**Квартиры**

**Откуда получена информация по стоимости:**

В рамках выполнения курсовых проектов в университете нашей группе предоставляли разные выгрузки данных, в том числе и данные полученные с Циана. Собственно эти данные я и взял для выполнения задания, так как просто значение средней стоимости квартиры в Москве или цены за ее 1 м2, которое можно найти в новостных и аналитических публикациях в открытом доступе, не отображает влияние ряда факторов, которые важны при ценообразовании: расположение, класс ЖК, транспортная доступность, квадратура, количество комнат, этажность, тип планировки, отделка и др.

В выгрузке с Циан содержится набор полей, который содержит информацию об адресе ЖК, площади квартиры, количестве комнат, этажности, стоимости. Так как в выгрузке тестового задания квартиры помимо идентифицирующих атрибутов и атрибутов связанных со статусом продажи еще включает и атрибуты, связанные с этажом, типом, размерностью, площадью и количеством комнат, то по этим 5 признакам можно попробовать написать модель для оценки стоимости представленных квартир и цены за 1 м2.

**Как можно еще получить информацию:**

1. Спарсить более актуальные данные с ЦИАН и аналогичных агрегаторов об открытых позициях, расширить обучающую выборку
2. Получить сведения из Росреестра о совершенных сделках на первичном рынке по покупке квартир в ЖК в Москве
3. Получить сведения из других источников (ЕИСЖС, банки и др.)

**Чего не хватало дополнительно в качестве данных / информации:**

* Прочих ключевых параметров, влияющих на ценообразование (локация ЖК, класс ЖК (но вероятно, что это бизнес-класс), отделка, транспортная доступность)
* Конкурентный анализ в схожем сегменте и/или локации
* Специфика ценообразования на рынке первичного жилья (в конечном результате получилось, что для одинаковых квартир по площади, типу, размерности и числу комнат, но с разной этажностью цены отличаются, хотя на практике они могут иметь одинаковый прайс и др.)

**Специфика в исходных данных тестового задания:**

1. Квартиры располагаются с 1-ого по 31-ый этаж
2. Минимальная площадь лота 16,30 м2
3. Максимальная площадь лота 197,10 м2
4. Тип квартир завязан на количество комнат в ней. Максимальное число комнат – 4
5. Каждый тип квартиры имеет свою разбивку по размерности в зависимости от площади без балкона
6. Для студий и однокомнатных квартир количество комнат указано одинаково (1)

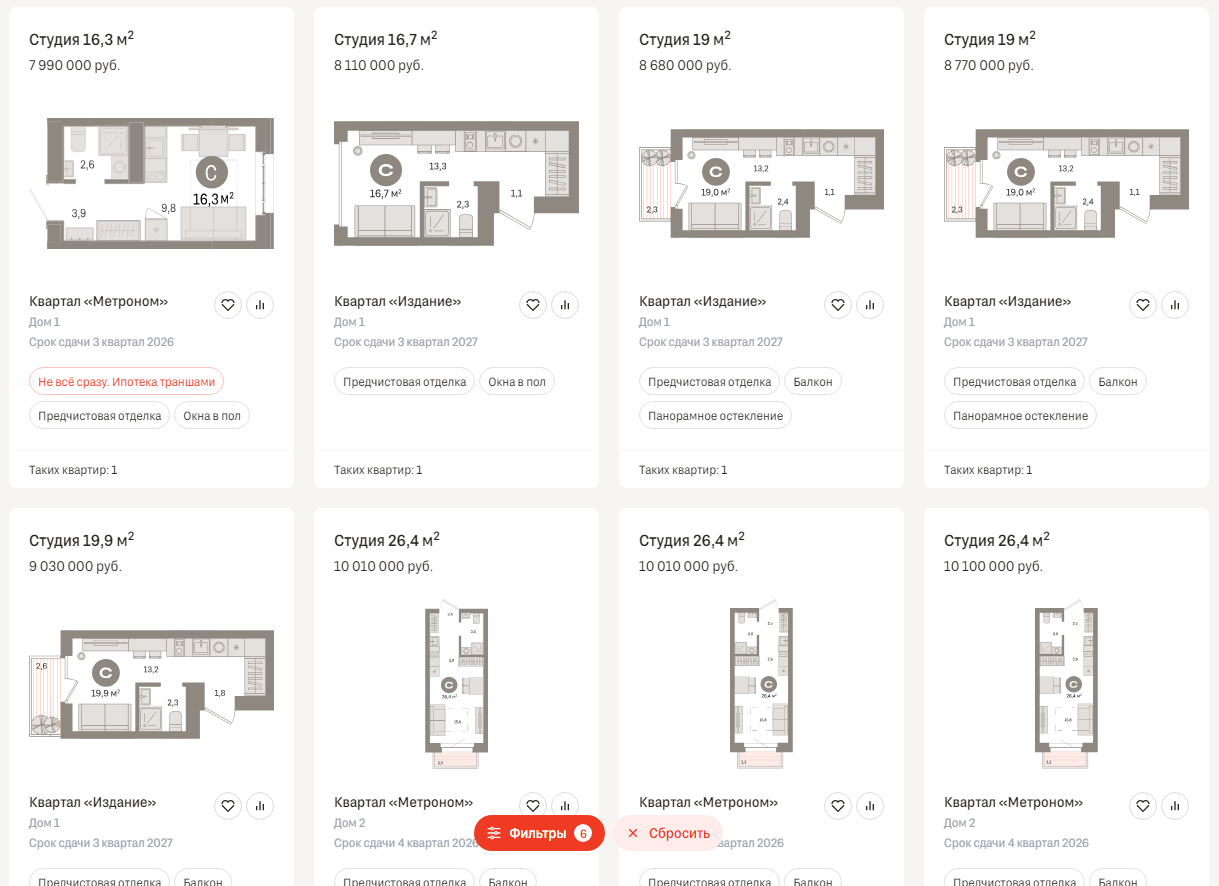
**Выполнение работы:**

Код выложил на GitHub: <https://github.com/smykovevgen/brusnika_jobtest>

1. Загрузка данных для Москвы с Циан
2. Очистка от лишних колонок, оставив только колонки со стоимостью, площадью, количеством комнат и этажностью
3. Очистка от лишних значений и замена разных вариантов одинаковых по смыслу значений в признаке про количество комнат (Например, очистить от записей с апартаментами и заменить текстовое наименование 2-комн. кв. на 2)
4. Добавление столбца с типом квартир по условию на основе преобразованных значений количества комнат
5. Добавление столбца с размерностью квартир по условию на основе типов квартир и их площадей
6. Замена категориальных значений на числовые в новых столбцах
7. Удаление строк с пустыми значениями в размерности
8. Замена типов данных в колонках на float
9. Обучение модели на основе обучающей выборки
10. Загрузка подготовленного датасета Брусники
11. Расчет с помощью модели стоимости квартир в тестовой выборке
12. Округление значений стоимости до тыс. руб.
13. Выгрузка результата
14. Перенос полученных значений стоимости в исходную таблицу с данными о квартирах и расчет цены за 1 м2 для каждой (стоимость / общая площадь)

**Проверка с открытыми для продажи позициями:**

Несложно открыть сайт и посмотреть стоимость открытых на продажу квартир у Брусники:



Изображение выглядит как текст, диаграмма, снимок экрана, План

Автоматически созданное описание

На примере студий можно выполнить сравнение значений по нескольким вариантам:

|  |  |  |  |
| --- | --- | --- | --- |
| **Лот** | **Стоимость на сайте, руб.** | **Мин. стоимость из модели** | **Макс. стоимость из модели** |
| Студия 16,3 м2 | 7 990 000 | 6 210 000 | 6 880 000 |
| Студия 26,4 м2 | 10 010 000 | 7 950 000 | 8 170 000 |
| 1-комнатная 39,4 м2 | 14 530 000 | 12 660 000 | 14 980 000 |
| 1-комнатная 39,5 м2 | 14 970 000 | 12 660 000 | 14 910 000 |
| 1-комнатная 42,3 м2 | 15 210 000 | 12 650 000 | 14 600 000 |
| 1-комнатная 43,8 м2 | 14 990 000 | 13 850 000 | 14 060 000 |

Из полученных значений видно, что разница между прогнозом модели и ценой на сайте может достигать 2,1 млн рублей, что, к сожалению, говорит о недостаточном качестве работы модели.

**Что не было учтено, но можно учесть при ценообразовании:**

1. Рыночные изменения стоимости квадратного метра квартир в Москве за 6-10 месяцев, так как данные обучающей выборки были получены как минимум полгода назад.
2. Инфляционные изменения
3. Дополнительные ценообразующие факторы, которые были уже озвучены
4. Динамика продаж, акции, стратегия продаж на 3-ий и 4-ый квартал 2024 года и др.
5. Продажи конкурентов
6. Популярность планировок

**Результат:**

|  |  |
| --- | --- |
| Заполненная таблица по стоимости и средней цене 1 м2 квартир |  |