

# Анализ рынка первичной недвижимости в Санкт-Петербурге

Миролюбов Вениамин Русланович



# Введение и цель проекта

Рынок первичной недвижимости Санкт-Петербурга представлен тысячами объявлений, размещенных на различных интернет-платформах. Потенциальные покупатели и аналитики сталкиваются с рядом трудностей:

1. **Децентрализация информации**
2. **Отсутствие единого формата**
3. **Нет публичной аналитики**
4. **Неудобство поиска**

Разработка автоматизированной системы сбора, структурирования и анализа информации о квартирах в новостройках Санкт-Петербурга, с последующей визуализацией данных и интеграцией в Telegram-бота для пользователей.



# Технологии и инструменты

## Python

- основной инструмент для сбора и очистки данных

## SQL

- использовался для хранения данных, подключения средств визуализации и Telegram-бота

## Yandex DataLens

- средство визуализации данных

## Notion

- ведение документации

## GitHub

- хранение кода и публичное портфолио

## Telegram Bot API

- для реализации Telegram-бота

# Источник данных

Источником данных была выбрана доска объявлений Avito

Для сбора данных был применен парсинг с ручной выгрузкой (библиотеки BeautifulSoup и Selenium)

Извлекались: название ЖК, цена, цена за кв/м, местоположение (метро или населенный пункт), количество комнат, площадь, ссылка

Позже был добавлен застройщик путем ручного ввода

```
html = card.get_attribute('innerHTML')
soup = BeautifulSoup(html, 'html.parser')

jk_tag = soup.find('a', href=re.compile('/catalog/novostroyki'))
residential_complex = jk_tag.get_text(strip=True) if jk_tag else None

title_tag = soup.find('a', {'data-marker': 'item-title'})
title = title_tag.get_text(strip=True) if title_tag else ""
rooms = re.search(r'(\d+|Студия)-к', title)
rooms = rooms.group(1) if rooms else None

area = re.search(r'(\d+(?:[.]\d+)?)\s*м²', title)
area = area.group(1).replace('.', '') if area else None

price_meta = soup.find('meta', itemprop='price')
price = price_meta['content'] if price_meta else None

price_per_m2_tag = soup.find('p', string=re.compile('р. за м²'))
price_per_m2 = price_per_m2_tag.get_text(strip=True).replace('\xa0', ' ') if price_per_m2_tag else None

metro_tag = soup.find('a', {'data-marker': 'metro_link'})
metro = metro_tag.get_text(strip=True) if metro_tag else None

url = "https://www.avito.ru" + title_tag['href'] if title_tag else None
```



# Очистка и преобразование данных

## Очистка

- удаление дублей по уникальному URL
- приведение строк к числовому формату
- приведение числового значения к строке
- удаление лишних символов

## Преобразование

- ручное обновление информации местоположения у некоторых ЖК
- ручное добавление застройщиков

После цикла сбора и очистки данных я получил информацию по 4838 объектам первичной недвижимости



# Ключевые метрики для анализа и визуализации

Основные метрики для анализа по проекту:

- Средняя цена за квартиру
- Средняя цена за квартиру у застройщиков
- Средняя цена за квартиру по местоположению
- Количество предложений по выбранной комнатности
- Цена за квадратный метр по определенной комнатности

Дополнительно были рассчитаны средняя, медианная, максимальная и минимальная цена для каждого застройщика и местоположения. На основе этих данных мы напрямую можем посмотреть рейдж и амплитуду по выбранной комнатности.

```
# Группировка по жилому комплексу и расчет статистик
developer_price_stats = df.groupby('developer')['price'].agg(
    developer_average_price='mean',
    developer_median_price='median',
    developer_min_price='min',
    developer_max_price='max'
).round().astype('Int64').reset_index()
```

# Визуализация в Yandex DataLens

Анализ рынка первичной недвижимости в Санкт-Петербурге

Количество застройщиков

56

Количество жилых комплексов

170

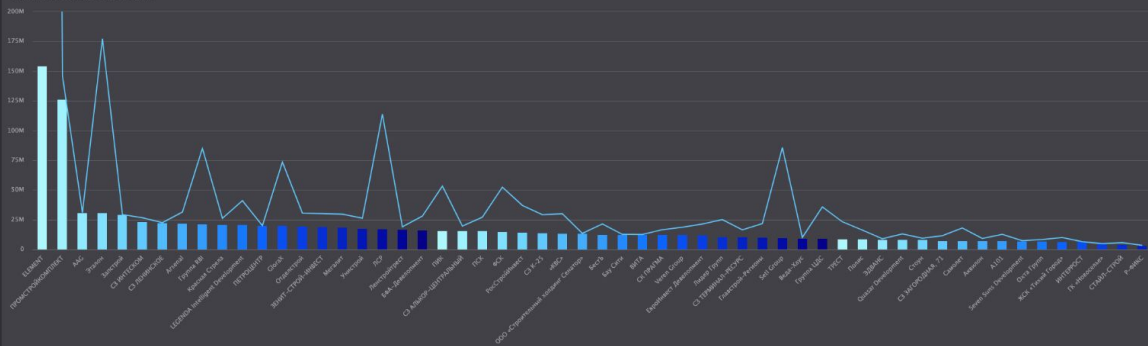
Средняя цена за квартиру

14 208 395 ₽

Средняя площадь квартир

49,9 кв/м

Цены на квартиры по застройщикам



Местоположение

Нет выбранных значений

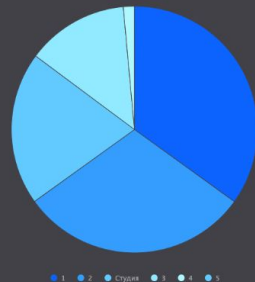
Застройщик

Нет выбранных значений

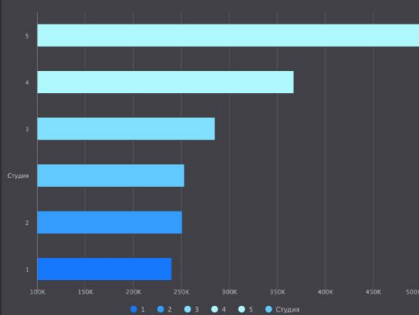
Количество комнат

Нет выбранных значений

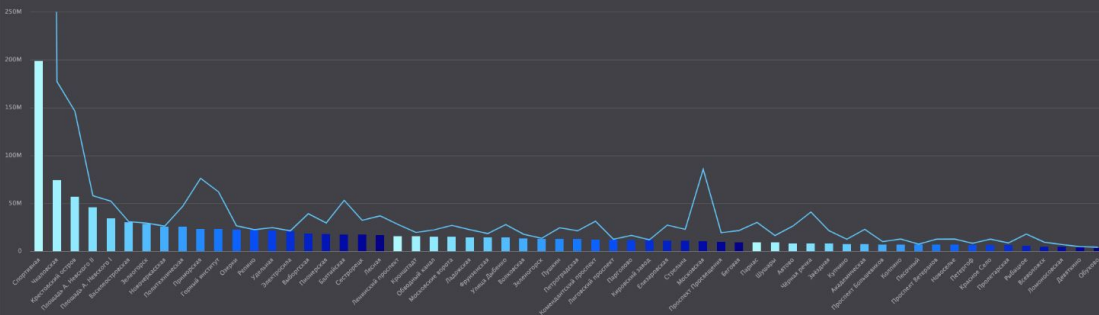
Количество предложений по комнатам



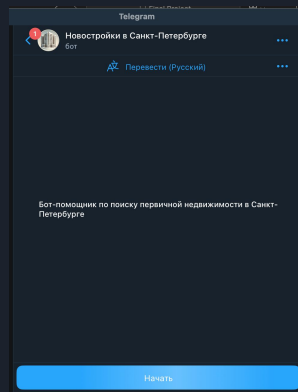
Распределение цен за квадратный метр



Цены на квартиры по местоположению



# Telegram-Бот



## Новостройки в Санкт-Петербурге

Выберите комнаты:

|              |   |   |
|--------------|---|---|
| Студия       | 1 | 2 |
| 3            |   |   |
| 4            |   |   |
| 5            |   |   |
| Продолжить ➔ |   |   |

21:45

Выберите район:

|                     |            |
|---------------------|------------|
| Приморский          | Выборгский |
| Калининский         |            |
| Красногвардейский   |            |
| Невский             |            |
| Центральный         |            |
| Василеостровский    |            |
| Адмиралтейский      |            |
| Петроградский       |            |
| Фрунзенский         |            |
| Московский          |            |
| Кировский           |            |
| Всеволожский        |            |
| Пригород            |            |
| Продолжить к цене ➔ |            |

Теперь выбери диапазон цен:

|           |
|-----------|
| до 5 млн  |
| 5-10 млн  |
| 10-15 млн |
| 15-25 млн |
| 25+ млн   |



ЖК: ЖК «Сандэй»

Площадь: 24.0 м²

Цена: 5934154 Р

Локация: Проспект Ветеранов

[Ссылка](#)

Авито

Квартира-студия, 24 м², 5/7 эт. на продажу в Санкт-Петербурге | Купить квартиру | Авито  
Квартира-студия, 24 м², 5/7 эт. продается в Санкт-Петербурге. Объявление на Авито. Предлагается квартира студия комфорт-класса 24,04 кв.м в корпусе 7.3 на 5 этаже, в ЖК «Сандэй» Квартира без отделки. Доступность опции "отделка" и возможность...

Выберите действие:

|              |
|--------------|
| Показать ещё |
| В начало     |

21:48

21:48





# Планы на будущее

1. Автоматизировать сбор, очистку и выгрузку данных
2. Интеграция с другими площадками объявлений
3. Добавление информации об отдаленности до ближайшей станции метро или ж/д станции
4. Поставить бота на непрерывную работу
5. Подключение к Telegram-боту ИИ-ассистента
6. Добавить интеграцию с картой
7. Добавить дополнительные (необязательные) фильтры
8. Добавить возможность ставить уведомления на обновление предложений по заданным фильтрам



# ССЫЛКИ

GitHub: [https://github.com/venyamir/project\\_github.git](https://github.com/venyamir/project_github.git)

Yandex DataLens: <https://datalens.yandex/arb8nzh3a00w>

Notion:

<https://inexpensive-jumbo-dd0.notion.site/1fc74626173c80d4af9bd7dc55ed21d5>

Telegram-бот: [https://t.me/spbapartspb\\_bot](https://t.me/spbapartspb_bot)