

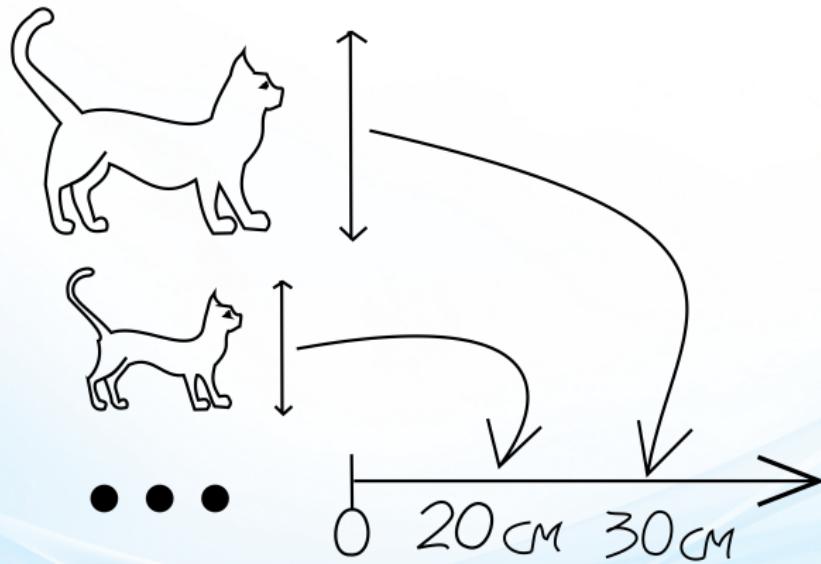
Обзор задач анализа данных



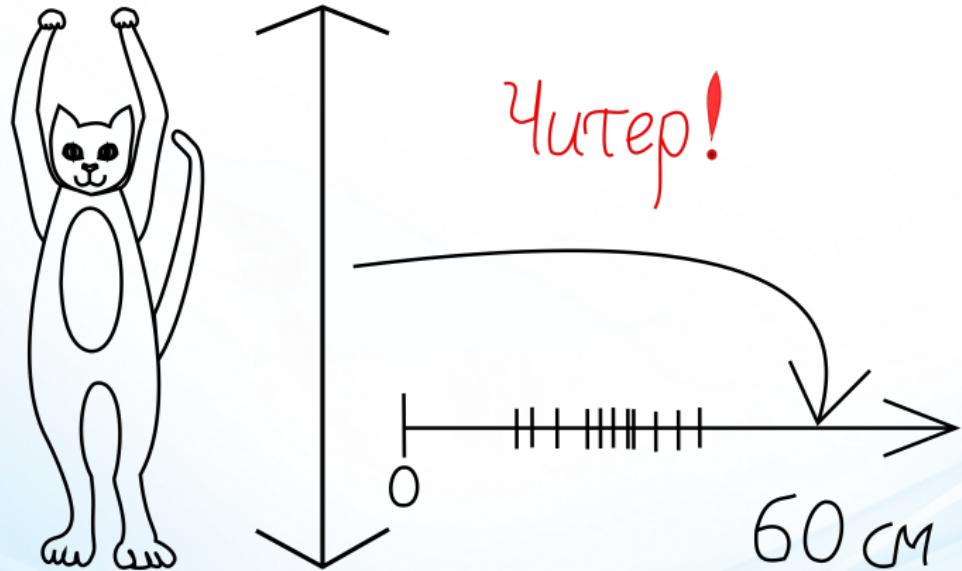
Мурмурландия



Каков средний рост котиков?



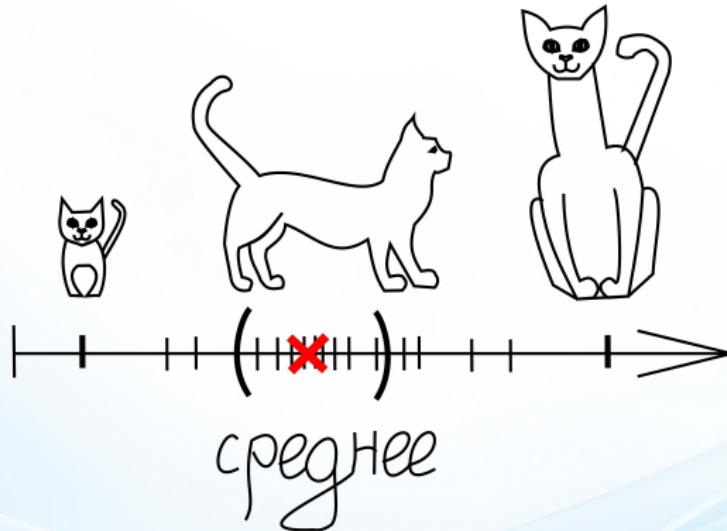
Точечное оценивание



Выбросы

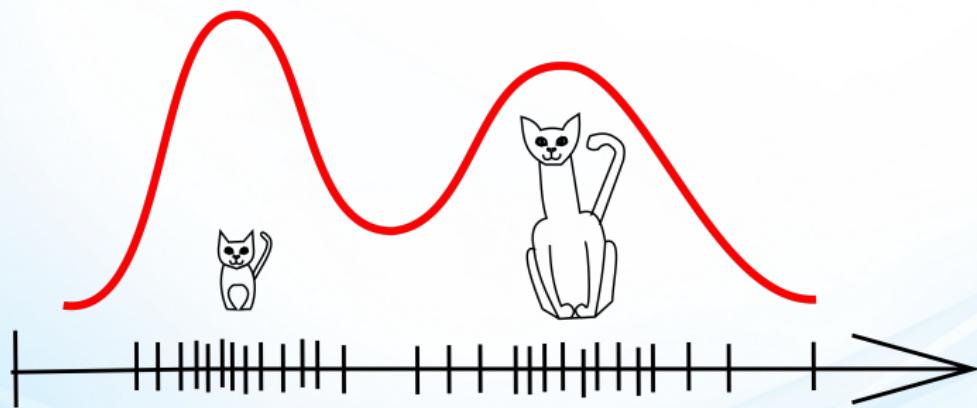


Среднее определяется неточно



Интервальное оценивание

Характер распределения



Непараметрическое оценивание



низкие



высокие



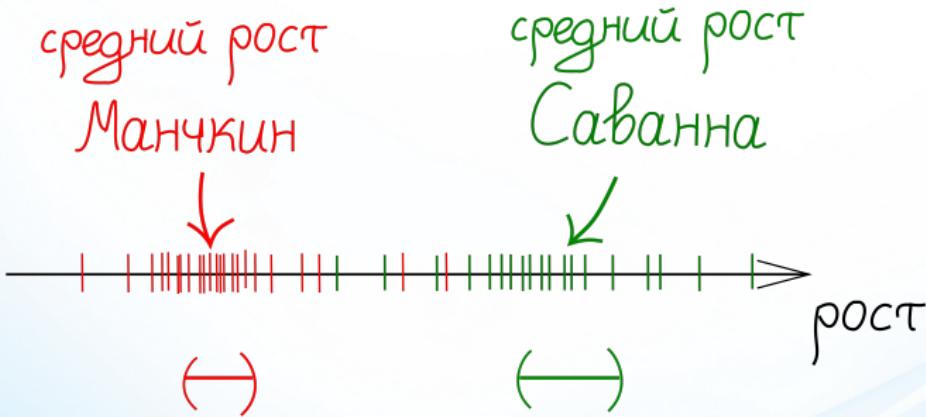
Саванна

Манчкин



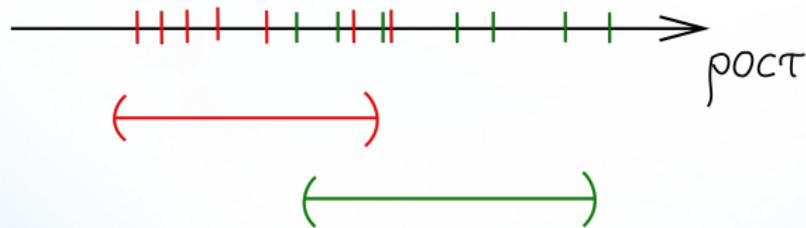
Отличается ли их средний рост?

Собираем данные



отличается

Если данных мало

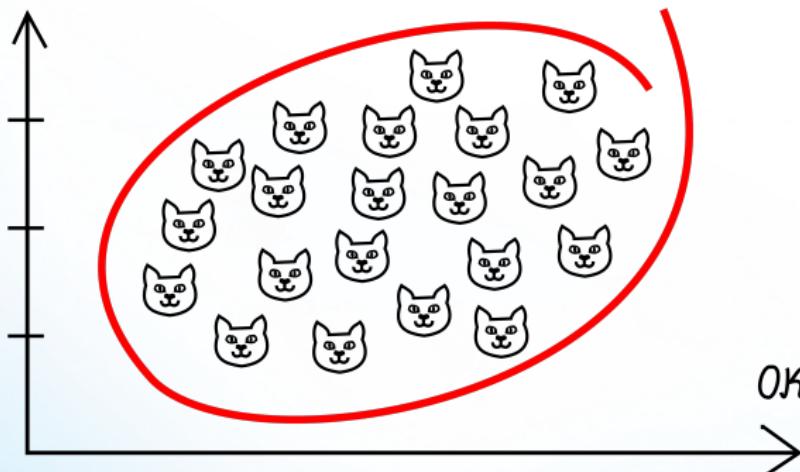


непонятно

Статистические гипотезы, АВ-тесты

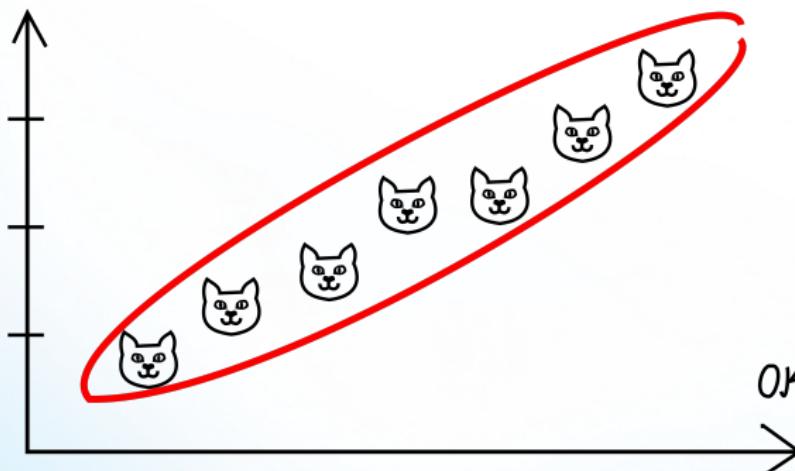


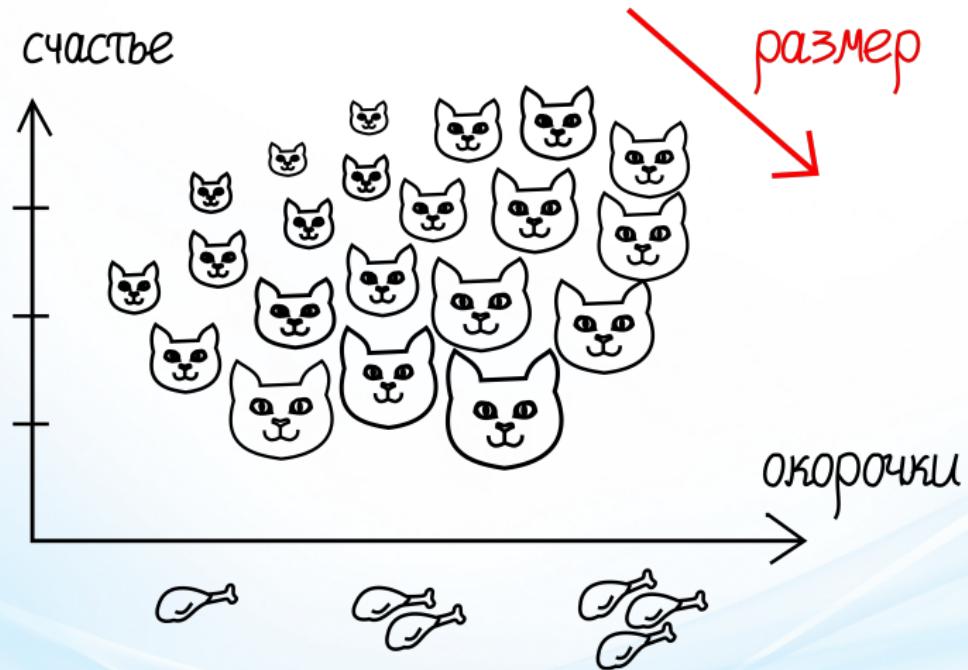
счастье





счастье





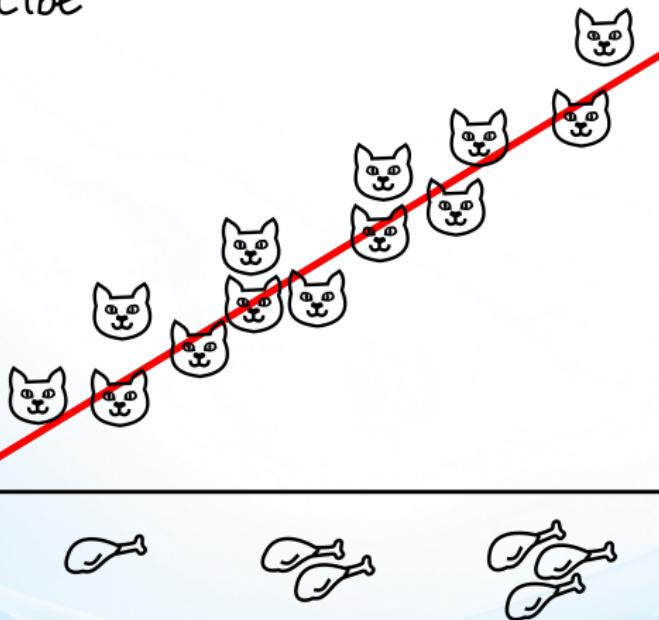
Корреляционный анализ



счастье



окорочки



Формула счастья


$$\text{счастье} = \theta_0 + \theta_1 \times \text{кал-бо} + \text{погрешности}$$



Больше окорочков

счастье



Формула счастья



$$= \theta_0 + \theta_1 \times \text{кал-бо} - \theta_2 \times (\text{кал-бо})^2$$

+ погрешности



Другие факторы

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

Регрессионный анализ

Классификация котиков

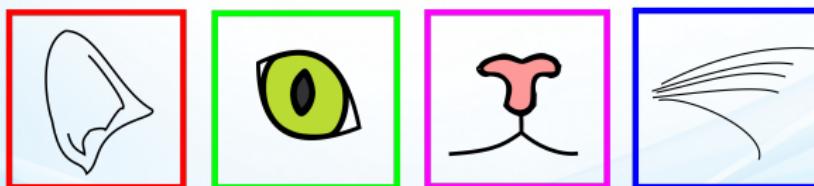
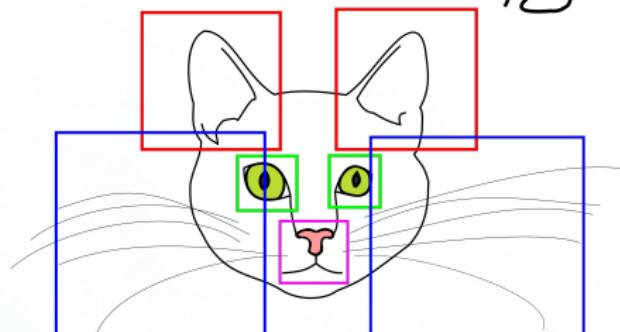


Классификация

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



Гаврилова



Книга с похожим содержанием

