

АНАЛИТИЧЕСКИЙ ОТЧЁТ О ВЫПОЛНЕНИИ ПЕРВОГО ИНТЕНСИВА

Отчёт составил: Филиппов Дмитрий Михайлович.

Группа: ИСП-23В.

Курс: III.

СОДЕРЖАНИЕ

1. Цели, задачи.
2. Методология и инструменты.
3. Ход работы и анализ данных.
4. Вывод.

ВВЕДЕНИЕ

В данном отчёте представлена информация о работе, проведённой в ходе выполнения первого интенсива.

ЦЕЛИ, ЗАДАЧИ

Цель работы: сбор и анализ данных по квартирам в Москве и Московской области для выделения параметров влияющих на их цены. А также подготовка данных для их использования в машинном обучении.

МЕТОДОЛОГИЯ И ИНСТРУМЕНТЫ

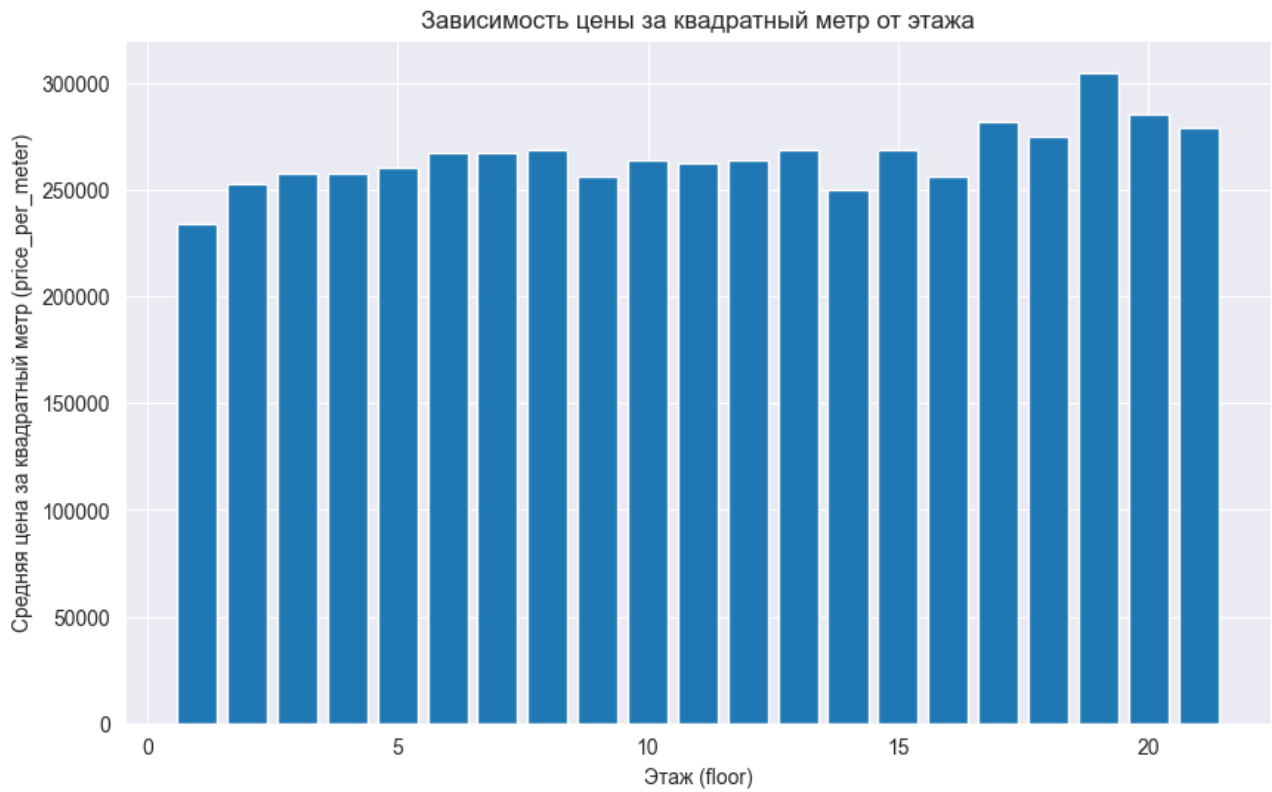
Методы выполнения: Для выполнения поставленной задачи используется язык программирования Python, для автоматизации процессов, библиотеки Numpy и Pandas для работы и анализа данных. Был написан парсер для сбора данных. За визуальную часть диаграмм и графиков отвечают библиотеки Matplotlib и Seaborn.

В качестве источника данных использовался cian.ru.

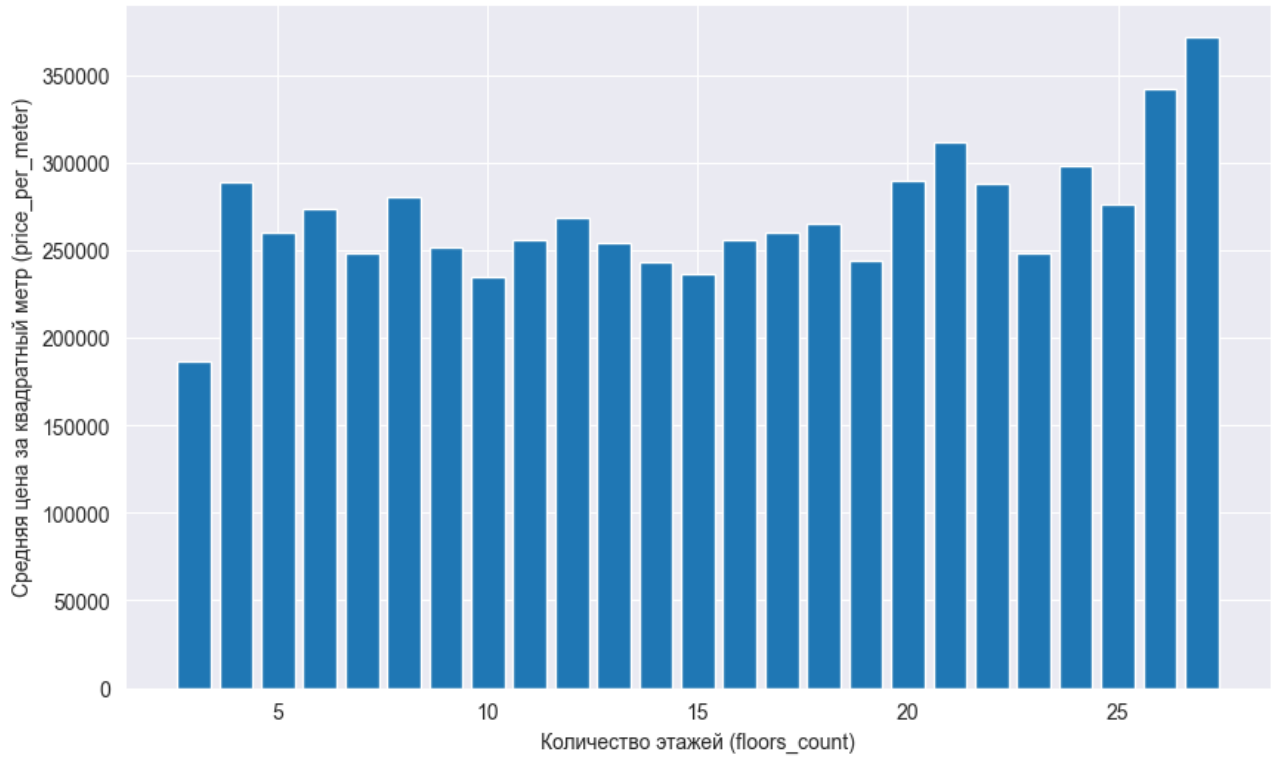
ХОД РАБОТЫ И АНАЛИЗ ДАННЫХ

Создан парсер для сбора данных. После — был создан датафрейм и была проведена очистка от ненужной информации и анализ уже имеющейся информации. При проведении корреляционного анализа данных, выяснилось: цена квартиры напрямую зависит от

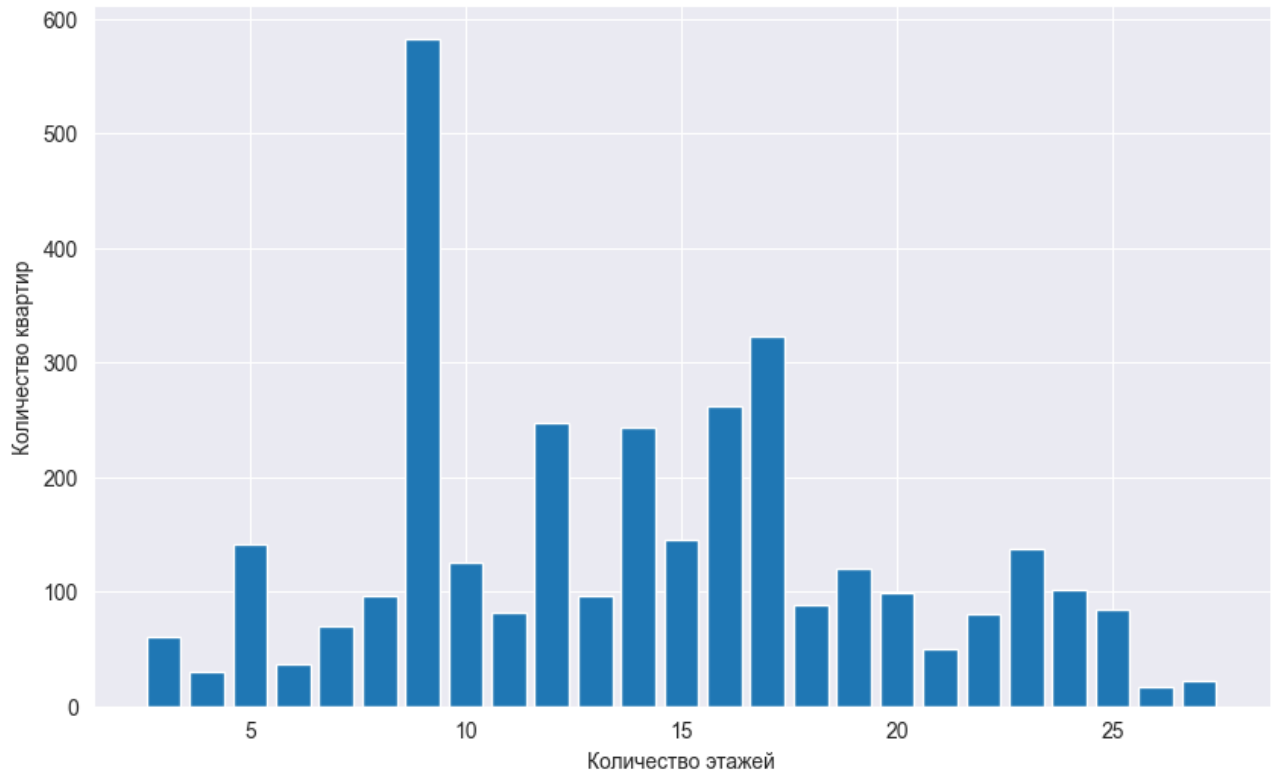
расположения относительно метро, этажа и количества этажей, а также от метража и количества комнат. В свою очередь метраж, количество комнат, и местоположение относительно метро сильно коррелируют между собой.



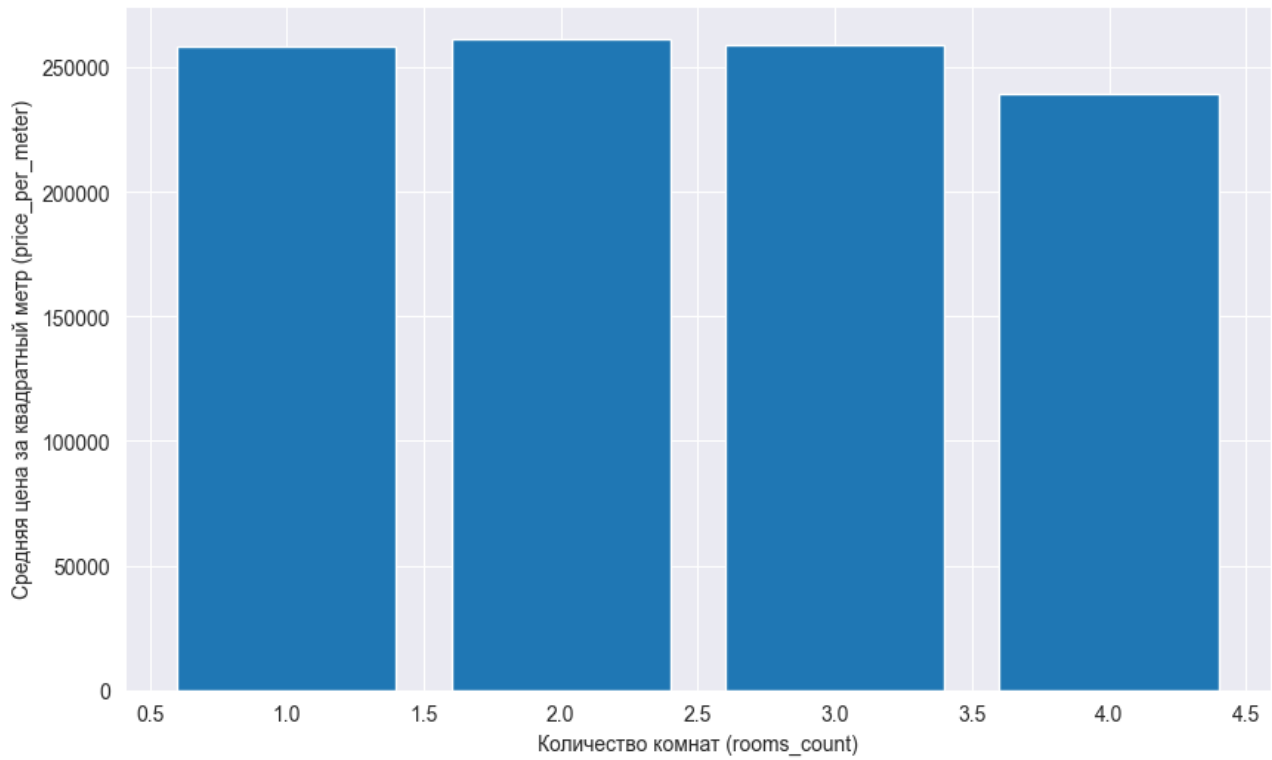
Зависимость цены за квадратный метр от количества этажей



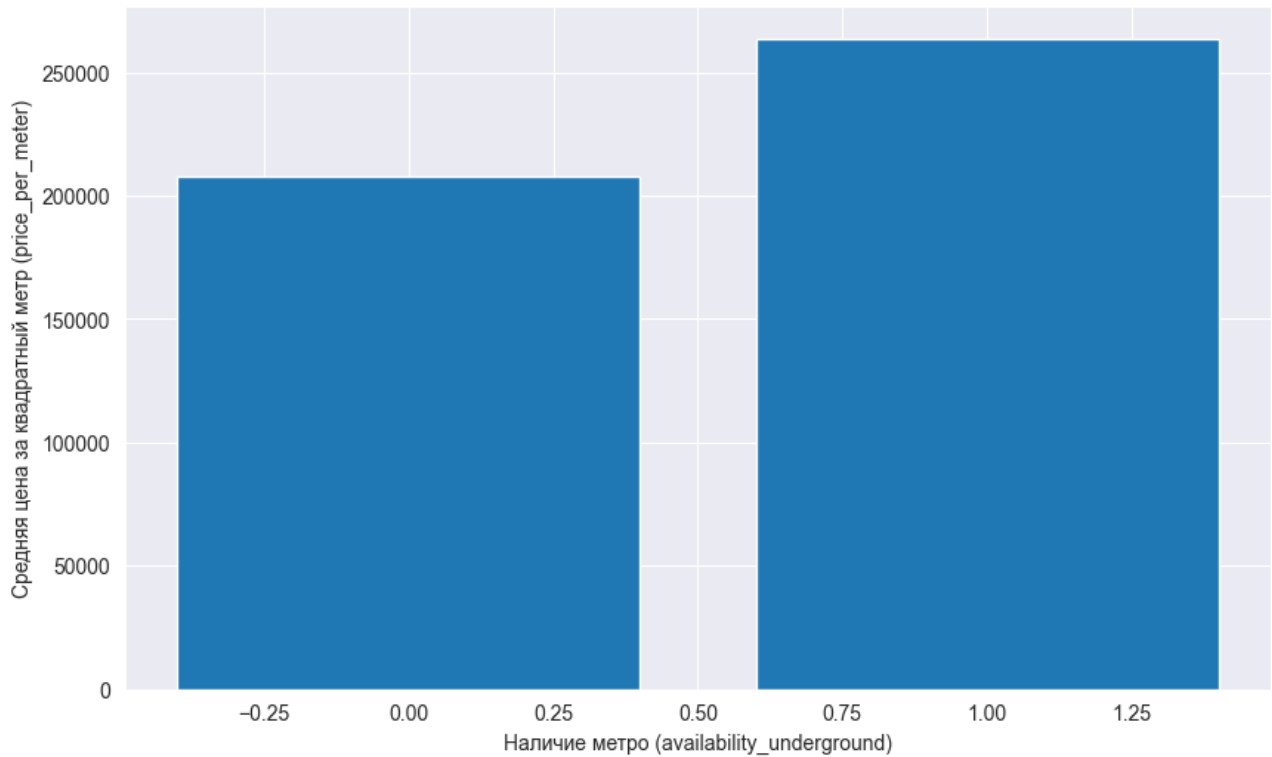
Распределение количества этажей



Зависимость цены за квадратный метр от количества комнат



Зависимость цены за квадратный метр от наличия метро





ВЫВОД

В ходе работы был проведён сбор, анализ, построение моделей и корреляционный анализ. Вышеперечисленные действия позволили получить данные и разобраться какие связи находятся между данными. А также разобраться в роли тех или иных параметров при работе с квартирами.