

```
In [27]: import pandas as pd
import numpy as np
```

Загружаем датасет

```
In [5]: data = pd.read_csv('products_sentiment_train.csv', delimiter=',')
data.head()
```

```
Out[5]:
```

	N	name	price	url	pages	size1	size2	size3
0	1	Чисто российское преступление: Самые громкие и...	540.0	https://www.chitai-gorod.ru/product/chisto-ros...	298.0	1.7	14.6	21.6
1	2	Безлюдное место: Как ловят маньяков в России	539.0	https://www.chitai-gorod.ru/product/bezlyudnoe...	264.0	1.4	14.0	21.0
2	3	Письма о добром и прекрасном	238.0	https://www.chitai-gorod.ru/product/pisma-o-do...	192.0	1.5	11.6	18.1
3	4	У холмов есть подкаст. 24 истории о серийных у...	715.0	https://www.chitai-gorod.ru/product/u-holmov-e...	448.0	2.7	13.0	20.7
4	5	Фундаментальные законы человеческой глупости	605.0	https://www.chitai-gorod.ru/product/fundamenta...	160.0	1.8	13.3	17.3



```
In [6]: data.columns
```

```
Out[6]: Index(['N', 'name', 'price', 'url1', 'pages', 'size1', 'size2', 'size3',  
              'weight', 'cover', 'year', 'description'],  
              dtype='object')
```

```
In [7]: data.dtypes
```

Out[7]:

0

N	int64
name	object
price	float64
url	object
pages	float64
size1	float64
size2	float64
size3	float64
weight	int64
cover	object
year	float64
description	object

dtype: object

Очищаем датасет от строк с неопределенной ценой книг

```
In [11]: data = data[data['price'].notna()]
```

```
In [12]: data
```

Out[12]:

	N	name	price	url	pages	size1	size2
0	1	Чисто российское преступление: Самые громкие и...	540.0	https://www.chitai-gorod.ru/product/chisto-ros...	298.0	1.7	14.0
1	2	Безлюдное место: Как ловят маньяков в России	539.0	https://www.chitai-gorod.ru/product/bezlyudnoe...	264.0	1.4	14.0
2	3	Письма о добром и прекрасном	238.0	https://www.chitai-gorod.ru/product/pisma-o-do...	192.0	1.5	11.0
3	4	У холмов есть подкаст. 24 истории о серийных у...	715.0	https://www.chitai-gorod.ru/product/u-holmov-e...	448.0	2.7	13.0
4	5	Фундаментальные законы человеческой глупости	605.0	https://www.chitai-gorod.ru/product/fundamenta...	160.0	1.8	13.0
...
1141	1142	Нечто: сборник статей	773.0	https://www.chitai-gorod.ru/product/nechto-sbo...	188.0	1.2	15.0
1142	1143	Корпорацию не нагнешь	844.0	https://www.chitai-gorod.ru/product/korporaciy...	288.0	2.0	13.0
1143	1144	Европа и Евразия	1149.0	https://www.chitai-gorod.ru/product/evropa-i-e...	346.0	2.0	14.0
1144	1145	Исповедь для побежденных : После шестнадцати м...	1989.0	https://www.chitai-gorod.ru/product/ispoved-dl...	346.0	2.0	14.0
1149	1150	Репетиция апокалипсиса. Тайная история атомных...	590.0	https://www.chitai-gorod.ru/product/repeticiya...	198.0	1.5	14.0

1145 rows × 12 columns



Расчитываем показатели удельной цены по объему книги (высота ширина - толщина), по весу, по количеству страниц

```
In [13]: data['volume'] = data['size1']*data['size2']*data['size3']
```

```
In [14]: data['volume_price']=data['price']/data['volume']
```

```
In [15]: data['weight_price']=data['price']/data['weight']
```

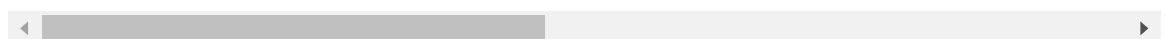
```
In [16]: data['page_price']=data['price']/data['pages']
```

```
In [17]: data
```

Out[17]:

	N	name	price	url	pages	size1	size2
0	1	Чисто российское преступление: Самые громкие и...	540.0	https://www.chitai-gorod.ru/product/chisto-ros...	298.0	1.7	14.0
1	2	Безлюдное место: Как ловят маньяков в России	539.0	https://www.chitai-gorod.ru/product/bezlyudnoe...	264.0	1.4	14.0
2	3	Письма о добром и прекрасном	238.0	https://www.chitai-gorod.ru/product/pisma-o-do...	192.0	1.5	11.0
3	4	У холмов есть подкаст. 24 истории о серийных у...	715.0	https://www.chitai-gorod.ru/product/u-holmov-e...	448.0	2.7	13.0
4	5	Фундаментальные законы человеческой глупости	605.0	https://www.chitai-gorod.ru/product/fundamenta...	160.0	1.8	13.0
...
1141	1142	Нечто: сборник статей	773.0	https://www.chitai-gorod.ru/product/nechto-sbo...	188.0	1.2	15.0
1142	1143	Корпорацию не нагнешь	844.0	https://www.chitai-gorod.ru/product/korporaciy...	288.0	2.0	13.0
1143	1144	Европа и Евразия	1149.0	https://www.chitai-gorod.ru/product/evropa-i-e...	346.0	2.0	14.0
1144	1145	Исповедь для побежденных : После шестнадцати м...	1989.0	https://www.chitai-gorod.ru/product/ispoved-dl...	346.0	2.0	14.0
1149	1150	Репетиция апокалипсиса. Тайная история атомных...	590.0	https://www.chitai-gorod.ru/product/repeticiya...	198.0	1.5	14.0

1145 rows × 16 columns



Определяем значения персентилей для удельных показателей (5%)

```
In [41]: v1 = np.percentile (data.volume_price , 95)
```

```
In [42]: v2 = np.percentile (data.page_price , 95)
```

```
In [43]: v3 = np.percentile (data.weight_price , 95)
```

```
In [44]: v1, v2, v3
```

```
Out[44]: (3.511460632659837, 7.244578313253011, 5.198758482030658)
```

Ищем книги, попадающие в выборку 5% с наивысшими значениями по всем трем
удельным показателям цены

```
In [45]: mask = (data['volume_price'].values > v1) & (data['page_price'].values > v2) & (
```

```
In [46]: df_new = data.loc[mask]
```

Выводим список самых "дорогих" книг, отсортированный по показателю цена/
объем, сохраняем в отдельный файл

```
In [47]: df_new.sort_values(by=['volume_price'], ascending=False)
```

Out[47]:

	N	name	price	url	pages	size1	size2	
1038	1039	Мир, счастье и богатство	889.0	https://www.chitai-gorod.ru/product/mir-schast...	36.0	0.2	14.7	
974	975	В потоке дней порой бывает сладко, но и случае...	3359.0	https://www.chitai-gorod.ru/product/v-potoke-d...	112.0	1.0	15.5	
699	700	Как я был риелтором	519.0	https://www.chitai-gorod.ru/product/kak-ya-byt...	30.0	0.2	14.5	
1012	1013	Как размножаются кентавры	1669.0	https://www.chitai-gorod.ru/product/kak-razmno...	125.0	0.7	14.0	
969	970	Живые сплетения	1669.0	https://www.chitai-gorod.ru/product/zhivye-spl...	96.0	0.8	12.0	
769	770	Сингапур. Восьмое чудо света. Том 28	5490.0	https://www.chitai-gorod.ru/product/singapur-v...	286.0	2.3	15.0	
985	986	У всякой охоты свои заботы	599.0	https://www.chitai-gorod.ru/product/u-vsyakoy-...	80.0	0.3	14.2	
707	708	Провокация. Воззвание и право на переворот	919.0	https://www.chitai-gorod.ru/product/provokaciy...	100.0	0.6	12.0	
844	845	Возвращение в Ивто	1249.0	https://www.chitai-gorod.ru/product/vozvrashch...	112.0	0.7	13.0	
849	850	Еретик, или пути к своей цели	1359.0	https://www.chitai-gorod.ru/product/eretik-ili...	136.0	0.7	14.5	
681	682	Команда хантер-киллер	3569.0	https://www.chitai-gorod.ru/product/komanda-ha...	100.0	1.0	21.1	
870	871	Сетевые Мир и Война. Новый Миропорядок или Нов...	3879.0	https://www.chitai-gorod.ru/product/setevye-mi...	302.0	2.2	13.5	
911	912	90 лет своим путем. Воспоминая и размышления о...	2729.0	https://www.chitai-gorod.ru/product/90-let-svo...	272.0	1.6	13.4	
1047	1048	Проза Парижской	1779.0	https://www.chitai-gorod.ru/product/proza-pari...	160.0	1.0	14.5	

	N	name	price	url	pages	size1	size2	size3
		Жизни						
772	773	Как работает мир, в котором я живу	869.0	https://www.chitai-gorod.ru/product/kak-rabota...	104.0	0.5	14.5	
827	828	Портретное сходство	5249.0	https://www.chitai-gorod.ru/product/portretnoe...	168.0	1.5	21.5	
528	529	Зачем смотреть на животных?	889.0	https://www.chitai-gorod.ru/product/zachem-smo...	112.0	0.8	10.8	
523	524	Образ без цели. Эссе о литературе, теологии, п...	1459.0	https://www.chitai-gorod.ru/product/obraz-bez-...	152.0	1.0	13.0	
1102	1103	Письма бунтующего сценариста. Заметки о сценар...	809.0	https://www.chitai-gorod.ru/product/pisma-bunt...	94.0	0.5	14.5	
916	917	Развитие государства и человека в противоречии...	319.0	https://www.chitai-gorod.ru/product/razvitie-g...	36.0	0.2	14.3	
698	699	Может, весной	2199.0	https://www.chitai-gorod.ru/product/mozhet-ves...	208.0	1.3	15.2	
1060	1061	У мужчин нет совести, у женщин нет стыда	919.0	https://www.chitai-gorod.ru/product/u-muzhchin...	104.0	0.6	14.3	
914	915	Марка с цепелином	8751.0	https://www.chitai-gorod.ru/product/marka-s-ce...	513.0	4.9	17.0	
913	914	"...И вы будете как боги, знающие добро и зло"...	5626.0	https://www.chitai-gorod.ru/product/i-vy-budet...	584.0	3.5	15.5	
869	870	Вечное Лето - обратная сторона успеха	2251.0	https://www.chitai-gorod.ru/product/vechnoe-le...	220.0	1.5	15.5	
550	551	Вечное Лето - обратная сторона успеха	2251.0	https://www.chitai-gorod.ru/product/vechnoe-le...	220.0	1.5	15.5	
680	681	Рейнджеры: торящие путь	3569.0	https://www.chitai-gorod.ru/product/reyndzhery...	208.0	1.6	21.2	

	N	name	price	url	pages	size1	size2	
527	528	Дорожи тем, что ценишь. Депеши о выживании и С...	889.0	https://www.chitai-gorod.ru/product/dorozhi-te...	112.0	1.2	11.0	
213	214	Эта война! Dieser Krieg!	413.0	https://www.chitai-gorod.ru/product/eta-voyna-...	50.0	0.4	14.0	

```
In [50]: count_row = df_new.shape[0]
```

```
In [51]: count_row
```

```
Out[51]: 29
```

```
In [49]: df_new.to_csv('out2.csv', index=False)
```