Описание проекта

В вашем распоряжении данные сервиса Яндекс Недвижимость — архив объявлений о продаже квартир в Санкт-Петербурге и соседних населённых пунктах за несколько лет. Нужно научиться определять рыночную стоимость объектов недвижимости. Ваша задача — установить параметры. Это позволит построить автоматизированную систему: она отследит аномалии и мошенническую деятельность.

По каждой квартире на продажу доступны два вида данных. Первые вписаны пользователем, вторые — получены автоматически на основе картографических данных. Например, расстояние до центра, аэропорта, ближайшего парка и водоёма.

Инструкция по выполнению проекта

**Шаг 1**. **Откройте файл с данными и изучите общую информацию**

**Шаг 2. Предобработка данных**

1. Определите и изучите пропущенные значения:
   * Для некоторых пропущенных значений можно предположить логичную замену. Например, если человек не указал число балконов — скорее всего, их нет. Такие пропуски правильно заменить на 0. Для других типов данных нет подходящего значения на замену. В этом случае правильно оставить эти значения пустыми. Отсутствие значения — тоже важный сигнал, который не нужно прятать.
   * Заполните пропуски, где это уместно. Опишите, почему вы решили заполнить пропуски именно в этих столбцах и как выбрали значения.
   * Укажите причины, которые могли привести к пропускам в данных.
2. Приведите данные к нужным типам:
   * Поясните, в каких столбцах нужно изменить тип данных и почему.

**Шаг 3. Посчитайте и добавьте в таблицу**

* цену квадратного метра;
* день недели, месяц и год публикации [объявления](https://www.academy317.ru/mod/forum/view.php?id=809);
* этаж квартиры; варианты — первый, последний, другой;
* соотношение жилой и общей площади, а также отношение площади кухни к общей.

**Шаг 4. Проведите исследовательский анализ данных и выполните инструкции**

* Изучите следующие параметры: площадь, цена, число комнат, высота потолков. Постройте гистограммы для каждого параметра.
* Изучите время продажи квартиры. Постройте гистограмму. Посчитайте среднее и медиану. Опишите, сколько обычно занимает продажа.
* Уберите редкие и выбивающиеся значения. Опишите, какие особенности обнаружили.
* Какие факторы больше всего влияют на стоимость квартиры? Изучите, зависит ли цена от площади, числа комнат, удалённости от центра. Изучите зависимость цены от того, на каком этаже расположена квартира: первом, последнем или другом. Также изучите зависимость от даты размещения: дня недели, месяца и года. Постройте графики.
* Выберите 10 населённых пунктов с наибольшим числом объявлений. Посчитайте среднюю цену квадратного метра в этих населённых пунктах. Выделите населённые пункты с самой высокой и низкой стоимостью жилья. Эти данные можно найти по имени в столбце locality\_name.

**Шаг 5. Напишите общий вывод**

**Оформление:** Выполните задание в Jupyter Notebook. Заполните программный код в ячейках типа code, текстовые пояснения — в ячейках типа markdown. Примените форматирование и заголовки.

Описание данных

* airports\_nearest — расстояние до ближайшего аэропорта в метрах (м)
* balcony — число балконов
* ceiling\_height — высота потолков (м)
* cityCenters\_nearest — расстояние до центра города (м)
* days\_exposition — сколько дней было размещено объявление (от публикации до снятия)
* first\_day\_exposition — дата публикации
* floor — этаж
* floors\_total — всего этажей в доме
* is\_apartment — апартаменты (булев тип)
* kitchen\_area — площадь кухни в квадратных метрах (м²)
* last\_price — цена на момент снятия с публикации
* living\_area — жилая площадь в квадратных метрах (м²)
* locality\_name — название населённого пункта
* open\_plan — свободная планировка (булев тип)
* parks\_around3000 — число парков в радиусе 3 км
* parks\_nearest — расстояние до ближайшего парка (м)
* ponds\_around3000 — число водоёмов в радиусе 3 км
* ponds\_nearest — расстояние до ближайшего водоёма (м)
* rooms — число комнат
* studio — квартира-студия (булев тип)
* total\_area — площадь квартиры в квадратных метрах (м²)
* total\_images — число фотографий квартиры в объявлении

Отчет предоставить в виде дополненного файла Данные и ссылки на гитхаб. В папке проекта должны быть файл с выполненным заданием в формате юпитер ноутбук и дополненный файл с данными.

Также в ответ приложить ответ о лабораторной работе.