

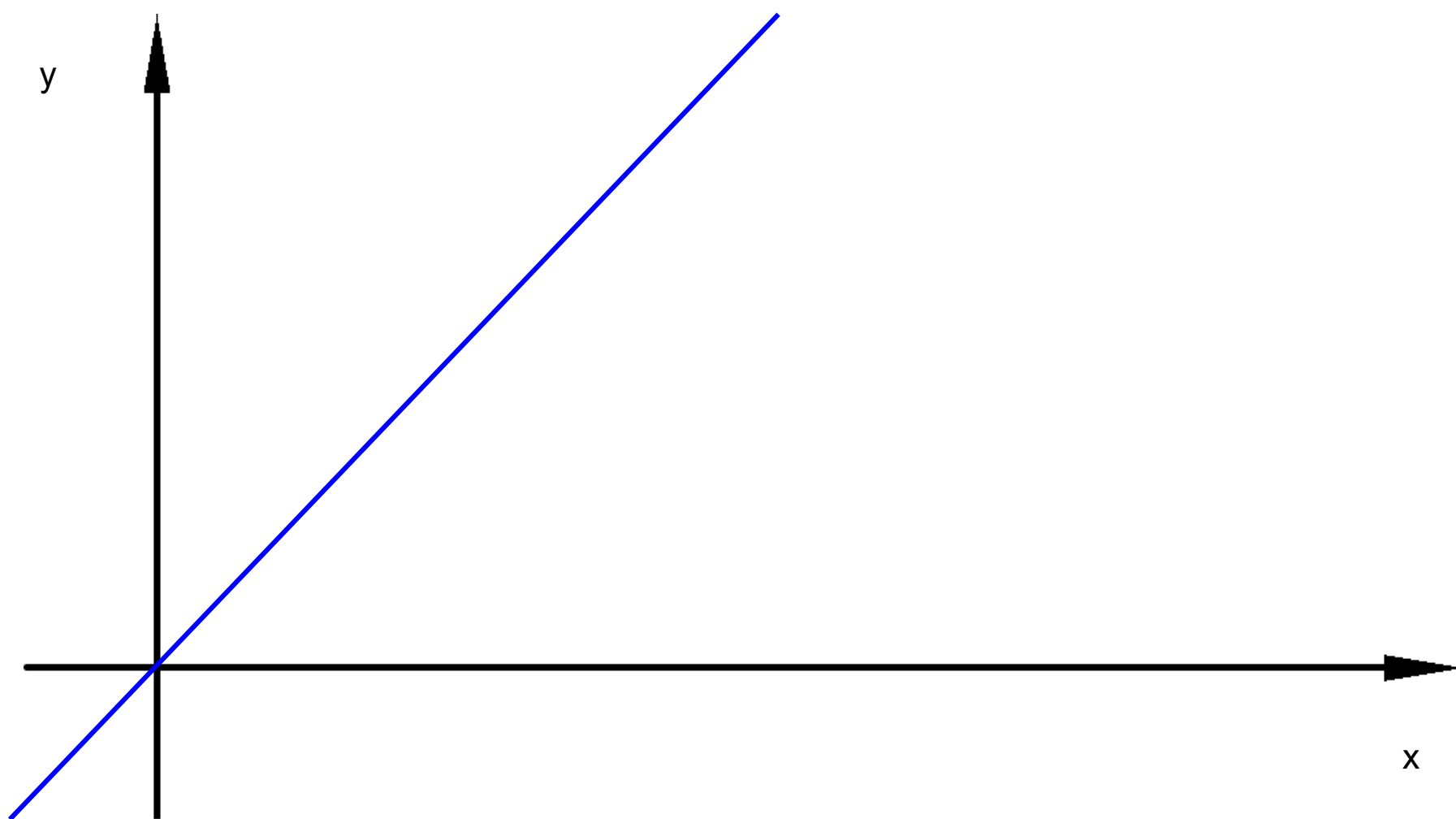
Линейная регрессия

Skillbox

образовательная платформа

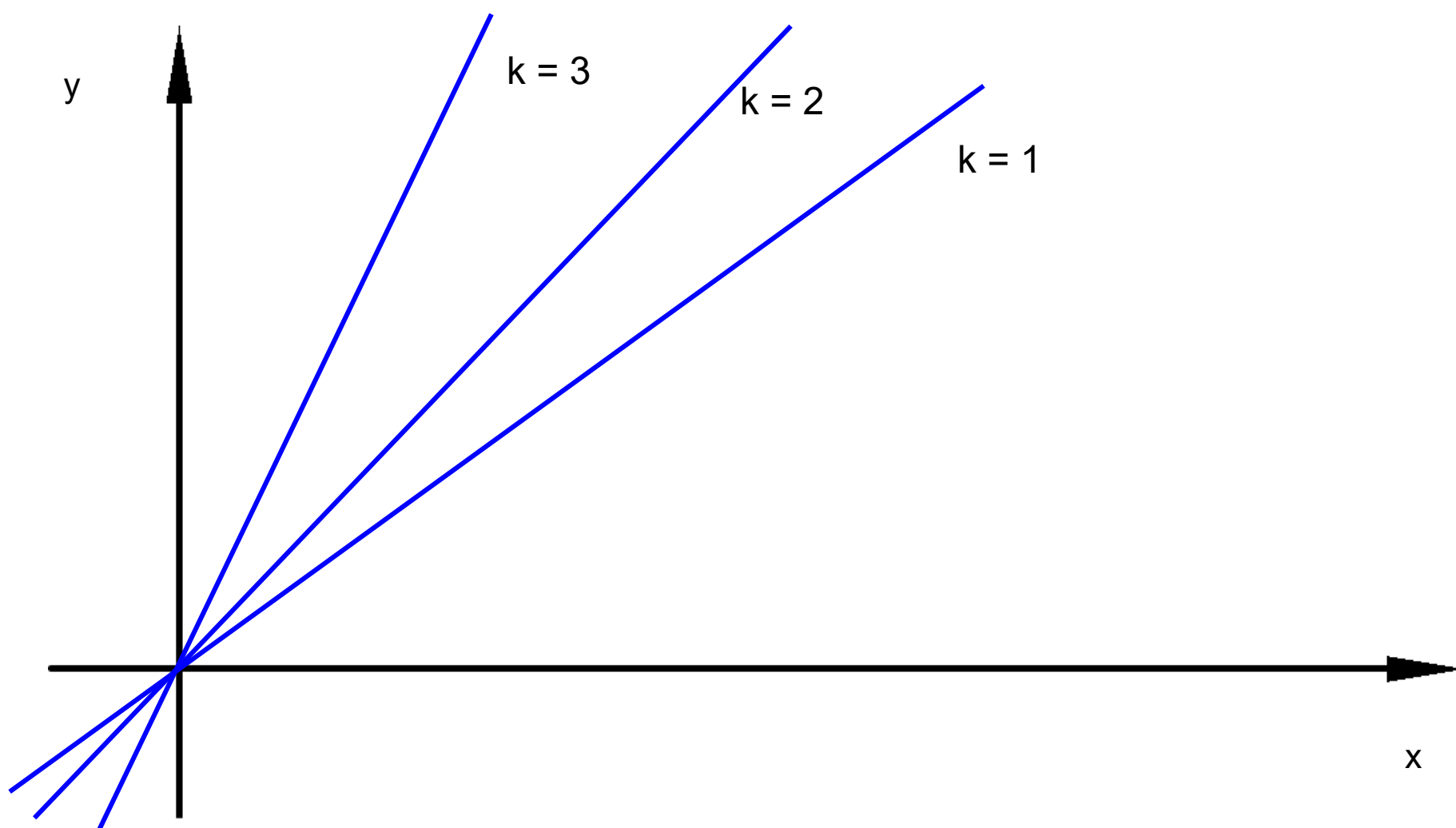
Линейная функция

$$y = kx + b$$

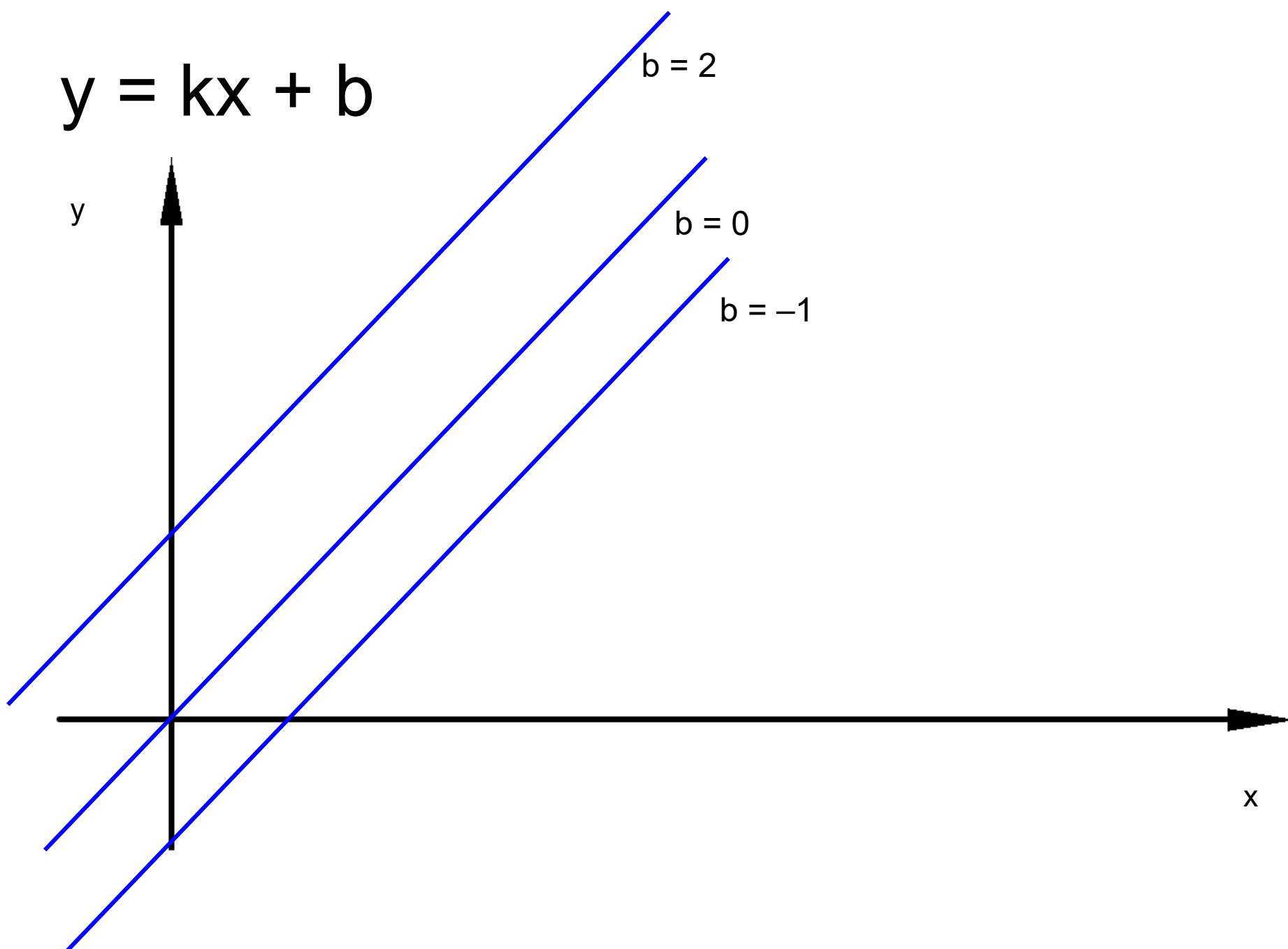


Линейная функция

$$y = kx + b$$



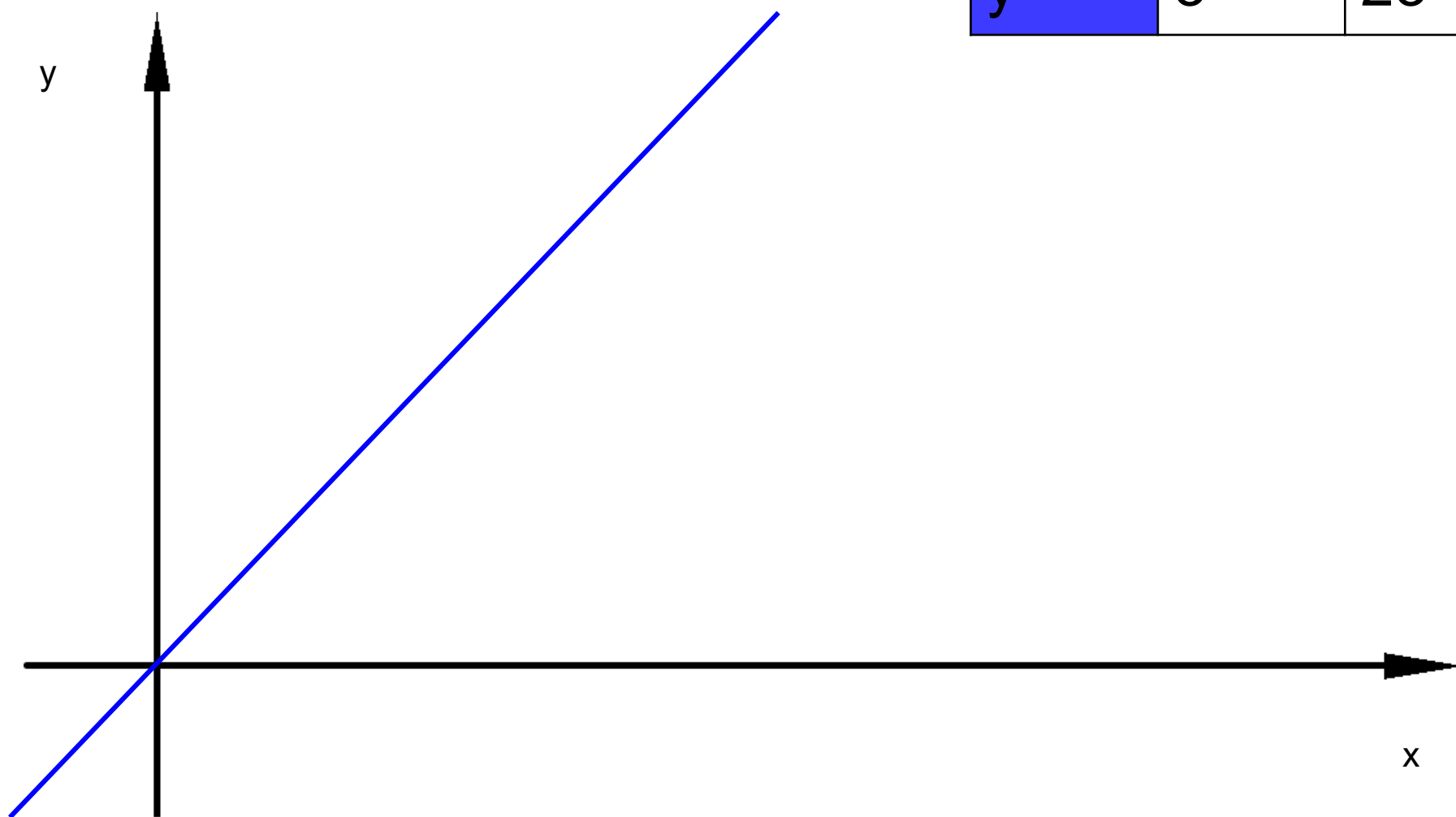
Линейная функция



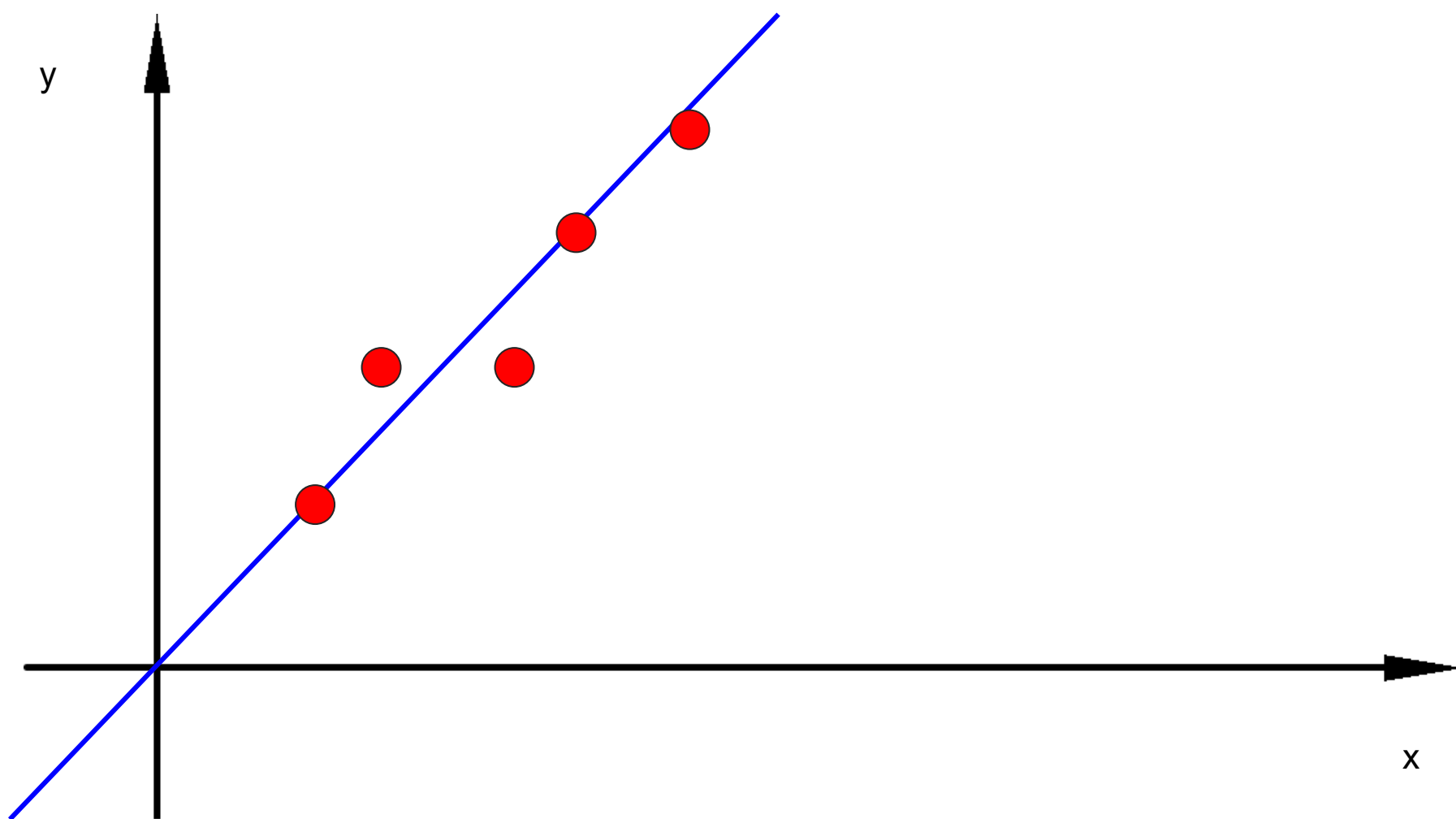
Линейная функция

$$y = 2x + 3$$

x	0	10
y	3	23



Линейная функция

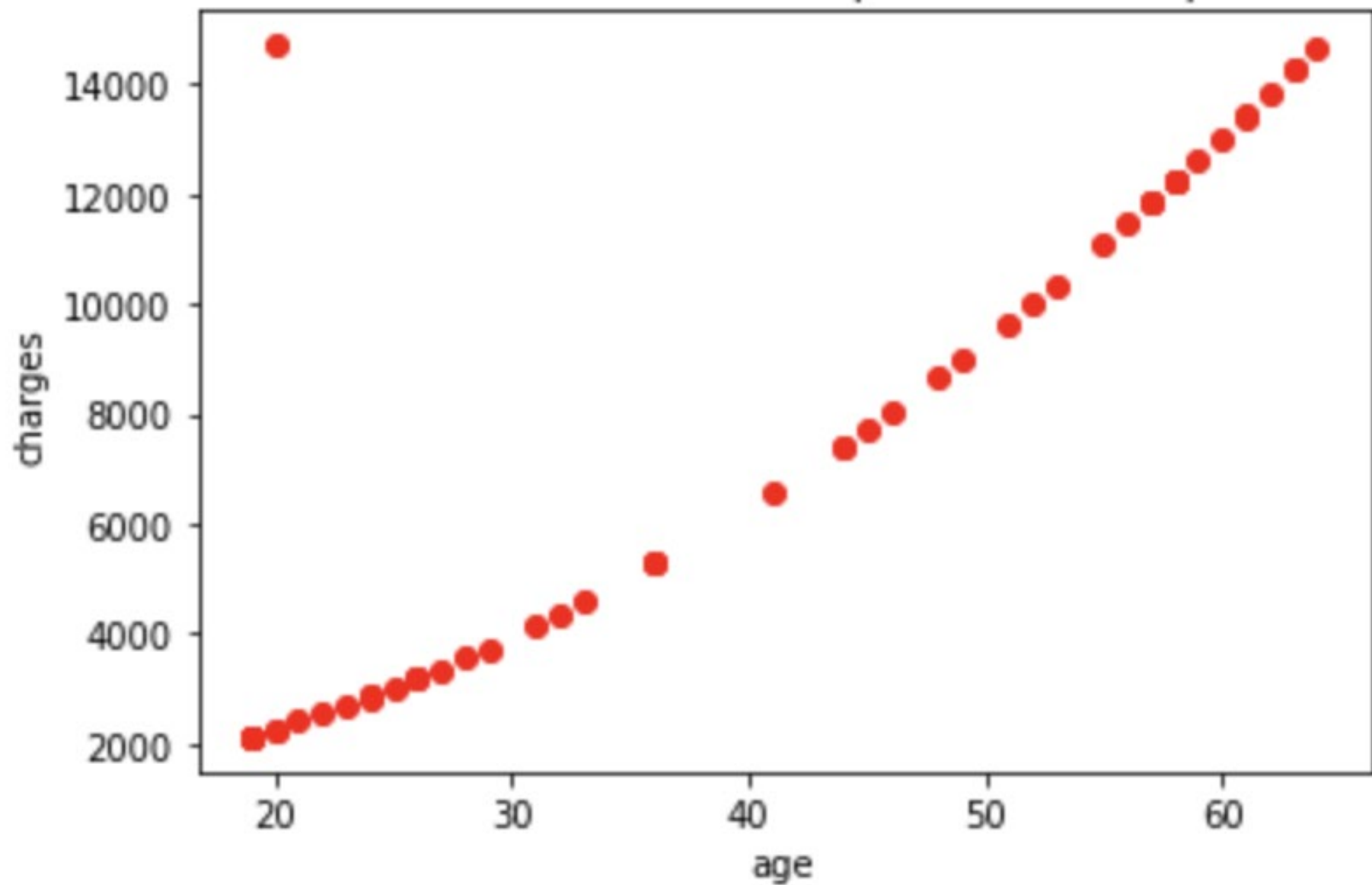


Постановка задачи

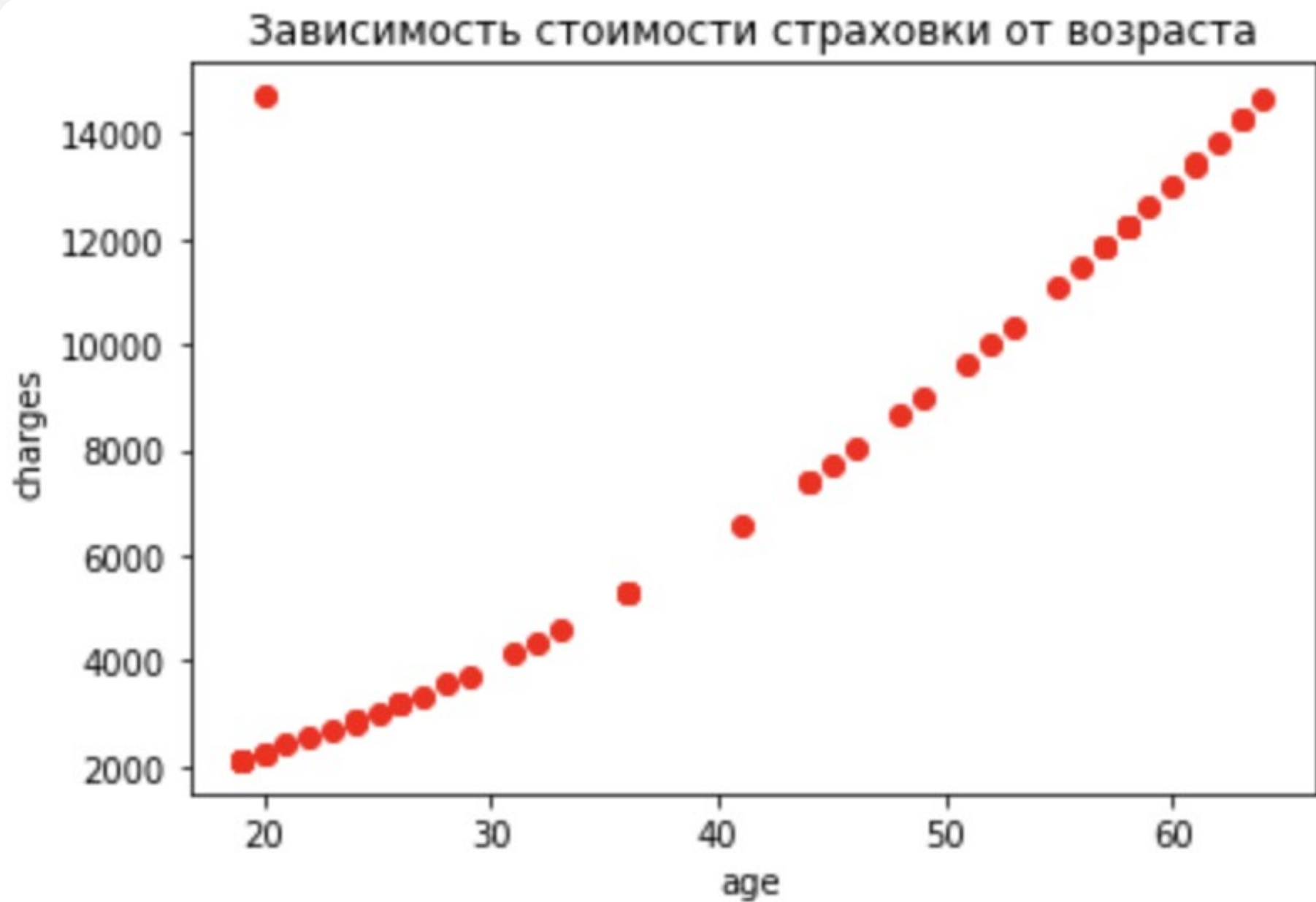
Спрогнозировать расходы на медицинское страхование по клиенту.

	age	sex	bmi	children	smoker	region	charges
0	19	female	27.900	0	yes	southwest	16884.92400
1	18	male	33.770	1	no	southeast	1725.55230
2	28	male	33.000	3	no	southeast	4449.46200
3	33	male	22.705	0	no	northwest	21984.47061
4	32	male	28.880	0	no	northwest	3866.85520
...
1333	50	male	30.970	3	no	northwest	10600.54830

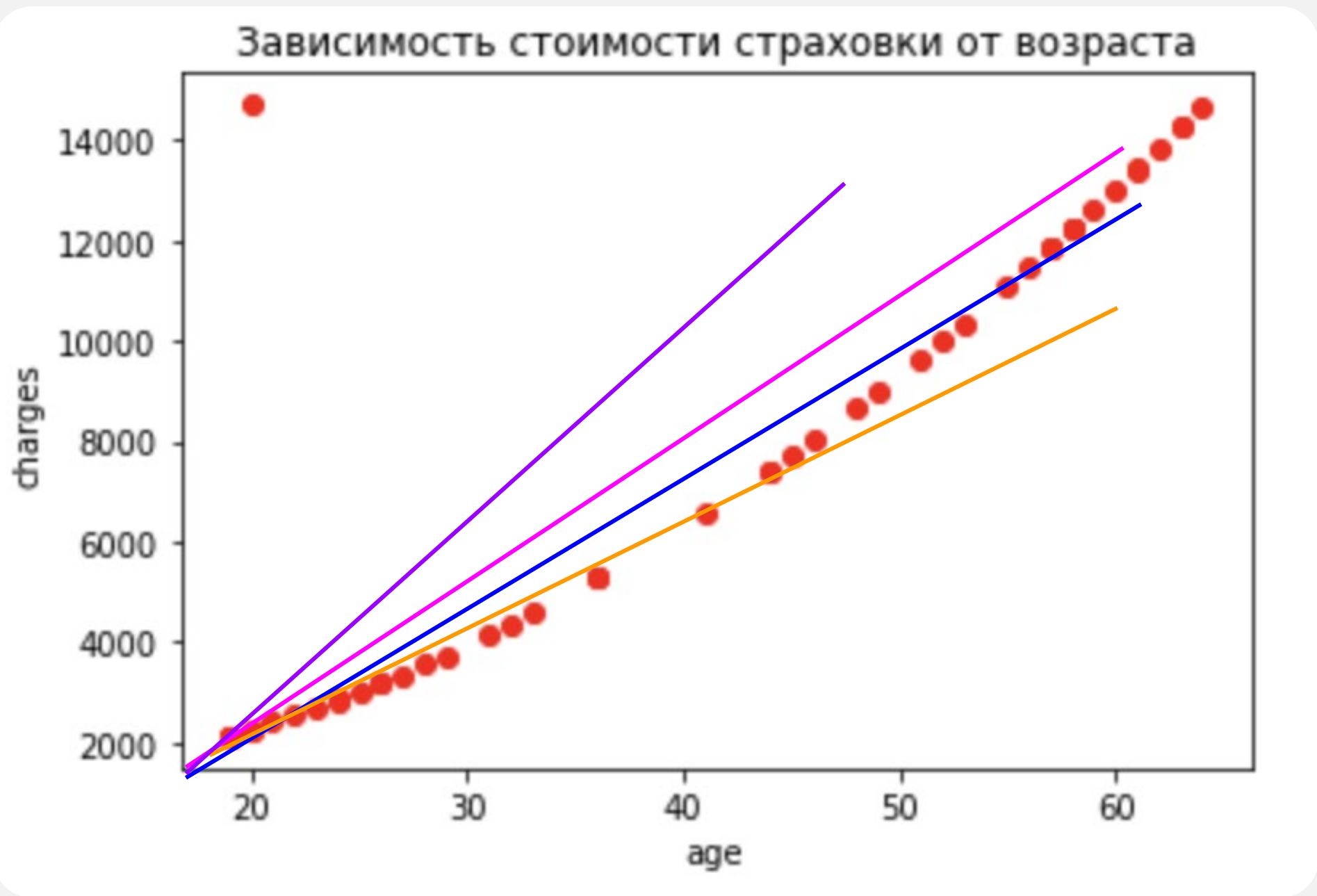
Зависимость стоимости страховки от возраста



* Женщины, без детей на иждивении, проживающие на северо-западе США, с суммой страховки менее \$15k

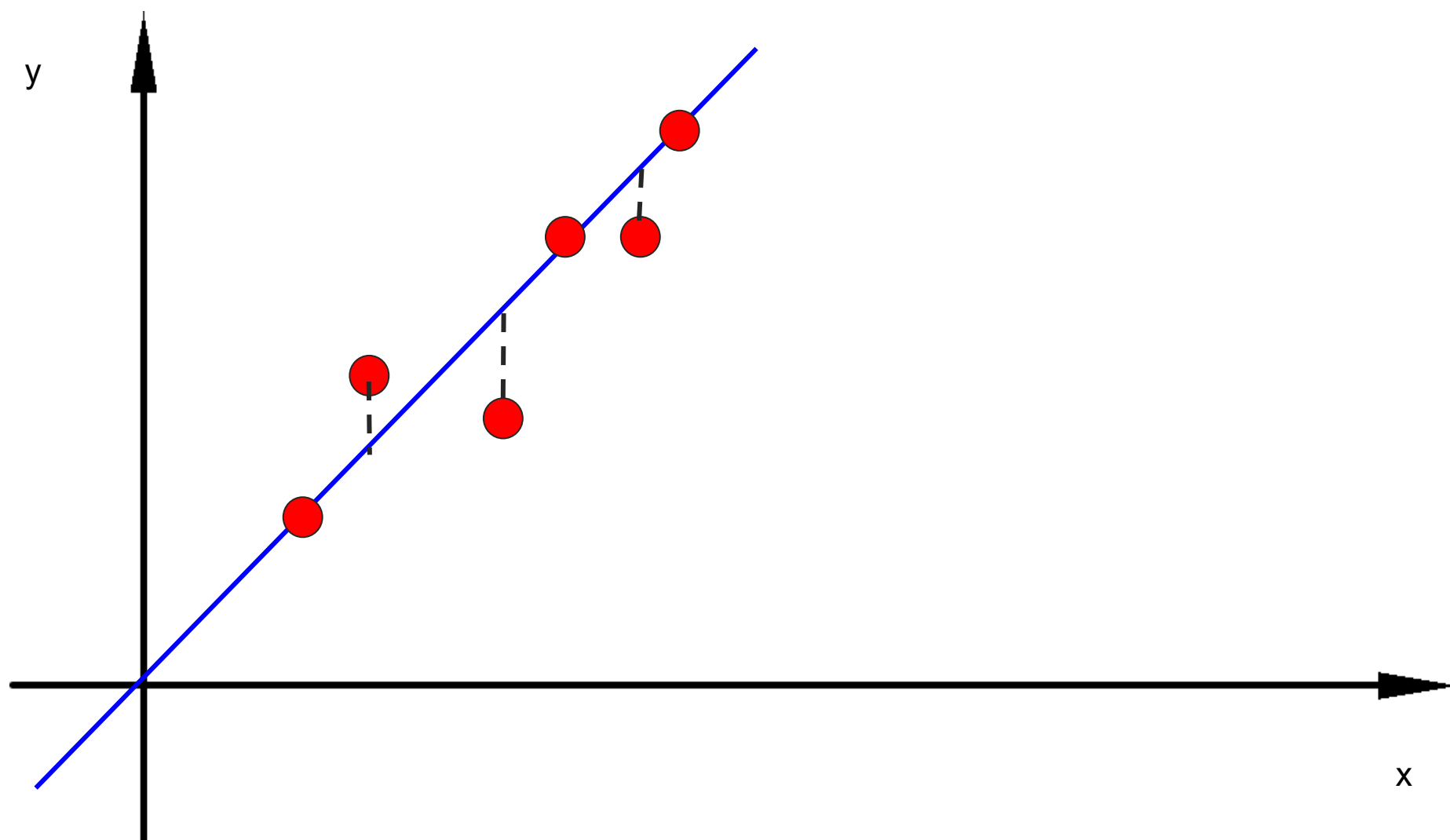


$$\text{charges} = k \times \text{age} + b$$

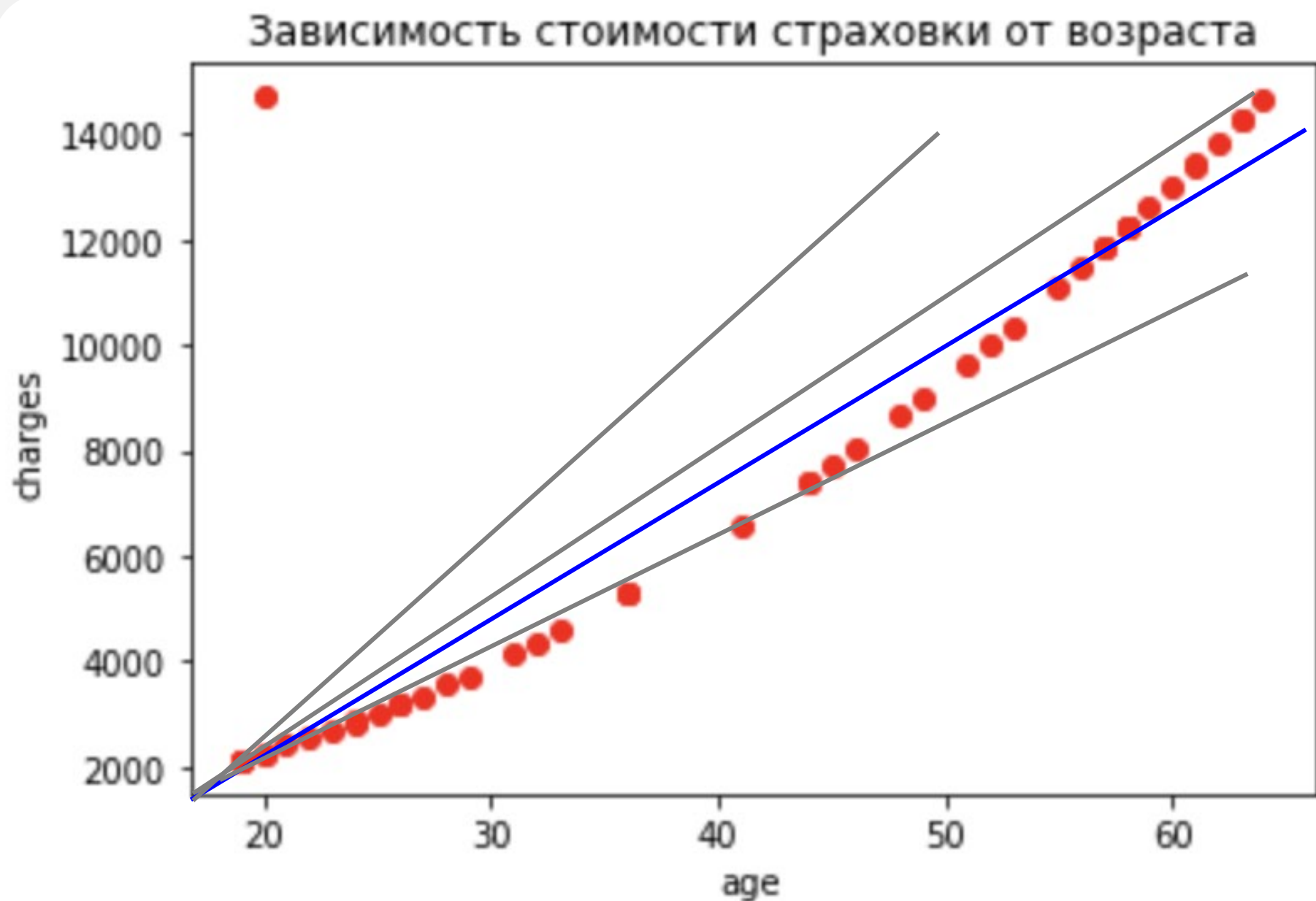


$$\text{charges} = k \times \text{age} + b$$

Расчёт ошибки



Расчёт ошибки



Для одной фичи

$$y(x) = kx + b$$

Для нескольких фичей

$$y(x_1, x_2 \dots x_n) =$$

$$w_1x_1 + w_2x_2 + \dots w_nx_n + w_0$$

Для одной фичи

$$y(x) = kx + b$$

Для нескольких фичей

$$y(x_1, x_2 \dots x_n) =$$

$$w_1 x_1 + w_2 x_2 + \dots w_n x_n + w_0$$

Наш случай

$$y(\text{age}, \text{bti}) =$$

$$w_1 \cdot \text{age} + w_2 \cdot \text{bti} + w_0$$

