

Рынок новостроек в Подмосковье

Внтар.pro

Random coders =>

Структура данных

Датасет об объектах
недвижимости

21273 строк

419 объектов

23 девелопера

Датасет о сделках с объектами
недвижимости

50955 строк

7 лет

420 объектов

23 девелопера

Данные агрегированы по времени покупки/выставления на продажу и характеристикам объектов.

Структура данных

Удаляем в обоих датасетах, строки с мусором в столбцах, а именно с данными из других столбцов и не правильными типами данных

Удаляем в датасете о сделках строки у которых средняя площадь/сумма одного лота выходят за $1.5 * \text{IQR}$

Датасет с проектными данными

было	стало	удаленно
21273	20774	499

Датасет со сделками

было	стало	удаленно
50955	49407	1548

Ввог новой метрики #1

ННІ (индекс Херфиндаля-Хиршмана) — показатель рыночной концентрации по округам.

Совершенная
конкуренция

0

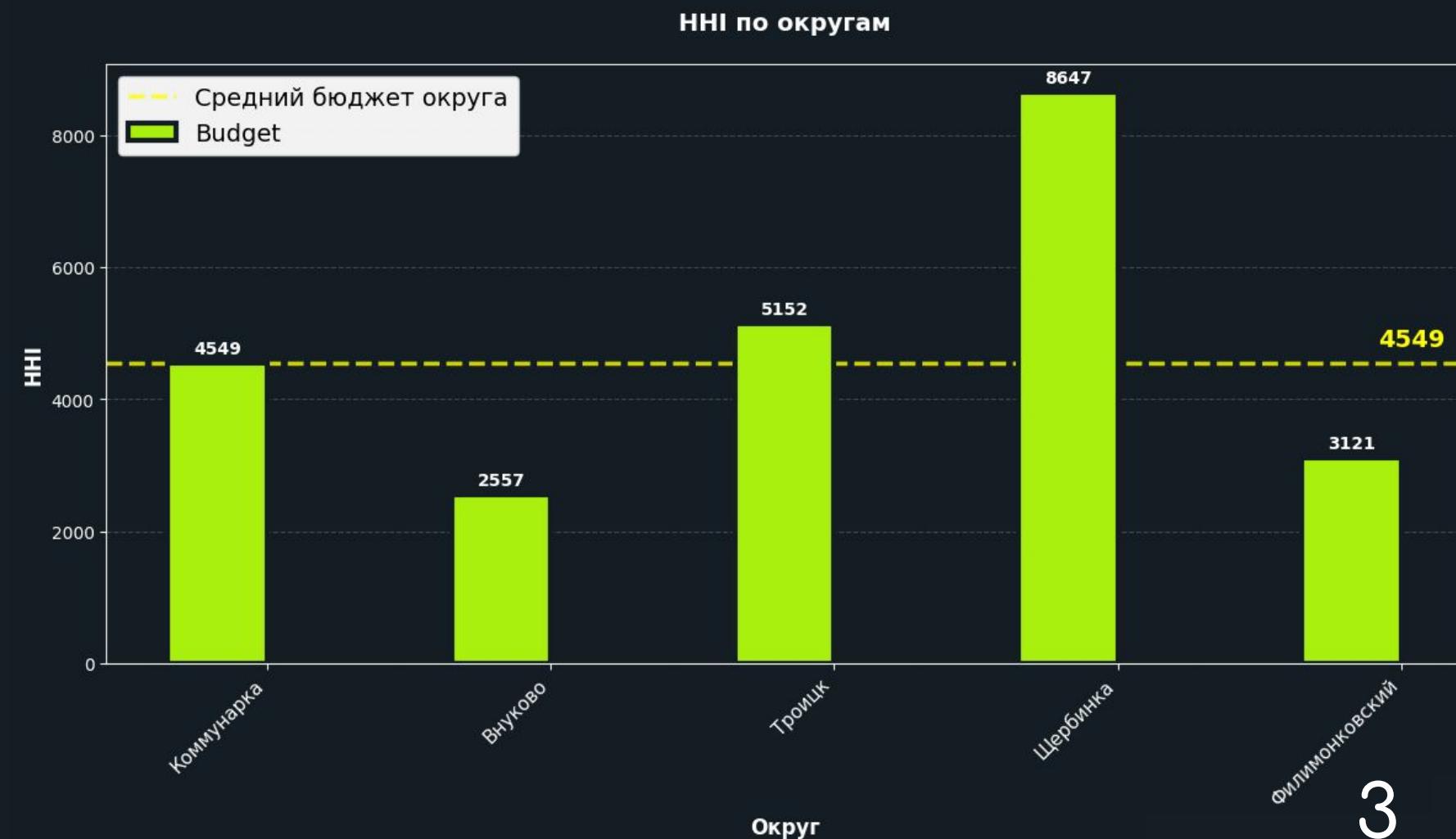
Монополия

10000

$$HHI = S_1^2 + S_2^2 + \dots + S_n^2$$

* S_i -доля рынка i -ой компании

У округа Щербинка гораздо выше монополизация по сравнению с другими округами



Ввод новой метрики #2

Цена за квадратный метр = суммарная площадь сделок / бюджет сделок

Среднее
244
тыс.

Медиана
240 тыс.

Мин.
122 тыс.

Макс.
367 тыс.



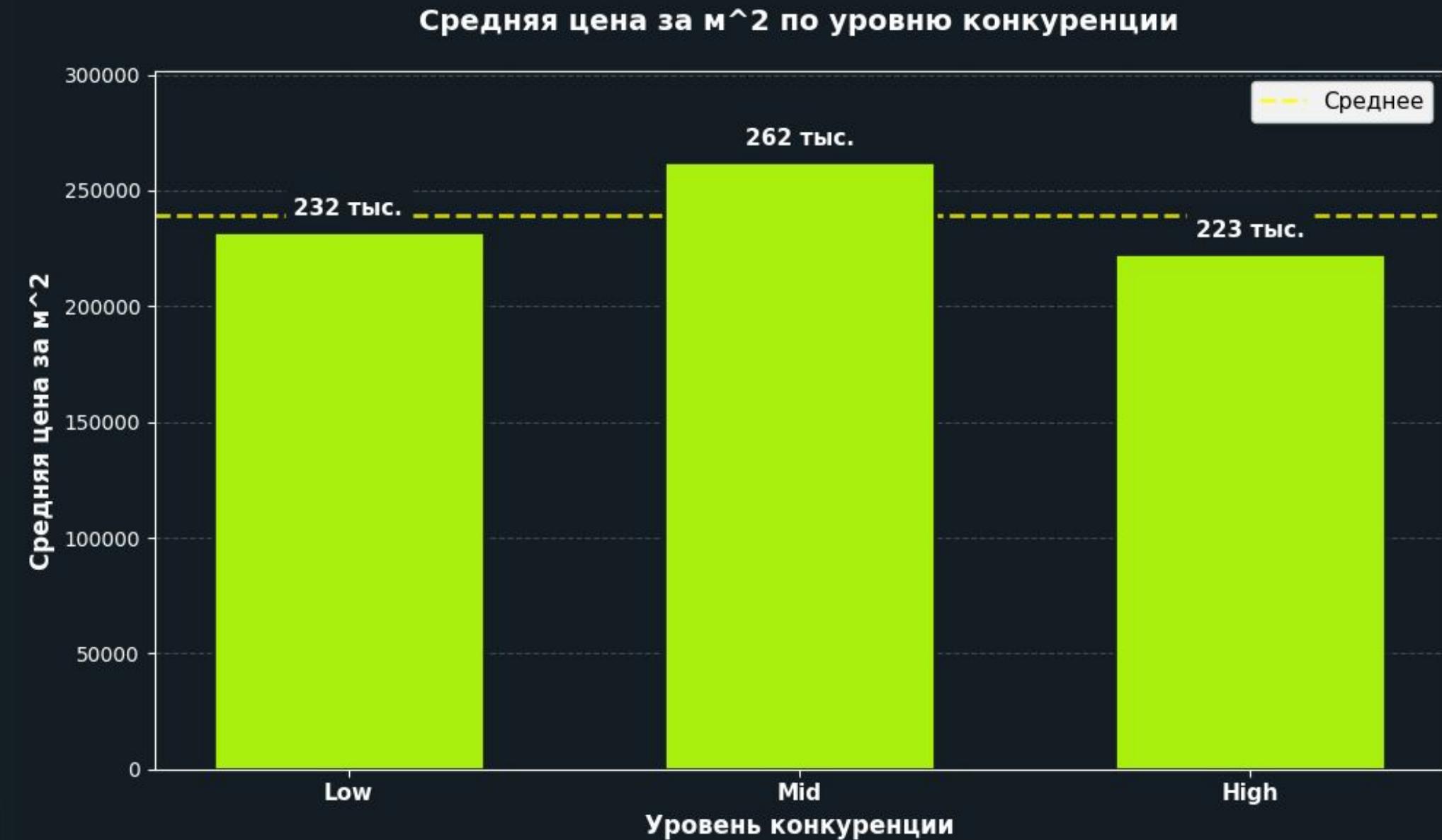
Ввог новой метрики #3

Уровень конкурентности округа - от низкого до высокого по тритилям ННІ

Кол-во корпусов в сегментах:

Low: 37
Mid: 267
High: 107

В округе со средним уровнем конкуренции цена за м² выше, чем у остальных...



Гипотеза и исследовательский вопрос

Исследовательский вопрос: Как уровень конкуренции влияет на цену за м²?

Гипотеза: В районах со средним уровнем конкуренции цена за м² в среднем выше, чем в районах с низким или высоким.

Механизм работы гипотезы

Конкуренция

Низкая

Уменьшение
стимулирования
инноваций,
эффективности,
экономического роста
(статья 1)

Средняя

Конкурирование по
качеству

Улучшение качества
(статья 2)

Рост цены за
квадратный метр

Высокая

Нарушение норм
ради клиента
(статья 3)

Математическая модель

Принимаемый уровень статистической значимости: 0.05

Размер выборки: 411

Подвыборка	Размер
Low	37
Mid	267
High	107

1 шаг

Понять, есть ли статистически значимые различия между подвыборками

2 шаг

Найти, какие именно выборки различаются

Особенности данных:

1. Не во всех подвыборках нормальное распределение¹
2. Дисперсии подвыборок не равны²
3. Подвыборки независимы³

Используемые инструменты:

1. Критерий Краскела-Уоллиса⁴
2. Критерий Данна⁵

*Подвыборки состоят из корпусов

1. Критерий Краскела-Уоллиса
2. Тест Данна с поправкой для множественных сравнений (подробнее в приложении)

1. Критерий Шапиро-Уилка
2. Критерий Левене
3. Нету корпусов, находящихся в нескольких округах. (подробнее в приложении)

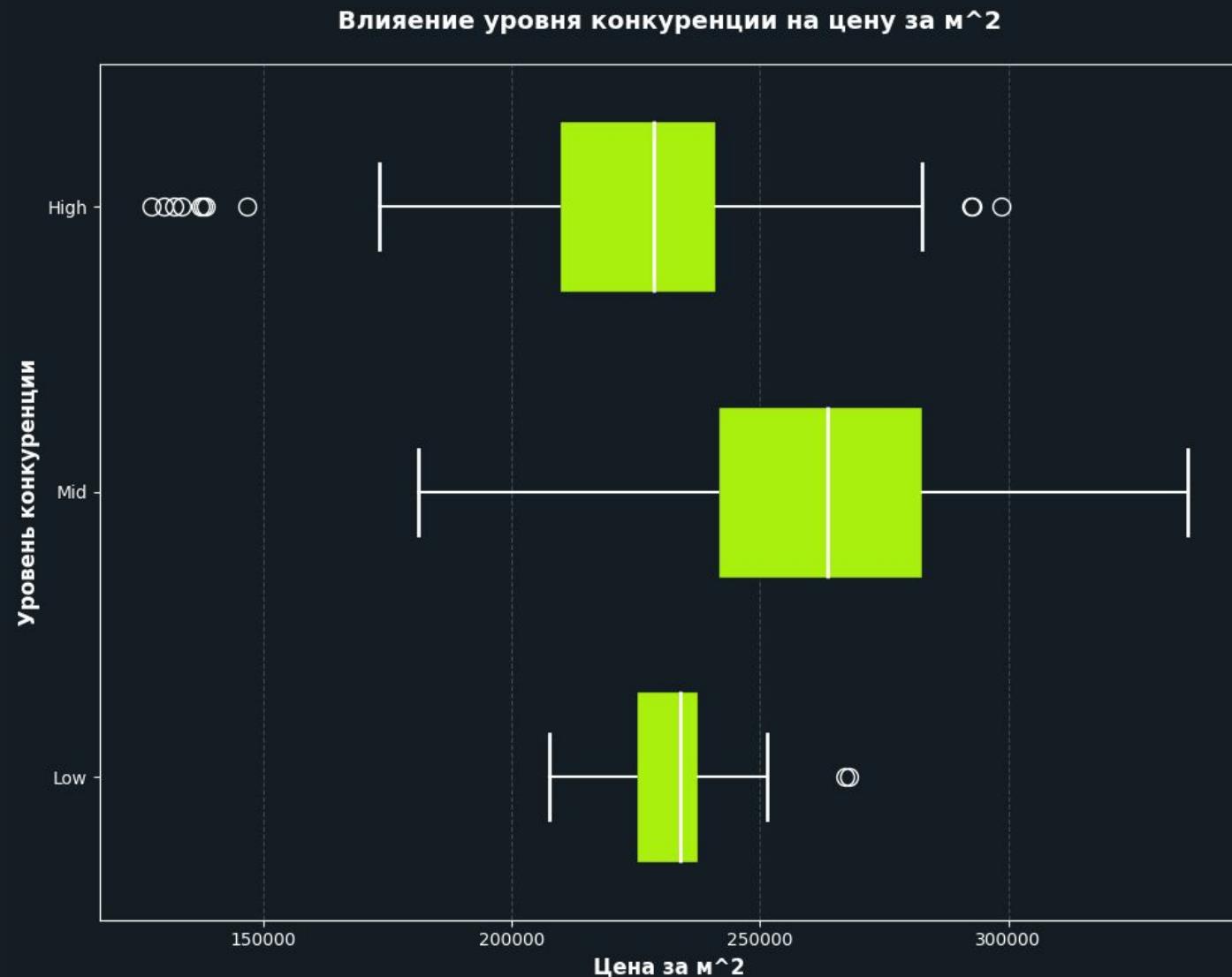
Проверка гипотезы

Критерий Краскела-Уоллиса
 $p\text{-value} < 0.01$

Определим какие именно подвыборки различаются

Подвыборка	Подвыборка	У кого больше
Mid	Low	Mid
Mid	High	Mid

Вывод: На нашей выборке гипотеза не отвергается.



Проверка устойчивости

размер наблюдений в подвыборке

конкурентность	Low	Mid	High	Краскел-Уоллис	Тест Данна
студия	88	188	59	Да	Да
1 комната	102	221	69	Да	Да
2 комнаты	101	218	63	Да	Да
3 комнаты	86	21	227	Да	Нет

*Четырехкомнатные квартиры не берем т.к. их очень мало и не удается собрать подвыборки



Делим не по тритилям, а по квартилям и получаем разницу между Mid и остальными



Нету разницы между округами с низким и средним уровнем конкуренции т.к. одна выборка < 30.

Интерпретация

Гипотеза подтвердилась

В районах со средней конкуренцией цена за м² в среднем выше, чем в районах с низкой или высокой.

Проверка на устойчивость

Устойчивость
подтвердилась
на всех подвыборках

Ограничения

1. Недостаток финансовых данных компаний

Перспективы дальнейшего развития

сбор финансовых данных компаний

2. Отсутствие данных об инфраструктуре

Сбор данных об инфраструктуре дома, окружении, экологии и т.д.

3. Неучтенные экономические показатели

Исследовать экономические факторы которые могли влиять на спрос на квартиры в исследуемый период

РЕКОМЕНДАЦИИ



Рекомендации для государства

Организовывать равномерное и контролируемое
освоение территорий для создания умеренной
конкуренции.

Рекомендация для девелоперов:

Стоит строиться в округах с умеренной
конкуренцией и избегать монопольной среды.



Спасибо за внимание!

Исходный код



github.com/Ushan4ik/Dano2025



КОМАНДА МЕЧТЫ

telegram

@Ushan41k @Ra1nDarr4ik @Aexxrr @dinl_09 @dreuls

Приложения

Здесь представлены дополнительные графики, таблицы и результаты анализа, которые подтверждают наши выводы и могут быть интересны для более глубокого изучения.



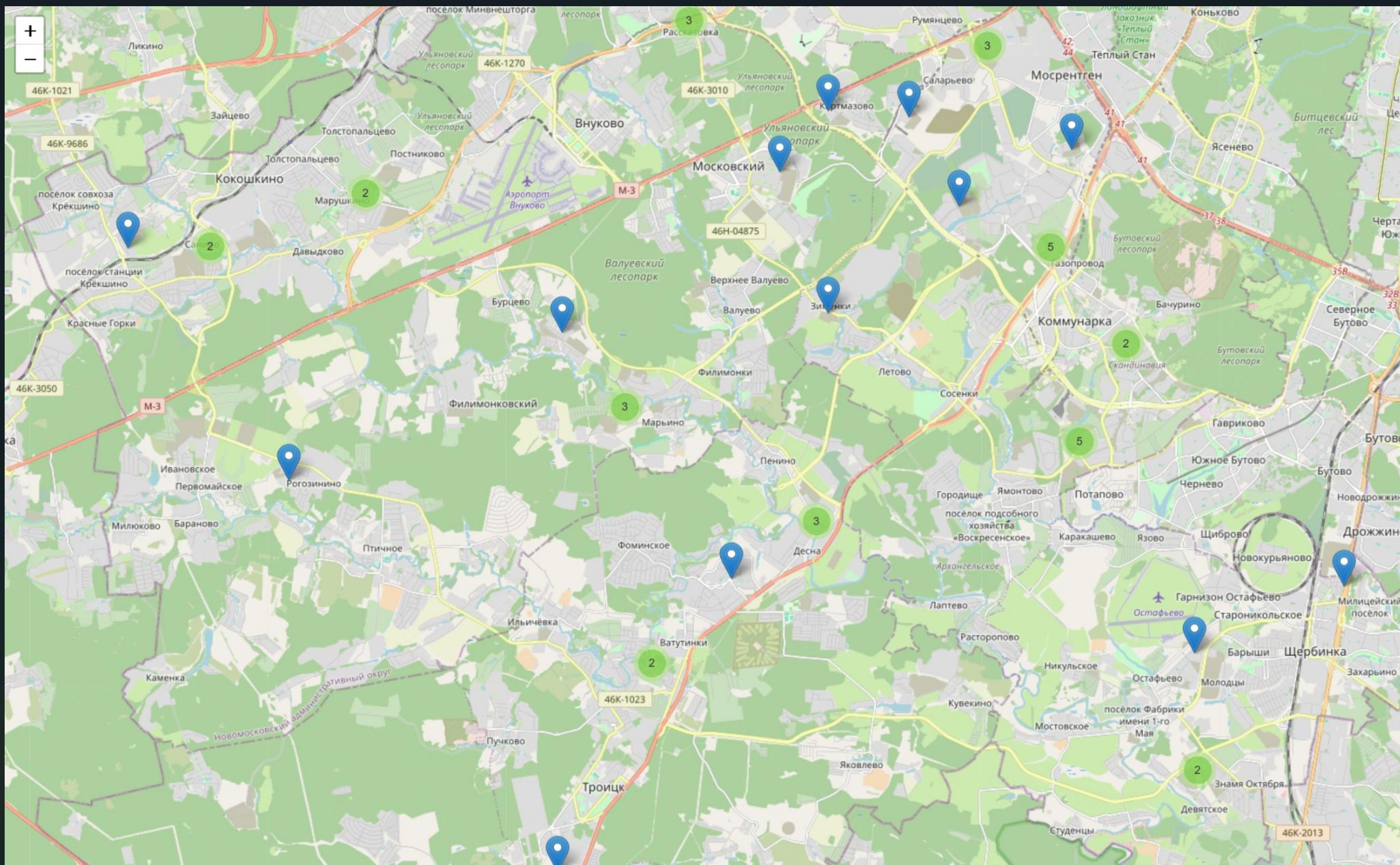
Ссылки на статьи

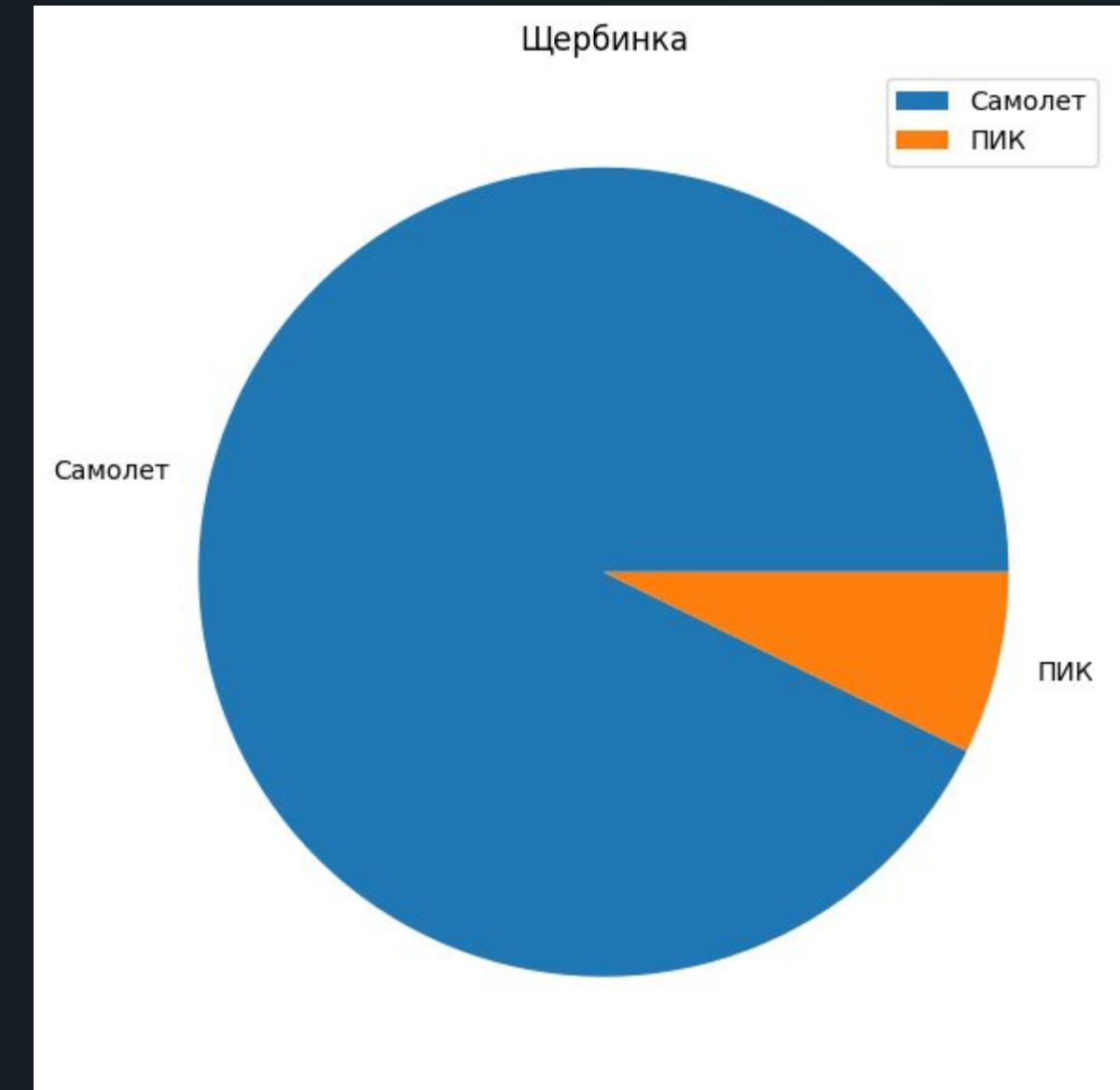
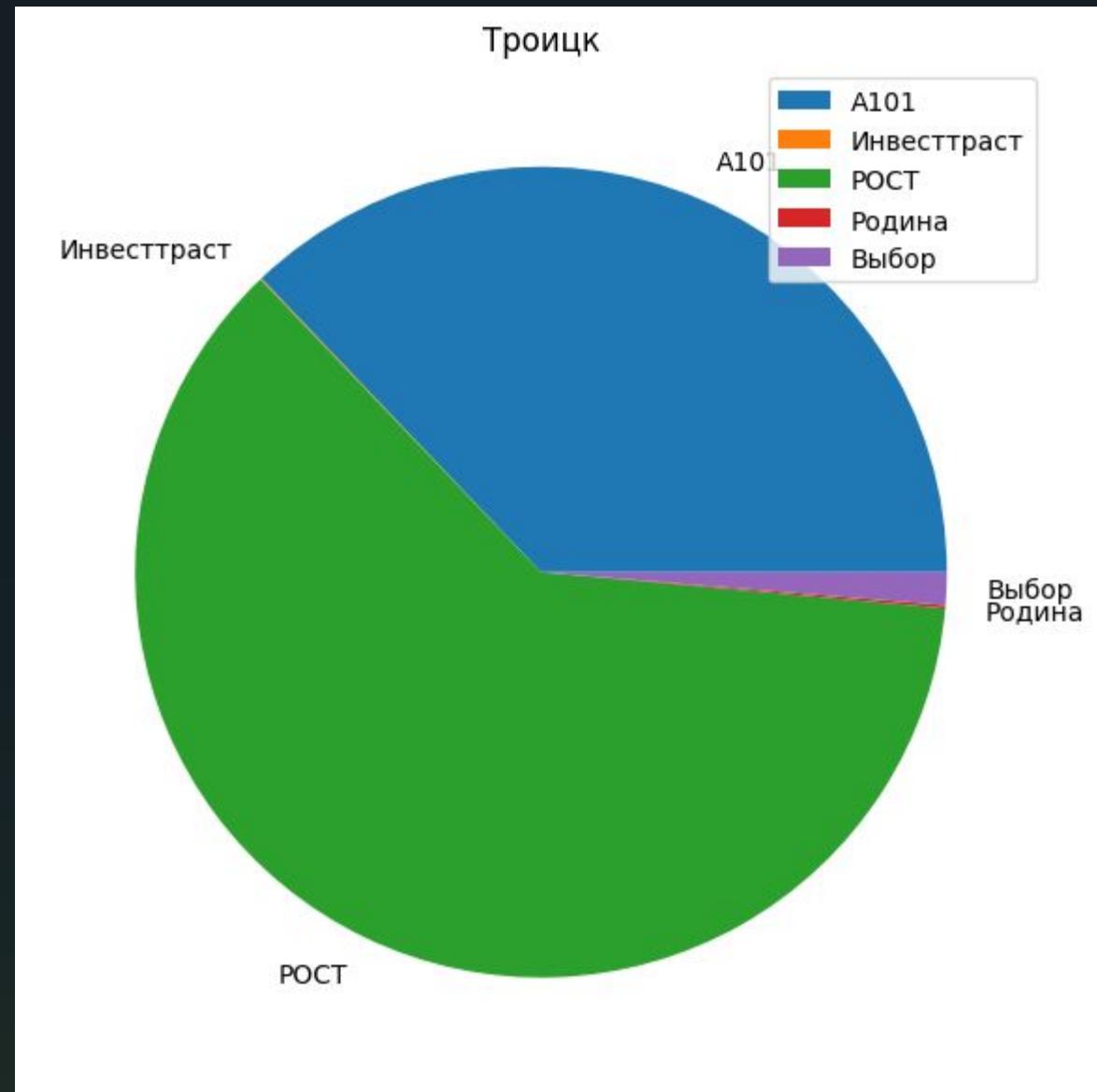


Статистика по районам

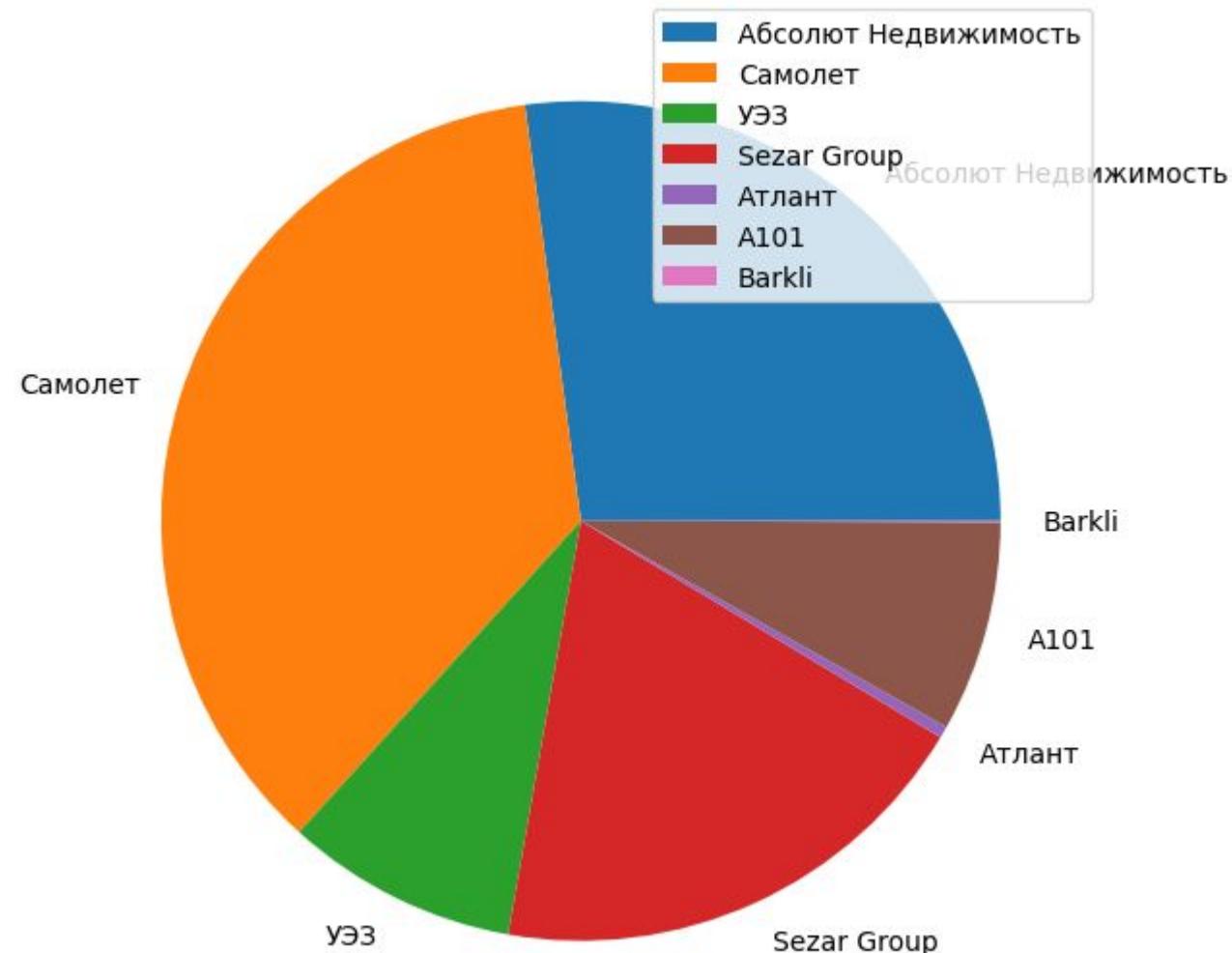


Карта расположения новостроек

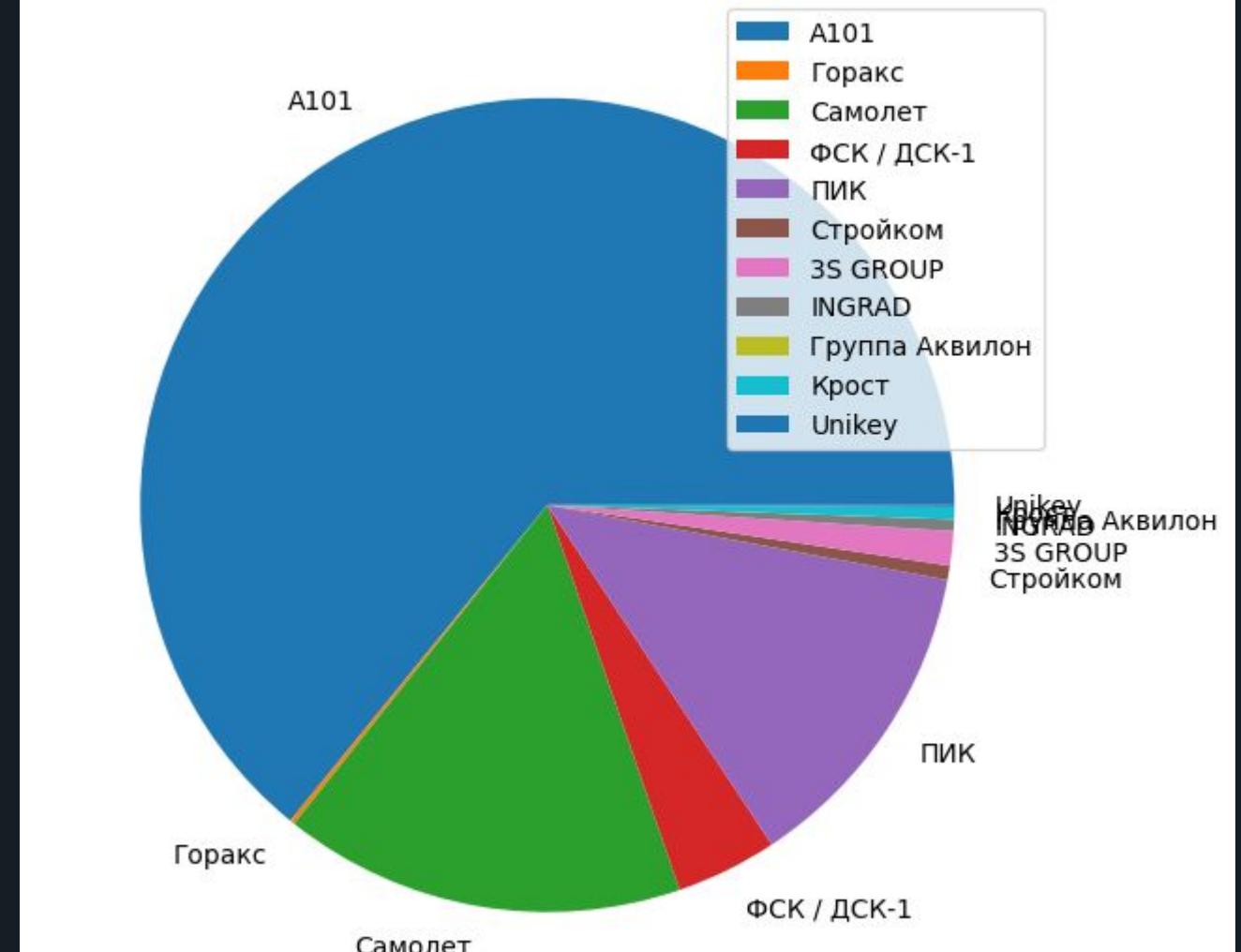


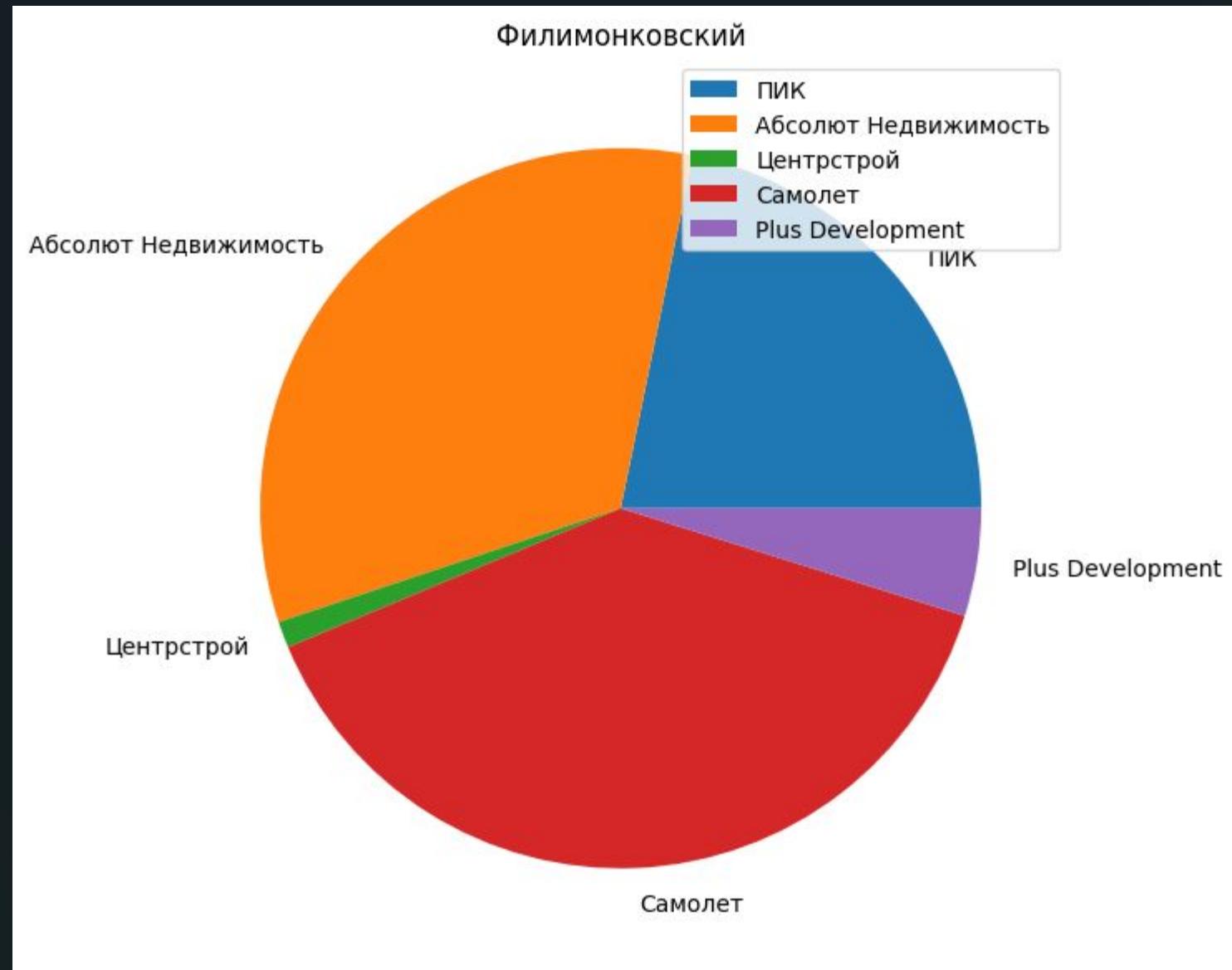


Внуково



Коммунарка





Статья 1:

https://www.researchgate.net/publication/381911420_THE_IMPORTANCE_OF_COMPETITION_IN_THE_FORMATION_OF_A_FAVORABLE_BUSINESS_ENVIRONMENT

Статья 2:

<https://www.tutorchase.com/answers/a-level/economics/what-is-the-effect-of-market-competition-on-product-quality>

Статья 3:

<https://www.library.hbs.edu/working-knowledge/when-business-competition-harms-society>

Проверка на устойчивость в блокноте