



Предсказание ценовой функции смартфонов на российском рынке

Тевс Анна БЭАД223



Данные

Сбор данных

Было решено использовать достаточно аккуратное и подробное описание характеристик телефонов с **Мвидео**. (Запросы по внутреннему **API сайта**)

Признаки:

- ID: name, brand
- Size: Экран (Дюймы), Габаритные размеры (В*Ш*Т/В*Ш*Г мм), Вес (г)
- Resolution: Камера (МПикс), Разрешение экрана, Разрешение видеосъемки (Пикс), Зум (x)
- Inside: Технология экрана, Яркость (кд/кв.м), Частота обновления экрана (Гц), Количество ядер (шт), Память (ГБ), Количество камер (шт), Степень защиты (IPXY), Мощность блока питания (Вт), Емкость аккумулятора (мАч)
- Outside/features: Безрамочный, Съёмка видео в портретном режиме, Оптическая стабилизация, Поддержка симкарт, Поддержка стандартов, Поддержка WiFi, Технология NFC, Сенсор распознавания лица, Сканер отпечатка пальца, Материал корпуса, Блок питания, Кабель, Чехол
- Others: Гарантия, Страна, Год релиза
- Targets: **salePrice**, **basePrice**

2

М.ВИДЕО
Главный эксперт по технике

Москва Магазины Установка и ремонт М.Комбо М.Клик М.Служб Юр. лицам Стать партнером 8-800-600-7775

Искать товары в Рассрочку

Статус заказа Войти Сравнение Избранное Корзина

Смартфон Apple iPhone 13 128GB nanoSim/eSim... 51 999 ₽ 74 999 Добавить в корзину

О товаре Характеристики Отзывы Наличие в магазинах

Все характеристики

Заводские данные

Гарантия 1 год
Гарантия предоставляется продавцом
Страна Китай
Год релиза 2021

Серия модели

Серия iPhone 13

Основные характеристики

Состояние новый
Внешний вид новый

Инструкции и сертификаты

PDF Инструкция по эксплуатации 229.7KB
PDF Скачать сертификат соответствия 83.9KB

ЕАЭС 6 10

М.ВИДЕО
Главный эксперт по технике

Искать товары в Рассрочку

Статус заказа Войти Сравнение Избранное Корзина

ВСЕ АКЦИИ ВЕСЕННИЕ СКИДКИ СМАРТФОНЫ ТЕЛЕВИЗОРЫ MINI LED КОНДИЦИОНЕРЫ ВЫКУП УСТРОЙСТВ НОУТЕБУКИ ОТ 1 050 Р В МЕС. ХОЛОДИЛЬНИКИ ВСТРАИВАЕМАЯ ТЕХНИКА

Главная — Смартфоны и связь — Смартфоны — iPhone — Apple

Рассрочка 0-0-25 -10% на чехол

Смартфон Apple iPhone 13 128GB nanoSim/eSim Blue

★★★★★ 436 отзывов

Цвет: Синий

Состояние: восстановленный, новый

Встроенная память (ROM): 128 ГБ, 256 ГБ

Поддержка SIM карт: 1 nano-SIM/1 eSim, 2 nano-SIM

Характеристики:

Экран 6.1"/2532x1170 Пикс
Технология экрана OLED
Тип процессора A15 Bionic
Встроенная память (ROM) 128 ГБ
Основная камера МПикс 12/12/0/0/0
Состояние новый

Код товара: 3006 3538

Авторизуйтесь, чтобы снизить цену

Финальная цена 51 999 ₽ 74 999
Посмотреть применённые скидки

Добавить в корзину

+1 560 Кешбэк

% Кредит от 5 233 ₽/мес
% Рассрочка от 2 080 ₽/мес

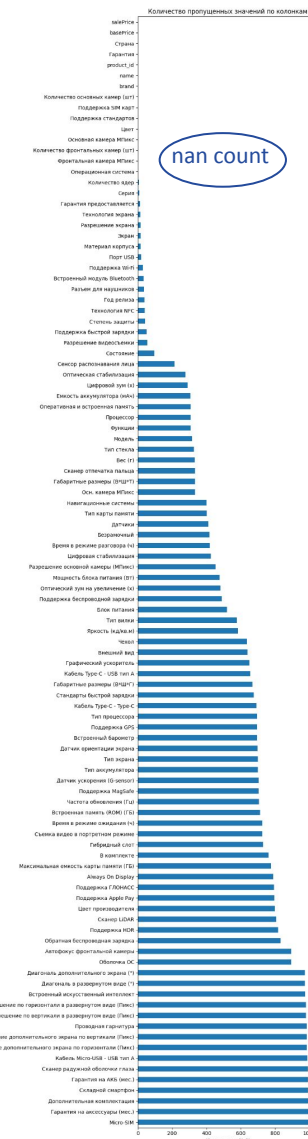
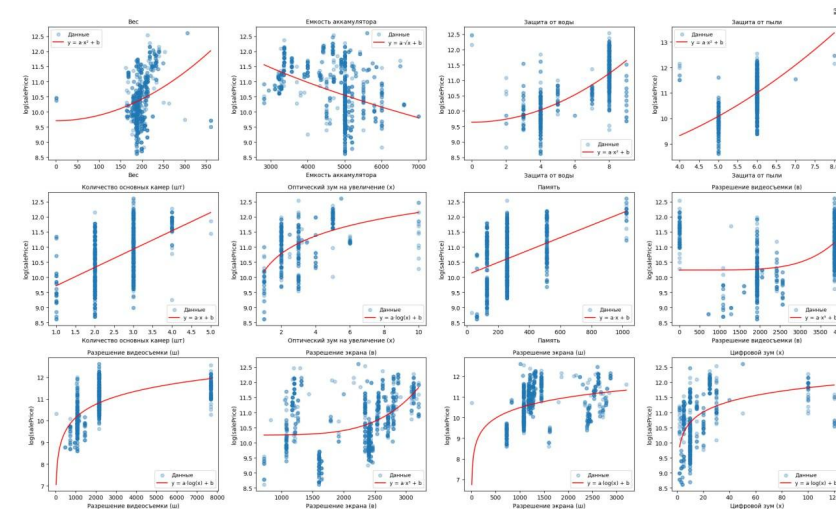
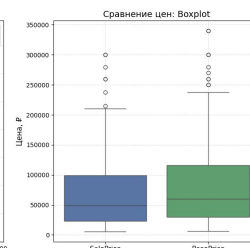
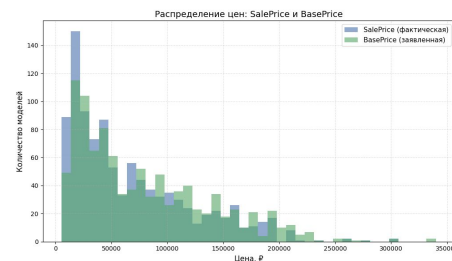
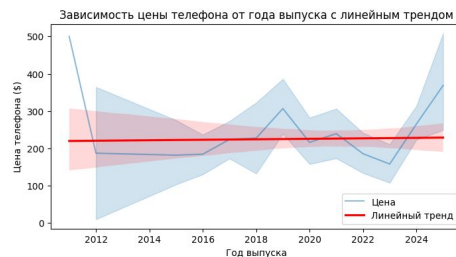
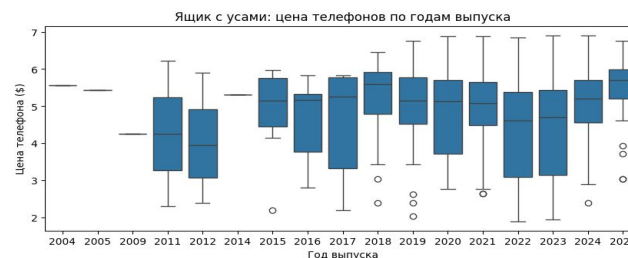
Рассчитать рассрочку

Обработка и анализ данных

Проблема: много характеристик

Решение:

- Анализ nan
- Графики
- a) для дамми в основном box-plot
- b) для количественных logreg, scatter, lineplot
- c) для категориальных всё ранее упомянутое + частота через histplot
- Построения “локальных” регрессия (для выявления незначимых на первых этапах)
- Здравый смысл! (+ ориентир на будущие предположения)

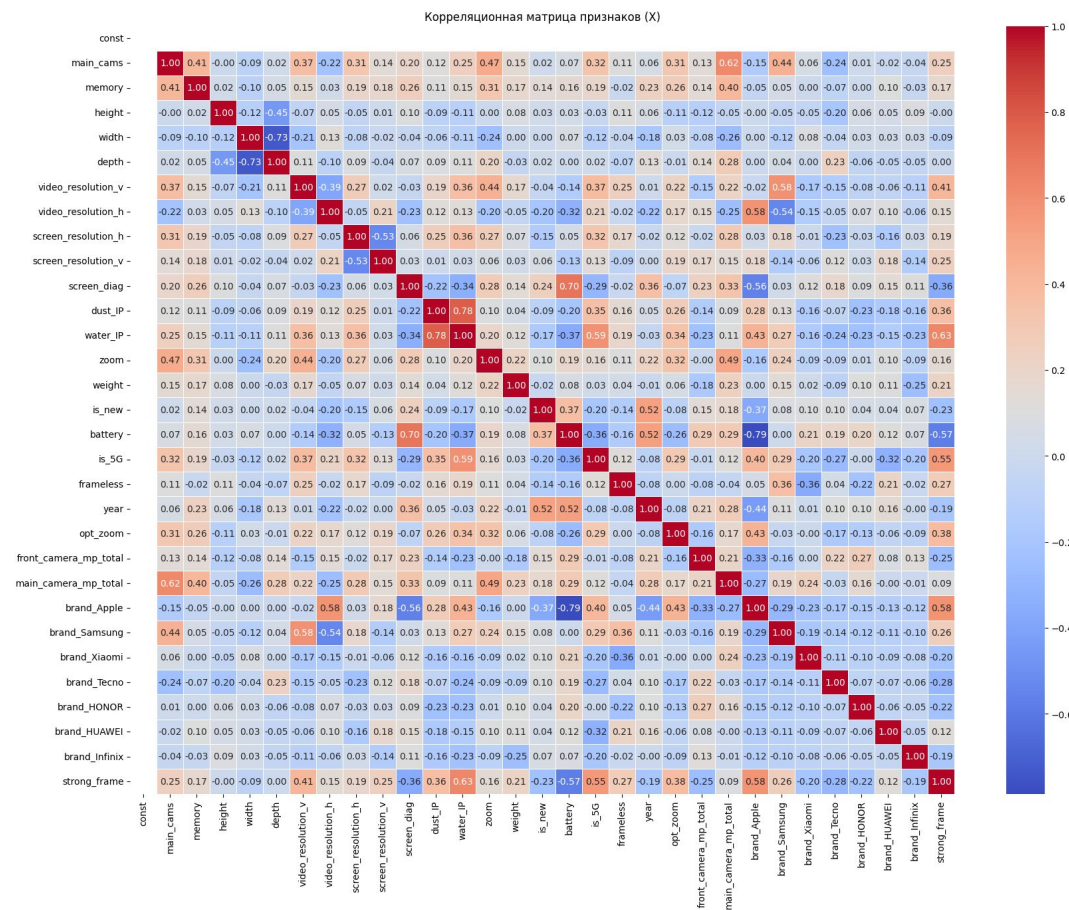




Correlation matrix
VIF

- новый признак (strong_frame = разность наличия металла и пластика в корпусе)
- РСА на разрешение видео (h/v)
- Периметр
- IP (сумма влагозащиты и защиты от пыли)

	feature	VI
0	brand_Apple	9.87058
1	depth	8.54269
2	width	6.54219
3	battery	5.46252
4	water_IP	5.37911
5	strong_frame	4.44206
6	brand_Samsung	4.15361
7	height	3.76765
8	dust_IP	3.29256
9	main_camera_mp_total	3.26793
10	video_resolution_h	3.22589
11	screen_resolution_h	3.10375
12	is_5G	3.09650
13	screen_resolution_v	3.05201
14	main_cams	2.98821
15	screen_diag	2.63117
16	video_resolution_v	2.43054
17	brand_HUAWEI	2.29305
18	zoom	2.16823
19	brand_Xiaomi	2.13435
20	year	2.13029
21	opt_zoom	2.05720
22	front_camera_mp_total	2.02223
23	brand_Tecno	1.77607
...		
27	memory	1.52618
28	brand_Infinix	1.41439
29	weight	1.40338
30	const	1.00000





Гетероскедастичность

Тесты:

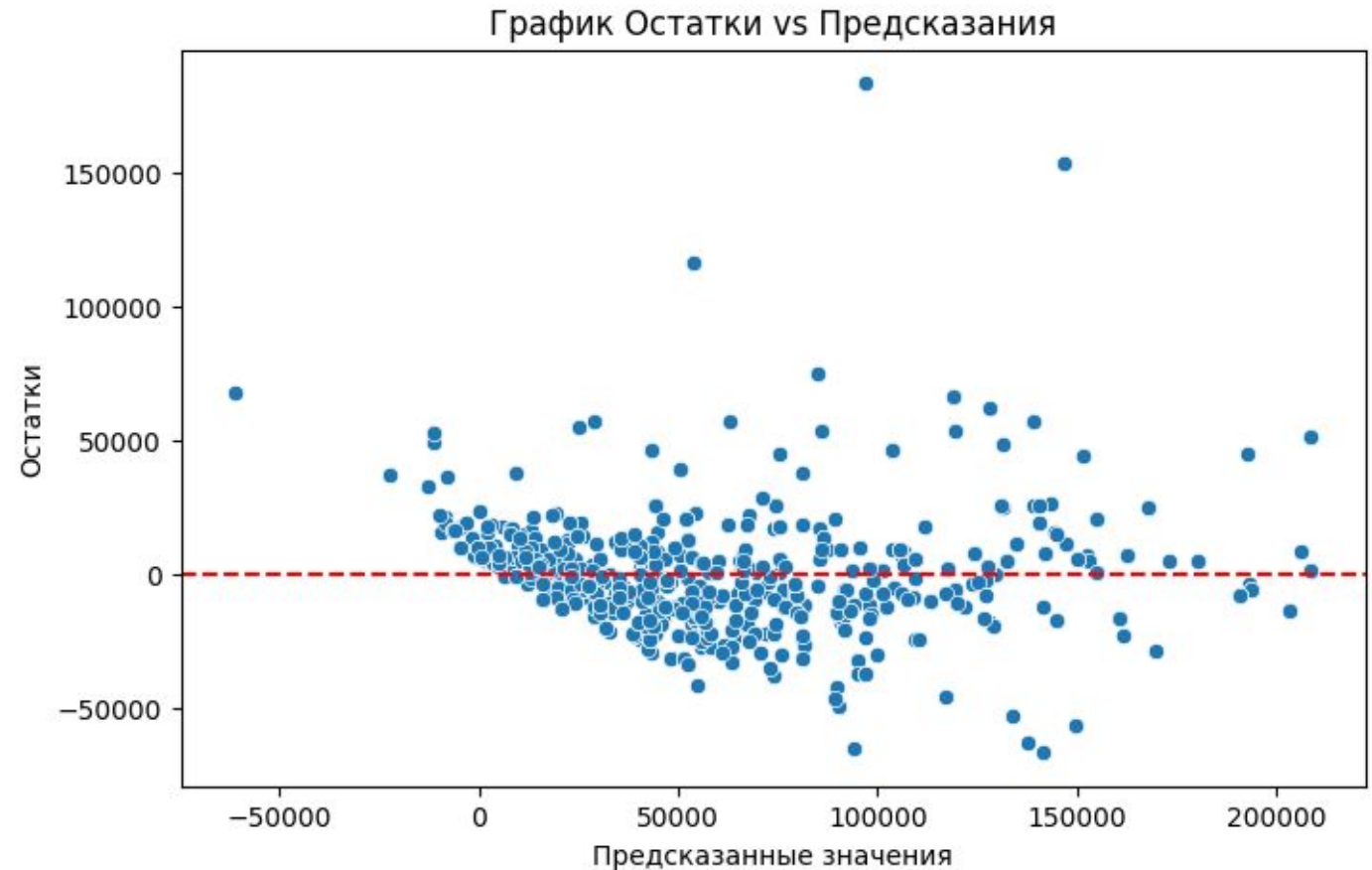
- Визуальный (облако точек)
- Уайта (F-Test p-value: 0.0000)

Проблема решена использование стандартных ошибок в форме Уайта

`(> ~) > (cov_type='HC3')`

Also tried:

- Изменить функциональную форму на этом этапе
- Полулогарифмическую модель





Выбор функциональной формы

Тест: Бокса-Кокса линейная vs полулогарифмическая

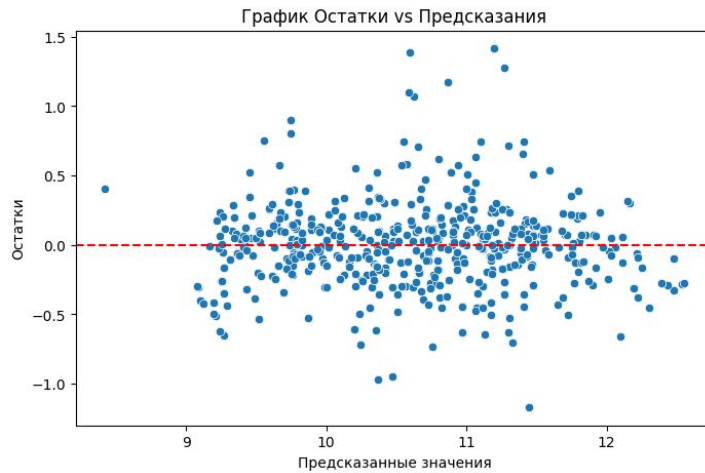
RSS (линейная модель): 374909571109.45

RSS (лог-преобразованная модель): 67.12

Статистика LM: 20850.042

P-значение: 0.0000

Отвергаем H0: полулогарифмическая модель лучше.



Dep. Variable:	salePrice	R-squared:	0.903
Model:	OLS	Adj. R-squared:	0.900
Method:	Least Squares	F-statistic:	434.9
Date:	Sun, 11 May 2025	Prob (F-statistic):	0.00
Time:	13:31:32	Log-Likelihood:	2083.9
No. Observations:	929	AIC:	-4106.
Df Residuals:	898	BIC:	-3956.
Df Model:	30		
Covariance Type:	HC3		

	coef	std err	z	P> z	[0.025	0.975]
const	-11.3353	3.946	-2.872	0.004	-19.070	-3.601
main_cams	0.0079	0.006	1.423	0.155	-0.003	0.019
memory	7.356e-05	8.52e-06	8.634	0.000	5.69e-05	9.03e-05
height	-0.0006	0.002	-0.268	0.788	-0.005	0.004
width	-0.0005	0.002	-0.209	0.835	-0.005	0.004
depth	-0.0006	0.002	-0.287	0.774	-0.005	0.004
video_resolution_v	2.569e-06	9.99e-07	2.573	0.010	6.12e-07	4.53e-06
video_resolution_h	1.317e-06	1.81e-06	0.727	0.467	-2.23e-06	4.87e-06
screen_resolution_h	1.676e-05	4.35e-06	3.854	0.000	8.24e-06	2.53e-05
screen_resolution_v	1.833e-05	3.87e-06	4.738	0.000	1.07e-05	2.59e-05
screen_diag	0.0073	0.019	0.391	0.696	-0.029	0.044
dust_IP	-0.0012	0.002	-0.709	0.478	-0.005	0.002
water_IP	0.0041	0.001	3.292	0.001	0.002	0.007
zoom	2.913e-05	6.68e-05	0.436	0.663	-0.000	0.000
weight	0.0001	0.000	0.706	0.480	-0.000	0.000
is_new	0.0421	0.005	8.293	0.000	0.032	0.052
battery	-2.727e-06	4.89e-06	-0.558	0.577	-1.23e-05	6.85e-06
is_5G	0.0409	0.004	10.073	0.000	0.033	0.049
frameless	0.0071	0.004	1.658	0.097	-0.001	0.015
year	0.0067	0.002	3.709	0.000	0.003	0.010
opt_zoom	0.0020	0.001	1.545	0.122	-0.001	0.005
front_camera_mp_total	0.0004	0.000	2.827	0.005	0.000	0.001
main_camera_mp_total	0.0001	5.42e-05	2.320	0.020	1.95e-05	0.000
brand_Apple	0.0687	0.015	4.553	0.000	0.039	0.098
brand_Samsung	0.0236	0.005	4.306	0.000	0.013	0.034
brand_Xiaomi	-0.0005	0.004	-0.120	0.904	-0.008	0.007
brand_Tecno	0.0008	0.005	0.142	0.887	-0.010	0.011
brand_HONOR	0.0075	0.008	0.904	0.366	-0.009	0.024
brand_HUAWEI	0.0153	0.009	1.676	0.094	-0.003	0.033
brand_Infinix	0.0052	0.007	0.762	0.446	-0.008	0.019
strong_frame	0.0159	0.005	3.383	0.001	0.007	0.025



Выбор функциональной формы

Тест: Рамсея

1. Оценивается исходная линейная модель:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \epsilon$$

2. Из неё получают предсказанные значения \hat{y} .

3. В расширенную модель добавляют **степени** \hat{y} (обычно \hat{y}^2 , \hat{y}^3):

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \alpha_1 \hat{y}^2 + \alpha_2 \hat{y}^3 + \epsilon$$

4. Проверяют **F-тестом**, значимы ли новые коэффициенты (α_1 , α_2):

- Если **p-value** < **0.05** → **отвергаем H_0** (модель некорректна).
- Если **p-value** > **0.05** → **нет оснований отвергать H_0** (форма модели адекватна).

Дополнительно добавили признаки с функциями:
log, sqrt(), ^2, ^3, ^4, ^5

После оценки незначимости коэффициентов перед огромным количеством переменных, а также сравнением моделей по R^2 регрессоры изменились

	Feature Transformation		R2
0	IP	x^4	0.495507
1	P	x^5	0.128859
2	brand_Apple	x	0.241322
3	brand_HONOR	x	0.008218
4	brand_HUAWEI	x	0.001738
5	brand_Infinix	x	0.033923
6	brand_Samsung	x	0.075613
7	brand_Tecno	x	0.072250
8	brand_Xiaomi	x	0.045041
9	dim_PC1	log(x)	0.470167
10	dim_PC2	x	0.343538
11	frameless	x	0.042148
12	is_5G	x	0.507242
13	is_new	x	0.000479
14	memory	x	0.215443
15	opt_zoom	x	0.266491
16	screen_diag	log(x)	0.041654
17	screen_resolution_h	x^5	0.200900
18	screen_resolution_v	log(x)	0.111842
19	strong_frame	x	0.583260
20	weight	x	0.038507
21	year	x^5	0.000311
22	zoom	log(x)	0.144068



Выбросы

Тесты и статистики:

- **Леверидж**

Леверидж h_{ii} — диагональные элементы матрицы

$H = X(X^T X)^{-1} X^T$ ("hat matrix").

$h_{ii} > \frac{2(k+1)}{n}$ — точка с высоким влиянием.

- **Стьюдентизированные остатки**

$$r_i = \frac{e_i}{s\sqrt{1-h_{ii}}}$$

Где:

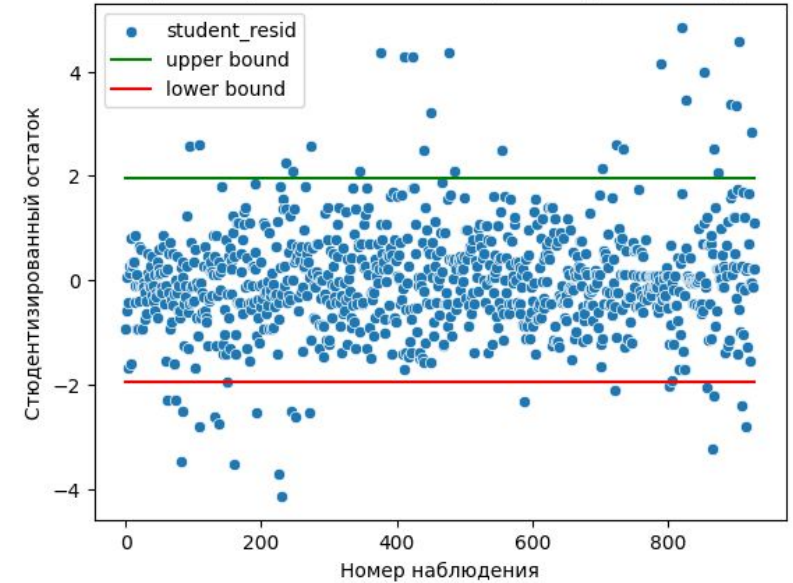
- $e_i = y_i - \hat{y}_i$ — обычный остаток для i-го наблюдения
- s — стандартная ошибка остатков всей модели (RMSE)
- h_{ii} — левверидж (диагональный элемент матрицы влияния H)

$|r_i| > 3$ — явный выброс

Удалено 5
выбросов



Стьюдентизированные остатки для каждого наблюдения



- **DFFITS**

Формула

$$DFFITS_i = t_i \sqrt{\frac{h_{ii}}{1-h_{ii}}}$$

где t_i — студентизированный остаток, h_{ii} — левверидж.

Интерпретация

- $|DFFITS_i| > 2\sqrt{\frac{k+1}{n}}$ — точка сильно влияет на модель.



Выводы

$$\ln(\text{price}) = -55.96 + 0.4 \ln(\text{memory}) + 0.7 \text{brand_Apple} + 0.47 \text{is_new} + 0.292 \text{strong_frame} + 0.42 \text{is_5G} + 0.2 \text{main_cams} + 0.099 \ln(\text{zoom}) + 0.000015 \text{year}^2 - 0.0004 (\text{brand_Apple} * \text{memory}) + 0.00005 (\text{frameless} * \text{year}) + 0.08 \text{dim_PC2} - 0.000005 P^2 + 0.001 IP^2 - 0.004 (\text{brand_HUWAEI} * \text{front_camera_mp_total}) + e$$

Попробовали предсказать цену телефона **Google Смартфон Google Pixel 7 Pro 12/128GB**, который стоит 49.990, у нас получилось 52.360

R-squared:	0.902
Adj. R-squared:	0.900
F-statistic:	925.3
Prob (F-statistic):	0.00
Log-Likelihood:	-106.57
AIC:	245.1
BIC:	322.4



Квантильная регрессия

Квантили: 0.1, 0.5, 0.75, 0.9

