



Обзор задач анализа данных

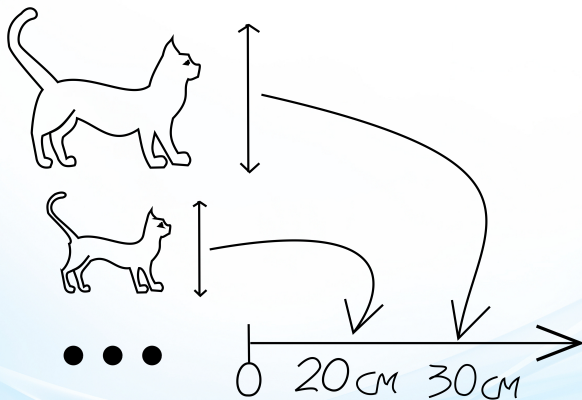


Мурмурландия

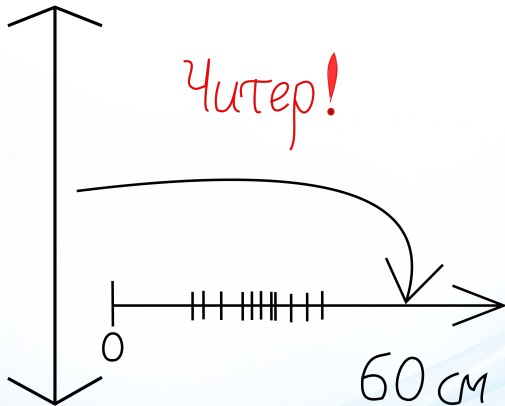




Каков средний рост котиков?



Точечное оценивание

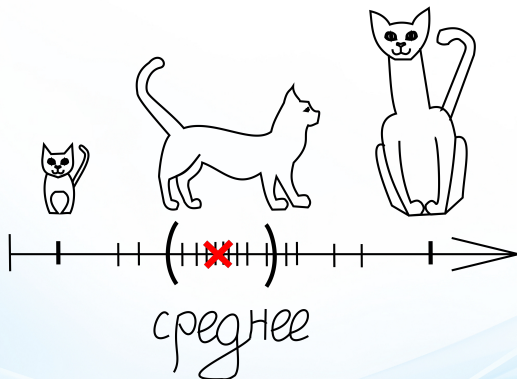


Выбросы





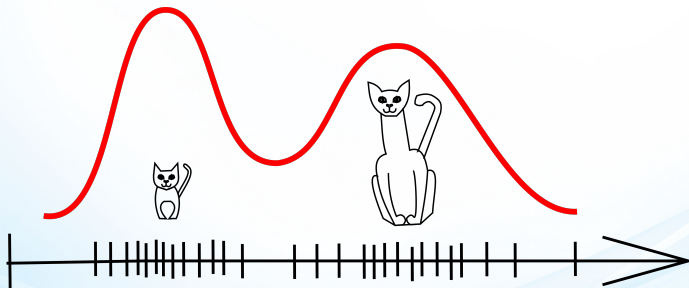
Среднее определяется неточно



Интервальное оценивание



Характер распределения



Непараметрическое оценивание



низкие



высокие

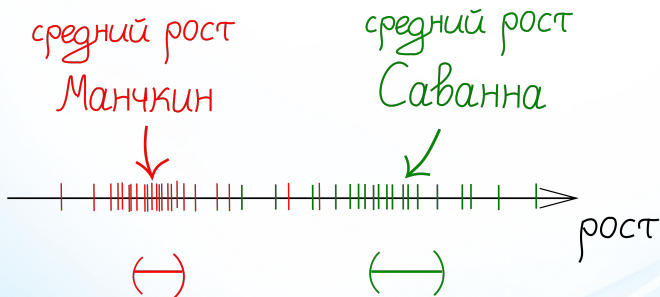




Отличается ли их средний рост?

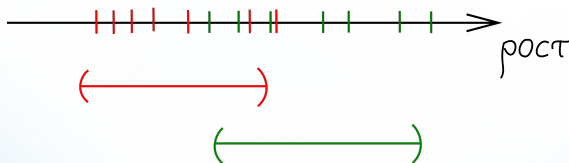


Собираем данные



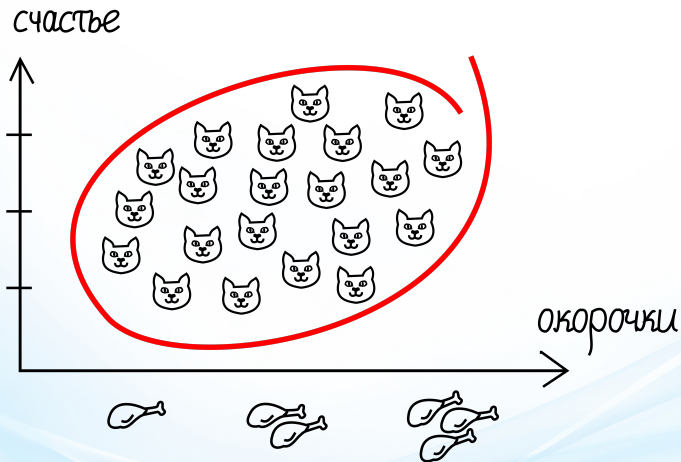
отличается

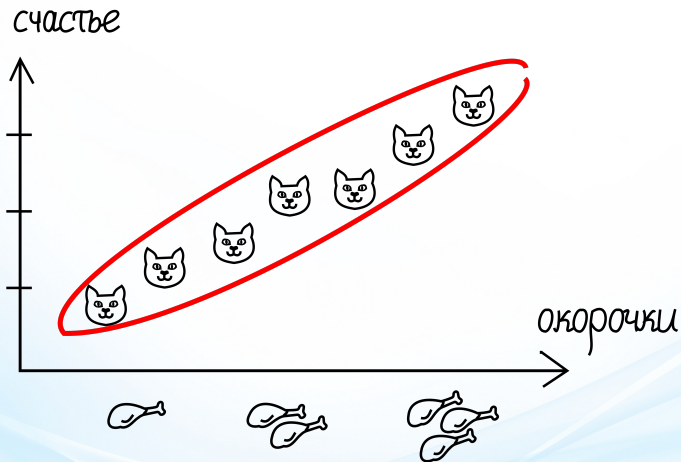
Если данных мало

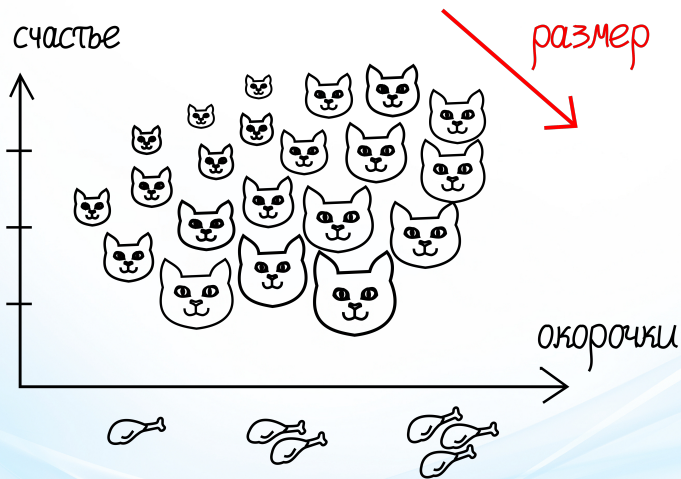


непонятно

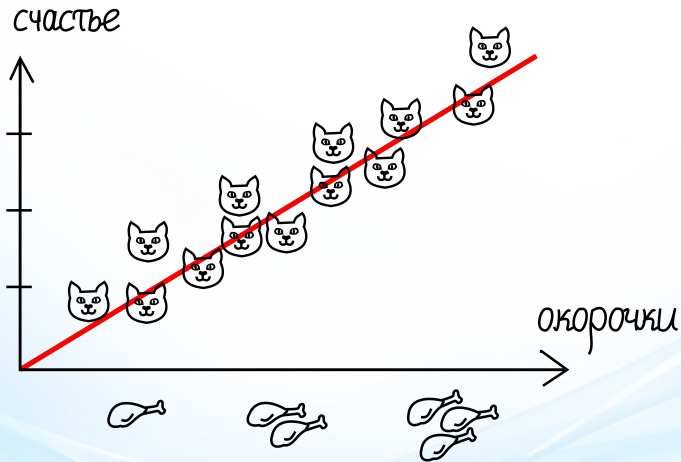
Статистические гипотезы, АВ-тесты








Корреляционный анализ





Формула счастья


$$= \theta_0 + \theta_1 \times \text{кол-во} \begin{array}{c} \text{курицы} \\ \text{и костей} \end{array} + \text{погрешности}$$



Больше окорочков

счастье





Формула счастья



$$= \theta_0 + \theta_1 \times \text{кол-во} \text{ (bone icon)} - \theta_2 \times (\text{кол-во} \text{ (bone icon)})^2 + \text{погрешности}$$



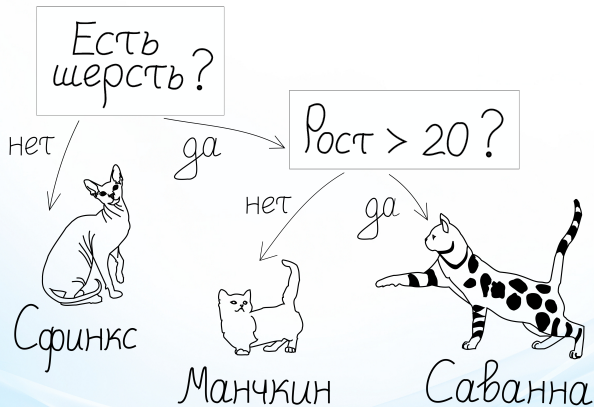
Другие факторы

$$\begin{aligned} &= \theta_0 + \theta_1 \times \text{курица} - \theta_2 \times (\text{курица})^2 \\ &\quad + \theta_3 \times \text{шарик} \\ &\quad + \theta_4 \times \text{диван} \\ &\quad + \text{погрешности} \end{aligned}$$

Регрессионный анализ







Классификация котиков



Классификация

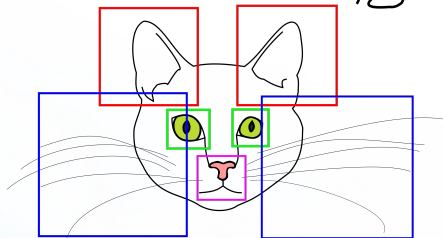


Собираем данные

котик	порода	рост	шерсть
	Саванна	50 см	да
	Сфинкс	30 см	нет
	Манчкин	15 см	да
	Саванна	40 см	да



Распознавание мордочек



Нейронные сети



Художник курса



Евгений



Книга с похожим содержанием

