

Исследование объявлений о продаже квартир в Санкт- Петербурге

Суспицына Елена
Yandex.Practicum 2023

Описание и задачи проекта

По данным **сервиса Яндекс.Недвижимость** (архив объявлений о продаже квартир в Санкт-Петербурге и соседних населенных пунктов за несколько лет) - надо научиться определять рыночную стоимость объектов недвижимости.

Наша задача — установить параметры. Это позволит построить автоматизированную систему: она отследит аномалии и мошенническую деятельность.

total_images	0.000000
last_price	0.000000
total_area	0.000000
first_day_exposition	0.000000
rooms	0.000000
ceiling_height	38.800000
floors_total	0.400000
living_area	8.000000
floor	0.000000
is_apartment	88.300000
studio	0.000000
open_plan	0.000000
kitchen_area	9.600000
balcony	48.600000
locality_name	0.200000
airports_nearest	23.400000
cityCenters_nearest	23.300000
parks_around3000	23.300000
parks_nearest	65.900000
ponds_around3000	23.300000
ponds_nearest	61.600000
days_exposition	13.400000

% процент пропусков данных

очень много пропусков(более 60%) - в столбце близость от парка, parks_nearest, близость от водоема ponds_nearest и апартаменты (не важно), is_apartment

много пропусков (от 20 до 59%)- близость до аэропорта, airports_nearest , близость до центра города, cityCenters_nearest, высота потолков, ceiling_height, кол-во балконов (не важно) balcony, кол-во парков в радиусе 3 км parks_around3000(не важно) и водоемов в радиусе 3 км ponds_around3000 (не важно)

мало пропусков, но больше статистической погрешности 3% (3-19%) - количество дней размещения объявления,days_exposition, площадь кухни, kitchen_area, жилая площадь living_area

Странности данных

минимальное кол-во комнат - ноль,

- минимальная стоимость квартиры 12190 рублей,

- минимальная высота потолков 1 метр,

- максимальная высота потолков 100 метров

- минимальная жилая площадь - 2 метра

- минимальная площадь кухни 1,3 метра

- расстояние до аэропорта 0 метров -

- минимальное расстояние до парка 1 метр

- минимальное расстояние до водоема 13 метров - набережные!

- максимальное количество дней размещения объявления 1580 - это 4 года и 4 месяца,

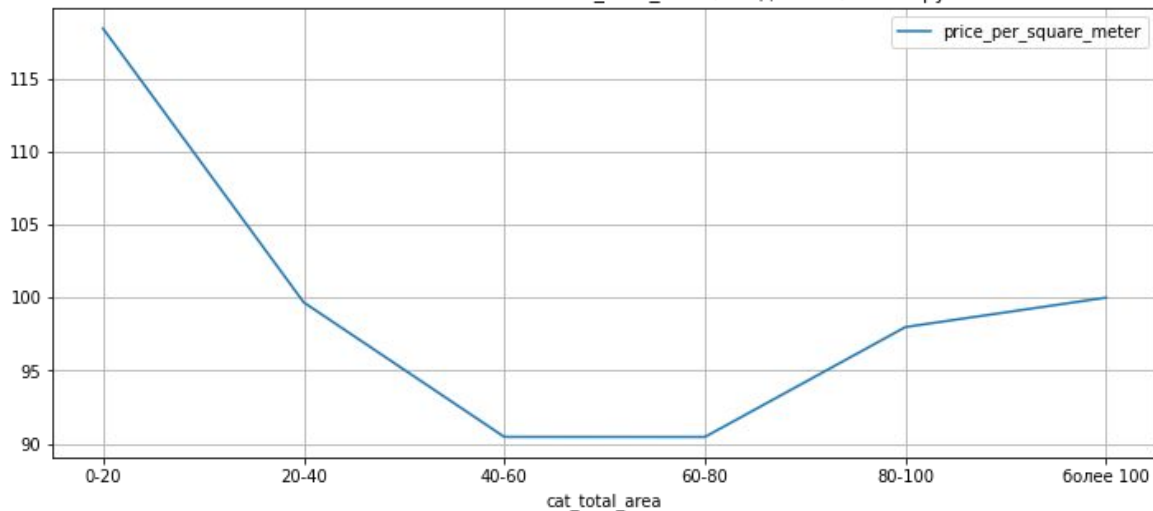
Преобразования (предобработка)

- 1) Замена названий населенных пунктов на уникальные (365-313 было-стало)
- 2) Минимальная жилая 6 кв метров, кухня 5 кв метров (согласно законодат.)
- 3) замена пропусков площ. кухни - медианными значениями в соотв. с кол-вом комнат.
- 4) замена пропусков общей этажности этажом квартиры
- 5) удаление строк с аномальными площадями кухонь
- 6) заполнение пропусков высоты потолка средним значением в СПб по СПб, остальные города средним без учета СПб
- 7) замена пропусков расстояний до центра медианным по СПб значением в СПб и 3000 м в городах кроме СПб
- 8) заполняем пропуски в длительности размещения объявления разницей даты подачи и последним днем подачи объявления в датасете
- 9) убираем редкие и выбивающиеся значения

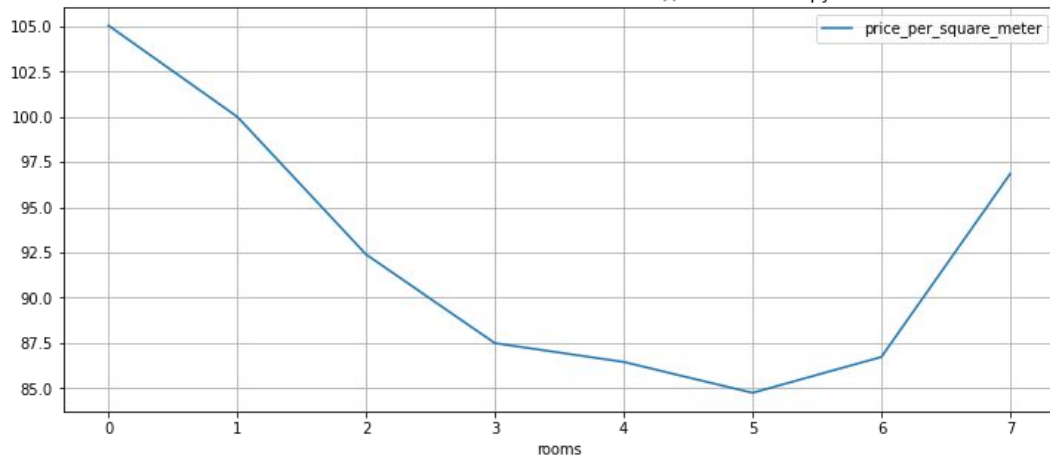
После преобразований мы потеряли в общей сложности менее 7% данных

Какие факторы больше всего влияют на стоимость квартиры?

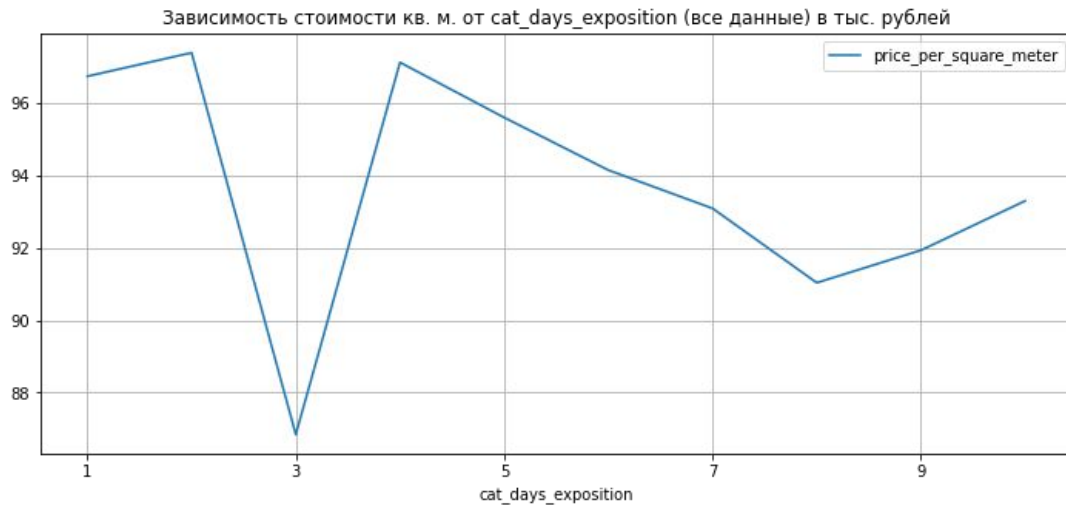
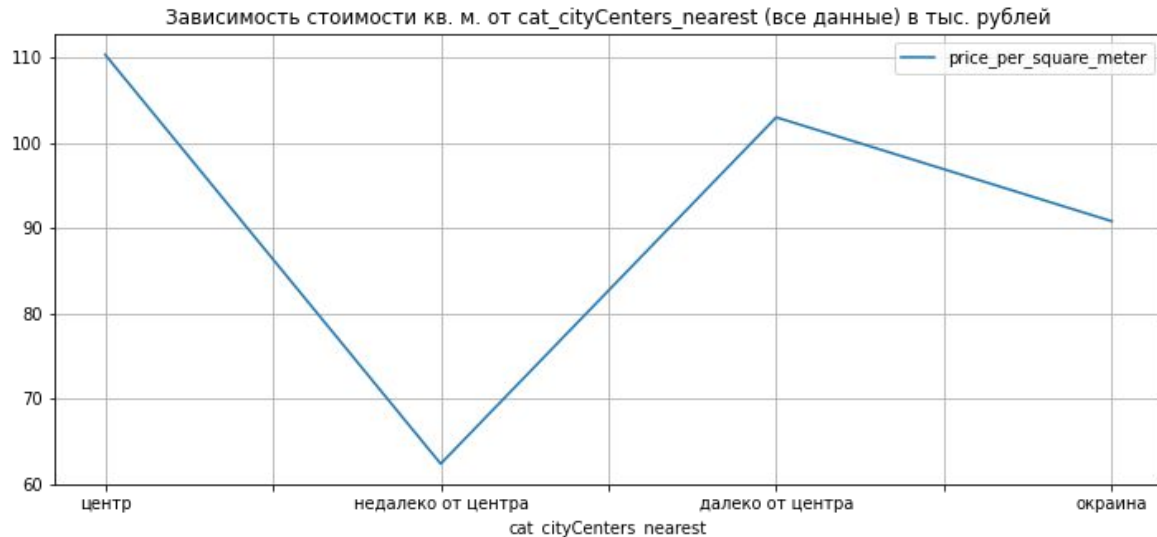
Зависимость стоимости кв. м. от cat_total_area (все данные) в тыс. рублей



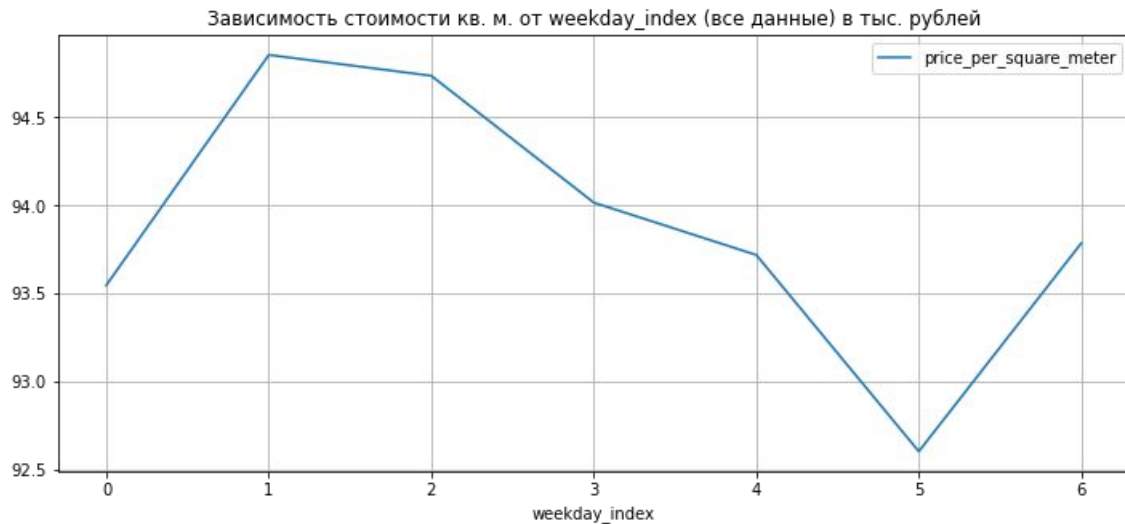
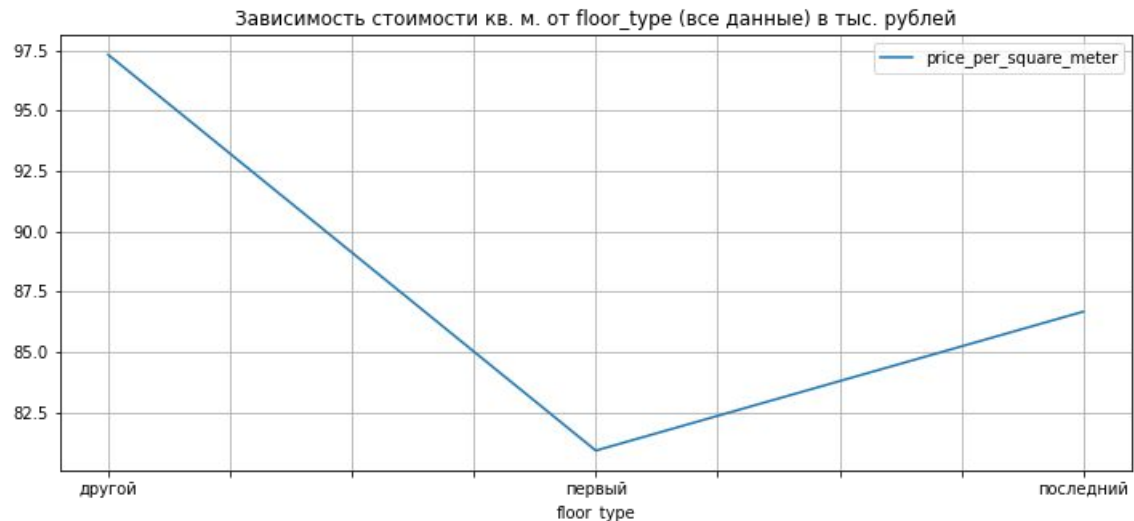
Зависимость стоимости кв. м. от rooms (все данные) в тыс. рублей



Какие факторы больше всего влияют на стоимость, в квартиры?

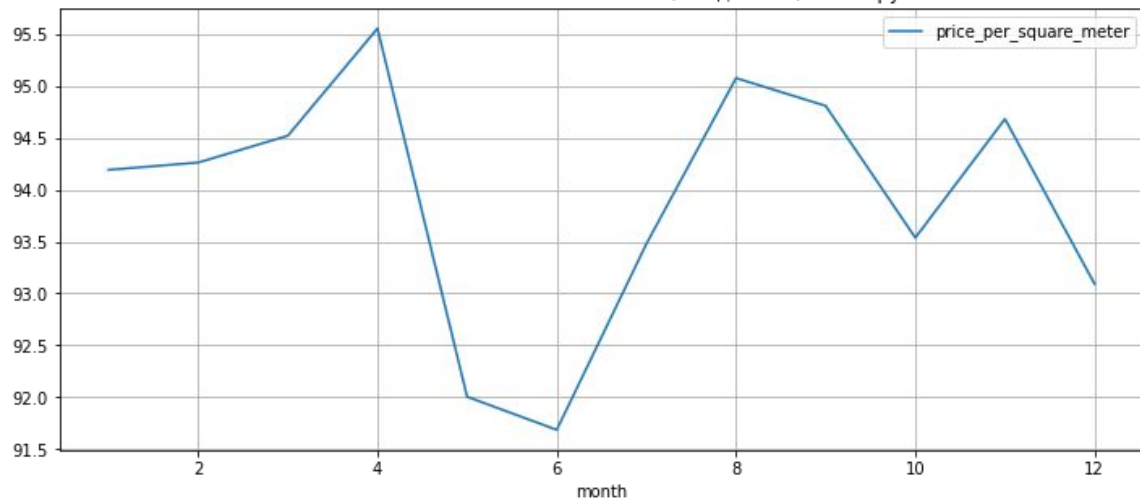


Какие факторы больше всего влияют на стоимость квартиры?

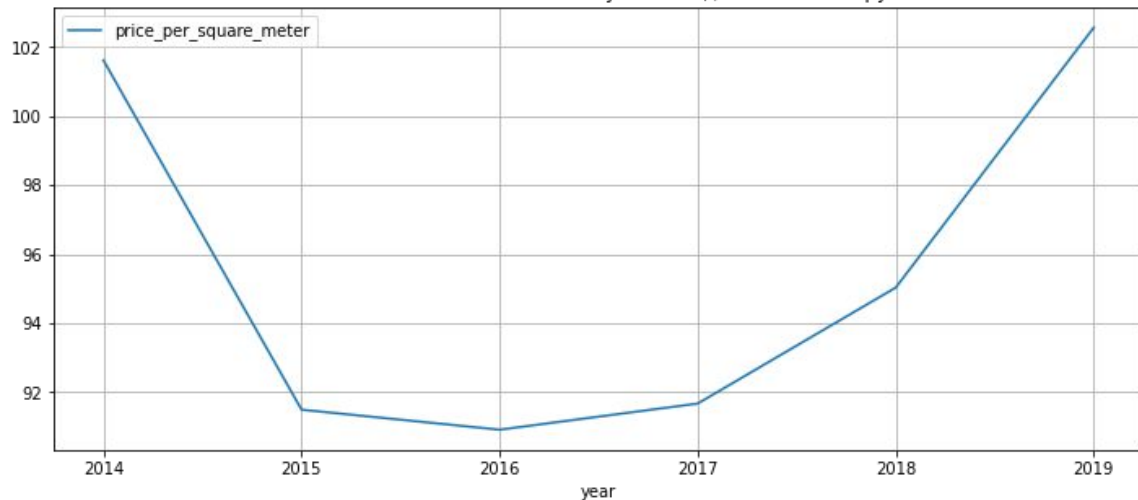


Какие факторы больше всего влияют на стоимость, в квартиры?

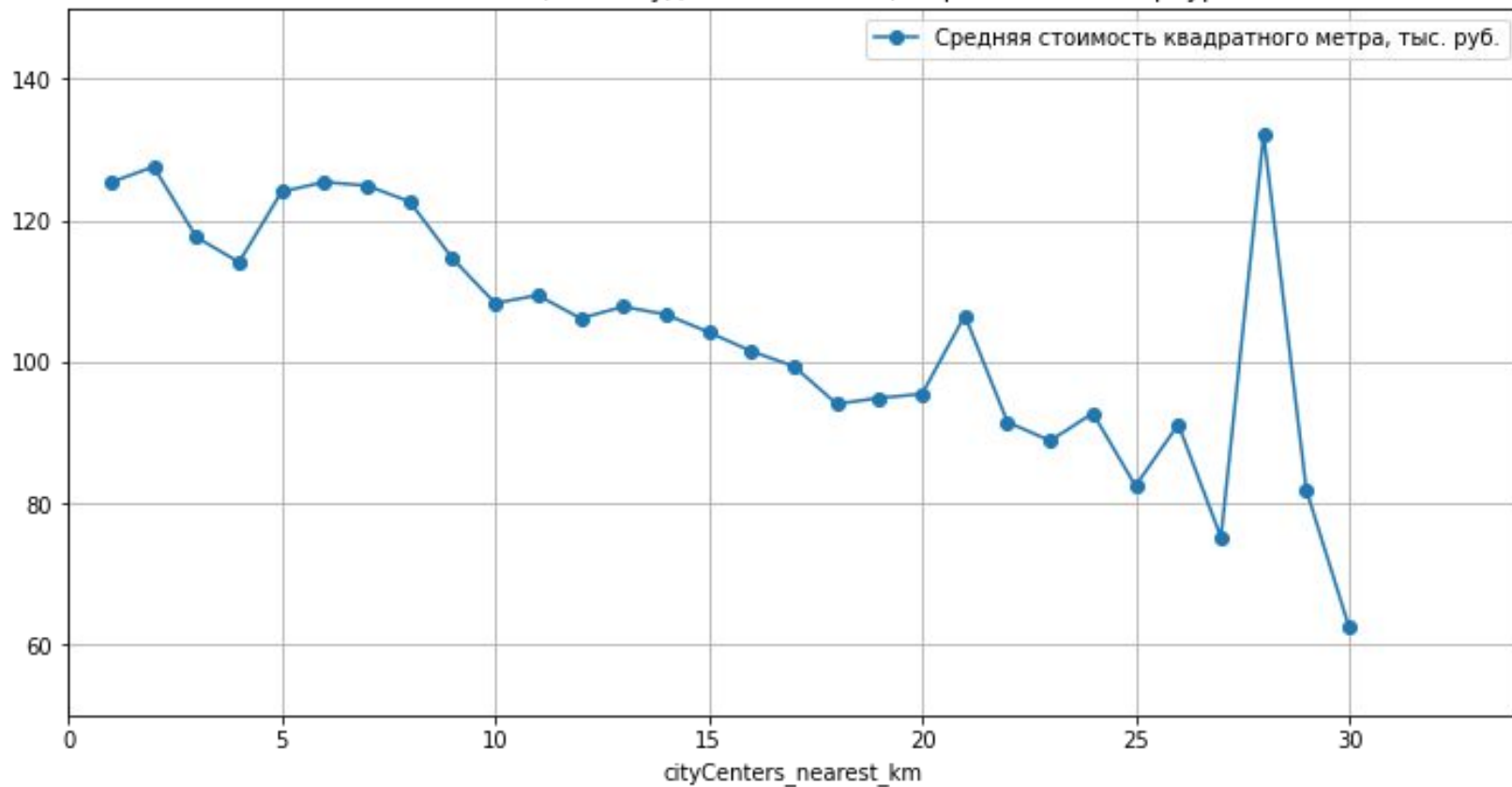
Зависимость стоимости кв. м. от month (все данные) в тыс. рублей



Зависимость стоимости кв. м. от year (все данные) в тыс. рублей

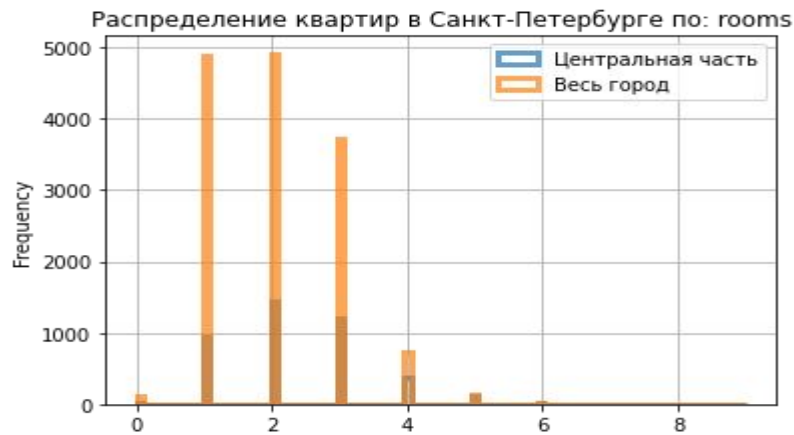
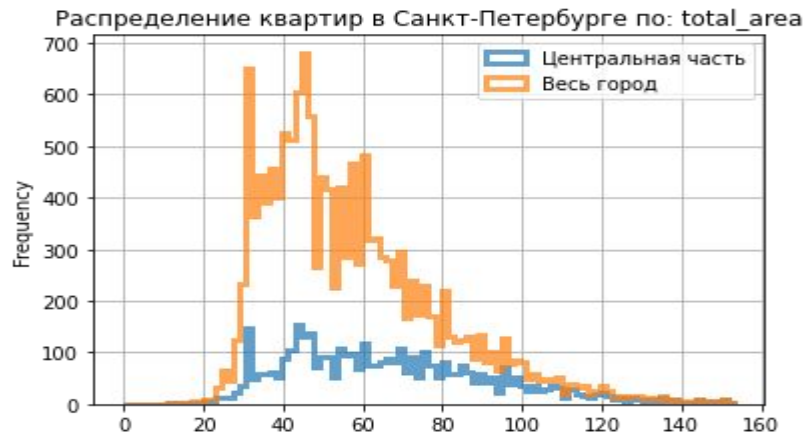


Зависимость цены от удалённости от центра (Санкт-Петербург)



каждый километр удаления от центра СПб снижает стоимость кв метра на 2000р

Сравнение распределения квартир в Санкт-Петербурге и его центре



ТОП 10 населённых пунктов

	count	price_per_square_meter
санкт-петербург	14524	103.145438
пушкин	349	99.257862
деревня кудрово	297	92.000000
поселок парголово	326	91.780224
поселок мурино	583	86.111111
поселок шушары	437	76.923077
колпино	337	74.722222
гатчина	303	67.796610
всеволожск	392	65.737874
выборг	232	58.133214

Самая высокая стоимость жилья: санкт-петербург
Самая низкая стоимость жилья: выборг

0.8 - максимальный вклад
в стоимость квадратного
метра вносит общая
площадь квартиры
'total_area';

0.6 - Далее идет - жилая
площадь 'living_area';

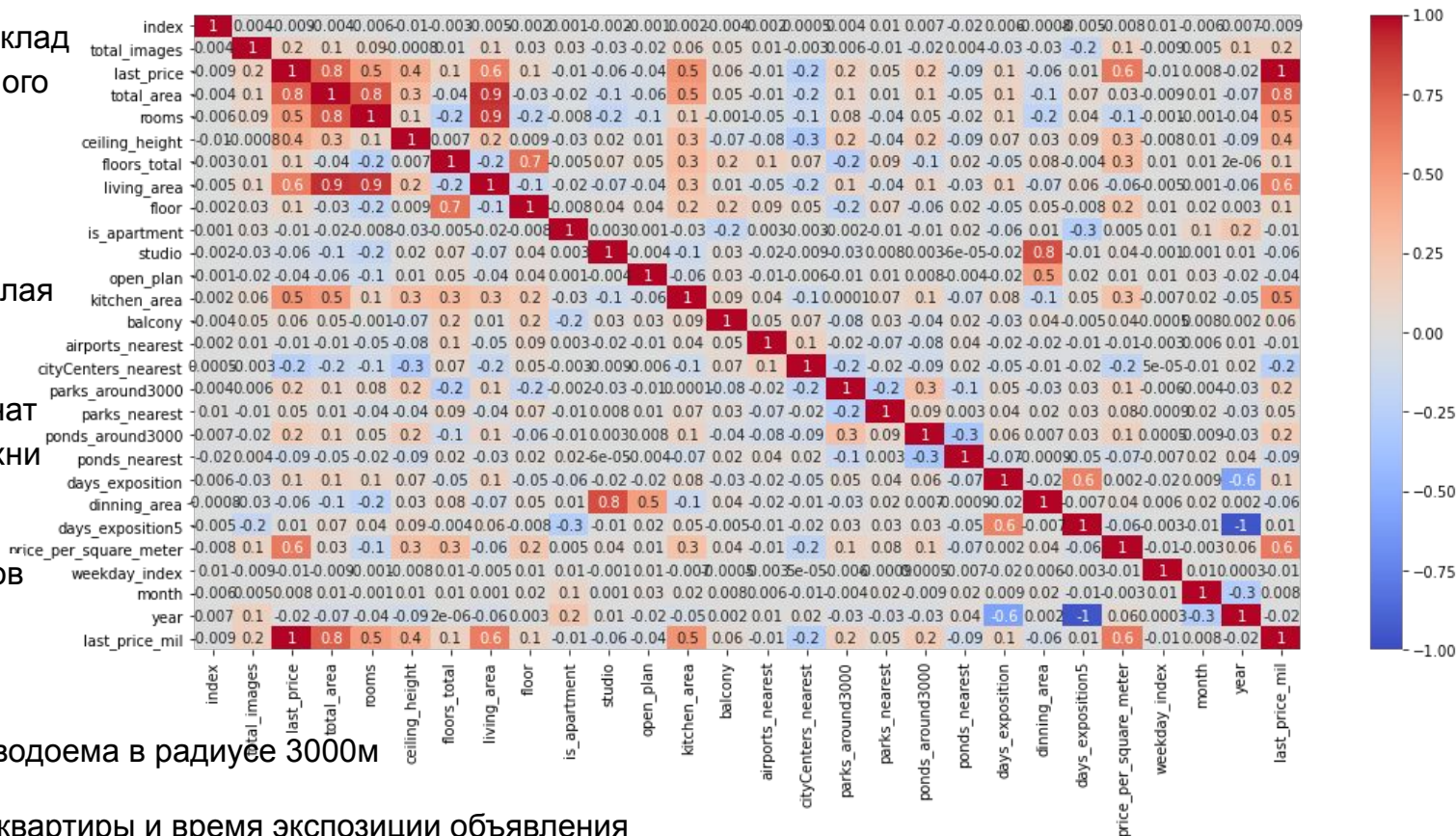
0.5 - Количество комнат
'rooms' и площадь кухни
'kitchen_area'

0.4 - Высота потолков
'ceiling_height'

0.2 - наличие парка и водоема в радиусе 3000м

0.1 - этаж дома, этаж квартиры и время экспозиции объявления

-0.2 - обратная корреляция с расстоянием до центра



Вывод для автоматизации параметров

- 1) ввести понятие квартир в формате "евро"
- 2) пояснить отличие студии(комната с окнами на одной стене), свободной планировки (пространство с определенными зонами, но без установленных перегородок, окна могут быть на нескольких стенах) и апартаментов(отсутствие возможности регистрации, возможность использования под офис)
- 3) ввести понятие "кухня-столовая" и разрешить указывать ее размер более чем 26-30 кв метров кухни
- 4) расстояние до центра города ограничить максимальным расстоянием до окраины Санкт-Петербурга, чтобы избежать лукавства при выставлении квартир в городах сателитах СПб как жилье в самом СПб. Также можно ввести поле "расстояние до центра СПб".
- 5) ограничить минимальный размер кухни в студиях 5 кв метра (добавить ссылку на законодательство)

Вывод для автоматизации параметров

- 6) ограничить максимальный размер кухни 26-30 кв метрами, исключение сделать только для апартаментов и формата евро с кухнями-столовыми.
- 7) ограничить минимум жилой площади 6 кв метрами (добавить ссылку на законодательство)
- 8) ограничить минимальную стоимость квадратного метра минус 20% от средней по населенному пункту с целью пресечения фактов мошенничества (или отправлять такие объявления на модерацию)
- 9) принудительно проставлять 0 балконов в объявлении, если человек не указал
- 10) ограничить высоту потолков от 2-х до 6 метров
- 11) Добавить сообщения владельцу объявления о среднем, медианном, максимальном и минимальном времени продажи недвижимости. А также предупреждение покупателям о снижении цены через 6 месяцев от даты подачи объявления.

Вывод для автоматизации параметров

12) в "мертвые месяцы" продажи выдавать предупреждение о снижении покупательской активности и предложить снизить цену владельцам объявлений

13) ограничить количество этажей в доме 55-ю, можно выдать сообщение о Лахта-Центре 87-этажном здании.

14) рассчитывать автоматически стоимость квадратного метра = максимальная цена в центре Санкт-Петербурга - $2000 \cdot \text{расстояние от центра в км}$.

15) установить предупреждение о том, что если не указать этажность дома, то система автоматически определит этаж квартиры как максимальный, а это на 10% снизит стоимость квартиры.