Государственное бюджетное профессиональное

образовательное учреждение Московской области

«Физико-технический колледж»

**Отчёт по кейсу «Самолёт»**

Работу выполнил:

Студент группы № ИСП-22

Рыжков Артём Русланович

Долгопрудный, 2024

# Введение:

В данном отчёте рассматриваются выводы, полученные с аналитической работы над данными в области «Квартиры в Московской области, и Москве».

# Цель:

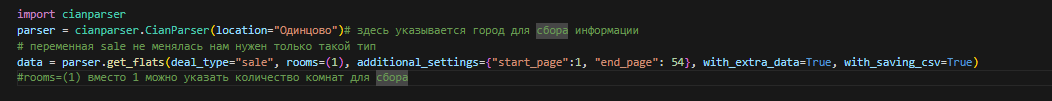
Собрать данные и произвести аналитическую работу над ними для будущих работ, например, создание модели на основе выводов.

# Задачи:

* Используя открытые источники собрать список данных.
* На основе полученной информации произвести удаление ненужных данных, дополнение необходимых, выявление аномалий и их отчистка.
* Визуализация данных при помощи, инструментов для подобных задач. Нахождение взаимосвязей между данными или их полное отсутствие, усреднённых показателей для уверенного отчёта.

# Процесс проделанной работы:

Для выполнения основной задачи, существует небольшой выбор источников, откуда собирать данные, при помощи скриптов, написанных на языке Python, было получено около десяти тысяч объявлений в нужных регионах (Рисунок 1-2).



Рисунок

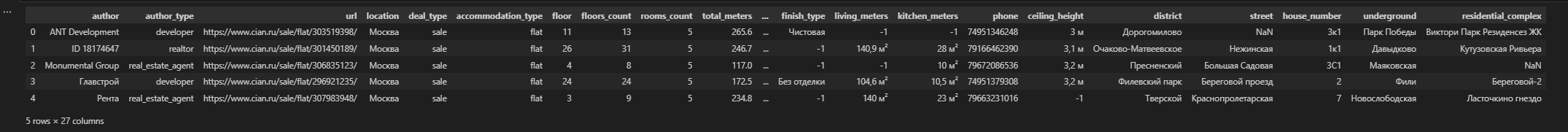
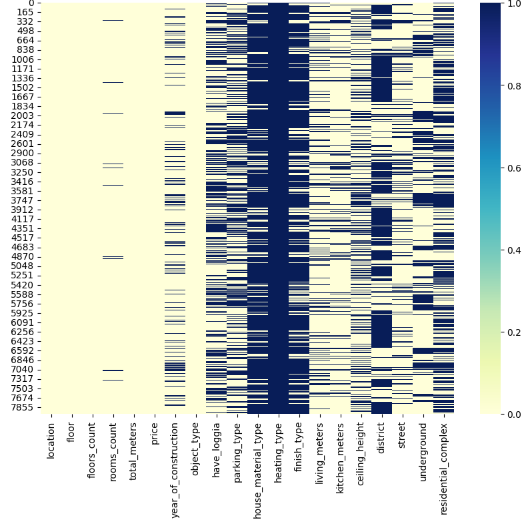


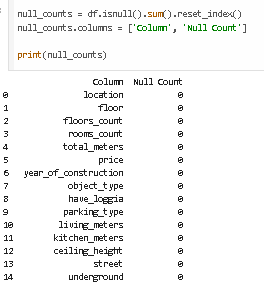
Рисунок 2

Далее мы убираем часть лишней информации из наших данных. Затем выводим количество отсутствующей информации визуально (Рисунок 3).

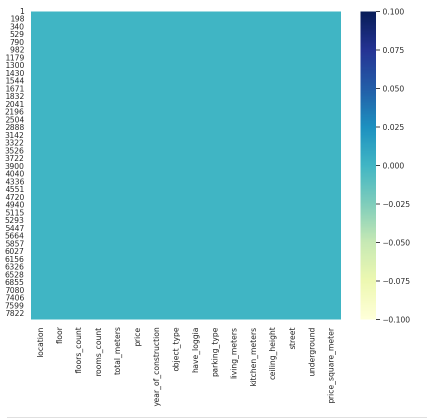


Рисунок

После этого убираем полностью пустые строки, повторяющуюся информацию и заполняем пустые значения 0, а затем выводим количество пустот в столбцах (Рисунок 4-5).

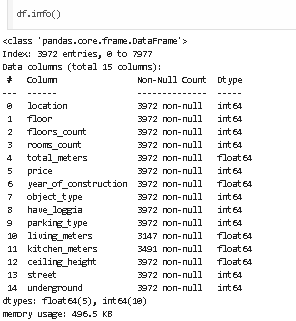


Рисунок



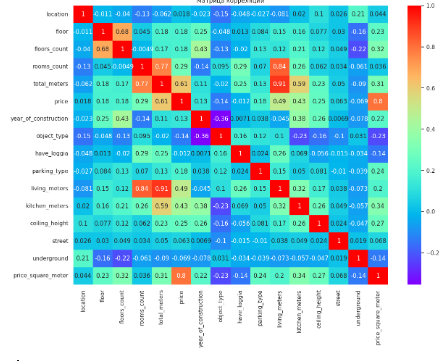
Рисунок

Меняем все возможные колонки с типа object на int или float и удаляем аномалии (Рисунок 6).



Рисунок

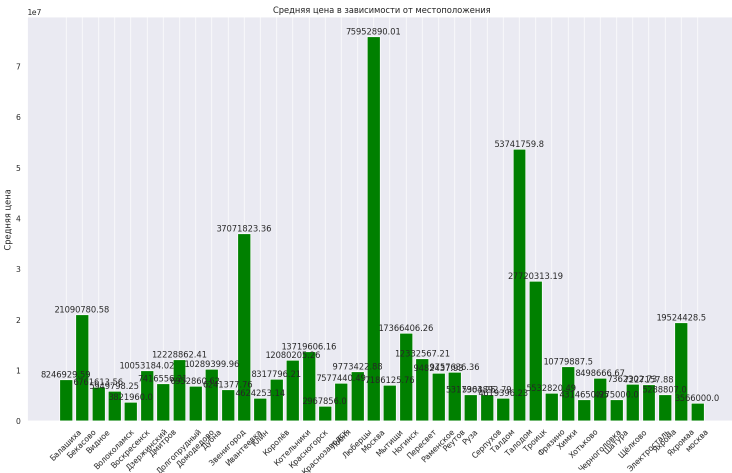
Окончательная часть создание матрицы корреляции (Рисунок 7)



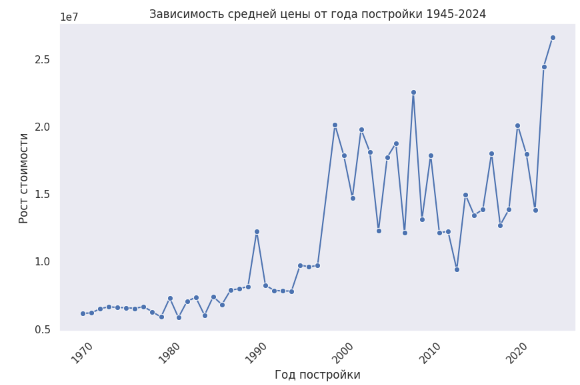
Рисунок

# Аналитика данных:

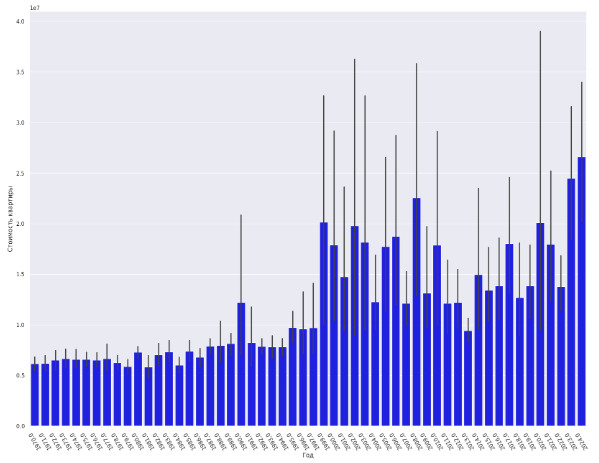
Благодаря выведенным графикам, можно сделать выводы, что цена в основном зависит от типа отделки, материала дома, города. От года постройки зданий зависит лишь их количество на рынке и количество комнат.



Рисунок



Рисунок



Рисунок

# Заключение:

В результате аналитической работы были собраны, отсортированы, почищены данные, простроены удобные для анализа графики, благодаря которым получилось выявить не маловажные критерии в оценивании стоимости недвижимости в Московской Области и Москве. Основными факторами, оказывающими влияние на стоимость, выявились общая площадь, жилая площадь, высоты потолков и количества комнат. Полученные данные могут быть использованы для дальнейшей разработки прогностических моделей.