# Проект первого модуля: анализ данных для агентства недвижимости

**Автор:** Василенко Егор

**Дата: 16.11.2024-17.11.2024**

**Дата исправления: 17.11.2024**

## **Часть 2. Решаем ad hoc задачи**

### Задача 1. Время активности объявлений

Чтобы спланировать эффективную бизнес-стратегию на рынке недвижимости, заказчику нужно определить — по времени активности объявления — самые привлекательные для работы сегменты недвижимости Санкт-Петербурга и городов Ленинградской области.

Проанализируйте результаты и опишите их. Ответы на такие вопросы:

1. Какие сегменты рынка недвижимости Санкт-Петербурга и городов Ленинградской области имеют наиболее короткие или длинные сроки активности объявлений?

*Напишите ваш ответ здесь*

*Было поделено на 4 сегмента: интервал [1, 30] дней с названием «до месяца», [31, 90] дней с названием «до трех месяцев», интервал [91, 180] дней с названием «до полугода» и интервал [181, +∞] c названием «более полугода». Также был создан отдельный сегмент для каждого региона «активная публикация», в которую попадают публикации со значением NULL по столбцу days\_exposition, что означает их продолжающуюся активность (они не были проданы и объявления ещё висят на сайте). Соответственно, исходя из этой сегментации, наиболее короткие строки выполнения имеет сегмент «до месяца», то есть товар продавался в срок от 1 дня до 30 месяцев включительно. Наиболее длинным является сегмент «более полугода», то есть публикация была снята с сайта (в контексте данной задачи квартира была продана) в срок более, чем 181 день включительно.*

2. Какие характеристики недвижимости, включая площадь недвижимости, среднюю стоимость квадратного метра, количество комнат и балконов и другие параметры, влияют на время активности объявлений? Как эти зависимости варьируют между регионами?

*Напишите ваш ответ здесь*

*Если говорить про Санкт-Петербург, то чем меньше средняя стоимость квадратного метра, тем быстрее объявление снималось с сайта (продавалось); в Ленинградской области ситуация немного другая — наблюдается некоторая симметрия со средними стоимостями квадратного метра. Быстрее всего продаются квартиры с одной из самых дорогих стоимостей за квадратный метр в регионе, однако это стоимость все равно значительно ниже стоимости в том же сегменте Санкт-Петербурга. В принципе, в Ленинградской области конкретная зависимость от цен за квадратный метр отсутствует.*

*Что качается средней площади недвижимости, средней жилой площади, среднего количества комнат в квартире, то, чем меньше значение каждого из них, тем быстрее продавалась квартира как в Санкт-Петербурге, так и в Ленинградской области.*

*С другими показателями конкретной зависимости ни для Санкт-Петербурга, ни для Ленинградской области нет. Вероятно, они являются либо менее предпочтительными для покупателей, либо представлены в меньшем количестве по тем или иным причинам.*

3. Есть ли различия между недвижимостью Санкт-Петербурга и Ленинградской области по полученным результатам?

*Напишите ваш ответ здесь*

*Основные различия, в первую очередь, заключаются в стоимости жилья. Об этом нам говорит средняя стоимость квадратного метра, которая примерно на 30-40% процентов в Ленинградской области ниже, относительно стоимости в Санкт-Петербурге. Также наблюдается разница в средних площади недвижимости и жилой площади, которые незначительно, но выше именно в Санкт-Петербурге (хотя, вероятно, в контексте проживания, это и значительно).*

### Задача 2. Сезонность объявлений

Заказчику важно понять сезонные тенденции на рынке недвижимости Санкт-Петербурга и Ленинградской области — то есть для всего региона, чтобы выявить периоды с повышенной активностью продавцов и покупателей недвижимости. Это поможет спланировать маркетинговые кампании и выбрать сроки для выхода на рынок.

Проанализируйте результаты и опишите их. Ответы на такие вопросы:

1. В какие месяцы наблюдается наибольшая активность в публикации объявлений о продаже недвижимости? А в какие — по снятию? Это показывает динамику активности покупателей.

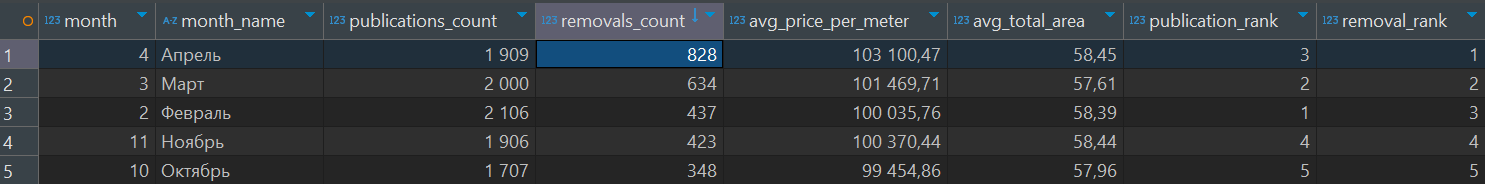
*Напишите ваш ответ здесь*

Наибольшая активность публикаций наблюдается в феврале, марте, апреле и ноябре (в остальные месяцы значения минимум на 200 меньше). Наибольшая активность по снятию публикаций наблюдается в апреле, марте, феврале ноябре, октябре и сентябре. Периоды активности приблизительно совпадают.

2. Совпадают ли периоды активной публикации объявлений и периоды, когда происходит повышенная продажа недвижимости (по месяцам снятия объявлений)?

*Напишите ваш ответ здесь*

*Периоды активности объявления и снятия не совпадают в точности, но совпадают приблизительно. В те месяцы, когда пользователи являются активными в плане публикаций, они также являются активными и в плане снятий объявлений (будь то либо продажа, либо простое снятие — в условии, по-моему, была продажа). Если мы возьмем два списка топ-5 по показателям публикации и снятию, то в них одинаковые месяцы.*

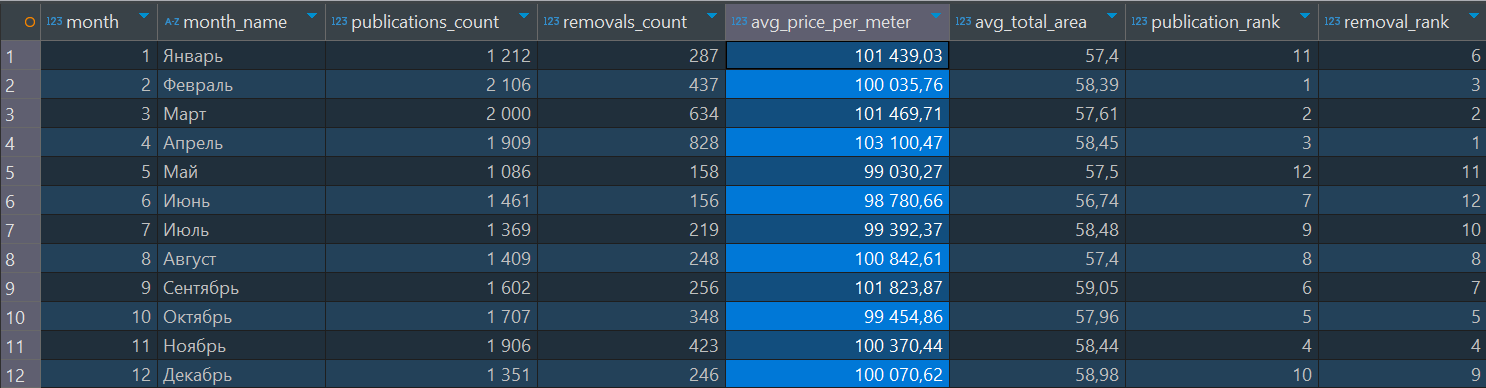
**

3. Как сезонные колебания влияют на среднюю стоимость квадратного метра и среднюю площадь квартир? Что можно сказать о зависимости этих параметров от месяца?

*Напишите ваш ответ здесь*

*Сезонные колебания влияют на среднюю стоимость квадратного метра и среднюю площадь квартир незначительно — значения находятся в пределах примерно одного значения с небольшими отклонениями. Для средней стоимости квадратного метра это где-то 100 тысяч рублей, для средней площади квартир это где-то 57-57,5 метров квадратных.*

*Это можно увидеть по выделенному столбцу и столбцу правее.*

**

### Задача 3. Анализ рынка недвижимости Ленобласти

Заказчик хочет определить, в каких населённых пунктах Ленинградской области активнее всего продаётся недвижимость и какая именно. Так он увидит, где стоит поработать, и учтёт особенности Ленинградской области при принятии бизнес-решений.

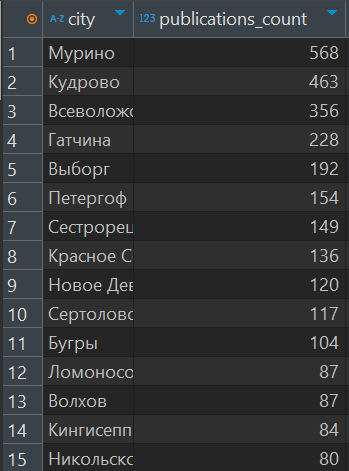
Проанализируйте результаты и опишите их. Ответы на такие вопросы:

1. В каких населённые пунктах Ленинградской области наиболее активно публикуют объявления о продаже недвижимости?

*Напишите ваш ответ здесь*

*Наиболее активно объявления публикуют в Мурино (568), Кудрово (463), Всеволожске (356), Гатчина (228), Выборг (192) наиболее активно публикуют объявления о продаже.*

*Весь топ 15:*

**

*Примечание: так как мне пришлось исправлять код, то я также учел дополнительно те населённые пункты, которые считаются частью Санкт-Петербурга, то есть к Колпино добавились Шушары, Пушкин, Парголово. Вероятно. Поэтому присутствует изменение списка лидеров, так как часть из них уехали по месту своей прописки.*

2. В каких населённых пунктах Ленинградской области — самая высокая доля снятых с публикации объявлений? Это может указывать на высокую долю продажи недвижимости.

*Напишите ваш ответ здесь*

*Самая высокая доля снятых с публикаций объявлений в Мурино (0,94), Кудрово (0,94), Ломоносове (0,92), Кингисеппе (0,92), Сестрорецке (0,9) и Красном Селе (0,9).*

3. Какова средняя стоимость одного квадратного метра и средняя площадь продаваемых квартир в различных населённых пунктах? Есть ли вариация значений по этим метрикам?

*Напишите ваш ответ здесь*

*Средняя стоимость одного квадратного метра значительно колеблется от одного населённого пункта у другому. В целом, в топ-15 после изменений и отправки нескольких населенных пунктов в Санкт-Петербург, значения стали ещё больше колебаться. Теперь это от 34 тысяч рублей в Волхове до 104 тысяч рублей в Сестрорецке.*

*Средняя площадь продаваемых квартир также может значительно колебаться в зависимости от населенного пункта. Значения лежат в интервале от 42,16 в Никольском до 62,45 в Сестрорецке, который также является первым по стоимости за квадратным метр.*

*Вариация значений по этим метрикам есть, но для конкретного анализа нужно, как минимум, найти по ним статистику.*

4. Среди выделенных населённых пунктов какие пункты выделяются по продолжительности публикации объявлений? То есть, где недвижимость продаётся быстрее, а где — медленнее.

*Напишите ваш ответ здесь*

*Для более удобного анализа я ввёл дополнительный столбце, который ранжируется показатели времени продажи по возрастанию и присваивает им ранги, также по возрастания. Опять же, после изменений он также изменился. Теперь на первом месте находится Кингисепп (75), именно в нём быстрее всего из топ-15 продаётся недвижимость. Дальше Мурино (90), Бугры (97), Кудрово (101) и Волхов (109). Медленно недвижимость, в рамках топ-15, продается в Никольском (167), Ломоносове (160), Сестрорецке (152), Красном Селе (147) и Петергофе (140).*

### Общие выводы и рекомендации

*Сделайте общий вывод-рекомендацию по полученным результатам.*

1. Стоит уточнить к чему конкретно относится населённый пункт, ибо из-за разделения ответы получаются не совсем корректными. Я постарался отделить те, что я знаю и заметил, однако, конечно, в таких ситуация, наверное, лучше иметь отдельную таблицу с регионом, которая связана с одной из таблиц;
2. Выбор сегмента недвижимости зависит от целей и позиционирования компании (предпринимателя). Для более дорого и элитного жилья подойдёт Санкт-Петербург, имеющий более высокие стоимости за квадратный метр недвижимости, что в теории может вылиться в доходы. В добавок, недвижимость в Санкт-Петербурге намного активнее продается, хотя явно и будет иметь больше конкурентов. Ленинградская область (и некоторые населенные пункты, занесенные туда) подойдут для развивающихся компаний, так как порог входа на рынок находится значительно ниже (задача 1);
3. Между месяцами не наблюдается особой разницы в цене, однако наибольшая активность происходит в феврале, марте, апреле и ноябре — именно они, вероятно, и буду оптимальным временем для выхода на рынок (задача 2);
4. В целом, на основе анализа можно рекомендовать сосредоточиться на развитии и продвижении бизнеса в таких населенных пунктах, как Мурино, Кудрово, Всеволожске и Гатчина, где наблюдает как высокая активность публикаций, так и высокая активность продаж недвижимости. Однако, стоит также учесть, что почти во всех населённых пунктах высокая доля снятия объявлений, что делает все эти пункты интересными точками для рассмотрения, за исключением, разве что, Волхова с долей в 0,78 (задача 3).