел	0.752710	0.000312
	РЧА - доля_района_с_недостаточным_развитием_инфраструктуры	0.732697	0.001890
	Атеросклероз - конц_ввозд_диоксид_азота	0.717768	0.000796
	Атеросклероз - площадное_загрязнение_возд_впроц	0.690402	0.001516
	НТГ - доля_района_очень_удобная_для_жизни	0.671464	0.006124
	нейропатия - число_преступлений_натысячучел	0.660816	0.002830
	Атеросклероз - м2_зеленых_насаждений_начел	-0.657379	0.003030
	нефропатия - площзагряз_возд_этилбензол_впроц	0.644971	0.003851
	РЧА - число_объектов_ос_сброс_всточныеводы	0.644531	0.005222
	диабетическая_нефропатия - площзагряз_возд_пыль_древесная_впроц	-0.628046	0.006942
	инфаркт - производственный_травматизм_натыс	-0.627602	0.005297
	инфаркт - конц_ввозд_пыль_абразивная	0.614101	0.006703
	Атеросклероз - конц_ввозд_этилбензол	0.606740	0.007588
	диабетическая_нефропатия - конц_ввозд_пыль_древесная	-0.600790	0.008371
	Атеросклероз - площзагряз_возд_диоксид_серы_впроц	0.586474	0.010522
	РЧА - число_медучереждений_натыс_чел	-0.583075	0.011092
	Атеросклероз - конц_ввозд_пыль_абразивная	0.581025	0.011448
	РЧА - опасная_почва_впроц	-0.578947	0.011818
	нефропатия - среднее_значение_радиации	0.574234	0.015922