

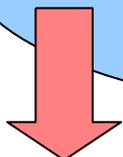
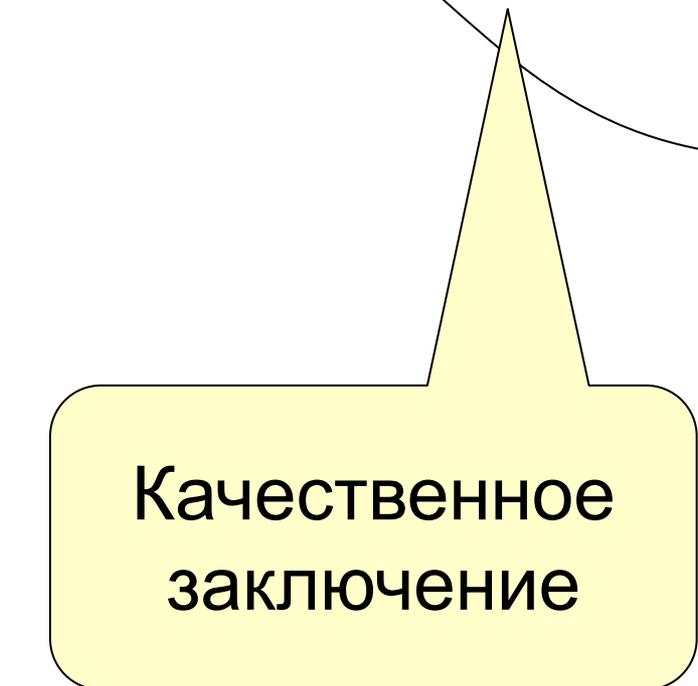
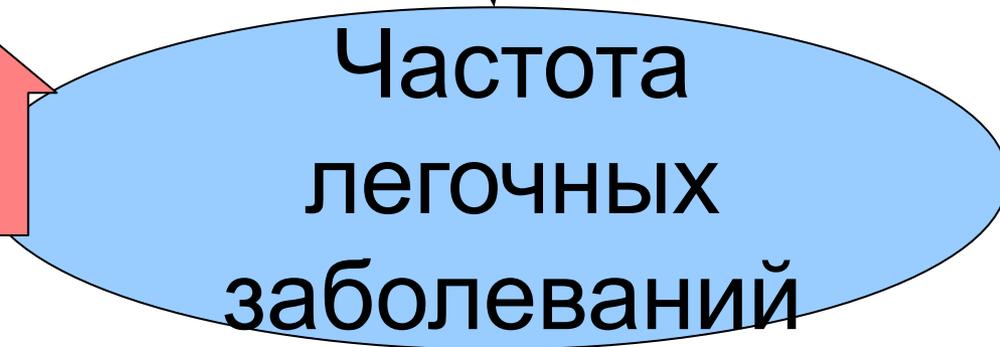
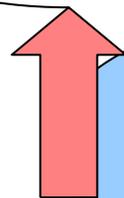
# **Модели статистического прогнозирования**

Определим  
характер  
зависимости

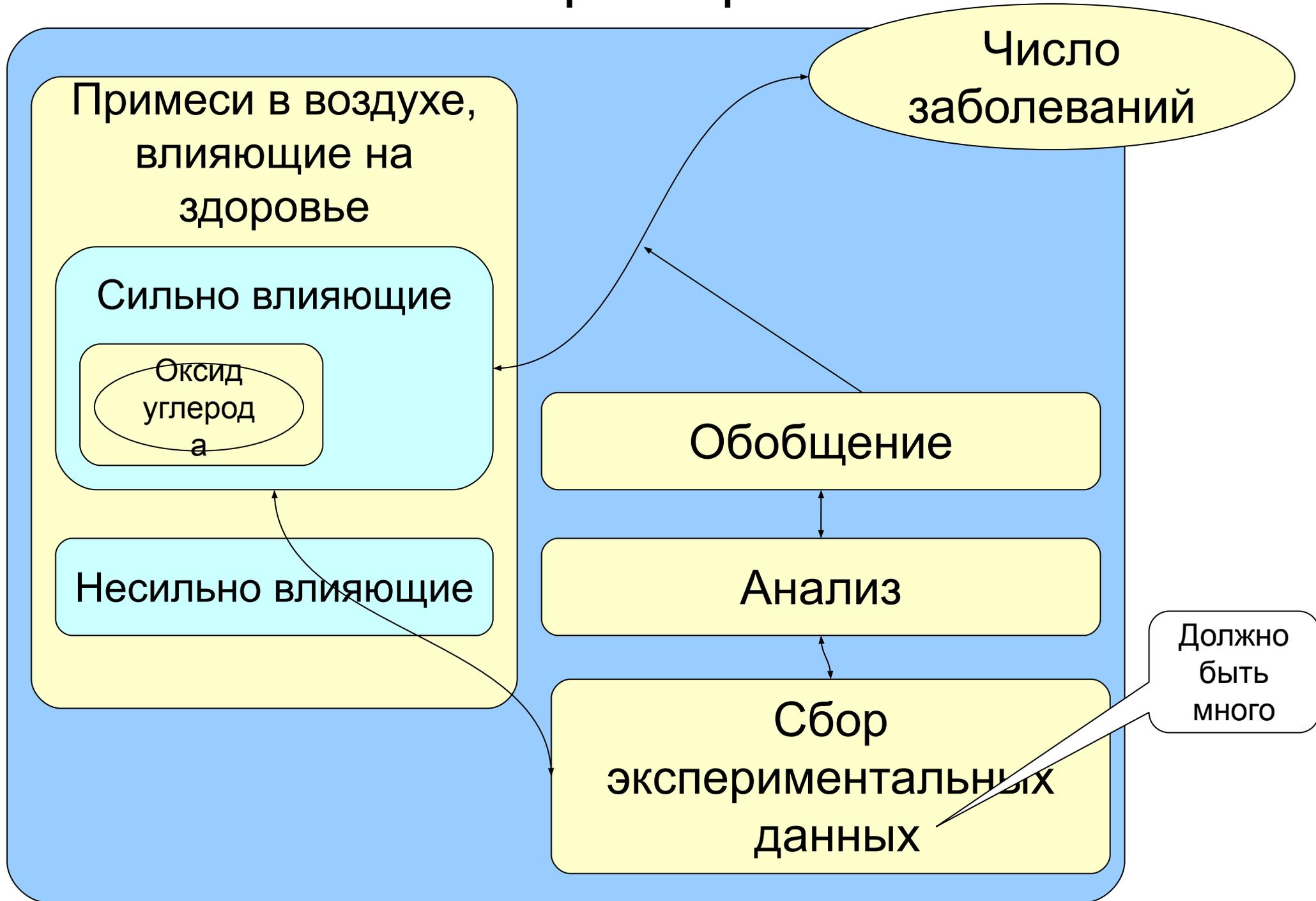
Качество  
воздуха  
в городе

Частота  
легочных  
заболеваний

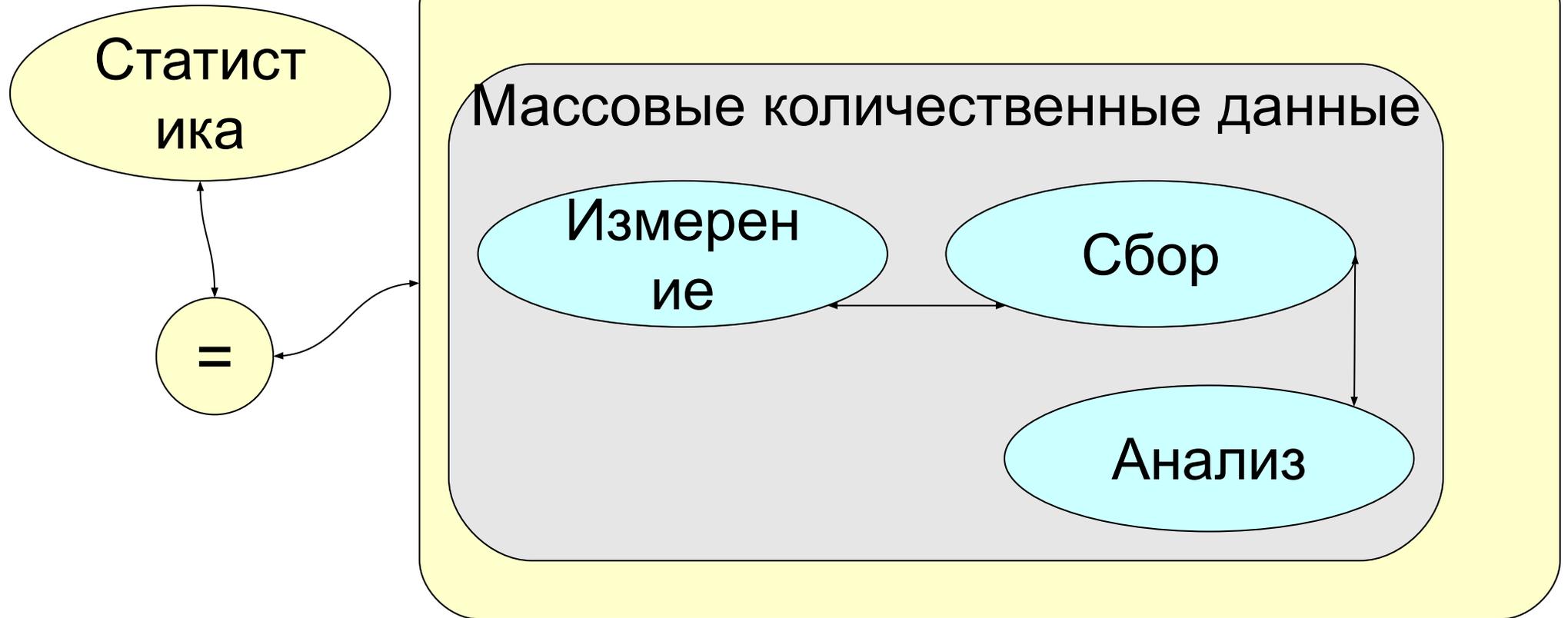
Качественное  
заключение



# Уточнение характера зависимости



# Наука



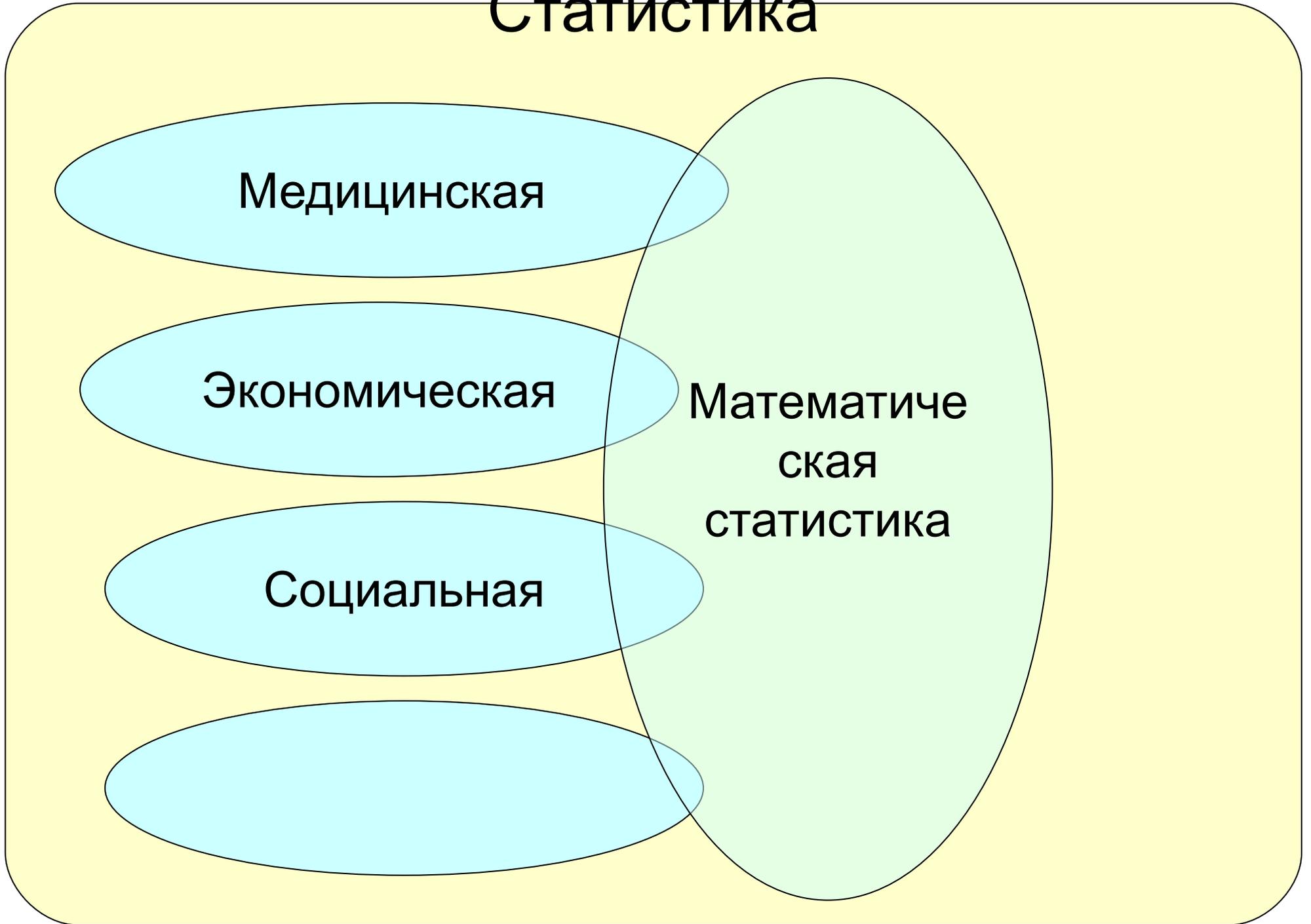
# Статистика

Медицинская

Экономическая

Социальная

Математическая  
статистика



# Пример из медиц. статистики:

## Представление экспериментальных данных

Средняя концентрация угарного газа

Число хронических больных на 1000 жителей

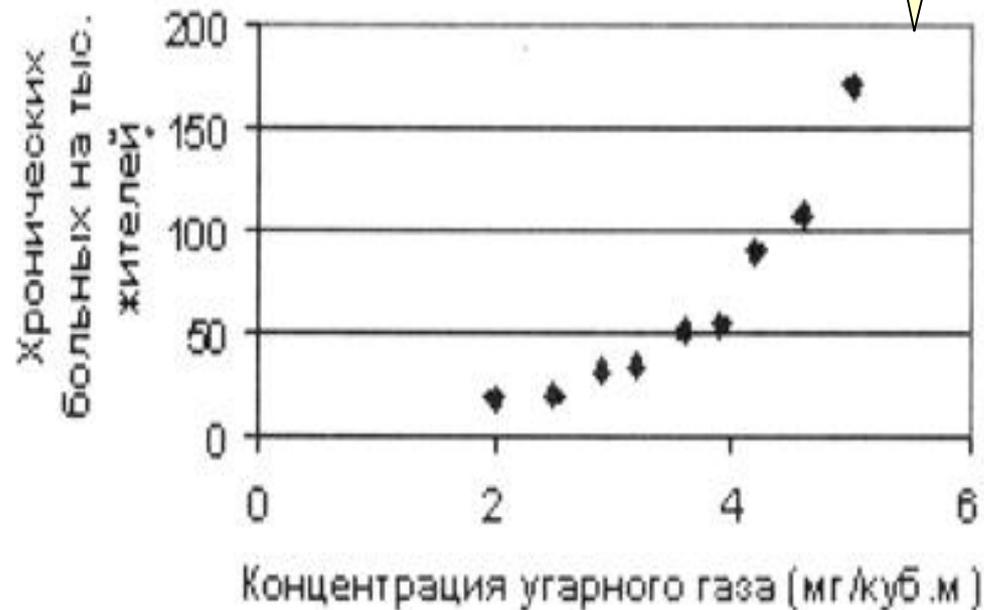
Графическое

$C,$ мг/м <sup>3</sup>	$P,$ бол./тыс.
2	19
2,5	20
2,9	32
3,2	34
3,6	51
3,9	55
4,2	90
4,6	108
5	171

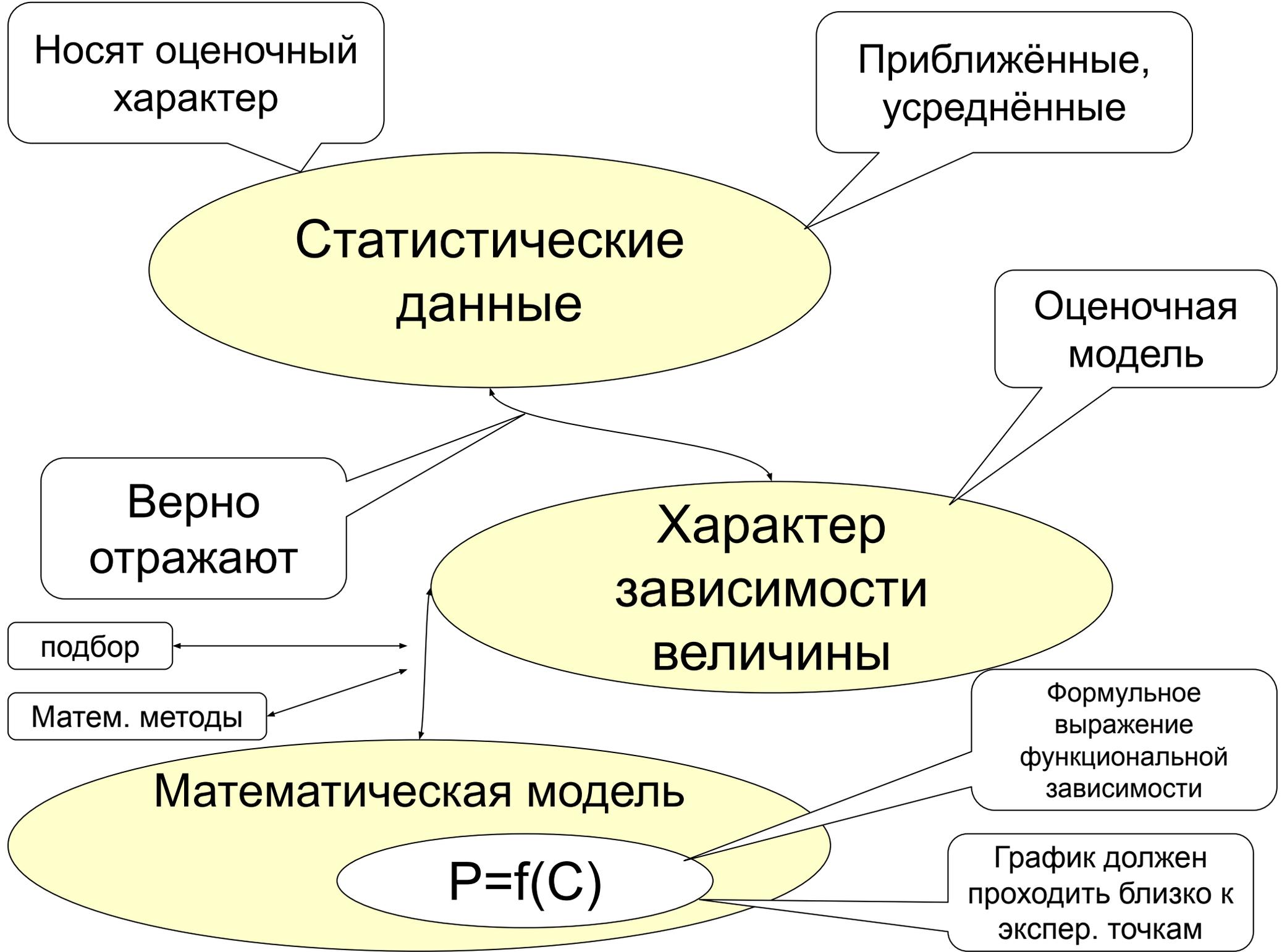
Табличное

Несильное влияние

Заболеваемость астмой



Резкий рост заболеваемости



Несут оценочный характер

Приближённые, усреднённые

Статистические данные

Оценочная модель

Верно отражают

Характер зависимости величины

подбор

Матем. методы

Математическая модель

Формульное выражение функциональной зависимости

$$P=f(C)$$

График должен проходить близко к экспер. точкам

Приближенные  
данные

# Искомая функция

График проходит  
через все  
эксперимента  
льные  
точки

## Основные требования

Достаточная простота

Отклонения точек от  
графика

Минимальн  
ы

Равномерн  
ы

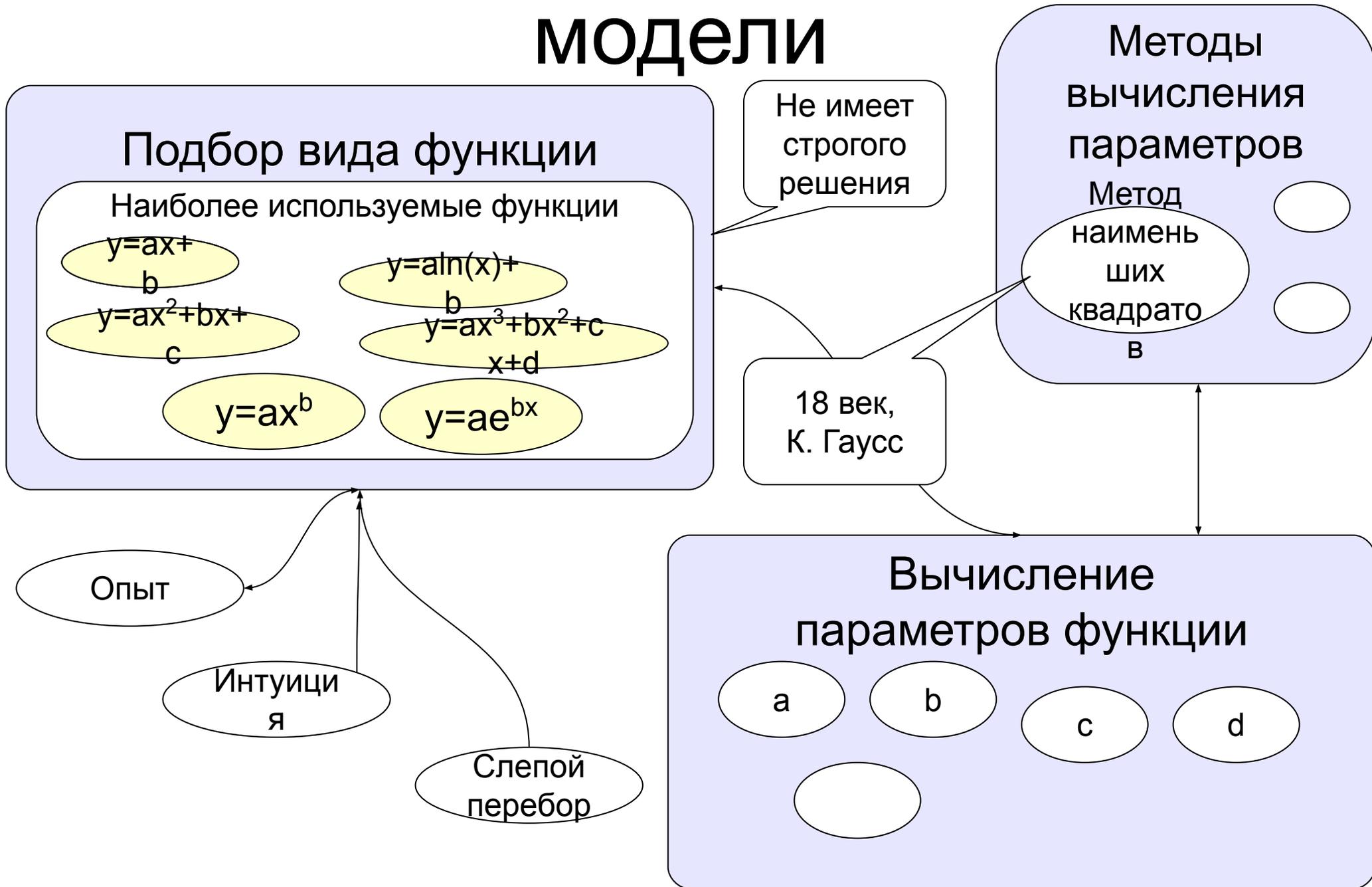
Нет смысла

Слишком  
сложный  
вид функции

Регрессионная  
модель

Удобно использовать  
в дальнейших вычислениях

# Получение регрессионной модели



$y = ax + b$  — линейная функция;

$y = ax^2 + bx + c$  — квадратичная

функция;

$y = a \ln(x) + b$  — логарифмическая

функция;

$y = ae^{bx}$  — экспоненциальная

функция;

$y = ax^b$  — степенная функция.

**Квадратичная функция называется в математике полиномом второй степени.**

**Иногда используются полиномы и более высоких степеней, например полином третьей степени имеет вид:**

$$y = ax^3 + bx^2 + cx + d.$$

**Во всех этих формулах:**

**$x$  — аргумент,**

**$y$  — значение функции,**

**$a, b, c, d$  — параметры функции,**

**$\ln(x)$  — натуральный логарифм,**

**$e$  — константа, основание  
натурального логарифма.**

МНК

Искомая функция должна быть построена так, чтобы сумма квадратов отклонений у-координат всех экспериментальных точек от у-координат графика функции была минимальной

Искомая функция

$$(y_{1э} - y_{1ф})^2$$

$$(y_{2э} - y_{2ф})^2$$

$$(y_{iэ} - y_{iф})^2$$

$\Sigma$

$$\sum_{i=1}^n (y_{iэ} - y_{iф})^2$$

min

# Статистическая обработка данных

Используемые  
математические  
пакеты  
программ

МНК

Критерии  
соответствия

Построение  
любой  
функции

Регрессионная  
модель

График

Тренд

**С первого взгляда хочется отбраковать вариант линейного тренда. График линейной функции — это прямая.**

**Полученная по МНК прямая отражает факт роста заболеваемости от концентрации угарного газа, но по этому графику трудно что-либо сказать о характере этого роста.**

**А вот квадратичный и экспоненциальный тренды правдоподобны.**

**Полученные функции:**

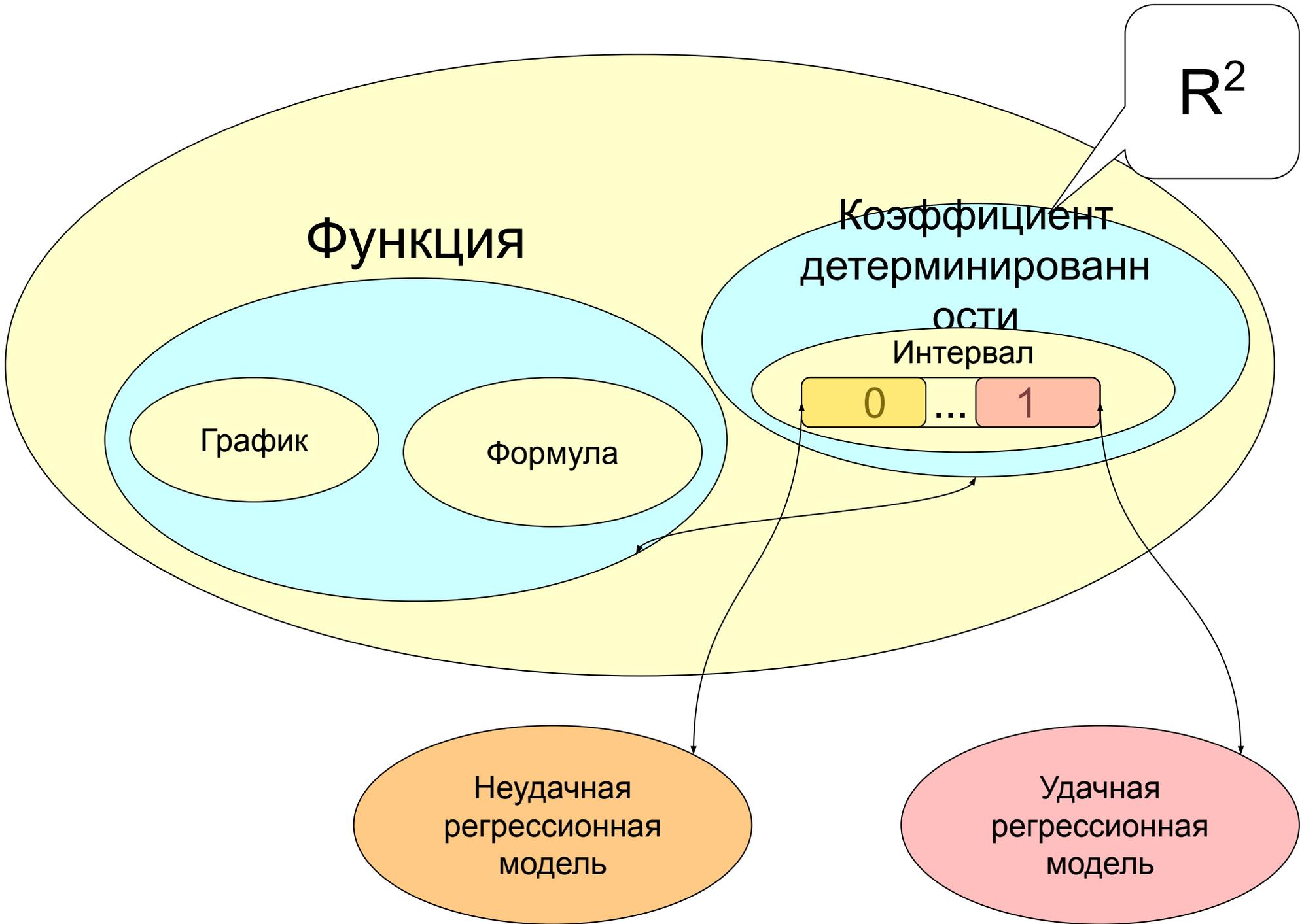
**линейная функция:  $y = 46,361x - 99,881$ ;**

**экспоненциальная функция:**

$$y = 3,4302 e^{0,7555x} ;$$

**квадратичная функция:**

$$y = 21,845x^2 - 106,97x + 150,21.$$



$R^2$

Функция

График

Формула

Коэффициент  
детерминированности

Интервал

0

...

1

Неудачная  
регрессионная  
модель

Удачная  
регрессионная  
модель

Значения,  
полученные  
путём  
измерений

В том числе  
с помощью ЭТ

Регрессионная  
математическая  
МОДЕЛЬ

В пределах  
экспериментальных  
значений

**Прогнозирование процесса**  
для других значений аргумента

За пределами  
экспериментальных  
данных

Восстановлени  
е  
значения

Экстраполяция

Графическ  
им  
способом

держится на  
гипотезе:  
предположим, что  
за пределами  
экспериментально  
й области  
закономерность  
сохраняется

**КОНЕЦ ФИЛЬМА**