



# «Оценка рынка недвижимости в Торонто»

---

Выполнили:

Цепьюк Анастасия

Аникина Анна

# Цель

- Проанализировать рынок недвижимости в Торонто, Канада

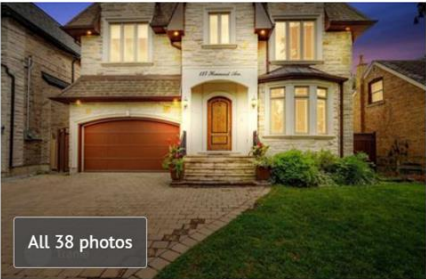
(найти и объяснить факторы, формирующие цены на жилье в городе)

# Данные:


- <https://tranio.com/canada/1892755/>


← → ↻

tranio.com/canada/adt1892755/



All 38 photos





Bedrooms Bathrooms

6 7

Check availability of the property

Send the listing to yourself

### Features

Car parks

Welcome To This Timeless Classic In West Willowdale. This Rare \*5 Bedroom\* 7 Bath Property Is Situated On A Large 50 X 139 Lot. Mainfloor Features 10 Ft Ceilings With Detailed Plaster Moulding Throughout. Chefs Delight Gourmet Kitchen Surrounds An Oversized Granite Island With A Kitchen-Aid Gas Cooktop And Miele Dishwasher. This Aaa Location Is A Short Walk To Yonge St, Shopping Centres, Goulding Park, And Newtonbrook School.


Townhome – North York, Toronto, Ontario, Canada

\$2,089,000

Bedrooms Bathrooms

6 7

Enquire about this townhome



Anna Boyarchukova  
Real Estate Expert Tranio in North York

Активация Windows

Чтобы активировать Windows, перейдите в раздел "Параметры".

Name

# Данные

## Зависимая переменная:

- Логарифм цены квартиры

## Ковариаты:

- Количество комнат в квартире, Количество ванных комнат
- Ковариаты, взятые из описания: 'Garage', 'Schools', 'Restaurants', 'Parking', 'Pool', 'Backyard', 'Closet', 'East', 'South'... - **(44 слова)**
- Улицы, где квартиры были выставлены на продажу: 'Richmond Street West', 'Roehampton Avenue', 'Saint Clair Avenue West', 'Scarborough'... - **(26 улиц)**

Всего наблюдений изначально: 12 548

Дата сбора данных:

01.12.2020

# Первоначальный анализ:

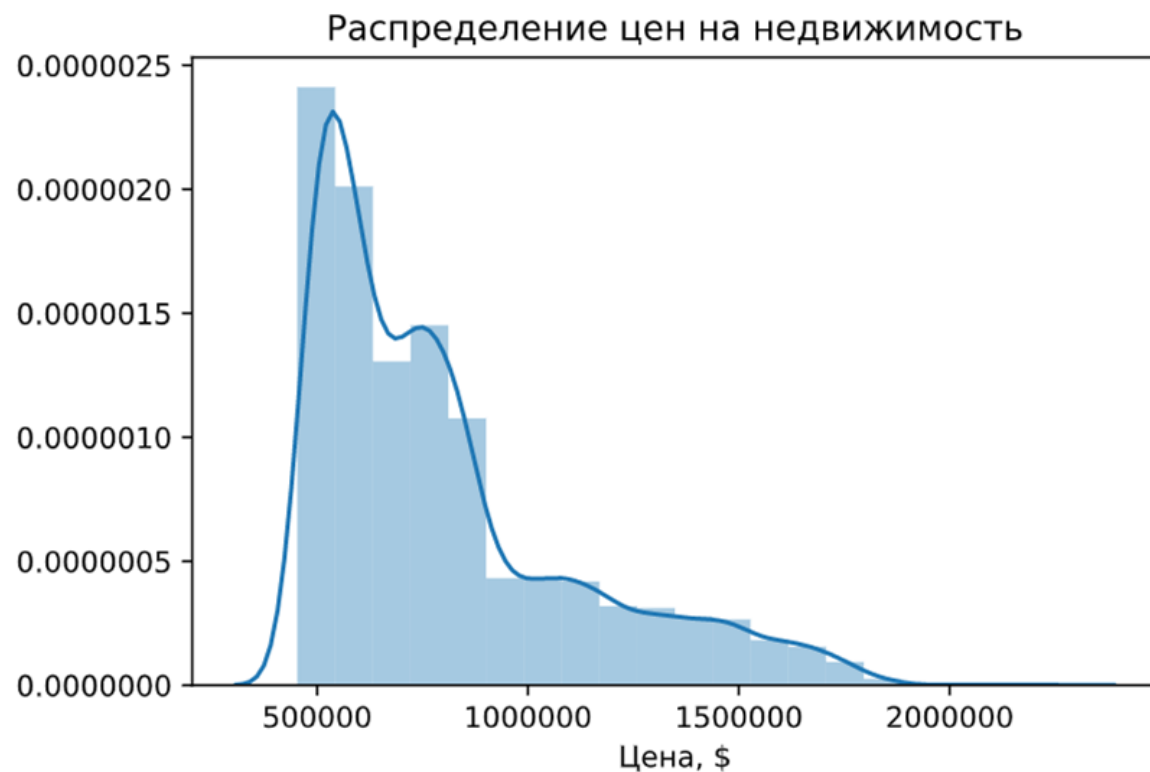


График сильно скошен вправо



График похож на сумму  
логнормальных распределений

# Первоначальный анализ:

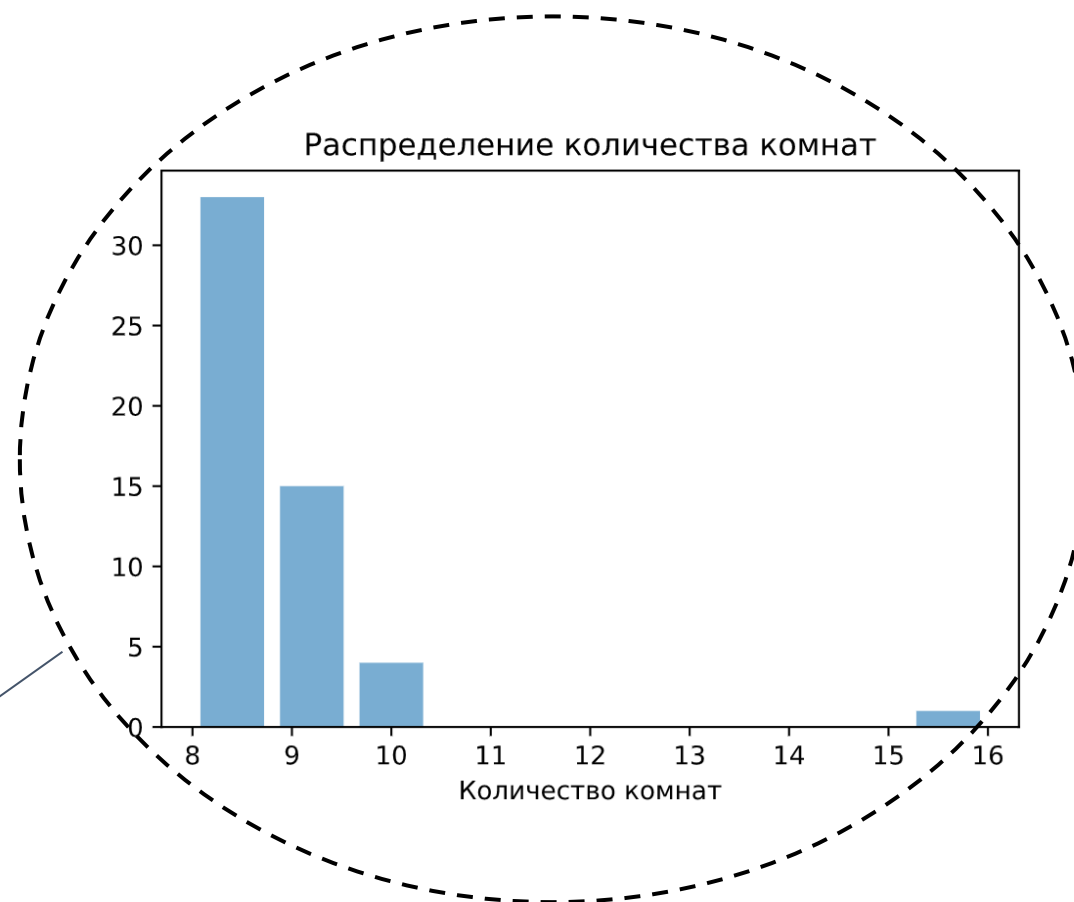
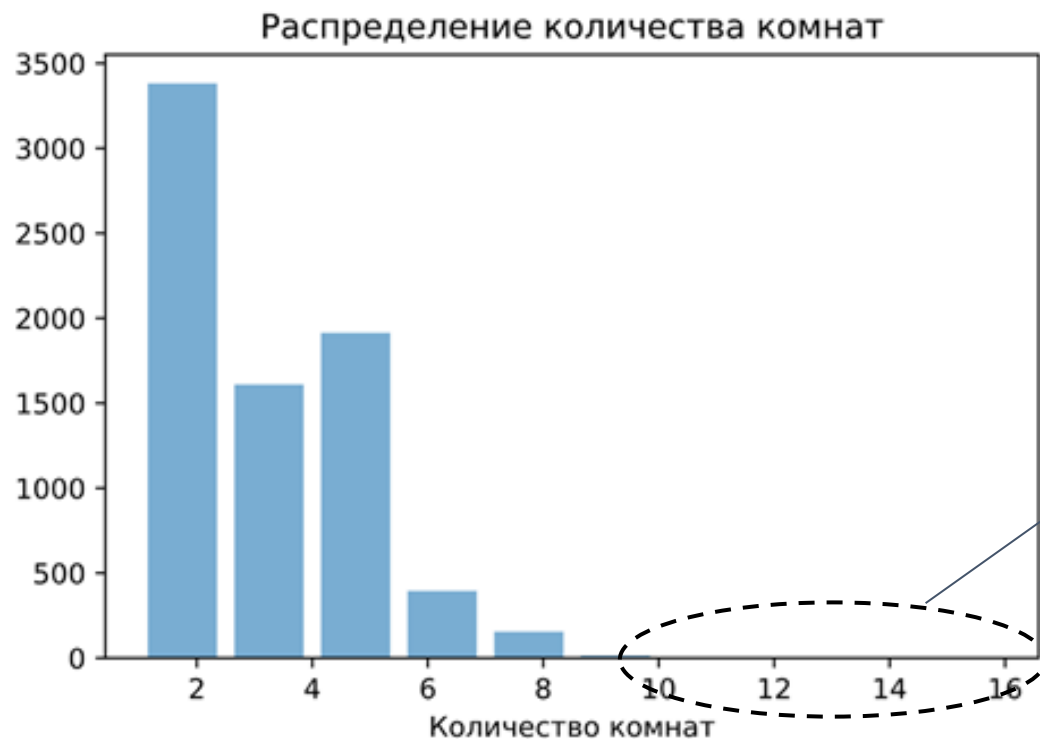
Описательная статистика для цены (в тысячах долларов)

	mean	std	min	25%	50%	75%	max
price	801.35342	315.292101	453.0	556.0	715.0	923.0	2244.0

Описательная статистика для логарифма цены

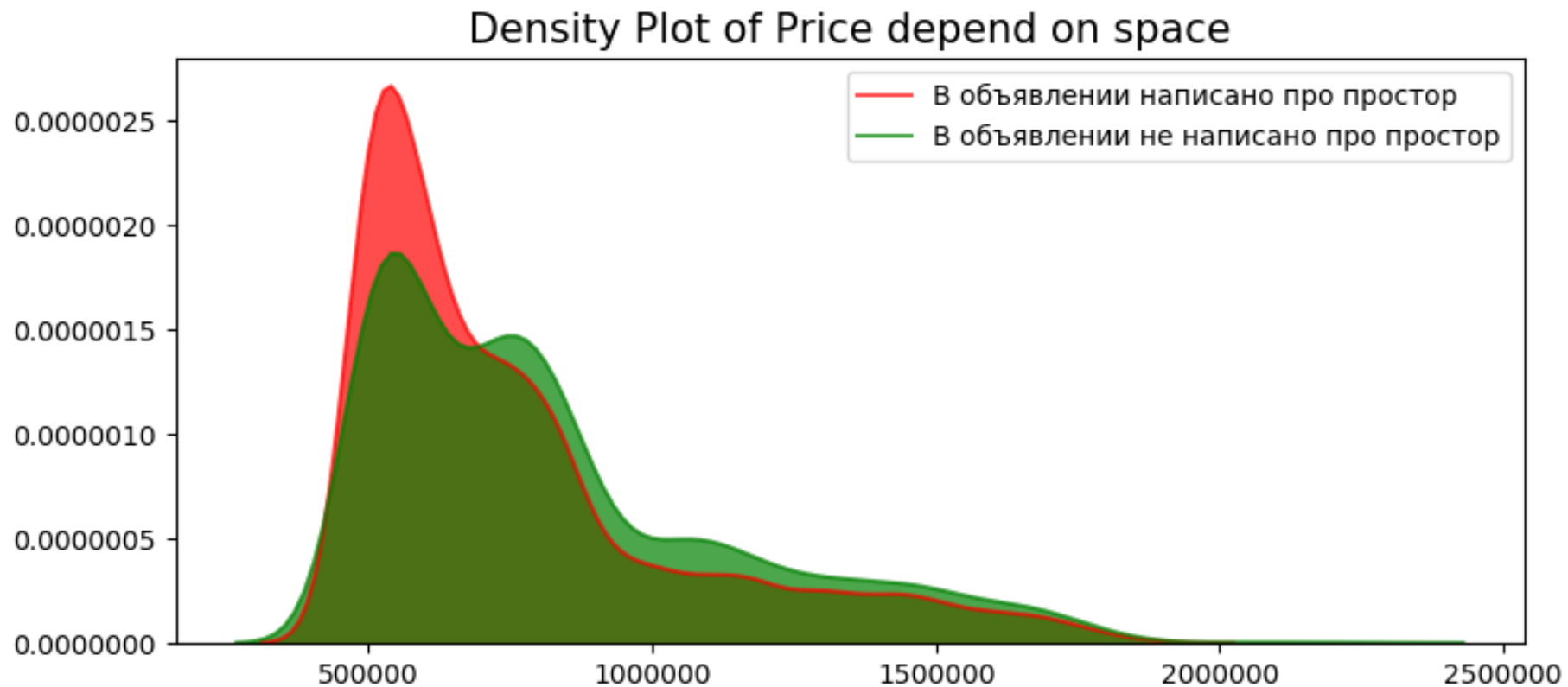
	mean	std	min	25%	50%	75%	max
log_price	13.528134	0.350736	13.023647	13.228524	13.480038	13.735385	14.623771

# Первоначальный анализ:



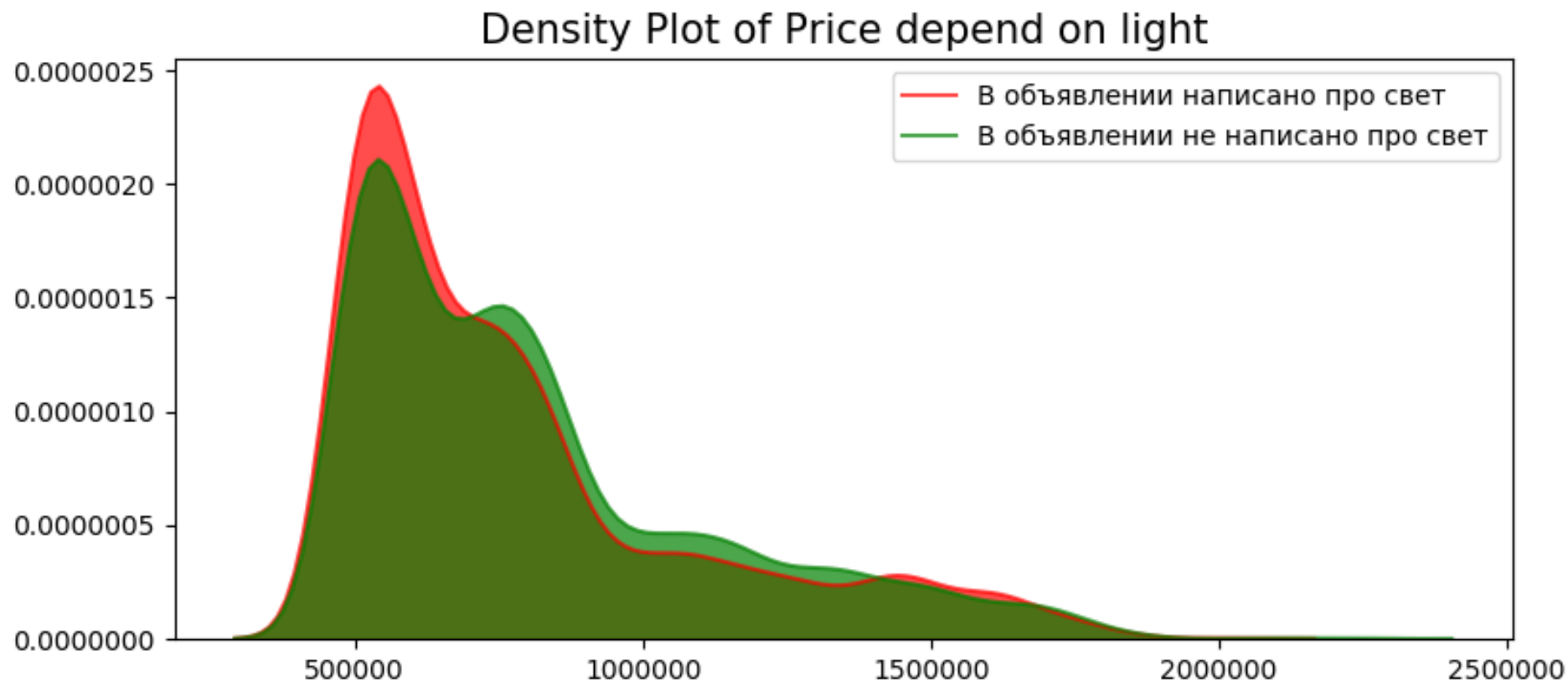
Были исключены редкие наблюдения

# Распределение цены в зависимости от слов в описании:

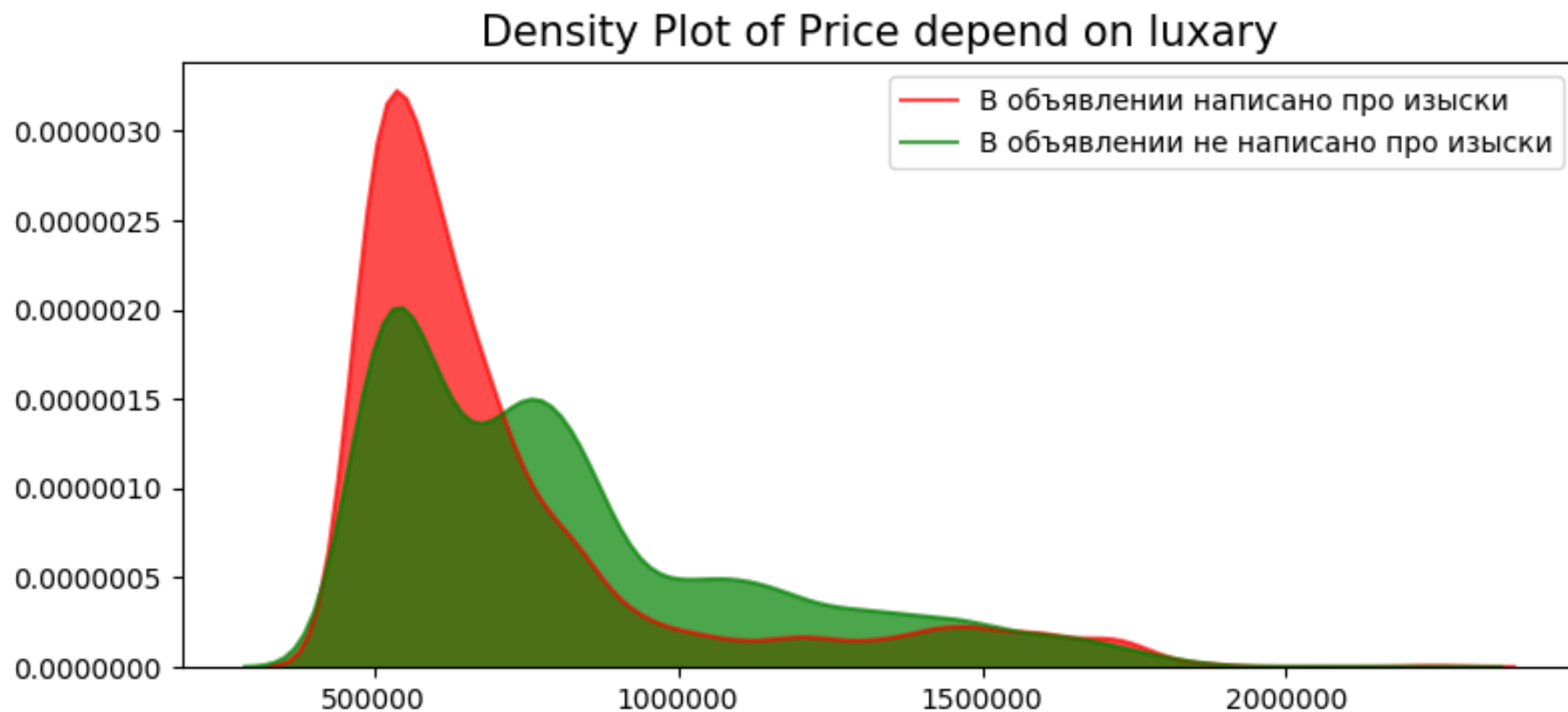




# Распределение цены в зависимости от слов в описании:



# Распределение цены в зависимости от слов в описании:



# Обработанные данные

	item_id	bathrooms	numberOfRooms	price	Amenities	Basement	Private	Garage	Schools	Transit	...	street_Roeampton Avenue	street_Saint Clair Avenue West	street_Sca
0	1855634	3	4	871000	1	1	1	1	0	0	...	0	0	
1	1855635	3	6	803000	0	1	0	0	1	1	...	0	0	
2	1855636	2	3	774000	0	0	1	0	0	0	...	0	0	
3	1855637	6	8	738000	1	0	0	0	0	0	...	0	0	
4	1855638	5	5	1809000	0	0	0	0	0	0	...	0	0	
5	1855639	5	5	1660000	0	1	0	0	0	0	...	0	0	
7	1855641	4	5	1614000	0	0	0	0	0	0	...	0	0	
8	1855642	4	3	1605000	0	0	1	0	0	1	...	0	0	
10	1855644	2	3	1692000	0	1	1	1	1	0	...	0	0	

Shape before preprocessing:

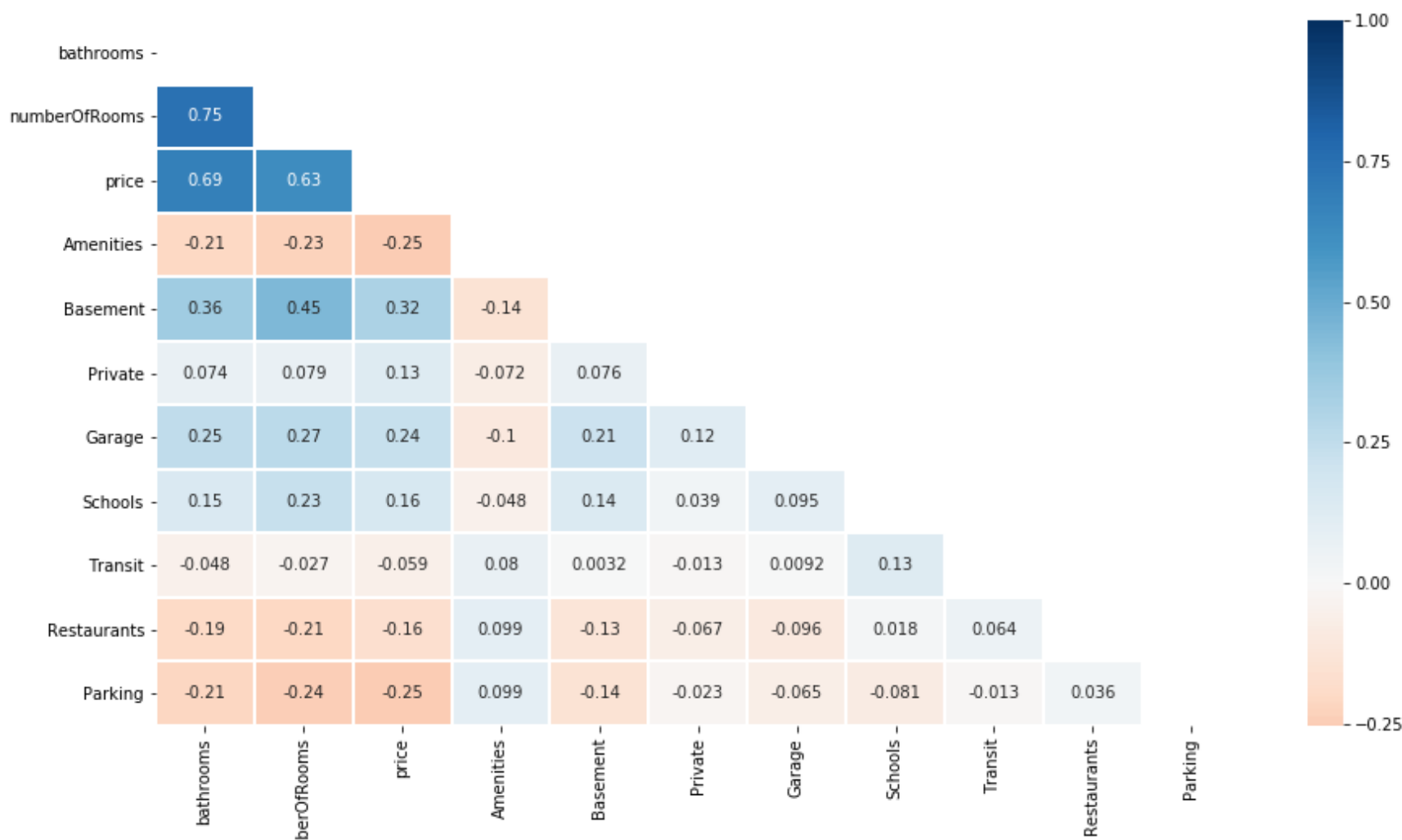
12548 rows × 127 columns



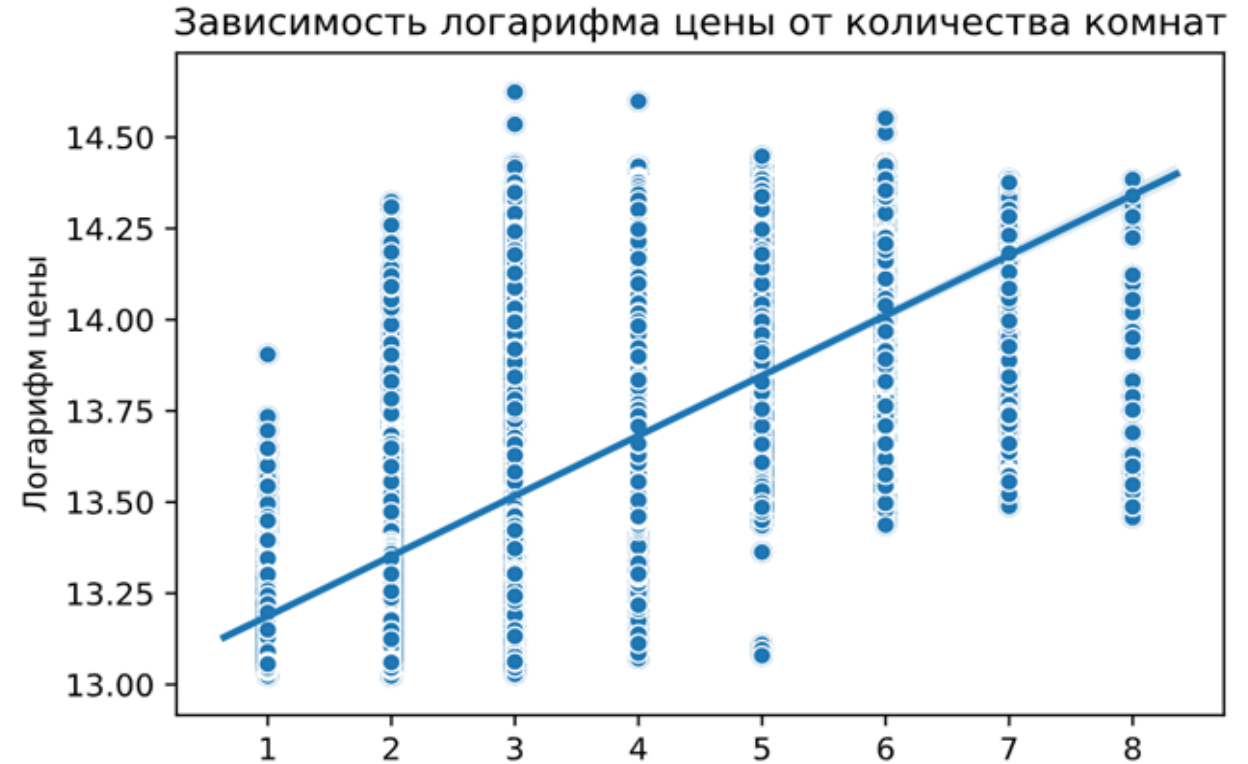
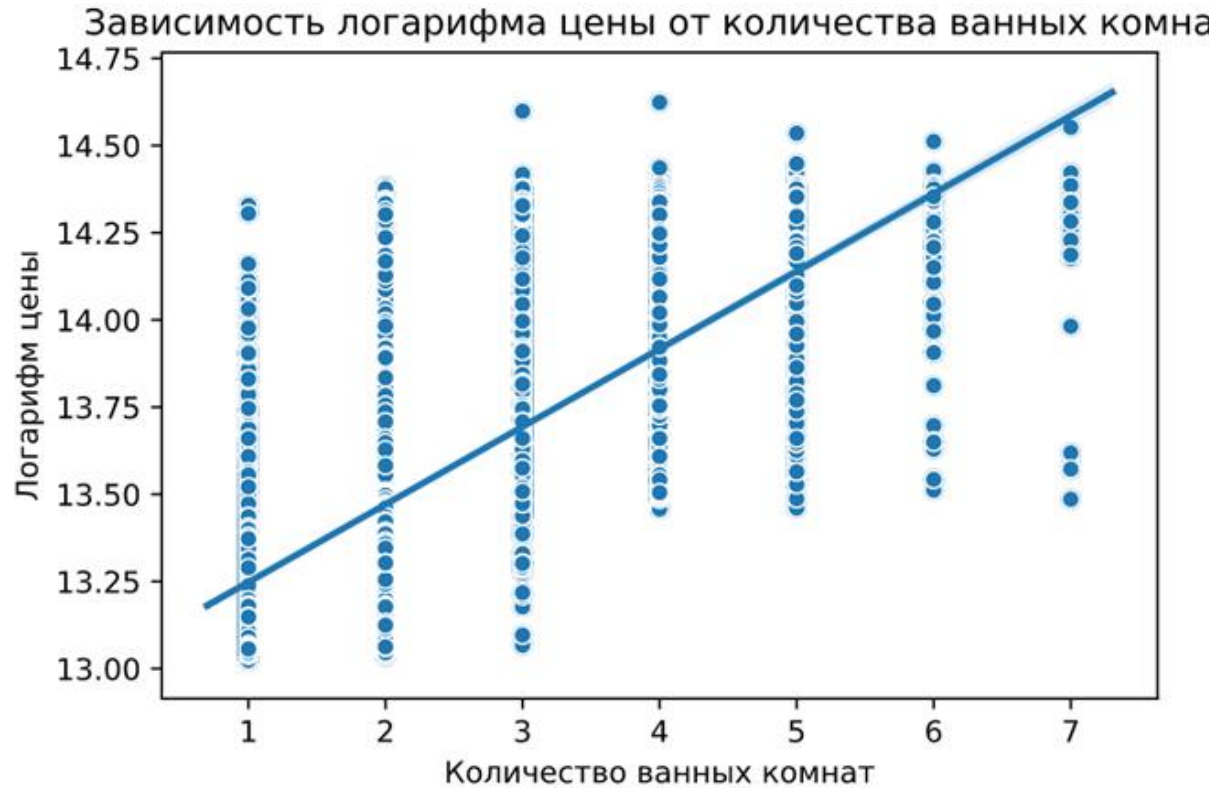
Shape after preprocessing:

7501 rows × 69 columns

# Корреляционная матрица :



# Зависимость логарифма цены от количества комнат и ванных комнат после предобработки и борьбы с выбросами:



## Модель:

В качестве эмпирической стратегии используется метод OLS:

$$\hat{Y}_i = \beta_0 + \sum_{j=1}^n \beta_j \cdot X_{j,i} + \epsilon_i, \text{ где}$$

$Y$  - логарифм цены на квартиру

$X$  - ковариаты

$\epsilon$  - случайная ошибка

## Результат оценивания методом OLS:

### OLS Regression Results

```
=====
Dep. Variable:          log_price    R-squared:                0.663
Model:                  OLS          Adj. R-squared:           0.660
Method:                 Least Squares  F-statistic:             206.8
Date:                  Fri, 11 Dec 2020  Prob (F-statistic):       0.00
Time:                  01:50:38       Log-Likelihood:          1260.0
No. Observations:      7328          AIC:                    -2380.
Df Residuals:          7258          BIC:                    -1897.
Df Model:               69
Covariance Type:       robust
=====
```

## Результаты оценивания после применения регуляризации LASSO:

Фактор	Коэффициент	Фактор	Коэффициент
Принадлежность к кондоминиуму	-0.060	Частная территория	+ 0.007
Наличие балкона	-0.053	Задний двор	+0.017
Улица Скарборо	-0.124	Наличие веранды	+0.018
“Отличный вид”	-0.049	Количество комнат	+0.073
Ламинат	-0.035	Количество ванных комнат	+0.131
“Пространство”	-0.009		
Тренажерный зал	-0.005		



## Дополнительные замечания:

- Учет фактора площади
- Учет макроэкономических факторов
- Учет пространственной связи: цены в соседних районах/улицах  
могут влиять друг на друга

Спасибо за внимание!