

```
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt

df = pd.read_csv('_data.csv', index_col=0)
df.head(10)
```

	ID	объявления	Количество	комнат	Тип	
Метро	\					
0	271271157		4	Квартира	м. Смоленская	(9 мин пешком)
1	271634126		4	Квартира	м. Смоленская	(8 мин пешком)
2	271173086	4, Оба варианта	Квартира	м. Смоленская	(7 мин пешком)	
3	272197456	4, Оба варианта	Квартира	м. Смоленская	(3 мин пешком)	
4	273614615		2	Квартира	м. Арбатская	(7 мин пешком)
5	274837728		3	Квартира	м. Смоленская	(5 мин пешком)
6	273643908	5, Оба варианта	Квартира	м. Арбатская	(4 мин пешком)	
7	274475342		3	Квартира	м. Смоленская	(3 мин пешком)
8	273973191		3	Квартира	м. Смоленская	(9 мин пешком)
9	272900409		4	Квартира	м. Арбатская	(10 мин пешком)

	Адрес	Площадь, м2	\
0	Москва, улица Новый Арбат, 27	200.0/20.0	
1	Москва, улица Новый Арбат, 27	198.0/95.0/18.0	
2	Москва, улица Новый Арбат, 27	200.0/116.0/4.0	
3	Москва, переулок Плотников, 21С1	170.0/95.0/17.0	
4	Москва, улица Новый Арбат, 15	58.0/38.0/5.0	
5	Москва, 1-й Смоленский переулок, 21	92.0	
6	Москва, переулок Романов, 3С1	213.0/140.0/20.0	
7	Москва, улица Арбат, 43С3	98.0/63.0/9.0	
8	Москва, Новинский бульвар, 18С1	120.0/95.0/10.0	
9	Москва, улица Арбат, 30/3С1	90.0/66.0/10.0	

	Дом	Парковка	\
0	5/16, Монолитный	подземная	
1	5/16, Монолитно-кирпичный	подземная	
2	5/16	подземная	
3	5/6	подземная	
4	12/26, Панельный	NaN	
5	3/7	NaN	

6	4/5	наземная
7	2/4, Монолитный	подземная
8	5/10, Сталинский	открытая
9	2/7, Сталинский	открытая

	Цена \	
0	500000.0 руб./	За месяц, Залог - 500000 руб., ...
1	500000.0 руб./	За месяц, Залог - 500000 руб., ...
2	500000.0 руб./	За месяц, Залог - 500000 руб., ...
3	400000.0 руб./	За месяц, Залог - 400000 руб., ...
4	225000.0 руб./	За месяц, Залог - 225000 руб., ...
5	470000.0 руб./	За месяц, Залог - 470000 руб., ...
6	350000.0 руб./	За месяц, Залог - 350000 руб., ...
7	250000.0 руб./	За месяц, Залог - 250000 руб., ...
8	130000.0 руб./	За месяц, Залог - 130000 руб., ...
9	210000.0 руб./	За месяц, Залог - 210000 руб., ...

	Телефоны ...		Окна \
0	+79166369231	...	NaN
1	+79850608590	...	На улицу и двор
2	+79672086536, +79099269384	...	На улицу и двор
3	+79660342340, +79099269384	...	На улицу и двор
4	+79852432860	...	На улицу и двор
5	+79684728732	...	NaN
6	+79175135482, +79099269384	...	На улицу и двор
7	+79152004882	...	Во двор
8	+79153429055	...	На улицу
9	+79850610614	...	На улицу и двор

	Санузел	Можно с детьми/животными
0	NaN	Можно с детьми, Можно с животными
1	Совмещенный (2), Раздельный (1)	Можно с детьми
2	Совмещенный (3)	Можно с детьми
3	Совмещенный (3)	Можно с животными
4	Совмещенный (2)	NaN
5	NaN	NaN
6	Совмещенный (2)	Можно с детьми, Можно с животными
7	Совмещенный (1), Раздельный (1)	Можно с детьми, Можно с животными
8	Совмещенный (1)	Можно с животными
9	Совмещенный (1), Раздельный (1)	Можно с детьми

	Дополнительно	Название
ЖК \		
0 Мебель в комнатах, Мебель на кухне, Ванна, Душ...	Новый Арбат ,	
2010		
1 Мебель в комнатах, Мебель на кухне, Ванна, Душ...	Новый	
Арбат		
2 Мебель в комнатах, Мебель на кухне, Ванна, Душ...	Новый	
Арбат		
3 Мебель в комнатах, Мебель на кухне, Ванна, Душ...		
NaN		
4 Мебель в комнатах, Мебель на кухне, Ванна, Душ...	The	
Book		
5 Мебель в комнатах, Мебель на кухне, Ванна, Душ...	Smolensky De	
Luxe		
6 Мебель в комнатах, Мебель на кухне, Ванна, Сти...		
NaN		
7 Мебель в комнатах, Мебель на кухне, Ванна, Душ...		
NaN		
8 Мебель на кухне, Ванна, Стиральная машина, Кон...		
NaN		
9 Мебель в комнатах, Мебель на кухне, Ванна, Сти...		
NaN		

	Серия дома	Высота потолков, м	Лифт	Мусоропровод	\
0	NaN	3.0	Пасс (4), Груз (1)	Да	
1	NaN	3.5	Пасс (1), Груз (1)	Нет	
2	NaN	3.2	Пасс (1)	NaN	
3	NaN	3.2	Пасс (1)	NaN	
4	NaN	3.9	Пасс (1), Груз (1)	Да	
5	NaN	NaN	NaN	NaN	
6	NaN	3.2	Пасс (1)	NaN	
7	NaN	3.2	Пасс (1)	Нет	
8	NaN	3.0	Пасс (1)	Нет	
9	NaN	310.0	Пасс (1)	Да	

	Ссылка на объявление
0	https://www.cian.ru/rent/flat/271271157
1	https://www.cian.ru/rent/flat/271634126
2	https://www.cian.ru/rent/flat/271173086
3	https://www.cian.ru/rent/flat/272197456
4	https://www.cian.ru/rent/flat/273614615
5	https://www.cian.ru/rent/flat/274837728
6	https://www.cian.ru/rent/flat/273643908
7	https://www.cian.ru/rent/flat/274475342
8	https://www.cian.ru/rent/flat/273973191
9	https://www.cian.ru/rent/flat/272900409

[10 rows x 24 columns]

Первый вывод dataset и сразу бросается в глаза тот факт (первые три записи), что Адрес различных объявлений одинаковый, также ближайшее метро по данному адресу одинаковое,
однако, время пути до метро в разных объявлениях отличается.

Пример

```
df.loc[:2, ['Метро', 'Адрес']]
```

Не будем предавать этому сильно влияющее значение, но признаку "Метро" стоит уделить некоторое время в Релизе2, чтобы осуществить обработку и вынести, например,
время до ближайшего метро (хотя сначала мы попытаемся узнать, действительно ли оно ближайшее) в отдельную feature, а также определиться с выбросами и различными
временными показателями пути до метро в случае одинакового адреса дома.

	Метро	Адрес
0	м. Смоленская (9 мин пешком)	Москва, улица Новый Арбат, 27
1	м. Смоленская (8 мин пешком)	Москва, улица Новый Арбат, 27
2	м. Смоленская (7 мин пешком)	Москва, улица Новый Арбат, 27

Узнаем объем данных, с которыми мы будем работать

```
count = df.shape
```

```
f'Всего мы имеем {count[0]} записей и {count[1]} признаков.'
```

```
'Всего мы имеем 23368 записей и 24 признаков.'
```

Выявим название наших features, которые мы имеем в DataFrame, чтобы понять, с чем мы работаем и что потенциально нам может пригодиться или, наоборот, не пригодится

```
df.columns.to_list()
```

Вывод: Уже можем понять, основные features, которые нам не будут нужны (не будут влиять на цену аренда квартиры), например, первый столбец "Unnamed: 0", можно будет

в Релизе2 сразу удалить. Также подобному критерию не соответствуют столбцы "Ссылка на объявление", "Телефоны", "Серия дома"

Также существует другая группа features, которая вызывает сомнения по поводу влияния на цену аренда квартиры, например, наличие Мусоропровода и других частей, причастных

к квартире (Окно, Балкон, Санузел, Высота потолков, Наличие и кол-во лифтов, длительность пути до Метро) и так далее

В Релизе 2 отберем нужные features, и займемся их очисткой

```
['ID объявления',  
 'Количество комнат',  
 'Тип',  
 'Метро',  
 'Адрес',
```

```
'Площадь, м2',
'Дом',
'Парковка',
'Цена',
'Телефоны',
'Описание',
'Ремонт',
'Площадь комнат, м2',
'Балкон',
'Окна',
'Санузел',
'Можно с детьми/животными',
'Дополнительно',
'Название ЖК',
'Серия дома',
'Высота потолков, м',
'Лифт',
'Мусоропровод',
'Ссылка на объявление']
```

```
# Посмотрим информацию о датасете
df.info()
```

```
# Уже можем увидеть, что количество записей во Фрейме 23368, а значит
# есть столбцы, с наименьшим количеством значений Non-null (не нулевое),
# что
# говорит о том, что в этих столбцах есть значения NaN
```

```
# Следующий вывод: тип объекта во всех столбцах (кроме трёх - один из
# них просто дополнительная строка индексации наших записей, её можно
# убрать) является object
# Этот тип данных является составным, то есть внутри него могут лежат
# различные типы данных, включая обычные str (строки)
# Значит мы уже должны понимать, что, например, столбец "Цена" мы
# должны обработать: избавиться от лишнего, либо какую-то важную
# информацию превратить в новую feature,
# которая сможет влиять на предсказание модели наилучшим образом
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 23368 entries, 0 to 23367
Data columns (total 24 columns):
```

#	Column	Non-Null Count	Dtype
0	ID объявления	23368 non-null	int64
1	Количество комнат	22327 non-null	object
2	Тип	23368 non-null	object
3	Метро	22053 non-null	object
4	Адрес	23368 non-null	object
5	Площадь, м2	23368 non-null	object
6	Дом	23368 non-null	object

7	Парковка	9951 non-null	object
8	Цена	23368 non-null	object
9	Телефоны	23368 non-null	object
10	Описание	23368 non-null	object
11	Ремонт	20613 non-null	object
12	Площадь комнат, м2	14458 non-null	object
13	Балкон	15390 non-null	object
14	Окна	16755 non-null	object
15	Санузел	20696 non-null	object
16	Можно с детьми/животными	17272 non-null	object
17	Дополнительно	23011 non-null	object
18	Название ЖК	5848 non-null	object
19	Серия дома	2163 non-null	object
20	Высота потолков, м	11206 non-null	float64
21	Лифт	17868 non-null	object
22	Мусоропровод	12846 non-null	object
23	Ссылка на объявление	23368 non-null	object

dtypes: float64(1), int64(1), object(22)

memory usage: 5.0+ MB

Выведем summary нашего датасета, однако, две наши величины не имеют никакого смысла в этом, глобальных представлений о датасете, они нам не дадут

`df.describe()`

В целом, что и требовалось доказать, единственная feature, которая потенциально сможет повлиять на цену аренды квартиры, является "Высота потолков"

	ID объявления	Высота потолков, м
count	2.336800e+04	11206.00000
mean	2.676265e+08	2.98345
std	1.876862e+07	7.61472
min	1.072986e+08	1.20000
25%	2.712663e+08	2.64000
50%	2.739030e+08	2.64000
75%	2.746872e+08	2.80000
max	2.750064e+08	320.00000

Попробуем вывести кол-во уникальных элементов каждой feature

`df.nunique()`

Как видим, ID-объявления полностью совпадает с общим количеством записей в датасете, значит каждая запись является уникальной, и мы не столкнемся с дубликатами.

Также feature "Количество комнат" заставляет нас обратить внимание на то, что там существуют странные элементы, так как по логике вещей, количество комнат

не должно превышать цифры 5-6. Далее обратим на это внимание.

ID объявления	23368
Количество комнат	24
Тип	1
Метро	5866
Адрес	12850
Площадь, м2	11255
Дом	2565
Парковка	5
Цена	3922
Телефоны	13172
Описание	23078
Ремонт	4
Площадь комнат, м2	3990
Балкон	18
Окна	3
Санузел	20
Можно с детьми/животными	3
Дополнительно	528
Название ЖК	1798
Серия дома	393
Высота потолков, м	95
Лифт	34
Мусоропровод	2
Ссылка на объявление	23368

dtype: int64

Посмотрим, какие существуют значения и их количество в каждой feature

```
for feat in df.columns[2:-1]:
    print(df[feat].value_counts(dropna=False))
    print()
    print()
```

Квартира 23368
Name: Тип, dtype: int64

NaN	1315
м. Академическая (10 мин пешком)	41
м. Водный стадион (5 мин пешком)	40
м. Приморская (None мин пешком)	35
м. Щелковская (15 мин пешком)	34
...	
м. Печатники (38 мин пешком)	1
м. Печатники (35 мин пешком)	1
м. Авиамоторная (1 мин пешком)	1
м. Фонвизинская (16 мин пешком)	1
м. Солнцево (5 мин на машине)	1

Name: Метро, Length: 5867, dtype: int64

Москва, Чапаевский переулок, 3
63
Москва, улица Василисы Кожиной, 13
46
Москва, улица 1-я Леонова, 18
42
Москва, Мосфильмовская улица, 8
37
Москва, Гостиничный проезд, 8К1
32

.
Краснодарский край, Сочи, Центральный мкр, улица Навагинская, 11А
1
Краснодарский край, Сочи, Центральный мкр, улица Нагорная, 19А
1
Краснодарский край, Сочи, Вишневая улица, 4
1
Краснодарский край, Сочи, Заречный мкр, улица Чайковского, 2
1
Москва, Боровский проезд, 11
1
Name: Адрес, Length: 12850, dtype: int64

40.0/20.0/10.0	224
40.0	116
45.0/30.0/6.0	110
60.0	91
45.0	88
...	
320.0/18.0	1
68.0/38.0/7.0	1
126.0/88.0/12.0	1
60.0/39.0/4.0	1
90.0/48.2/15.5	1

Name: Площадь, м2, Length: 11255, dtype: int64

3/5, Кирпичный	322
4/5, Кирпичный	296
2/5, Кирпичный	255
1/5, Кирпичный	232
5/5, Кирпичный	231
...	
19/35, Монолитный	1
8/35, Монолитно-кирпичный	1
12/49, Монолитный	1

13/34, Монолитный 1
12/25, Блочный 1
Name: Дом, Length: 2565, dtype: int64

NaN 13417
наземная 6043
подземная 2772
открытая 1017
многоуровневая 118
на крыше 1
Name: Парковка, dtype: int64

40000.0 руб./ За месяц, Залог - 40000 руб., Коммунальные услуги
включены, Срок аренды - Длительный, Предоплата 1 мес
1113

45000.0 руб./ За месяц, Залог - 45000 руб., Коммунальные услуги
включены, Срок аренды - Длительный, Предоплата 1 мес
1072

35000.0 руб./ За месяц, Залог - 35000 руб., Коммунальные услуги
включены, Срок аренды - Длительный, Предоплата 1 мес
959

50000.0 руб./ За месяц, Залог - 50000 руб., Коммунальные услуги
включены, Срок аренды - Длительный, Предоплата 1 мес
911

55000.0 руб./ За месяц, Залог - 55000 руб., Коммунальные услуги
включены, Срок аренды - Длительный, Предоплата 1 мес
678

...

100000.0 руб./ За месяц, Залог - 115000 руб., Коммунальные услуги
включены, Срок аренды - Длительный, Предоплата 1 мес
1

88000.0 руб./ За месяц, Сумма коммунальных платежей - 11000.0 руб.,
Залог - 88000 руб., Срок аренды - Длительный, Предоплата 1 мес 1

109999.0 руб./ За месяц, Залог - 119999 руб., Коммунальные услуги
включены, Срок аренды - Длительный, Предоплата 1 мес
1

77000.0 руб./ За месяц, Залог - 77000 руб., Коммунальные услуги
включены, Срок аренды - Длительный, Предоплата 2 мес
1

55000.0 руб./ За месяц, Залог - 50000 руб., Коммунальные услуги
включены, Срок аренды - Длительный, Предоплата 2 мес
1

Name: Цена, Length: 3922, dtype: int64

+79819217172 302
+74954760059 264

+79652883064	253
+79681933625	165
+79651095872	118

...

+79652256939	1
+79651093325	1
+79661963142	1
+79031261307	1
+79623641533	1

Name: Телефоны, Length: 13172, dtype: int64

0

108

Квартира сдаётся с Яндекс Арендой, а значит всё для вас:\n- без залога;\n- без единовременной комиссии;\n- с поддержкой от наших специалистов в процессе проживания.\n\nМы можем показать вам квартиру онлайн это так же детально, как вживую, только вы не тратите время на дорогу и запись возможна на ближайшее время., 14 фото

10

Квартира сдаётся с Яндекс Арендой, а значит всё для вас:\n- без залога;\n- без единовременной комиссии;\n- с поддержкой от наших специалистов в процессе проживания.\n\nМы можем показать вам квартиру онлайн это так же детально, как вживую, только вы не тратите время на дорогу и запись возможна на ближайшее время., 16 фото

9

Квартира сдаётся с Яндекс Арендой, а значит всё для вас:\n- без залога;\n- без единовременной комиссии;\n- с поддержкой от наших специалистов в процессе проживания.\n\nМы можем показать вам квартиру онлайн это так же детально, как вживую, только вы не тратите время на дорогу и запись возможна на ближайшее время., 15 фото

9

Квартира сдаётся с Яндекс Арендой, а значит всё для вас:\n- без залога;\n- без единовременной комиссии;\n- с поддержкой от наших специалистов в процессе проживания.\n\nМы можем показать вам квартиру онлайн это так же детально, как вживую, только вы не тратите время на дорогу и запись возможна на ближайшее время., 17 фото

9

...

Двухкомнатная квартира с качественным ремонтом. Кухня со всем необходимым, включая посудомоечную машину. 2 отдельные комнаты и санузел (ванна). Кровать можем убрать. Кондиционер. Телевизор. Интернет. Балкон. Красивый вид из окон. Окна во двор. Детская площадка и Британская международная школа во дворе. Магазин Перекрёсток. Парк., 15 фото

1

Покажу в любой день , две комнаты изолированные квартира трехкомнатная, 1 комната закрыта, но можно по желанию открыть.. Есть

балкон. Современный кухонный гарнитур. . В одной комнате двуспальная кровать, в другой угловой диван разборный. Залог делится., 5 фото

1

Квартира свободна к заселению. Изолированные комнаты. Прекрасное состояние. Балкон из большой комнаты, не остеклен. Залог можем разбить на 2 части., 18 фото

1

Ранее не сдавалась \nСдаётся 1 комнатная квартира на длительный срок в спальном районе города с развитой инфраструктурой\n\nКвартира обустроена всей современной мебелью и техникой для комфортного проживания, а именно:\nКухонный гарнитур, с плитой и духовым шкафом, обеденная группа, холодильник, микроволновая печь, стиральная машина автомат, кондиционер, телевизор, угловой диван-кровать, шкаф-гардероб зеркальный (большой шкаф, разными функциями хранения). Интернет-WIFI\n\nХороший ремонт, делали для себя. На окнах стеклопакеты, на полу плитка и ламинат. Металлическая сейфовая дверь. Хорошие соседи \n\nДом расположен в шаговой доступности 10 минут от метро, в тихом уголке рядом с парком. По пути всевозможные магазины, торгово развлекательный центр, кинотеатр, аптеки, банки, салоны красоты и спорта, кафе, общественные заведения, учебные заведения.\n\nУдобный быстрый выезд на МКАД, на Калужское шоссе, Симферопольское и Киевское шоссе\n\nКвартира готова к аренде с 2 июля\nДепозит можно разделить \nМосква, ..., 13 фото

1

Сдается просторная двухкомнатная квартира ВПЕРВЫЕ! КОММУНАЛЬНЫЕ ПЛАТЕЖИ СКИДКА 50%. Квартира находится в шаговой доступности от двух станций метро: Говорово и Солнцево. Имеется все необходимое для жизни! Кухня оснащена всем жизненно важным для комфортного быта: большой вместительный холодильник, посудомоечная машина, измельчитель для отходов, индукционная плита, духовка, посуда. В квартире в каждой комнате есть кондиционер. В ванной установлен бойлер, за счет чего можно не переживать за сезонное отключение горячей воды. Так же кроме вместительных шкафов для вещей имеется кладовая/гардеробная!

Просторная лоджия с большими окнами выходит в тихий и спокойный двор. Рядом с домом: несколько продуктовых магазинов, школы, детские сады и парк с прудом. Удобный съезд на Боровское шоссе. Удобная транспортная доступность до ЛУЧШИХ ВУЗОВ СТРАНЫ: РАНХИГС, МГИМО, РУДН, МГУ!, 16 фото

1

Name: Описание, Length: 23078, dtype: int64

Косметический 8499

Евроремонт 8470

Дизайнерский 3474

NaN 2755

Без ремонта 170

Name: Ремонт, dtype: int64

NaN 8910

20	1620
18	952
19	765
17	284
...	
17-10-12	1
30-30-25-30	1
13,6+18,9	1
22-10,5	1
26,2+15,5-22,0	1

Name: Площадь комнат, м2, Length: 3991, dtype: int64

NaN	7978
Балкон (1)	7428
Лоджия (1)	6007
Балкон (1), Лоджия (1)	716
Лоджия (2)	568
Балкон (2)	474
Балкон (3)	55
Лоджия (3)	45
Балкон (2), Лоджия (2)	25
Балкон (1), Лоджия (2)	24
Балкон (2), Лоджия (1)	20
Балкон (4)	6
Балкон (1), Лоджия (3)	5
Лоджия (4)	5
Балкон (3), Лоджия (1)	5
Балкон (2), Лоджия (3)	3
Балкон (1), Лоджия (4)	2
Балкон (3), Лоджия (3)	1
Балкон (4), Лоджия (4)	1

Name: Балкон, dtype: int64

Во двор	10870
NaN	6613
На улицу и двор	3295
На улицу	2590

Name: Окна, dtype: int64

Совмещенный (1)	10078
Раздельный (1)	7158
NaN	2672
Совмещенный (2)	1437
Совмещенный (1), Раздельный (1)	812
Раздельный (2)	534
Совмещенный (3)	241
Совмещенный (2), Раздельный (1)	188

Совмещенный (4)	77
Раздельный (3)	52
Совмещенный (1), Раздельный (2)	30
Совмещенный (3), Раздельный (1)	27
Совмещенный (2), Раздельный (2)	25
Раздельный (4)	15
Совмещенный (3), Раздельный (3)	6
Совмещенный (4), Раздельный (1)	6
Совмещенный (4), Раздельный (2)	4
Совмещенный (1), Раздельный (3)	2
Совмещенный (2), Раздельный (3)	2
Совмещенный (2), Раздельный (4)	1
Совмещенный (3), Раздельный (2)	1

Name: Санузел, dtype: int64

Можно с детьми	10134
Можно с детьми, Можно с животными	6899
NaN	6096
Можно с животными	239

Name: Можно с детьми/животными, dtype: int64

Мебель в комнатах, Мебель на кухне, Ванна, Стиральная машина, Телевизор, Холодильник, Интернет
2896

Мебель в комнатах, Мебель на кухне, Ванна, Стиральная машина, Холодильник, Интернет
1318

Мебель в комнатах, Мебель на кухне, Ванна, Стиральная машина, Кондиционер, Посудомоечная машина, Телевизор, Холодильник, Интернет
1042

Мебель в комнатах, Мебель на кухне, Ванна, Стиральная машина, Кондиционер, Телевизор, Холодильник, Интернет
967

Мебель в комнатах, Мебель на кухне, Ванна, Стиральная машина, Телевизор, Холодильник, Интернет, Телефон
887

...

Мебель на кухне, Ванна, Душевая кабина, Кондиционер, Посудомоечная машина, Интернет	1
Мебель в комнатах, Ванна, Душевая кабина, Кондиционер, Посудомоечная машина, Интернет	1
Интернет, Телефон	1
Мебель на кухне, Ванна, Стиральная машина, Кондиционер, Посудомоечная машина, Холодильник, Телефон	1
Мебель на кухне, Телевизор, Холодильник, Интернет	1

Name: Дополнительно, Length: 529, dtype: int64

NaN	17520
Символ, 2019	61
Водный, 2015	45
Квартал на Ленинском	37
Метрополия, 2021	36

...	
Помидор, 2000	1
Талисман на Дмитровском, 2022	1
LIFE-Волжская, 2012	1
Грани, 2020	1
В Солнцево	1

Name: Название ЖК, Length: 1799, dtype: int64

NaN	21205
индивидуальный проект	615
Индивидуальный проект	139
П44	89
П44Т	63

...	
I-510 (I-510/23БИ)	1
КПД-4572А	1
Монолитный	1
П44т	1
П 44-Т	1

Name: Серия дома, Length: 394, dtype: int64

NaN	12162
2.64	4467
3.00	1322
2.70	1040
2.48	676

...	
265.00	1
4.15	1
9.00	1
260.00	1
3.02	1

Name: Высота потолков, м, Length: 96, dtype: int64

Пасс (1)	5911
NaN	5500
Пасс (2)	4326
Пасс (1), Груз (1)	3962
Пасс (2), Груз (1)	1224

Пасс (2), Груз (2)	654
Пасс (3)	636
Пасс (4)	346
Пасс (3), Груз (1)	187
Пасс (1), Груз (2)	185
Груз (1)	95
Пасс (4), Груз (2)	65
Пасс (4), Груз (1)	64
Пасс (3), Груз (3)	45
Пасс (3), Груз (2)	44
Пасс (4), Груз (4)	28
Груз (4)	25
Груз (2)	20
Груз (3)	15
Пасс (4), Груз (3)	7
Пасс (6)	6
Пасс (1), Груз (3)	6
Пасс (2), Груз (3)	3
Пасс (60)	2
Пасс (50)	2
Пасс (1), Груз (12)	1
Пасс (5), Груз (1)	1
Пасс (8), Груз (8)	1
Пасс (7)	1
Пасс (1), Груз (4)	1
Пасс (5), Груз (3)	1
Пасс (2), Груз (4)	1
Груз (6)	1
Груз (8)	1
Пасс (5)	1

Name: Лифт, dtype: int64

Да	10897
NaN	10522
Нет	1949

Name: Мусоропровод, dtype: int64

```
d_nan = {}

for feat in df.columns[2:-1]:
    if df[feat].isna().sum() != 0:
        d_nan[feat] = df[feat].isna().sum()

d_nan
```

```
{'Метро': 1315,
 'Парковка': 13417,
 'Ремонт': 2755,
```

```
'Площадь комнат, м2': 8910,  
'Балкон': 7978,  
'Окна': 6613,  
'Санузел': 2672,  
'Можно с детьми/животными': 6096,  
'Дополнительно': 357,  
'Название ЖК': 17520,  
'Серия дома': 21205,  
'Высота потолков, м': 12162,  
'Лифт': 5500,  
'Мусоропровод': 10522}
```

Построим гистограмму для данного словаря, чтобы визуально увидеть feature и соответствующее им кол-во значений NaN

```
from matplotlib.ticker import MultipleLocator
```

```
fig = plt.figure(figsize=(10, 7))
```

```
ax = fig.add_subplot()
```

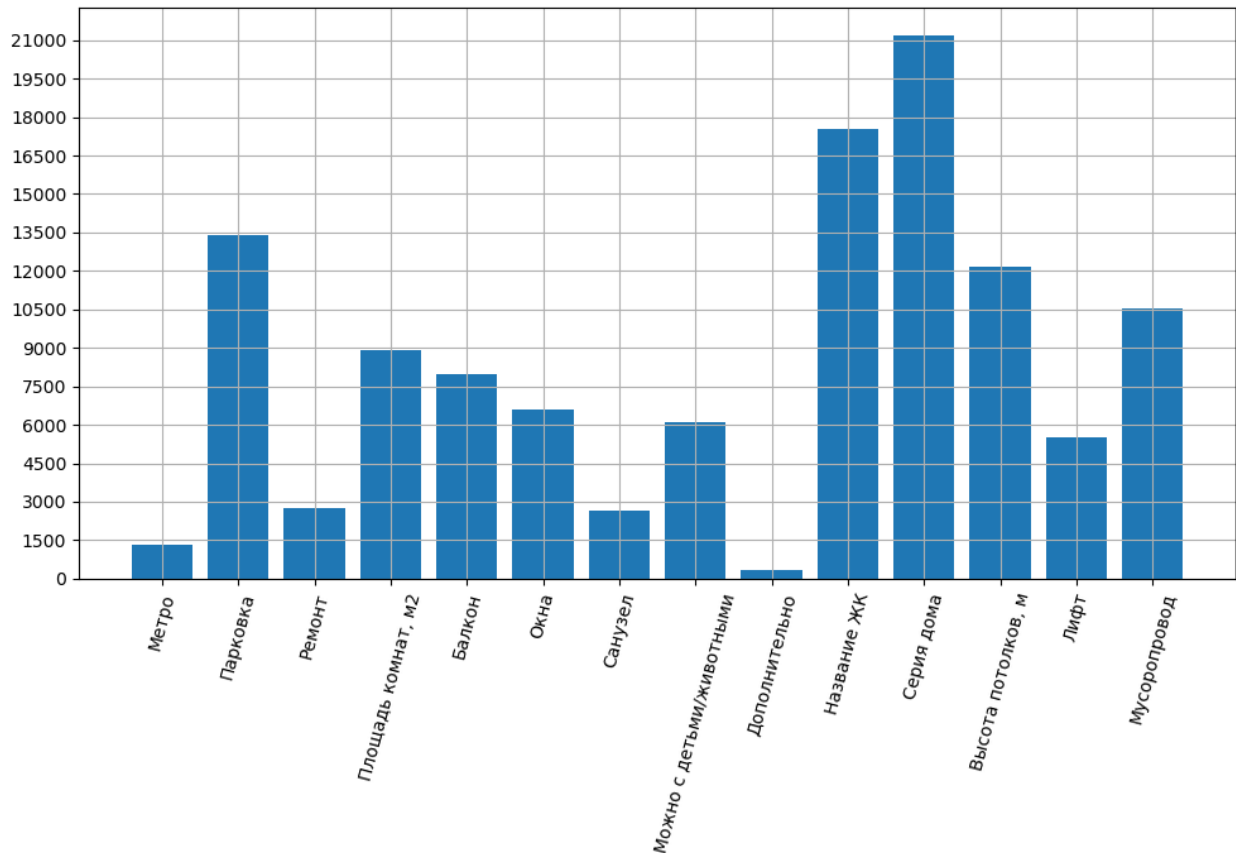
```
ax.bar(d_nan.keys(), d_nan.values());
```

```
ax.grid()
```

```
ax.yaxis.set_major_locator(MultipleLocator(base=1500))
```

```
plt.xticks(rotation=75)
```

```
plt.tight_layout();
```

```
# И другой вариант визуализации пропущенных данных
plt.figure(figsize=(12, 7))
sns.heatmap(df.isnull(), cbar=False, cmap='viridis',
yticklabels=False)
plt.title("Пропущенные значения по признакам")
plt.show()
```

Пропущенные значения по признакам

