

Государственное бюджетное профессиональное  
образовательное учреждение Московской области  
«Физико-технический колледж»

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

Работу выполнил:  
студент группы № ИСП-22  
Крылов Савва  
Проверил:  
Преподаватель информатики  
Базяк Г.В. Коновалов И. В.

Долгопрудный, 2024

## СОДЕРЖАНИЕ

СОДЕРЖАНИЕ .....	2
ВВЕДЕНИЕ .....	3
Основная часть .....	5
1.1 Методология .....	5
1.2 Анализ данных.....	7
Заключение .....	9

## **ВВЕДЕНИЕ**

### **Актуальность задачи**

### Оценка стоимости недвижимости — ключевая часть девелоперской деятельности. Данные о реальной рыночной цене квартиры интересуют множество участников рынка: покупателей, продавцов, девелоперов, риэлторов и других.

Эти данные могут использоваться для различных целей, включая лидогенерацию. Так, потенциальный клиент заходит на сайт, желая узнать стоимость своей квартиры, и вводит параметры: этаж, количество комнат, качество ремонта и т.д. Нажимая кнопку «узнать цену», перед получением ответа клиенту предлагается оставить номер телефона или адрес электронной почты.

Для этого требуется модель, способная оценить стоимость квадратного метра квартиры в Московском регионе, охватывая Москву, Новую Москву и Московскую область.

### **### Постановка задачи**

**\*\*Цель:\*\*** Собрать и проанализировать данные, чтобы построить модель, оценивающую стоимость квадратного метра жилья в Московском регионе (Москва, Новая Москва, Московская область).

## **\*\*Основные задачи:\*\***

### **\*\*Базовый уровень (обязательный)\*\***

- Определить перечень характеристик, значительно влияющих на стоимость квадратного метра жилой площади, используя открытые источники и собственные знания.
- С учетом выявленных факторов осуществить парсинг данных по квартирам, выставленным на продажу, используя разные источники. Источники данных могут включать такие сайты с объявлениями, как Циан, Авито, ДомКлик и другие.
- Подготовить данные для анализа: проверить их на наличие пропусков, выбросов и ошибок. Обработать аномалии путем удаления или заполнения данных.
- Провести разведочный анализ данных (EDA). Построить распределения ключевых параметров, визуализировать взаимосвязи между ними, а также выявить признаки, которые оказывают наибольшее влияние на целевую переменную.

## **Основная часть**

### **1.1 Методология**

Для сбора данных использовался `scraparser`, с его помощью удалось собрать 4131 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:

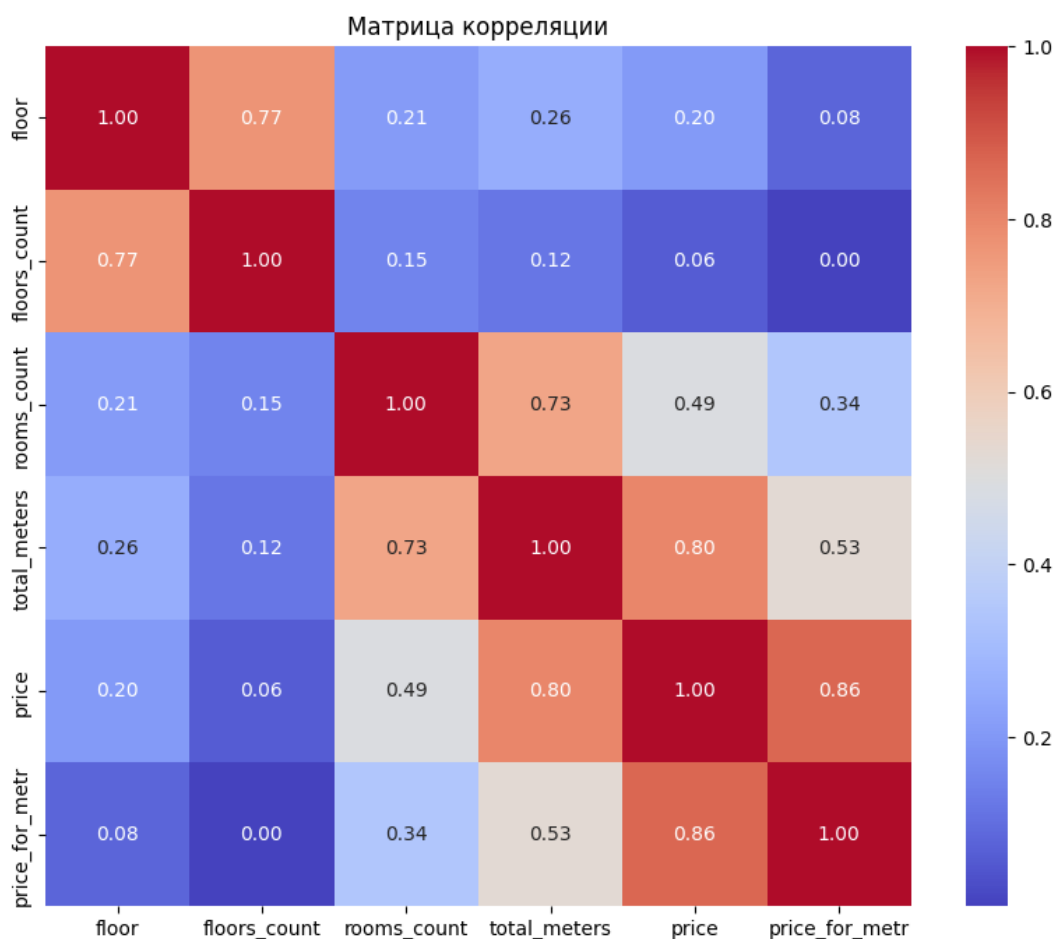
Колонка	Значение
url	ссылка
location	город
deal_type	тип объявления (продажа/аренда)
accommodation_type	тип жилья
floor	этаж
floors_count	количество этажей
total_meters	площадь квартиры
price	стоимость
seller.agent.is_agent	является ли продавец агентом
year_of_construction	год постройки здания
object_type	тип жилья
have_loggia	наличие лоджий
house_material_type	тип дома (монолитный, панельный и т.д.)
heating_type	тип отопления
finish_type	отделка
living_meters	жилая площадь
kitchen_meters	площадь кухни
phone	телефон
ceiling_height	высота потолка
district	район
street	улица
house_number	номер дома
underground	метро
residential_complex	жилой комплекс

Для анализа данных использовался метод регрессии, позволяющий глубже исследовать влияние различных характеристик квартиры на её стоимость. Это помогает определить основные факторы, формирующие рыночную цену недвижимости.

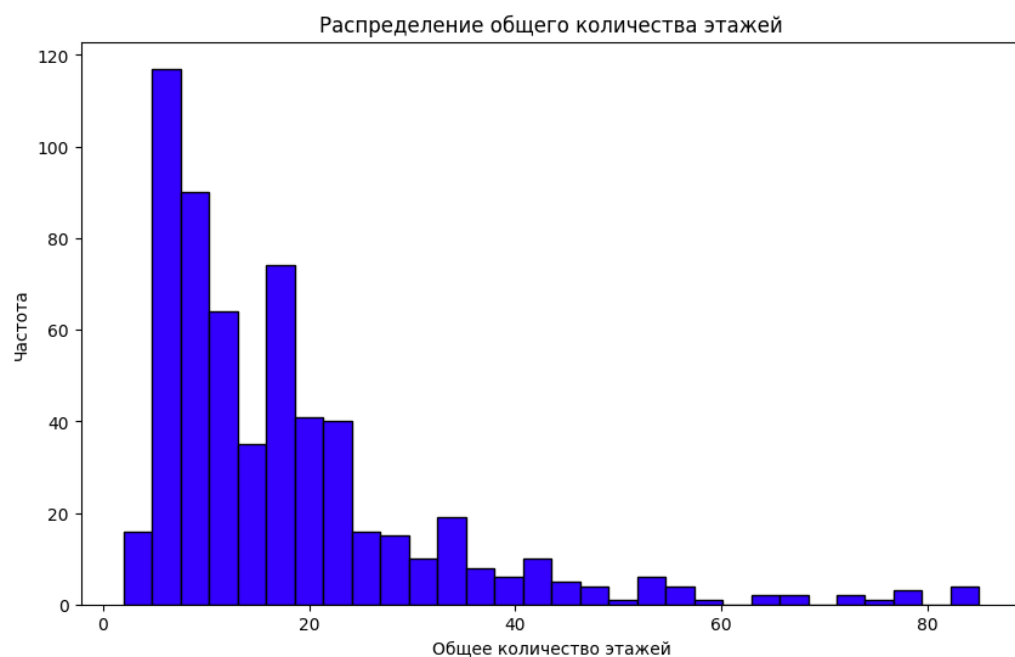
## 1.2 Анализ данных

После сбора данных мы очищаем их, удаляем ненужные столбцы, проверяем на наличие аномалий и исключаем строки с ними. Затем преобразуем все данные в числовой формат.

Для анализа зависимости между характеристиками квартир и их стоимостью был построен корреляционный график. Он показывает, что на цену сильнее всего влияет общая площадь квартиры (коэффициент 0,80), а затем -- количество комнат (коэффициент 0,49).



На данном графике мы видим преимущество малоэтажной застройки над многоэтажной





## **Заключение**

В результате проведенного анализа были определены ключевые параметры, влияющие на целевую переменную — стоимость квартиры. Наибольший вклад в формирование цены вносят следующие факторы:

- Площадь квартиры: основной параметр, оказывающий прямое влияние на стоимость.
- Количество комнат: значимо влияет на итоговую цену объекта.

Эти выводы формируют базу для создания модели, способной прогнозировать стоимость квартиры, опираясь на указанные ключевые характеристики.