Государственное бюджетное профессиональное

образовательное учреждение Московской области

«Физико-технический колледж»

**«Анализ цены квартиры на вторичном рынке по Московскому региону: Москва, Новая Москва, Московская область»**

Работу выполнил:

студент группы № ИСП-22

Богомяков Максим

Проверил:

Преподаватель информатики

Базяк Г.В. Коновалов И. В.

Долгопрудный, 2024

# СОДЕРЖАНИЕ

[СОДЕРЖАНИЕ 1](#_Toc1112863883)

[ВВЕДЕНИЕ 2](#_Toc1811315824)

[Основная часть 5](#_Toc1267612975)

[1.1 Методология 5](#_Toc854807823)

[1.2 Анализ данных 6](#_Toc32484602)

[Рефлексия 8](#_Toc783427107)

[Заключение 9](#_Toc1118852660)

# ВВЕДЕНИЕ

**Актуальность задачи**

Оценка недвижимости - важная составляющая девелоперского бизнеса. Информация, о реальной цене квартиры исходя из рынка, интересна для покупателей продавцов, застройщиков, агентов и др.

Эту информацию можно использовать по-разному, в частности, у бизнеса есть потребность использовать такую модель для лидогенерации клиентов. Потенциальный клиент заходит на сайт, хочет оценить свою квартиру, вводит параметры: квартира, на 3 этаже, 3-х комнатная, дизайнерский ремонт. Затем нажимает кнопку «узнать цену», и, прежде чем дать ответ, мы хотим запросить у клиента номер телефона или адрес почты.

Для этого нужна модель оценки цены квартиры по Московскому региону: Москва, Новая Москва, Московская область.

**Постановка задачи**

**Цель**: Собрать данные и провести разведочный исследовательский анализ данных (EDA) для построения модели, которая будет оценивать цену квадратного метра недвижимости в Московском регионе (Москва, Новая Москва, Московская область).

**Задачи:**

***Задание базового (обязательного) уровня)***

- Используя открытые источники и личный опыт, составить список параметров, значительно влияющих на цену квадратного метра жилой площади.

- С учётом выявленных выше факторов произвести парсинг данных по квартирам на продажу, используя различные парсеры. Данные получаем, используя различные сайты с объявлениями о продаже недвижимости: Циан, Авито, ДомКлик и др.

- Произвести подготовку данных для анализа: проверка на пропуски, выбросы и ошибки. Обработать выявленные аномалии (удалить / заполнить)

- Проведите Исследовательский Анализ Данных (EDA). Постройте распределение основных параметров; визуализируйте взаимосвязи между ними; определите признаки, оказывающие наиболее сильное влияние на целевую переменную.

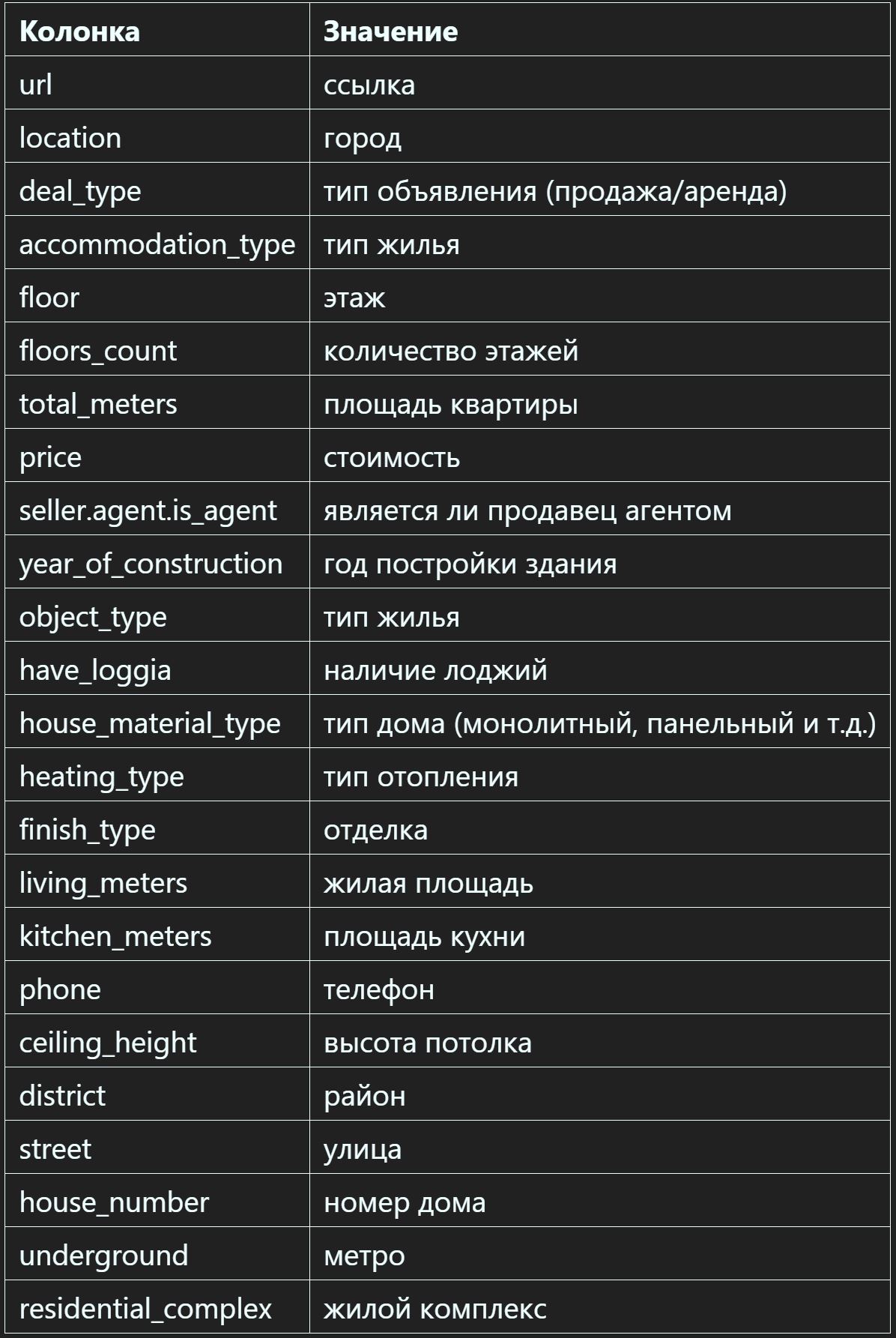
# Основная часть

# 

# Методология

Для сбора данных использовался cianparser, с его помощью удалось собрать 8393 квартиры. Данные были собраны с платформы Циан, поскольку она предоставляет более обширные и полезные для анализа показатели по сравнению с другими платформами, такими как ДомКлик.

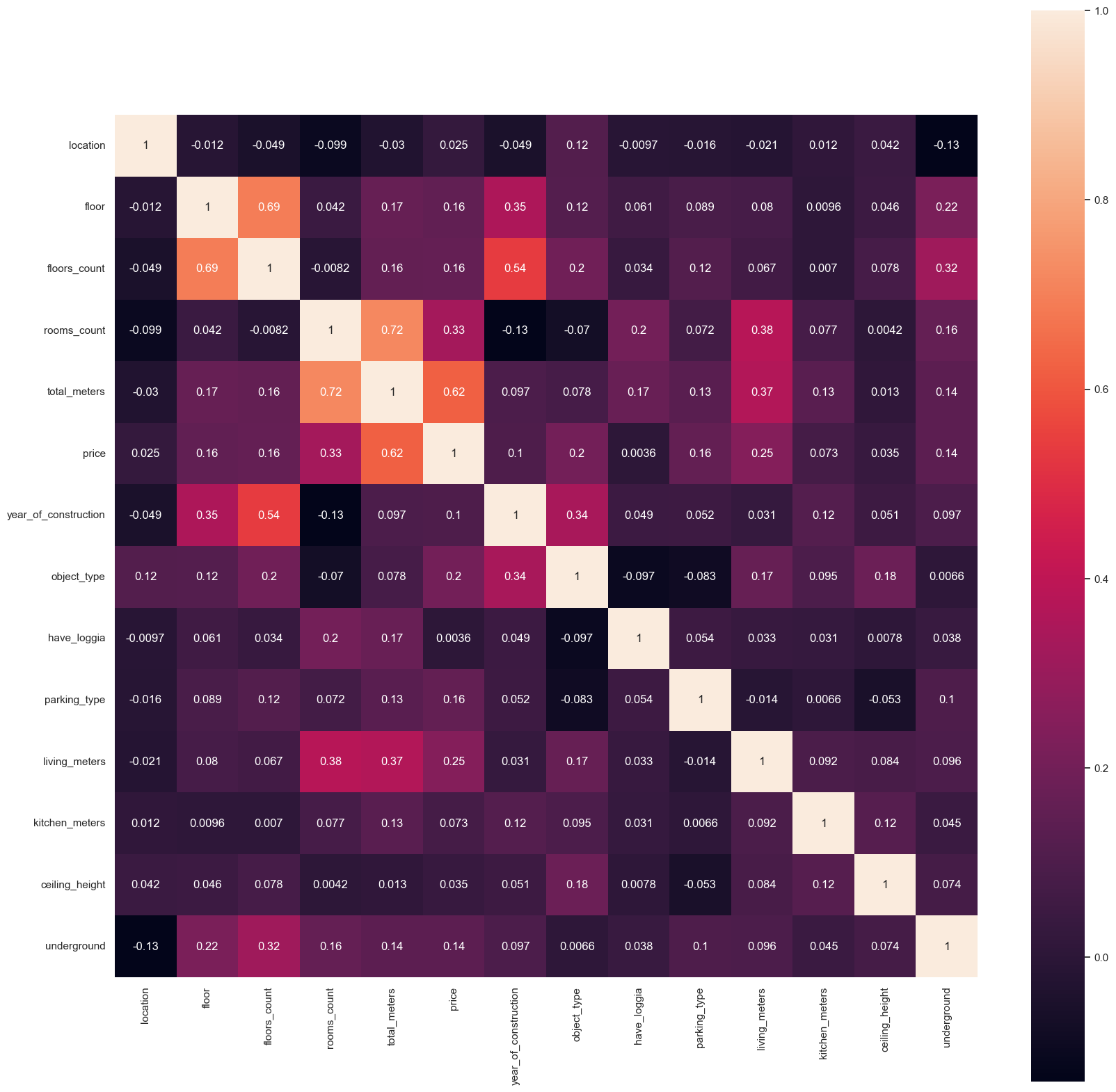
Основные колонки данных:



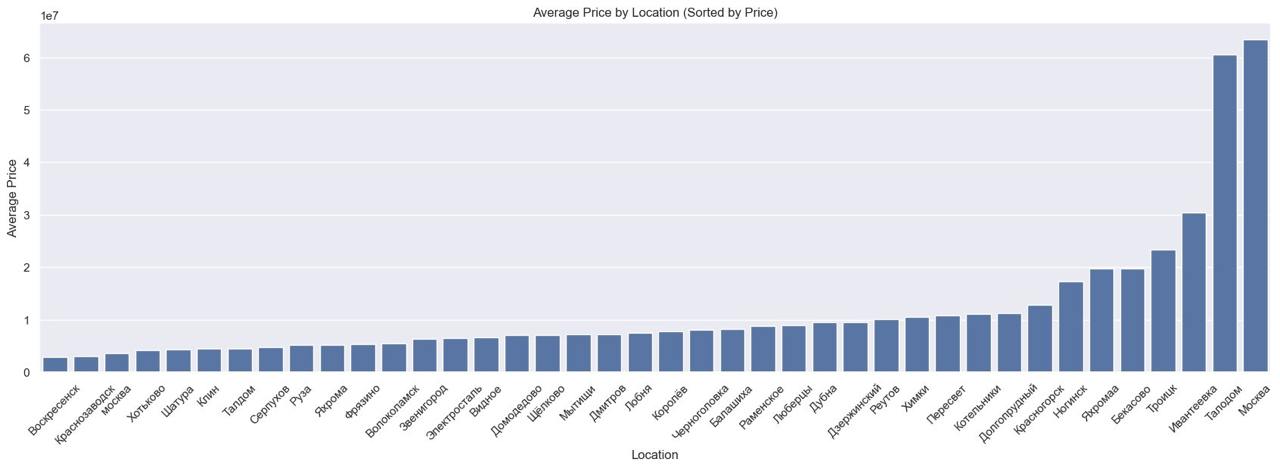
# 1.2 Анализ данных

После получения данных мы их очищаем, удаляем лишние столбцы, проверяем на аномальные значения и удаляем строки с аномалиями. Затем превращаем все наши данные в числовые.

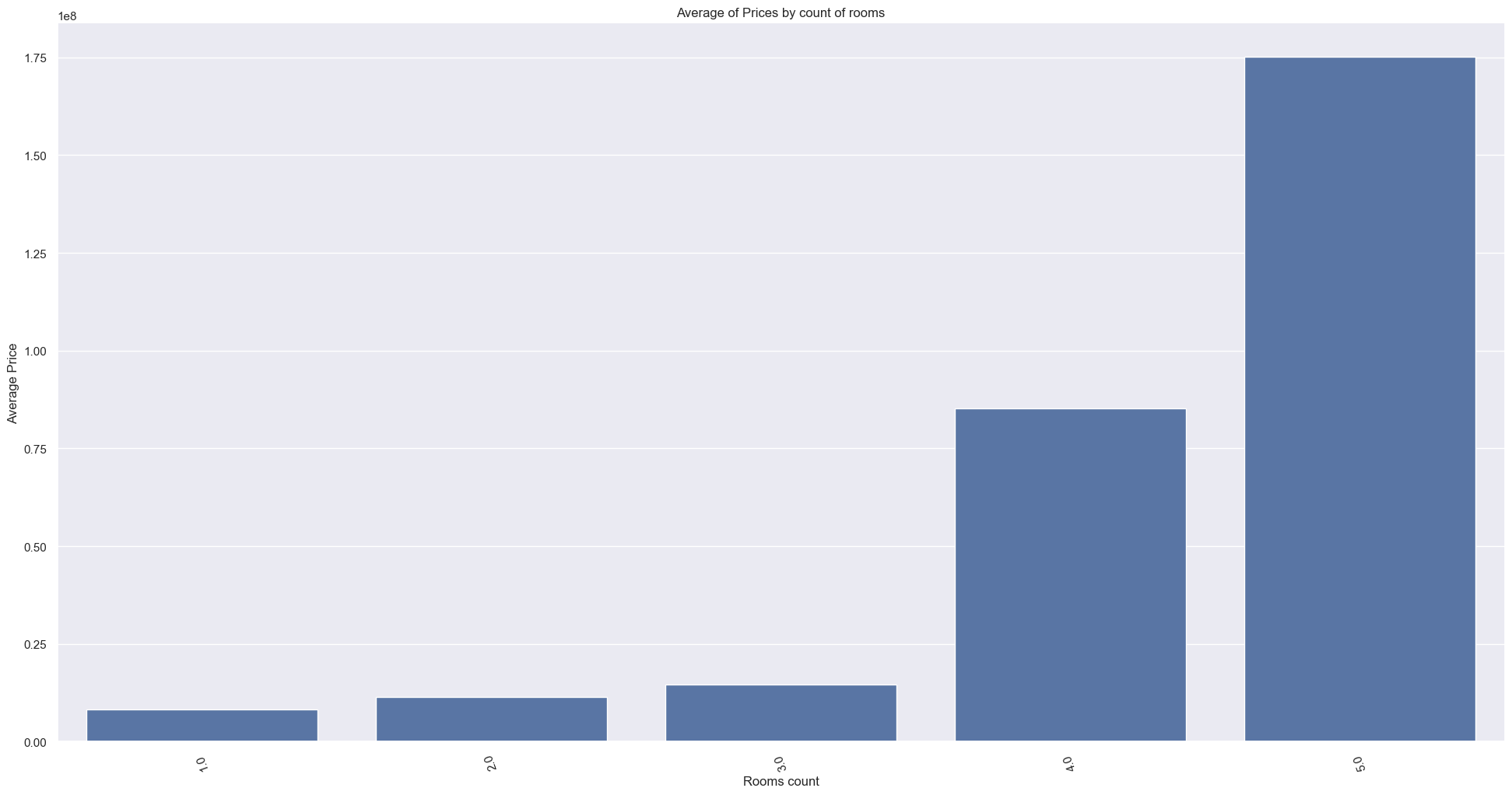
Для выявления зависимостей между характеристиками квартир и их ценой был построен график корреляций. По графику видно, что на цену больше всего влияет общее число метров квартиры (0,62), затем уже идет количество комнат (0,33)



Так же по этому графику мы видим, что локация тоже значительно влияет на цену



На данном графике видно, что цена так же зависит от количества комнат, что довольно логично



# Рефлексия

Что получилось:

* Выявить основные параметры наиболее сильно влияющие на нашу целевую.

Что можно улучшить:

* Избавиться от остальных выбросов
* Колонку "have\_loggia" заполнить не нулями и единицами, а другим способом

# Заключение

В ходе проведенного анализа были выявлены ключевые факторы, влияющие на целевой показатель — цену квартиры. Наибольшее влияние оказывают следующие параметры:

* Площадь квартиры — основной критерий, прямо пропорционально влияющий на стоимость.
* Локация — важный фактор, определяющий уровень цен в зависимости от расположения объекта недвижимости.
* Количество комнат — также оказывает значительное влияние на цену.

Данные результаты создают основу для разработки модели, способной прогнозировать цену квартиры на основе этих ключевых характеристик.