Отчет по интенсиву

**1. Введение**

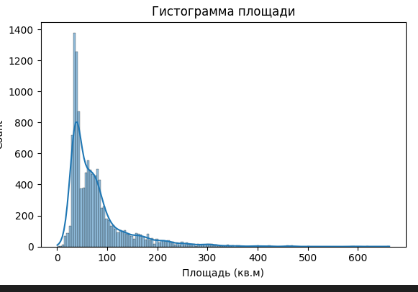
Данный отчёт посвящен анализу данных, полученных в результате парсинга объявлений о продаже квартир на сайте Циан. Целью исследования является построение модели, предсказывающей цену квадратного метра недвижимости на вторичном рынке Москвы, Новой Москвы и Московской области.

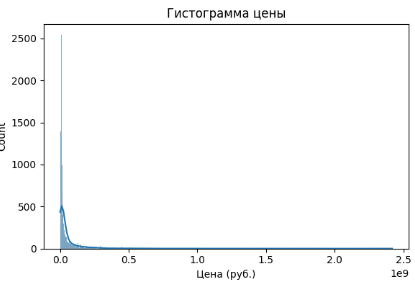
**2. Методы и данные**

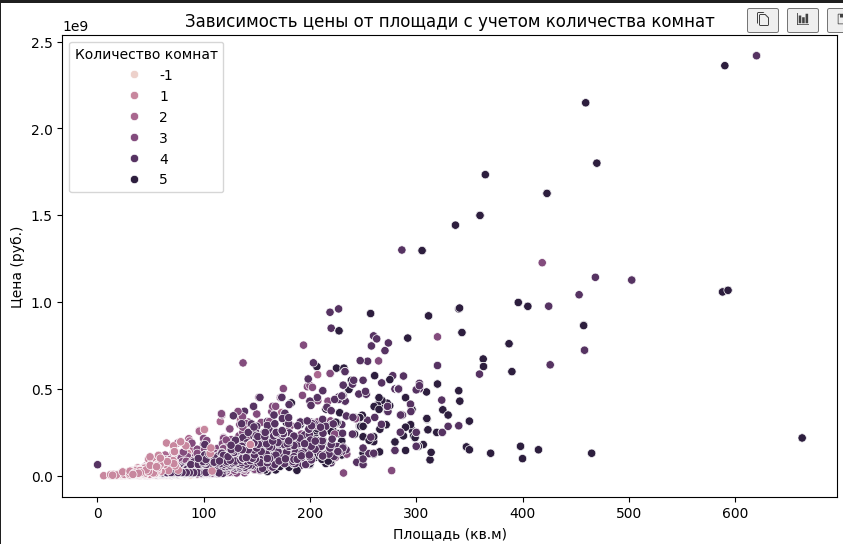
**2.1. Методы:**

* **Парсинг:** Для сбора данных использовалась библиотека cianparser **Обработка данных:** Данные обрабатывались с помощью библиотеки pandas для создания DataFrame
* **Визуализация:** Для визуализации данных использовались библиотеки matplotlib.pyplot и seaborn.

**2.2. Данные:**

* Данные были собраны с сайта Циан за период с 2023-04-28 по 2023-04-31 и включают в себя информацию о:
  + застройщик
  + Площадь
  + Количество комнат
  + Тип жилья (новостройка, вторичка)
  + Адрес
  + Район
  + Цена
  + Этаж
  + Тип дома
  + Станция метро
  + **3. Анализ данных**
  + **3.1. Исследовательский анализ данных (EDA)**
  + **3.2. Гистограмма этажности:** 
  + **3.3 Гистограмма площади:**
  + **3.4 Гистограмма цены**



**3.5 Зависимость цены от площади с учетом количества комнат**

**3.6 KDE распределение площади жилья**



**5. Рефлексия**

**5.1. Достижения:**

* Получение данных по продажам квартир в Московском регионе.
* Успешный анализ с использованием гистограмм.
* Визуализация данных с использованием различных графиков и диаграмм.

**5.2. Проблемы:**

* Были проблемы с созданием таблицы
* После определенного момента при парсинге, получил блокировку что замедлило сбор данных

**5.3. Дальнейшие шаги:**

* Исследовать влияние дополнительных факторов, таких как наличие парковки, отделки и т.д.
* Разработать более сложную модель, которая более точно отражает зависимость цены от факторов.