

# Работа в Excel 2007

1. Основы
2. Диаграммы
3. Численные методы
4. Статистика
5. Восстановление зависимостей
6. Моделирование

# Работа в Excel 2007

## Тема 1. Основы

# Электронные таблицы

---

**Основная задача** – автоматические вычисления с данными в таблицах.

## Кроме того:

- хранение данных в табличном виде
- представление данных в виде диаграмм
- анализ данных
- составление прогнозов
- поиск оптимальных решений
- подготовка и печать отчетов

## Примеры:

- *Microsoft Excel* – файлы \*.xls, \*.xlsx 
- *OpenOffice Calc* – файлы \*.ods – **бесплатно** 

# Электронные таблицы

активная  
ячейка

имена столбцов

	A	B	C	D
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				
11				

номера  
строк

строка

неактивная  
ячейка


- текст
- числа
- **формулы**
- время
- дата

столбец

# Начало работы с *Microsoft Excel*



Программы – Microsoft Office – Excel 2007

Файлы:  \*.xlsx (старая версия – \*.xls)

Вася.xlsx рабочая книга

Лист 1

Лист 2

План  
по валу

Вал  
по

переходы  
по листам

ЛКМ

ПКМ

Вставить...

Удалить

Переименовать

Переместить/скопировать...

 Исходный текст

 Защитить лист...

Цвет ярлычка

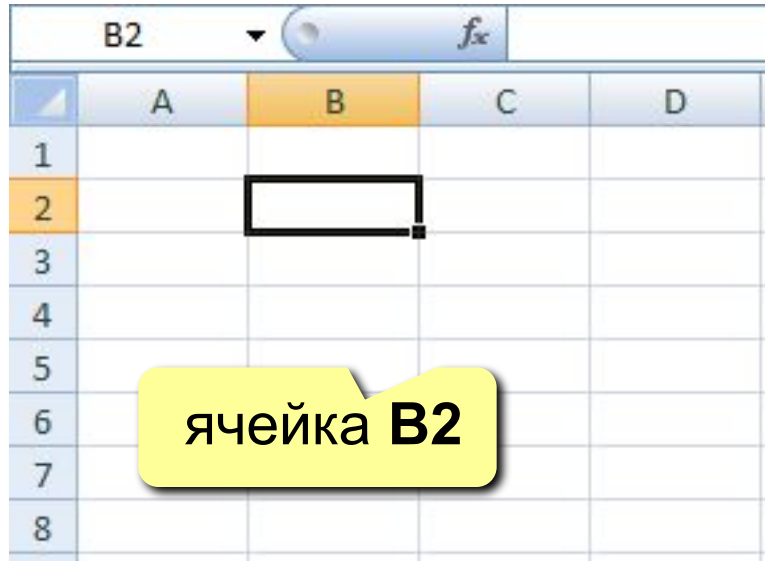
Скрыть

Отобразить...

Выделить все листы

# Адреса

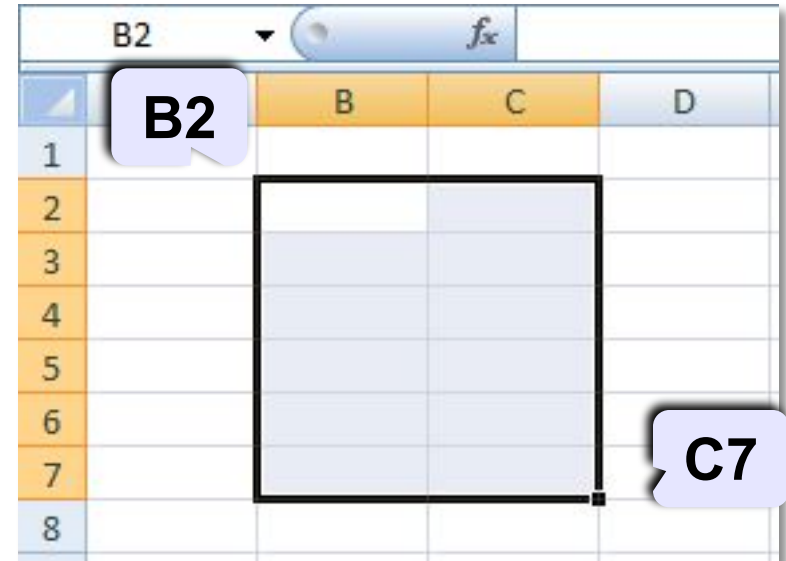
адрес активной ячейки



The screenshot shows an Excel spreadsheet with columns A, B, C, and D, and rows 1 through 8. The active cell is B2, which is highlighted with a thick black border. A yellow callout bubble points to the cell with the text "ячейка B2".

	A	B	C	D
1				
2				
3				
4				
5				
6				
7				
8				

диапазон B2:C7



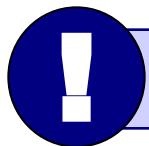
The screenshot shows an Excel spreadsheet with columns A, B, C, and D, and rows 1 through 8. The range B2:C7 is selected, indicated by a thick black border and a light blue fill. Two callout bubbles are present: one labeled "B2" pointing to the top-left corner of the range, and another labeled "C7" pointing to the bottom-right corner.

	A	B	C	D
1				
2				
3				
4				
5				
6				
7				
8				

## Ссылки в формулах:

$=B2+2*C3$

$=A2+2*СУММ(B2:C7)$



Формула всегда начинается знаком «=»!

# Ввод данных

адрес активной ячейки

отменить (*Esc*)

принять (*Enter*)

В2

Привет!

строка редактирования

ЛКМ

	A	B	C
1			
2		Привет!	
3			
4			
5			

**F2** – редактировать прямо в ячейке

# Выделение данных

ячейка:

	A	B
1		
2		
3		

диапазон:

	A	B	C	D
1				
2				
3				
4				
5				
6				
7				

строки:

	A	B	C	D
1				
2				
3				
4				

несвязанные диапазоны:

**+Ctrl** и выделять второй

столбцы:

	A	B	C	D
1				
2				
3				
4				

вся таблица:

	A	B	C	D
1				
2				
3				
4				



# Операции со строками и столбцами

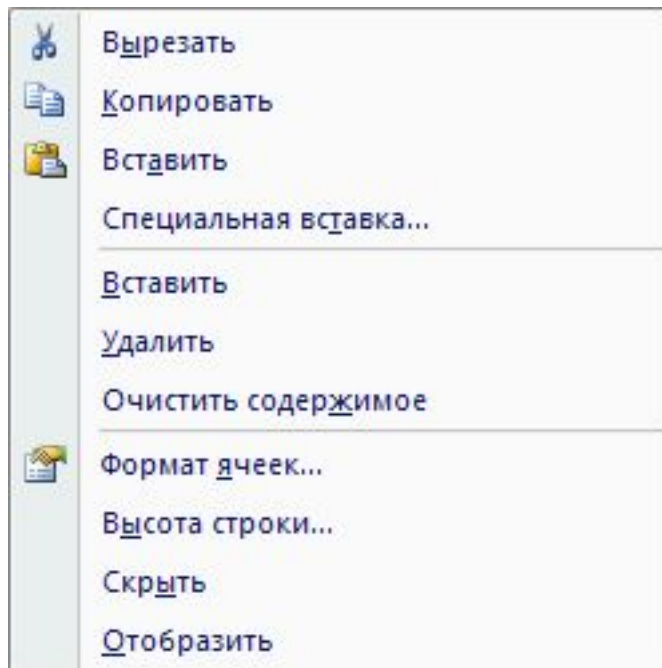
## размеры

высота  
строк

	A	B	C
1			
2			
3			

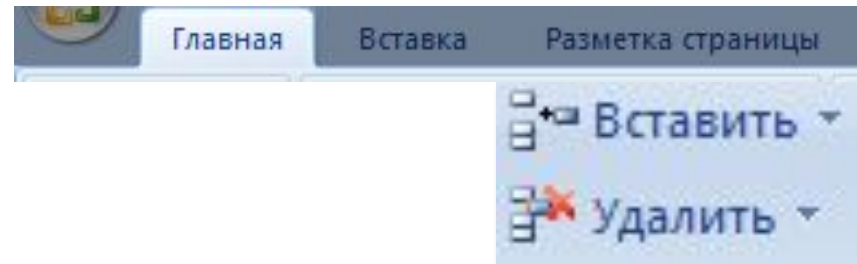
ширина  
столбцов

## добавление, удаление

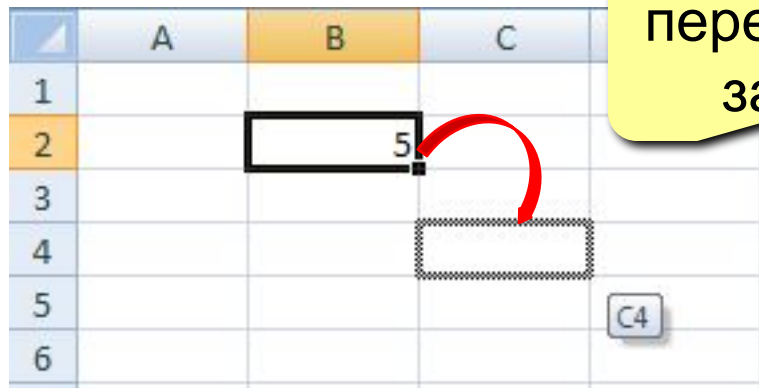


ПКМ

	A	B	C
1	1		
2	4		
3	5		
4	6		



# Перемещение и копирование

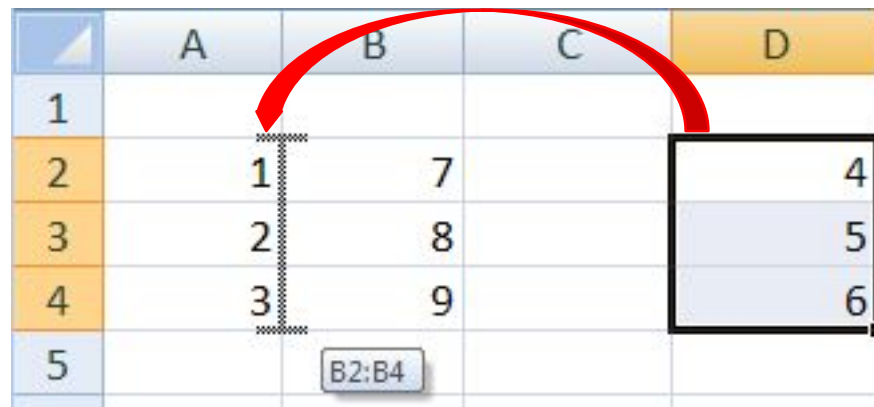


	A	B	C
1			
2		5	
3			
4			
5			
6			

перетащить ЛКМ за рамку (!)

**+Ctrl** = копирование  
**+Alt** = на другой лист

**перемещение со сдвигом (+Shift)**



	A	B	C	D
1				
2		1	7	4
3		2	8	5
4		3	9	6
5				

B2:B4

# Типы ссылок

**относительные** (меняются так же, как и адрес формулы )

	A	B	C
1			
2		=B5+C8	
3			=C6+D9

формула «переехала» на один столбец вправо и на одну строку вниз;  
**имя столбца** ↑ на 1  
**номер строки** ↑ на 1

## абсолютные

(не меняются)

	A	B
1	=\$B\$5+\$C\$8	=\$B\$5+\$C\$8
2	=\$B\$5+\$C\$8	=\$B\$5+\$C\$8
3	=\$B\$5+\$C\$8	=\$B\$5+\$C\$8

## смешанные

(меняется только относительная часть)

	A	B	C
1	=\$B4+B\$8	=\$B4+C\$8	=\$B4+D\$8
2	=\$B5+B\$8	=\$B5+C\$8	=\$B5+D\$8
3	=\$B6+B\$8	=\$B6+C\$8	=\$B6+D\$8

# Заполнение рядов

арифметическая прогрессия

	A
1	1
2	2
3	3
4	4
5	5
6	6
7	

ма  
запо

ЛКМ

копирование формул

	A	B
1	1	=A1^2
2	2	=A2^2
3	3	=A3^2
4	4	=A4^2
5	5	=A5^2
6	6	=A6^2
7		

ЛКМ

даты

	A
1	02.02.2009
2	05.02.2009
3	08.02.2009
4	11.02.2009
5	14.02.2009
6	

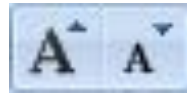
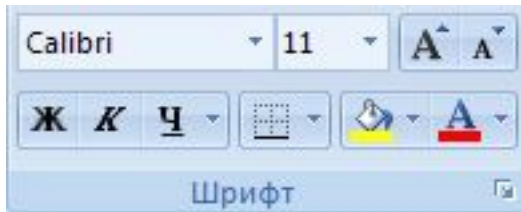
время

	A
1	12:00:00
2	12:20:00
3	12:40:00
4	13:00:00
5	13:20:00
6	

СПИСКИ

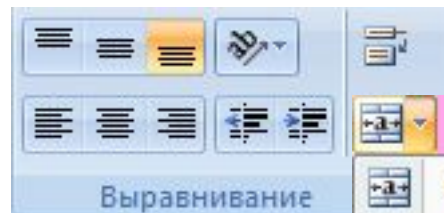
	A	B
1	январь	
2	февраль	
3	март	
4	апрель	
5	май	
6		

# Оформление ячеек



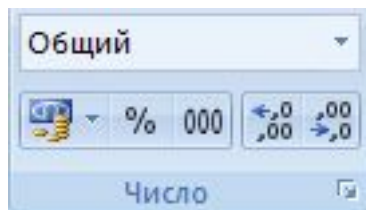
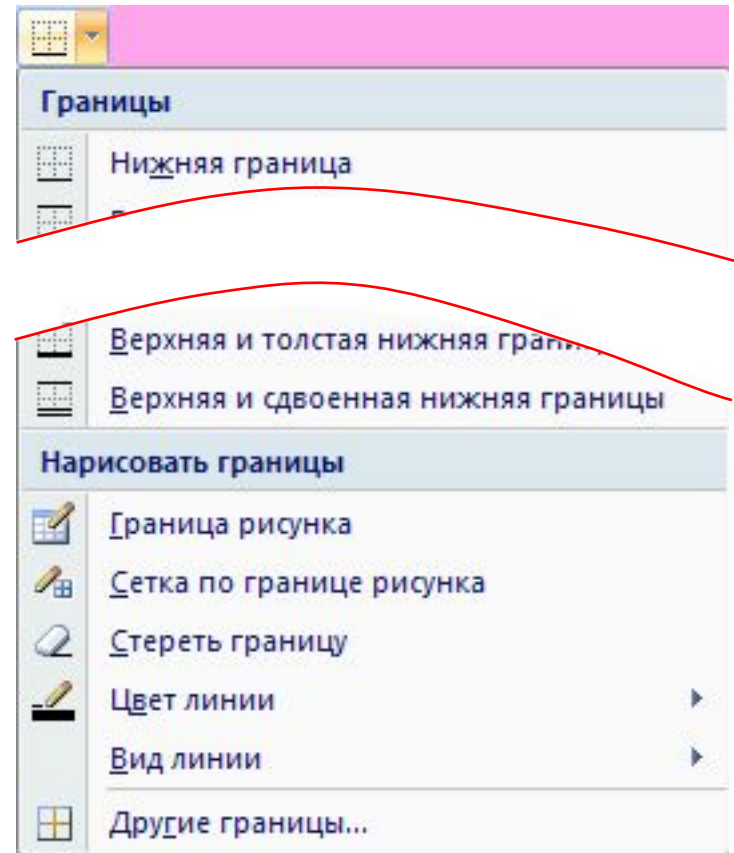
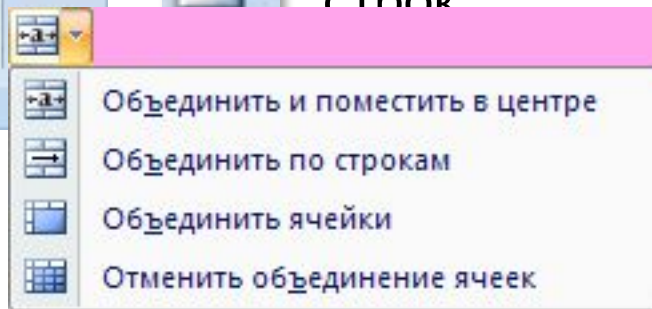
↑↓ размер

все свойства



направление

в несколько  
строк



денежный  
формат



количество знаков  
в дробной части

# Функции

ввод в строке редактирования

	A	B	C
1	1	2	5
2	3	4	6
3			=СУММ(A1:B2;C1)
4			

изменение  
диапазона

ввод в ячейке

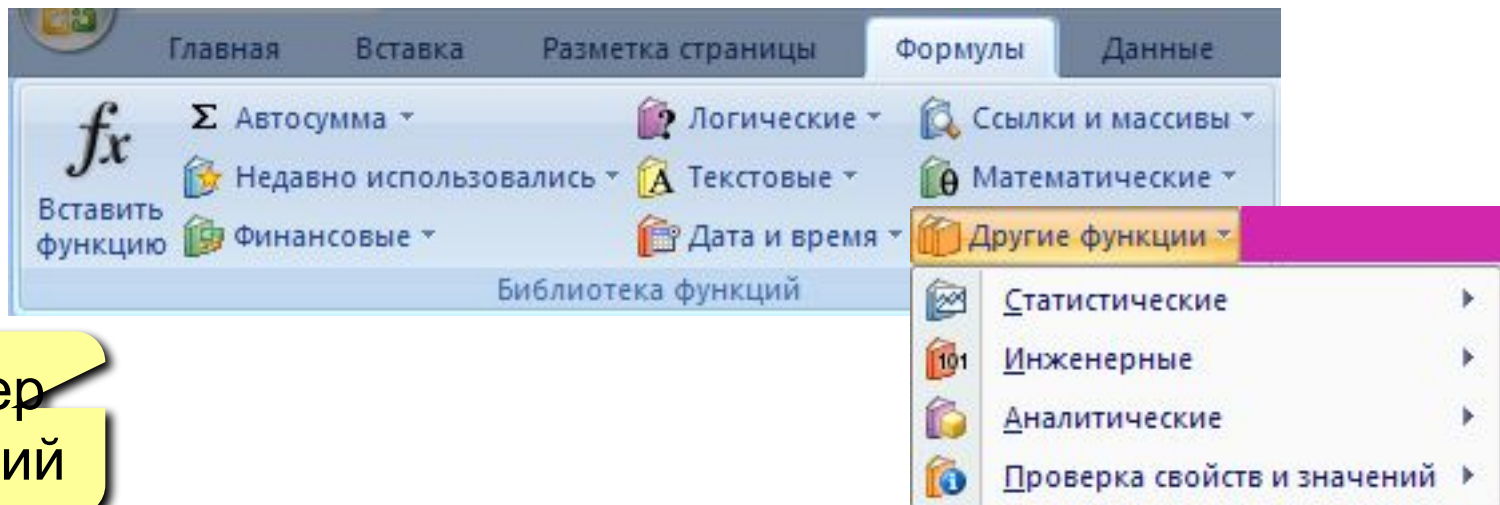
диапазон

ячейка



**Можно мышкой!**

мастер  
функций



# Некоторые функции

**СУММ** – сумма значений ячеек и диапазонов

**СРЗНАЧ** – среднее арифметическое числовых ячеек

**МИН** – минимальное значение

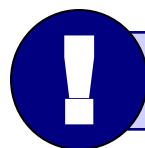
**МАКС** – максимальное значение

	A	B	C	D
1	1	3	=СУММ(A1:B2)	=МИН(A1:B2)
2	2	4	=СРЗНАЧ(A1:B2)	=МАКС(A1:B2)

	A	B	C	D
1	1	3	10	1
2	2	4	2,5	4

	A	B	C
1	1		=СРЗНАЧ(A1:B2)
2		3	

	A	B	C
1	1		2
2		3	



**Пустые и нечисловые ячейки не учитываются!**

# Некоторые функции

Сумма значений нескольких диапазонов  
(СУММ, СРЗНАЧ, МИН, МАКС)

	A	B	C	D	E
1	1	2	3		9
2					10
3	6	7	8		11
4					
5	=СУММ(A1:C1;A3:C3;E1:E3)				



# Некоторые функции

**СУММПРОИЗВ** – сумма произведений двух массивов

	A	B	C
1	<b>Товар</b>	<b>Цена за 1 кг</b>	<b>Вес, кг</b>
2	Бананы	50р.	1,5
3	Апельсины	65р.	2,3
4	Киви	80р.	0,7
5	Абрикосы	120р.	0,5
6			
7	<b>Всего:</b>	<b>=СУММПРОИЗВ(B2:B5;C2:C5):5</b>	

# Некоторые функции

**СЧЁТ** – количество числовых ячеек

**СЧЁТЕСЛИ** – количество ячеек, удовлетворяющих заданному условию (2-ой параметр)

	A	B	C
1	1	1	=СЧЁТ(A1:B3)
2	2	2	=СЧЁТЕСЛИ(A1:B3;"<3")
3	4		=СЧЁТЕСЛИ(A1:B3;1)

	A	B	C
1	1	1	5
2	2	2	4
3	4		2

# Некоторые функции

**СУММЕСЛИ** – сумма ячеек диапазона, удовлетворяющих условию, которое накладывается на **другой диапазон**

Найти общий вес учащихся 9-А класса.

	А	В	С	Д
1	Класс	Фамилия	Вес	
2	9-А	Иванов	45	
3	9-Б	Петров	52	
4	9-А	Сидоров	48	
5	9-Б	Васильев	55	
6	9-А	Семенов	51	
7				
8				=СУММЕСЛИ(A2:A6;"9-А";C2:C6)

диапазон для проверки условия

условие

диапазон для суммирования

# Функция ЕСЛИ

**ЕСЛИ** – выбор из двух вариантов

	А	В
1	Баллы	
2	68	=ЕСЛИ(A2>80;"сдал";"не сдал")
3	75	=ЕСЛИ(A3>=70;"сдал";"не сдал")
4	37	
5	88	

условие

если «да»

если «нет»

=ЕСЛИ (B2="сдал" ; ЕСЛИ (A2>80 ; 5 ; 4) ; "-")

	А	В
1	Баллы	Результат
2	68	не сдал
3	75	сдал
4	37	не сдал
5	88	сдал

# Логические операции

**НЕ** – обратное условие,  $\text{НЕ} (B2 < 10) \Leftrightarrow B2 \geq 10$

**И** – одновременное выполнение всех условий

	A	B	C	D
1	Фамилия	Год рождения	Рост	Принят
2	Алексеев	1995	176	=ЕСЛИ(И(B2>1994;C2>175);"да";"-")
3	Березин	1995	167	=ЕСЛИ(И(B3>1994;C3>175);"да";"-")
4	Викторов	1994	180	=ЕСЛИ(И(B4>1994;C4>175);"да";"-")

=ЕСЛИ ( И (B2>1994 ; C2>175) ; "да" ; "-")

	A	B	C	D
1	Фамилия	Год рождения	Рост	Принят
2	Алексеев	1995	176	да
3	Березин	1995	167	-
4	Викторов	1994	180	-

# Логические операции

**ИЛИ** – выполнение хотя бы одного из условий

	A	B	C	D
1	Фамилия	Математика	Физика	Принят
2	Алексеев	100	67	=ЕСЛИ(ИЛИ(B2=100;C2=100;B2+C2>=180);"да";"-")
3	Березин	98	98	=ЕСЛИ(ИЛИ(B3=100;C3=100;B3+C3>=180);"да";"-")
4	Викторов	90	80	=ЕСЛИ(ИЛИ(B4=100;C4=100;B4+C4>=180);"да";"-")

=ЕСЛИ( ИЛИ (B2=100; C2=100; B2+C2>=180) ; "да" ; "-" )

	A	B	C	D
1	Фамилия	Математика	Физика	Принят
2	Алексеев	100	67	да
3	Березин	98	98	да
4	Викторов	90	80	-

# Подсчёт числовых значений

**СЧЁТ** – считает ячейки с числами или формулами, которые дают числа

**=A1+1**

	A	B	C
1	1	2	
2		Вася	
3			<b>=СЧЁТ(A1:B2)</b>

**2**

# Подсчёт значений по условию

**СЧЁТЕСЛИ** – считает ячейки, удовлетворяющие условию

	A	B	C
1	1	2	=СЧЁТЕСЛИ(A1:B4;1)
2	3	1	
3	яблоко	дыня	
4	apple	дыня	

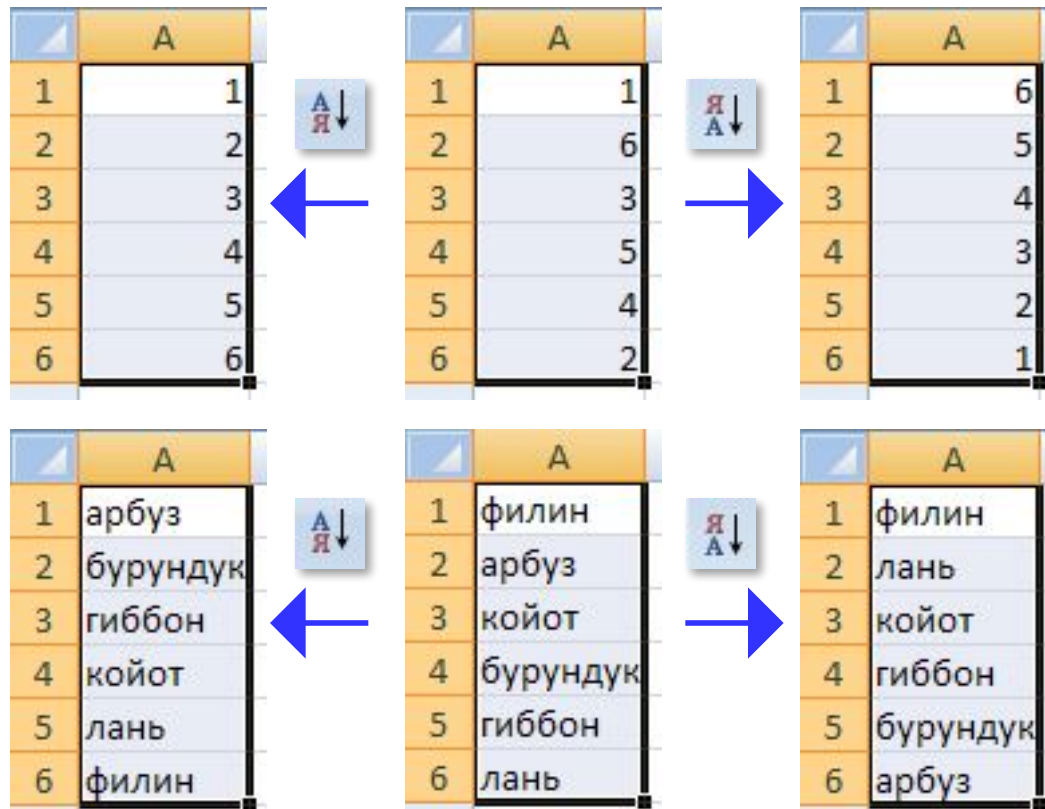
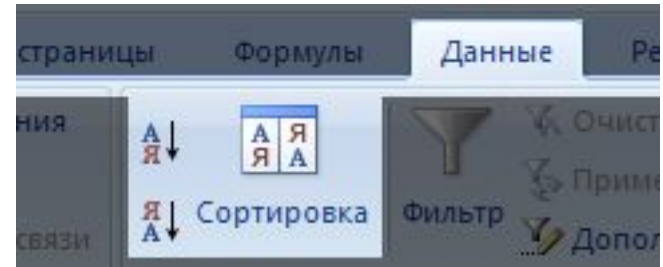
Diagram illustrating the COUNTIF function. A spreadsheet shows a range of cells A1:B4 containing values 1, 2, 3, 1, 'яблоко', 'дыня', 'apple', 'дыня'. Cell C1 contains the formula =СЧЁТЕСЛИ(A1:B4;1). Callout boxes with numbers 1, 2, and 3 point to the range A1:B4, the value 1, and the formula bar respectively.



# Сортировка

**Сортировка** – это расстановка элементов в заданном порядке.

## Сортировка одного столбца



# Сортировка связанных данных

	А	В	С
1	Фам	А	В
2	Иван	1	Фамилия
3	Петр	2	Иванов
4	Сидо	3	Иванов
5	Сидо	4	Петров
6	Петр	5	Петров
7	Иван	6	Петров
8	Петр	7	Сидоров
		8	Сидоров

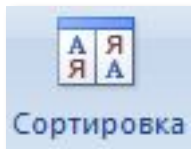


Почему нельзя  
сортировать по  
столбцу?

критерий

строки или  
столбцы

первая  
строка – это  
заголовки



**Сортировка**

Добавить уровень   
  Удалить уровень   
  Копировать уровень   
 Параметры...   
 Мои данные содержат заголовки

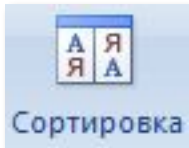
Столбец	Сортировка	Порядок
Сортировать по	Фамилия	Значения
	Фамилия Имя Вес	От А до Я

# Многоуровневая сортировка

	А	В	С
1	Фамилия	Имя	Вес
2	Иванов	Иван	
3	Петров	Георгий	
4	Сидоров	Степан	
5	Сидоров	Матвей	
6	Петров	Денис	
7	Иванов	Кузьма	
8	Петров	Семен	

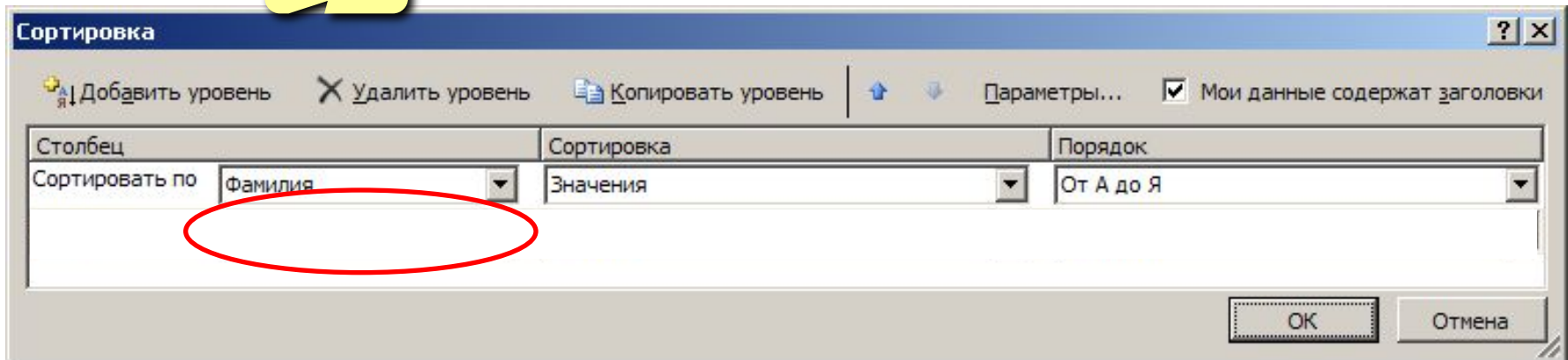
**Задача:** расставить фамилии по людям *с одинаковыми* расставить в порядке *по именам*.

	А	В	С
1	Фамилия	Имя	Вес
2	Иванов	Иван	54
3	Иванов	Кузьма	55
4	Петров	Георгий	75
5	Петров	Денис	62
6	Петров	Семен	68
7	Сидоров	Матвей	67
8	Сидоров	Степан	64



Сортировка

ЛКМ



# Имена ячеек и диапазонов

## Присвоить имя

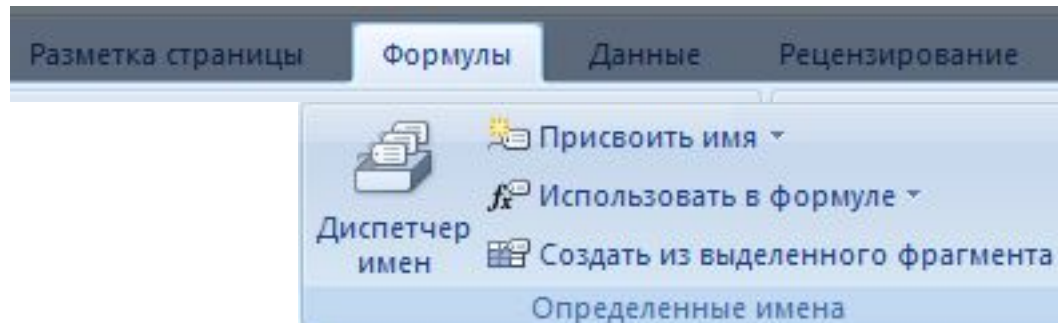
ВВЕСТИ ИМЯ

Таблица	A	B
1	1	2
2	3	4
3		

## Имена в формулах

СРЗНАЧ	A	B	C	D
1	1	2		
2	3	4		
3			=СРЗНАЧ(Таблица)	

## Работа с именами



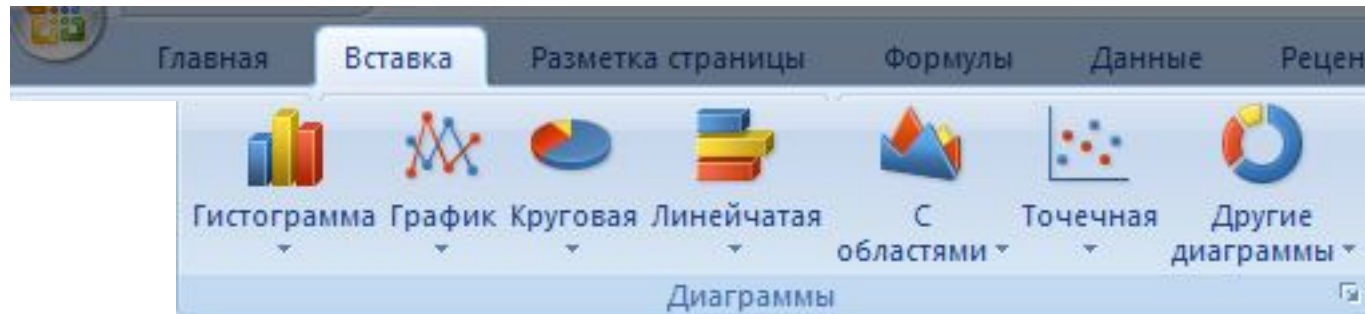
# Работа в Excel 2007

## Тема 2. Диаграммы

# Общий подход

---

- диаграммы строятся на основе данных таблицы
- проще всего сначала выделить все нужные данные, а потом...

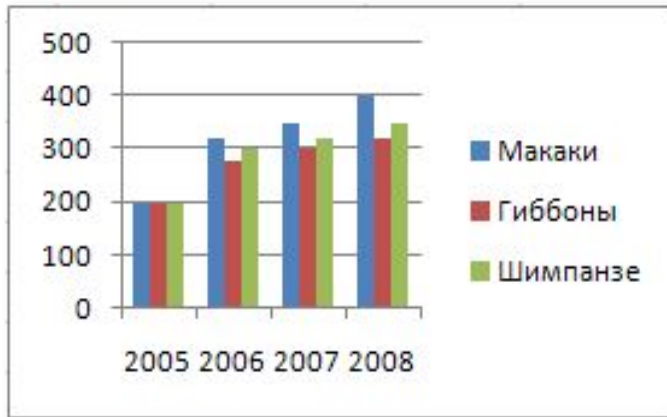


- все данные, которые должны обновляться автоматически, нужно выделить
- для выделения несвязанных диапазонов используем **+Ctrl**

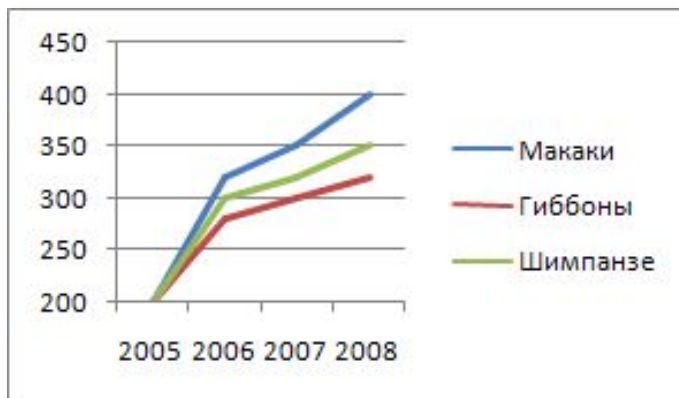
# Основные типы диаграмм

**Гистограмма (столбчатая диаграмма):**

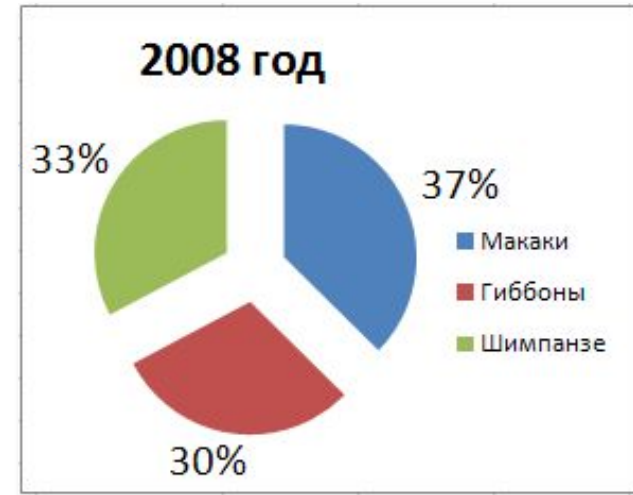
сравнение значений одного или нескольких рядов данных



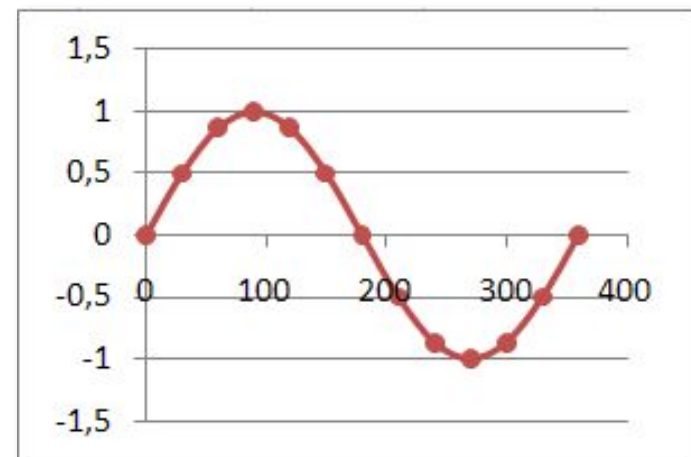
**График:** показывает изменение процесса во времени (равномерные отсчеты)



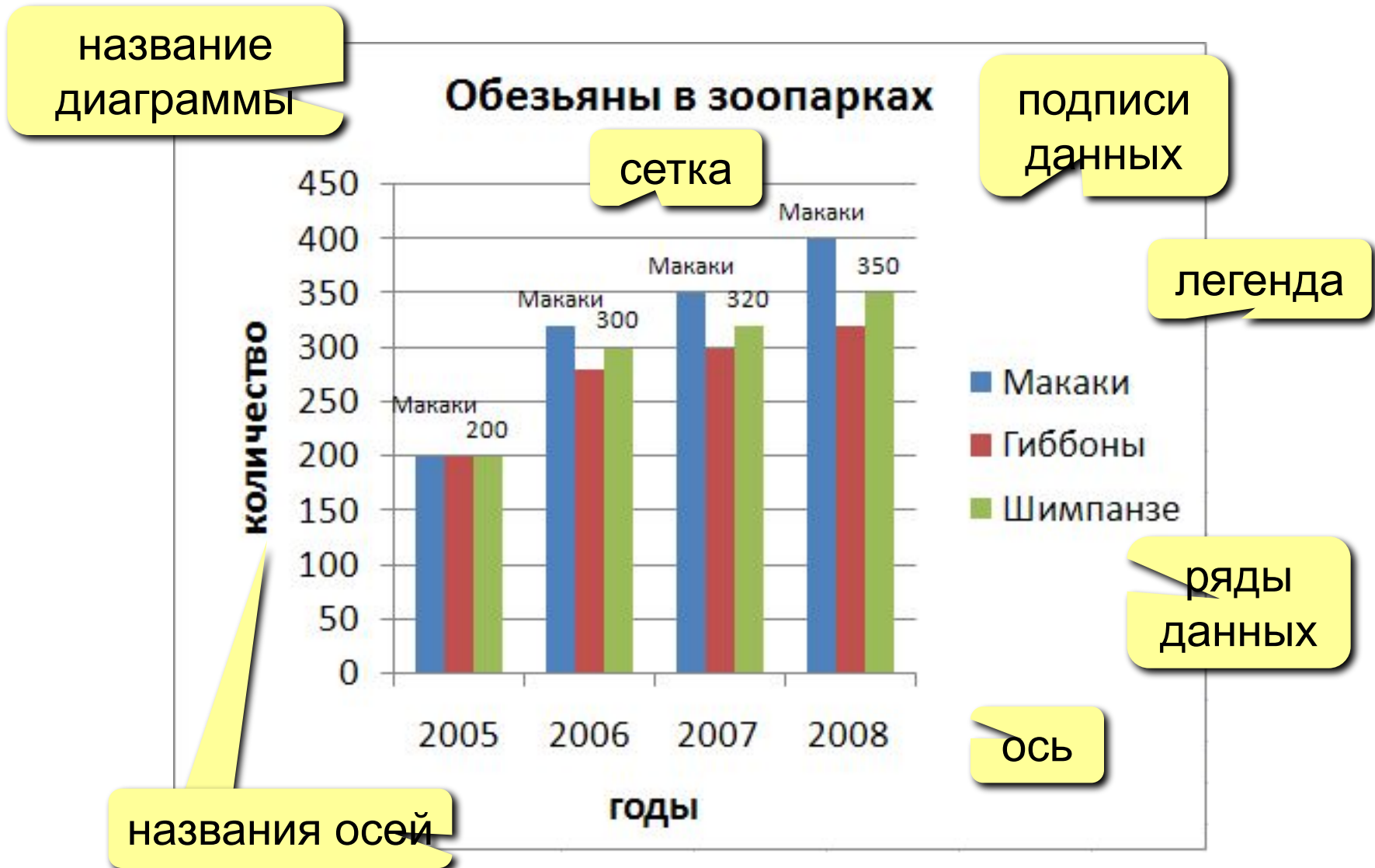
**Круговая:** доли в сумме



**Точечная:** связь между парами значений (график функции)

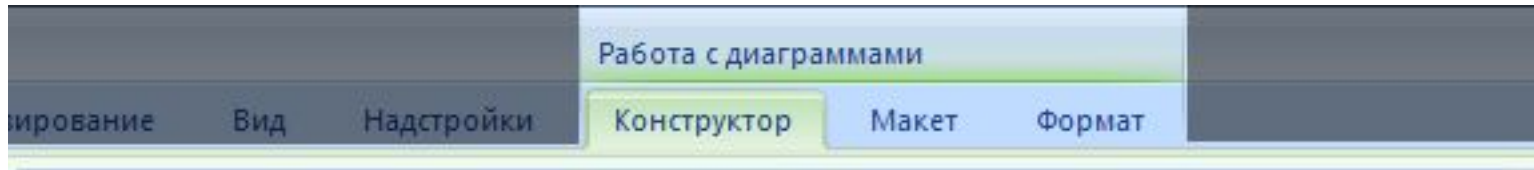


# Элементы диаграмм

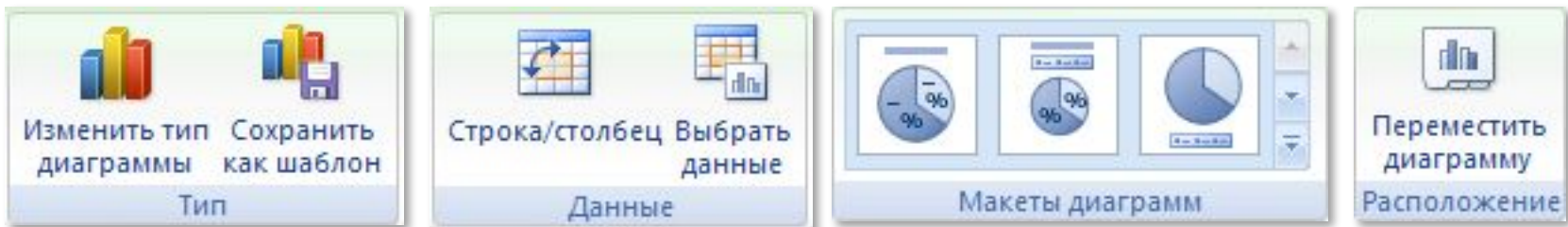




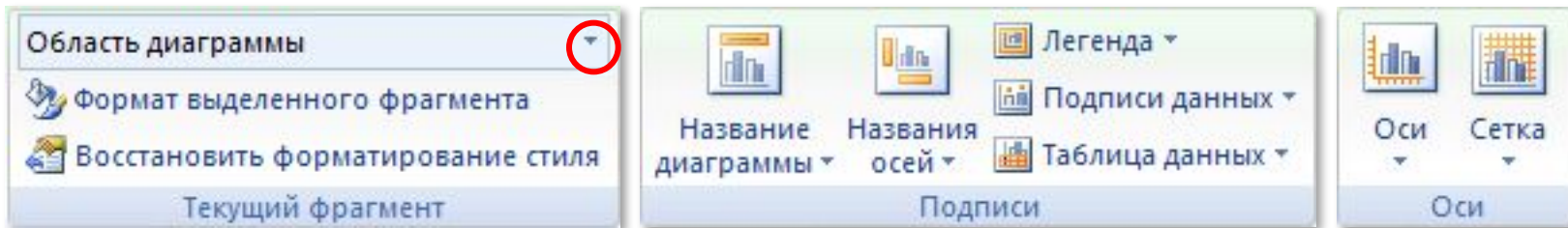
# Настройка диаграммы и ее элементов



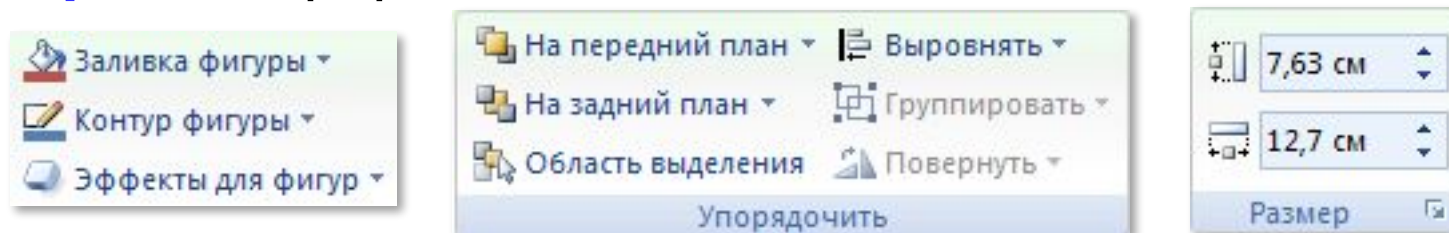
## Конструктор: общие свойства



## Макет: настройка свойств отдельных элементов



## Формат: оформление отдельных элементов



# Графики функций

**Задача:** построить график функции  $y = x^2$  для  $-5 \leq x \leq 5$ .

**Таблица значений функции:** шаг 0,5

	A	B
1	x	y
2	-5	
3	-4,5	
4	-4	
5	-3,5	
6	-3	
7	-2,5	
8	-2	

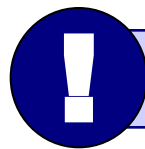
	A	B
1	x	y
2	-5	
3	-4,5	
4	-4	
5	-3,5	
6	-3	
7	-2,5	
8	-2	

	A	B
1	x	y
2	-5	=A2^2
3	-4,5	
4	-4	
5	-3,5	
6	-3	
7	-2,5	
8	-2	

	A	B
1	x	y
2	-5	=A2^2
3	-4,5	=A3^2
4	-4	=A4^2
5	-3,5	=A5^2
6	-3	=A6^2
7	-2,5	=A7^2
8	-2	=A8^2



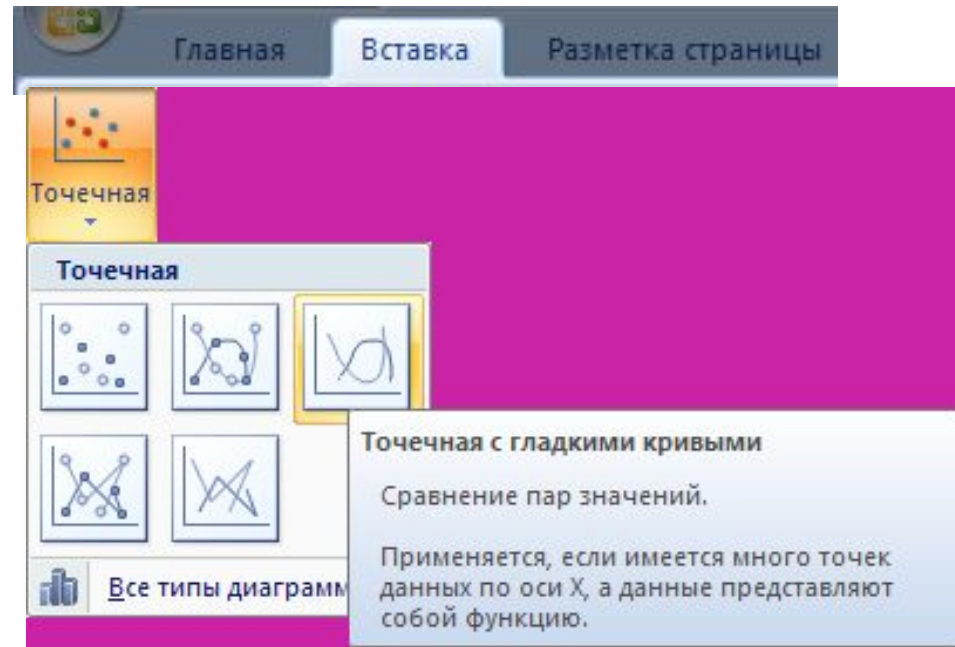
Что зависит от шага?

# Графики функций

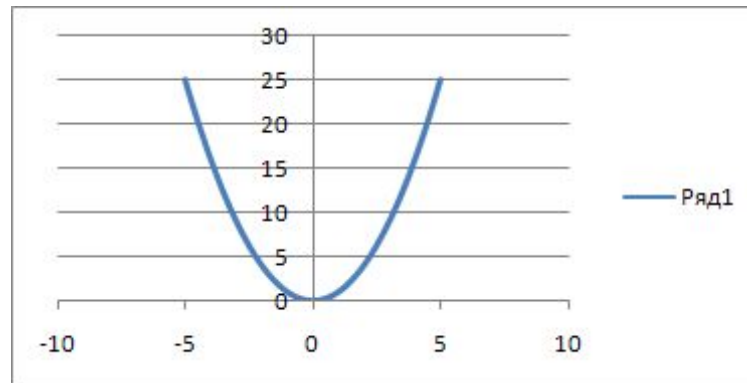
## Вставка диаграммы «Точечная»:

выделить данные

	A	B
1	x	y
2	-5	25
3	-4,5	20,25
4	-4	16
5	-3,5	12,25
6	-3	9
7	-2,5	6,25
8	-2	4
9	-1,5	2,25



результат:



# Работа в Excel 2007

## Тема 3. Численные методы

# Решение уравнений

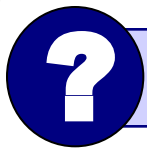
**Задача:** найти все решения уравнения  $x^2 = 5 \cos x$  на интервале  $[-5, 5]$



**Как решить математическими методами?**

**Методы решения уравнений:**

- **аналитические:** решение в виде формулы  $x = \dots$
- **численные:** *приближенное* решение, число
  - 1) выбрать *начальное приближение*  $x_0$  «рядом» с решением



**Как выбрать начальное приближение?**

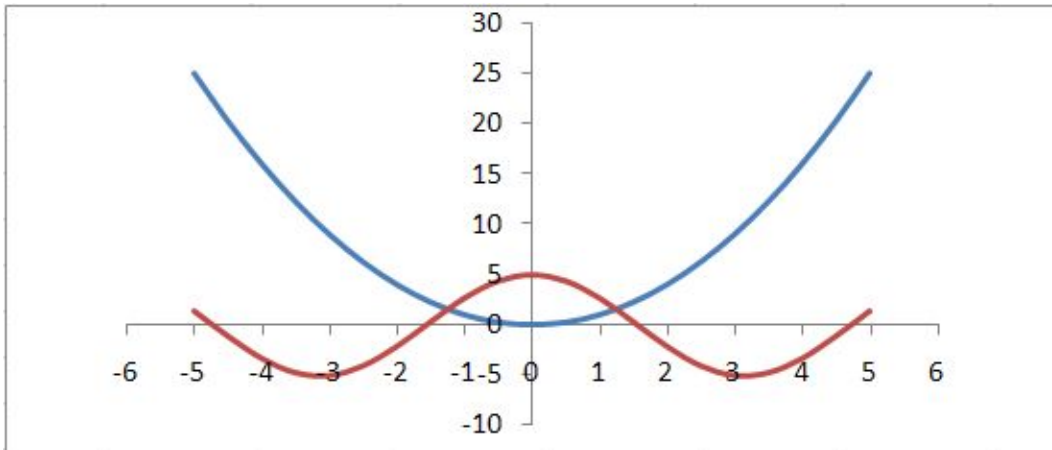
- 2) по некоторому алгоритму вычисляют первое приближение, затем – второе и т.д.  $x_0 \rightarrow x_1 \rightarrow x_2 \rightarrow \dots$
- 3) вычисления прекращают, когда значение меняется очень мало (метод сходится)  $x_0 \rightarrow \dots \rightarrow x_{15} \rightarrow x_{16} \approx x^*$

# Решение уравнения $x^2 = 5 \cos x$

## 1. Таблица значений функций на интервале $[-5, 5]$

	A	B	C	D
1	x	f1	f2	
2	-5	=A2^2	=5*COS(A2)	
3	-4,5			
4				

## 2. Графики функций (диаграмма «Точечная»)



**2 решения:**

начальные приближения

$$x_0 = -1,5$$

$$x_0 = 1,5$$

# Решение уравнения $x^2 = 5 \cos x$

## 3. Подготовка данных

начальное  
приближение

	E	F	G	H
1	x	f1	f2	f2-f1
2	-1,5	=E2^2	=5*COS(E2)	=F2-G2
3				

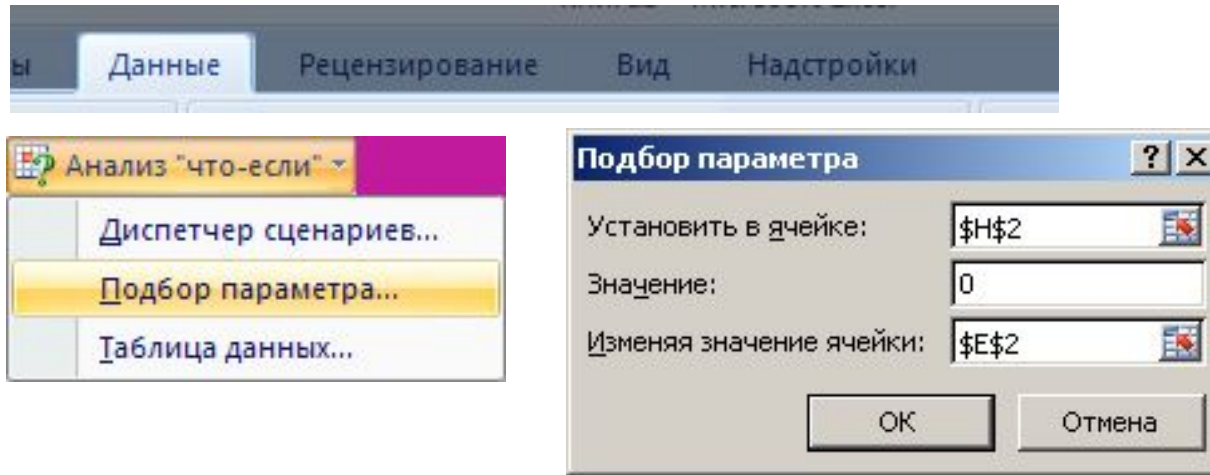
целевая  
ячейка

Цель: H2=0

 Зачем нужна разность?

# Решение уравнения $x^2 = 5 \cos x$

## 4. Подбор параметра



решение  
уравнения

	E	F	G	H
1	x	f1	f2	f2-f1
2	-1,252	1,568	1,567	0,00053

ошибка



Почему  
не нуль?



Как найти второе решение?



# Оптимизация

---

**Оптимизация** – это поиск оптимального (наилучшего) варианта в заданных условиях.

**Оптимальное решение** – такое, при котором некоторая заданная функция (*целевая функция*) достигает минимума или максимума.

## Постановка задачи:

- **целевая функция**

$$f(x) \rightarrow \min \quad (\text{расходы, потери, ошибки})$$

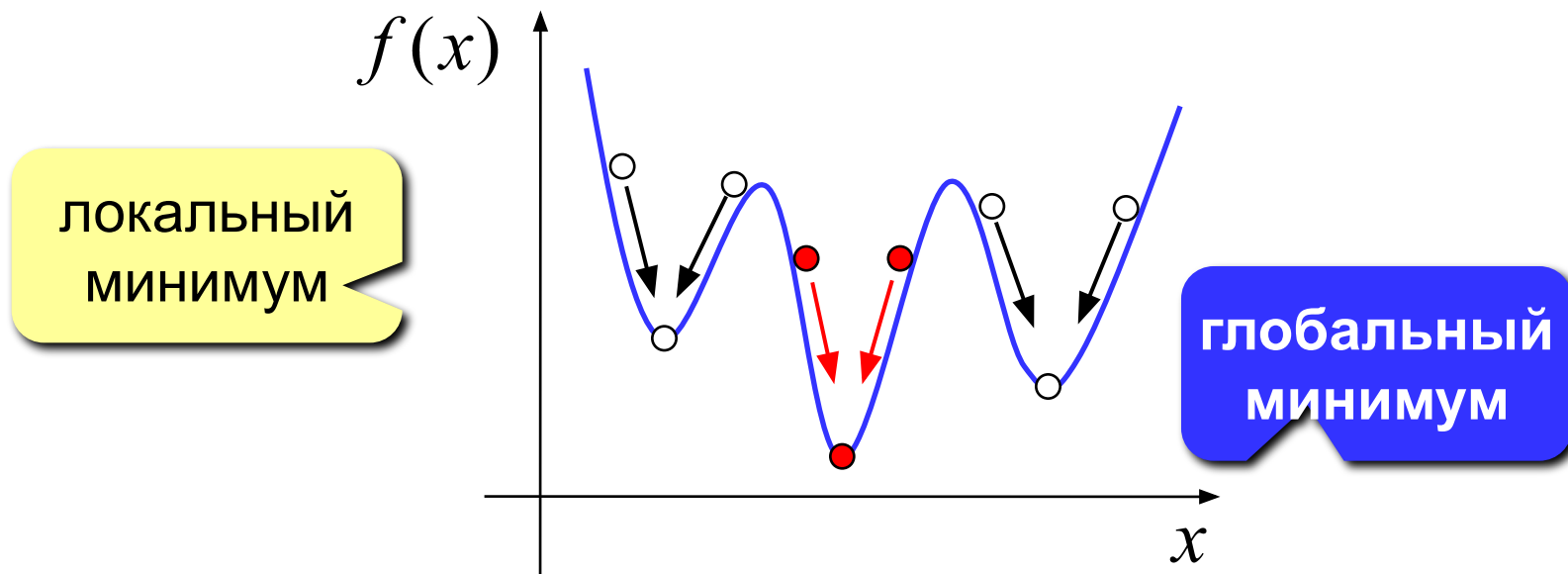
$$f(x) \rightarrow \max \quad (\text{доходы, приобретения})$$

- **ограничения**, которые делают задачу осмысленной

*Задача без ограничений:* построить дом  
при минимальных затратах.

*Решение:* не строить дом вообще.

# Оптимизация

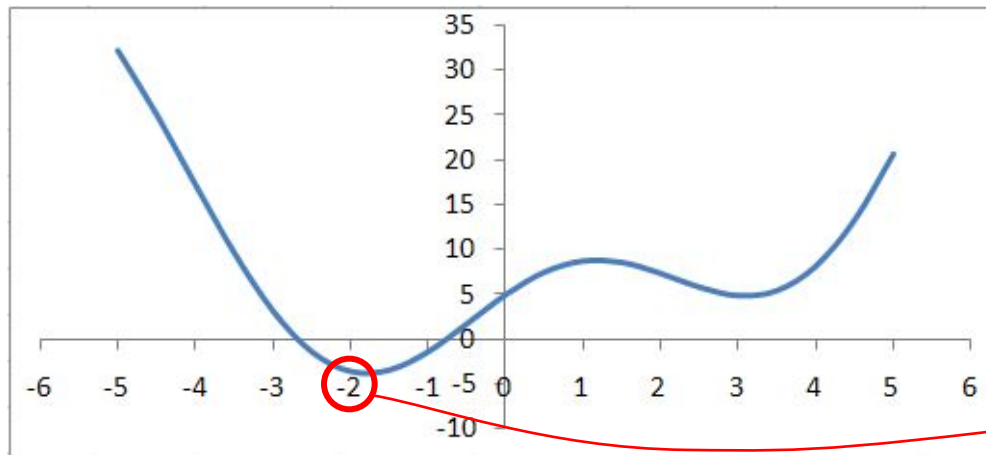


- обычно нужно найти **глобальный минимум**
- большинство численных методов находят только **локальный минимум**
- минимум, который найдет *Excel*, зависит от выбора начального приближения («шарик на горке скатится в ближайшую ямку»)

# Поиск минимума функции

$$y = x^2 + 6 \sin x + 5 \cos x$$

## 1. Строим график функции (диаграмма «Точечная»)



Зачем нужен график?

начальное приближение

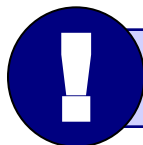
$$x_0 = -2$$

## 2. Подготовка данных

начальное приближение

	E	F
1	x	f
2	-2	=E2^2+6*SIN(E2)+5*COS(E2)

целевая ячейка



Изменение E2 должно влиять на F2!

# Поиск минимума функции

	E	F
1	x	f
2	-2	=E2^2+6*SIN(E2)+5*COS(E2)

## 3. Настройка «Поиск решения»

целевая  
ячейка

изменяемые  
ячейки:

E2  
D2:D6  
D2:D6; C5:C8

ограничения

A1  $\leq$  20  
B2:B8  $\geq$  5  
A1 = целое

Поиск решения

Установить целевую ячейку:

Равной:  максимальному значению  значению:

минимальному значению

Изменяя ячейки:

Ограничения:

# Параметры оптимизации

**Параметры поиска решения** [X]

Максимальное время:  секунд

Предельное число итераций:

Относительная погрешность:

Допустимое отклонение:  %

Сходимость:

Линейная модель       Автоматическое масштабирование

Неотрицательные значения       Показывать результаты итераций

Оценки

линейная

квадратичная

Разности

прямые

центральные

Метод поиска

Ньютона

сопряженных градиентов

OK

Отмена

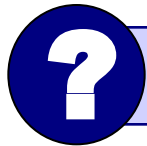
Загрузить модель...

Сохранить модель...

Справка

# Оптимизация

---



**Подбор параметра – это оптимизация?**

**Надстройка «Поиск решения» позволяет:**

- искать минимум и максимум функции
- использовать несколько изменяемых ячеек и диапазонов
- вводить ограничения ( $\leq$ ,  $\geq$ , целое, двоичное)



**Как влияет ограничение «А1-целое» на сложность решения задачи?**

# Работа в Excel 2007

## Тема 4. Статистика

# Ряд данных и его свойства

---

**Ряд данных** – это упорядоченный набор значений

$$x_1, x_2, \dots, x_n$$

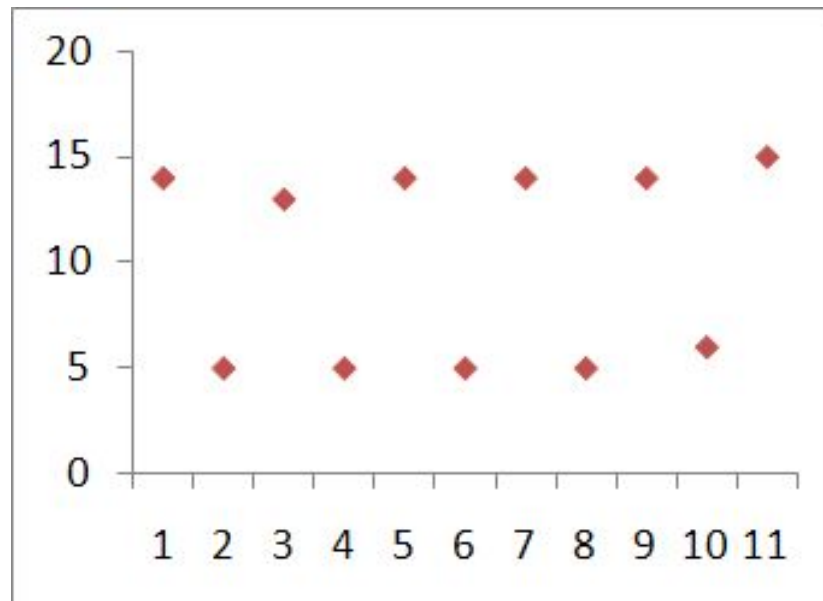
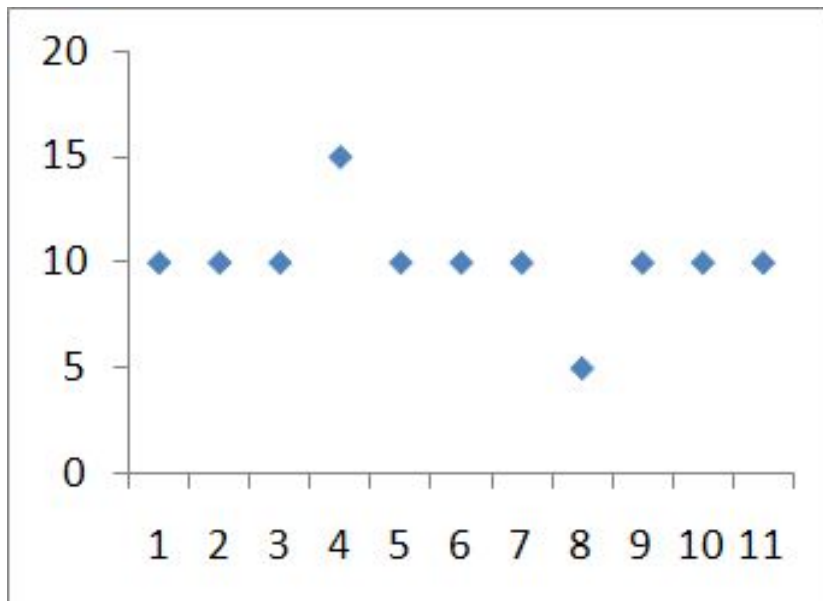
**Основные свойства** (ряд A1 : A20):

- количество элементов =СЧЕТ (A1 : A20)
- количество элементов, удовлетворяющих некоторому условию:  
= СЧЕТЕСЛИ (A1 : A20 ; "<5" )
- минимальное значение =МИН (A1 : A20)
- максимальное значение =МАКС (A1 : A20)
- сумма элементов =СУММ (A1 : A20)
- среднее значение =СРЗНАЧ (A1 : A20)



# Дисперсия

Для этих рядов одинаковы МИН, МАКС, СРЗНАЧ



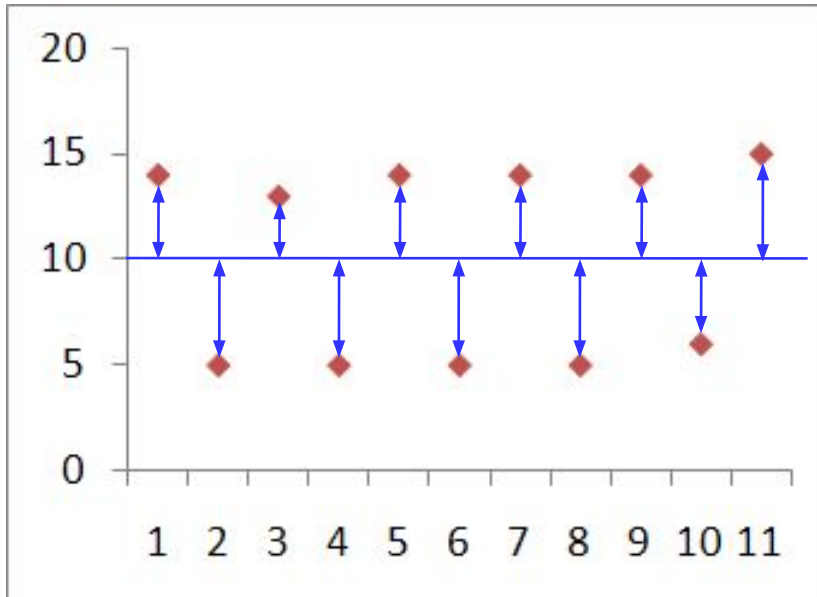
В чем различие?

**Дисперсия** («разброс») – это величина, которая характеризует разброс данных относительно среднего значения.

# Дисперсия

$$D_x = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{n}$$

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} \quad \text{среднее арифметическое}$$



$(x_1 - \bar{x})^2$  квадрат  
отклонения  $x_1$   
от среднего

$D_x$  *средний* квадрат  
отклонения от  
среднего значения

# Дисперсия и СКВО

---

## Стандартная функция

=ДИСПР (A1 : A20)

*Функции – Другие – Статистические*

## Что неудобно:

если  $x$  измеряется в метрах,  
то  $D_x$  – в  $m^2$



В каких  
единицах  
измеряется?

## СКВО = среднеквадратическое отклонение

$$\sigma_x = \sqrt{D_x}$$

=СТАНДОТКЛОНП (A1 : A20)

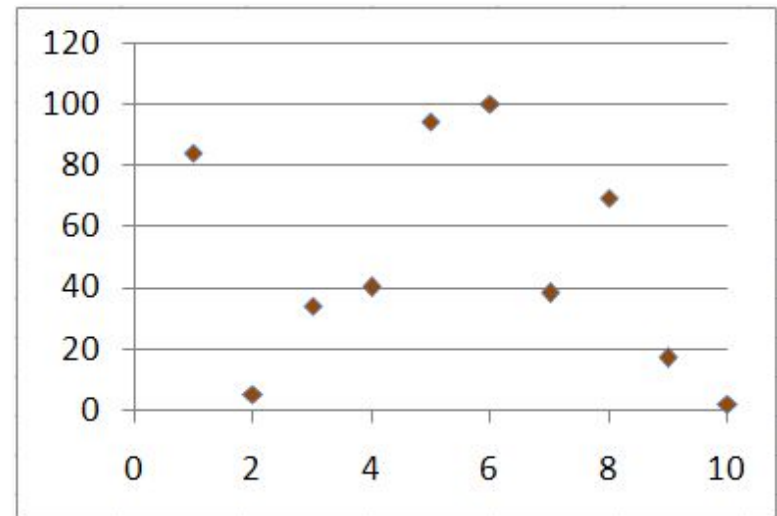
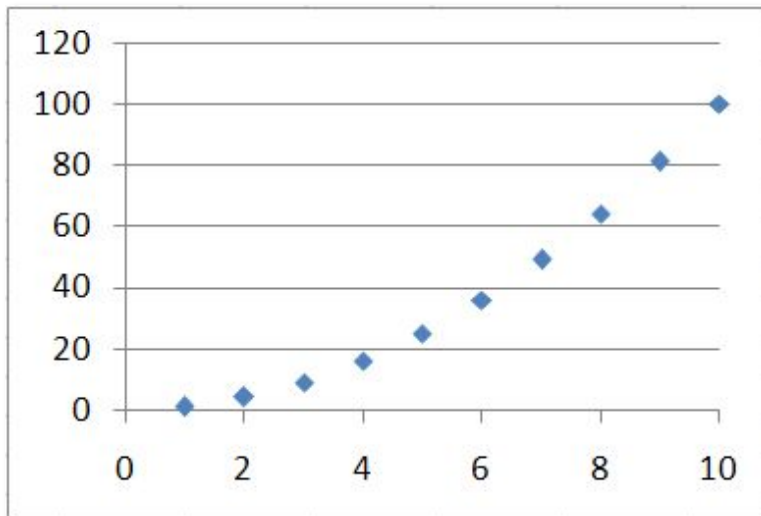
# Взаимосвязь рядов данных

## Два ряда одинаковой длины:

$$x_1, x_2, \dots, x_n \quad y_1, y_2, \dots, y_n$$

## Вопросы:

- есть ли связь между этими рядами (соответствуют ли пары  $(x_i, y_i)$  какой-нибудь зависимости  $y = f(x)$ )
- насколько сильна эта связь?



# Взаимосвязь рядов данных

Ковариация:

$$K_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{n}$$



Если  $x$  и  $y$  – один и тот же ряд?

$$K_{xx} = D_x$$

в среднем!

Как понимать это число?

- если  $K_{xy} > 0$  увеличение  $x$  приводит к увеличению  $y$
- если  $K_{xy} < 0$  увеличение  $x$  приводит к уменьшению  $y$
- если  $K_{xy} \approx 0$  связь обнаружить не удалось

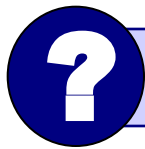
Что плохо?

- единицы измерения: если  $x$  в метрах,  $y$  в литрах, то  $K_{xy}$  – в м·л
- $K_{xy}$  зависит от абсолютных значений  $x$  и  $y$ , поэтому ничего не говорит о том, насколько сильна связь

# Взаимосвязь рядов данных

## Коэффициент корреляции:

$$\rho_{xy} = \frac{K_{xy}}{\sigma_x \cdot \sigma_y} \quad \sigma_x, \sigma_y \text{ – СКВО рядов } x \text{ и } y$$



**Какова размерность?** безразмерный!

$$-1 \leq \rho_{xy} \leq 1$$

## Как понимать это число?

- если  $\rho_{xy} > 0$  : увеличение  $X$  приводит к увеличению  $Y$
- если  $\rho_{xy} < 0$  : увеличение  $X$  приводит к уменьшению  $Y$
- если  $\rho_{xy} \approx 0$  : связь обнаружить не удалось

**=КОРРЕЛ (А1 : А20 ; В1 : В20)**

# Взаимосвязь рядов данных

## Как понимать коэффициент корреляции?

$0 < |\rho_{xy}| \leq 0,2$  : очень слабая корреляция

$0,2 < |\rho_{xy}| \leq 0,5$  : слабая

$0,5 < |\rho_{xy}| \leq 0,7$  : средняя

$0,7 < |\rho_{xy}| \leq 0,9$  : сильная

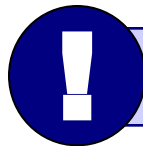
$0,9 < |\rho_{xy}| \leq 1$  : очень сильная

$\rho_{xy} = 1$  : линейная зависимость  $y = ax + b$ ,  $a > 0$

$\rho_{xy} = -1$  : линейная зависимость  $y = ax + b$ ,  $a < 0$



Если  $\rho_{xy} \approx 0$ , то связи нет?



Метод для определения линейной зависимости!

# Работа в Excel 2007

## Тема 5. Восстановление зависимостей

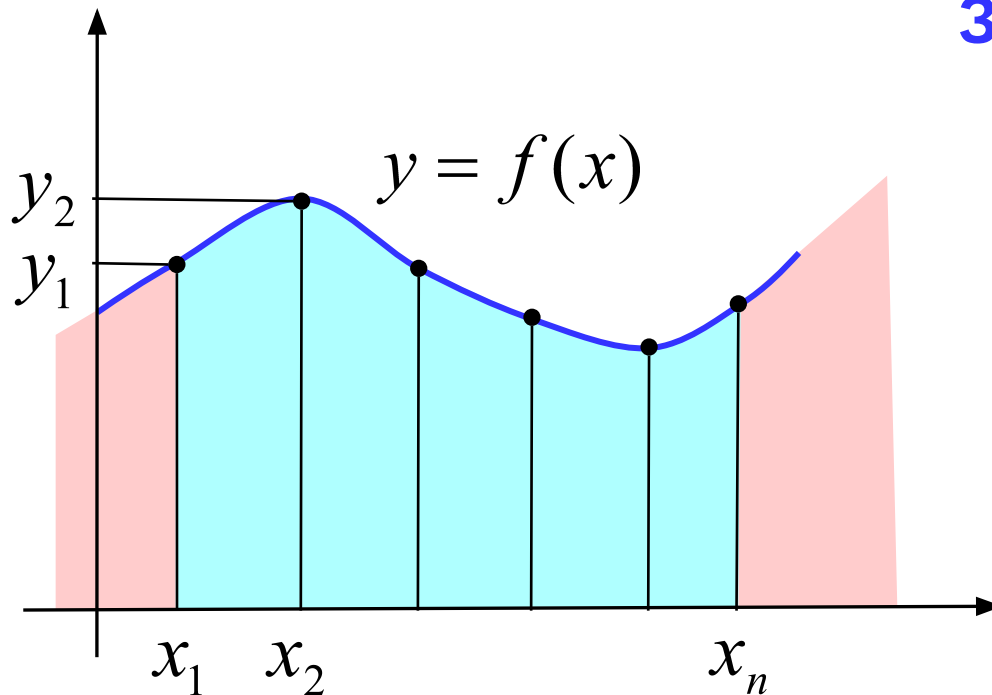


# Восстановление зависимостей

Два ряда одинаковой длины:

$$x_1, x_2, \dots, x_n \quad y_1, y_2, \dots, y_n$$

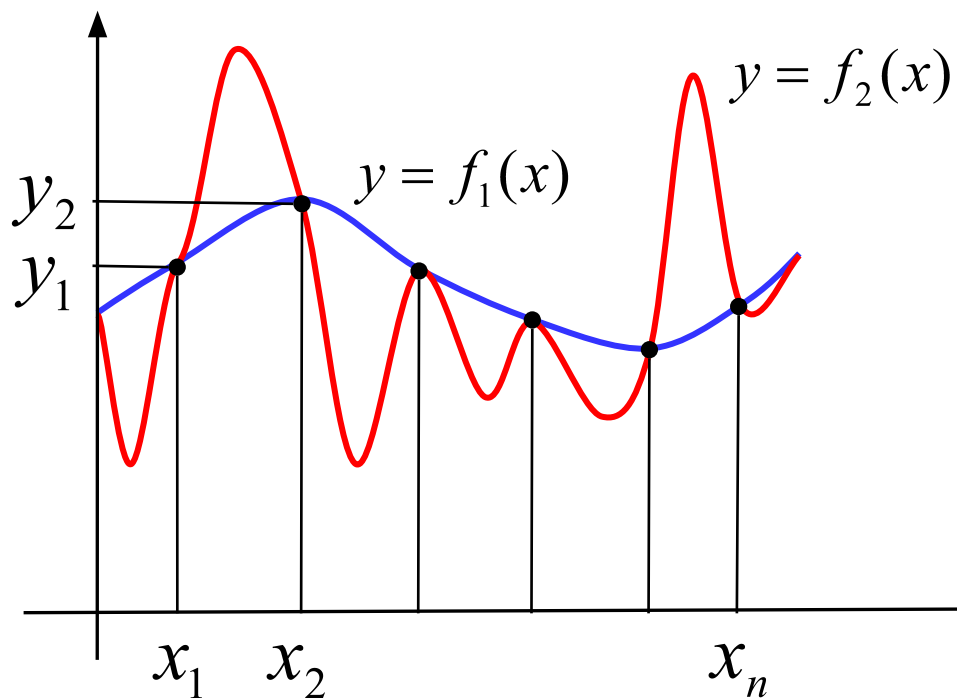
задают некоторую неизвестную функцию  $y = f(x)$



**Зачем:**

- найти  $y$  в промежуточных точках (интерполяция)
- найти  $y$  вне диапазона измерений (экстраполяция, прогнозирование)

# Какое решение нам нужно?

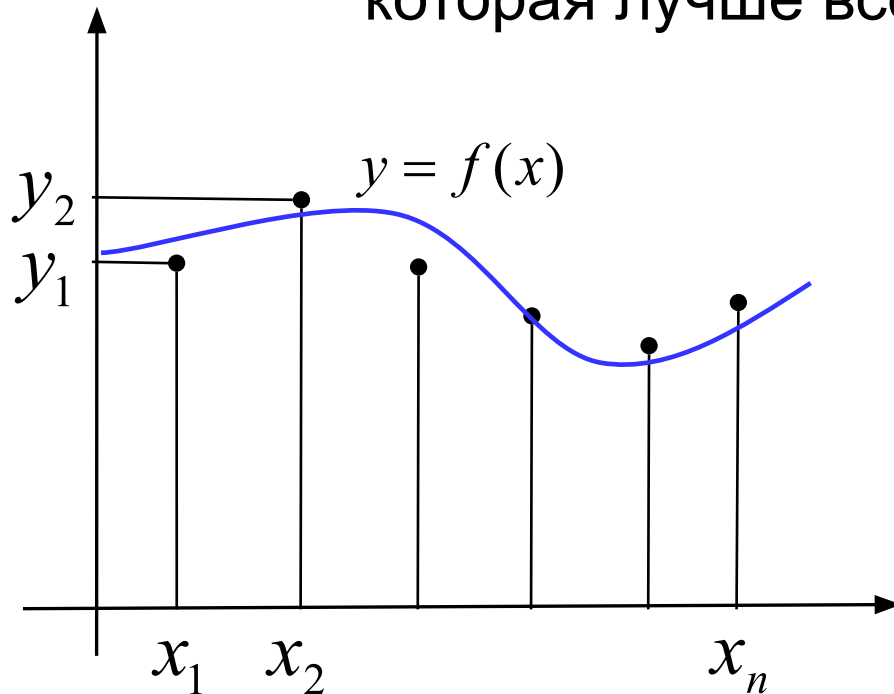


Через заданный набор точек проходит бесконечно много разных кривых!

**Вывод:** задача некорректна, поскольку решение неединственно.

# Восстановление зависимостей

**Корректная задача:** найти функцию заданного вида, которая лучше всего соответствует данным.



**Примеры:**

• линейная  $y = a \cdot x + b$

• полиномиальная  

$$y = a_3x^3 + a_2x^2 + a_1x + a_0$$

• степенная  $y = a \cdot x^b$

• экспоненциальная

$$y = a \cdot e^{bx}$$

• логарифмическая

$$y = a \cdot \ln x + b$$



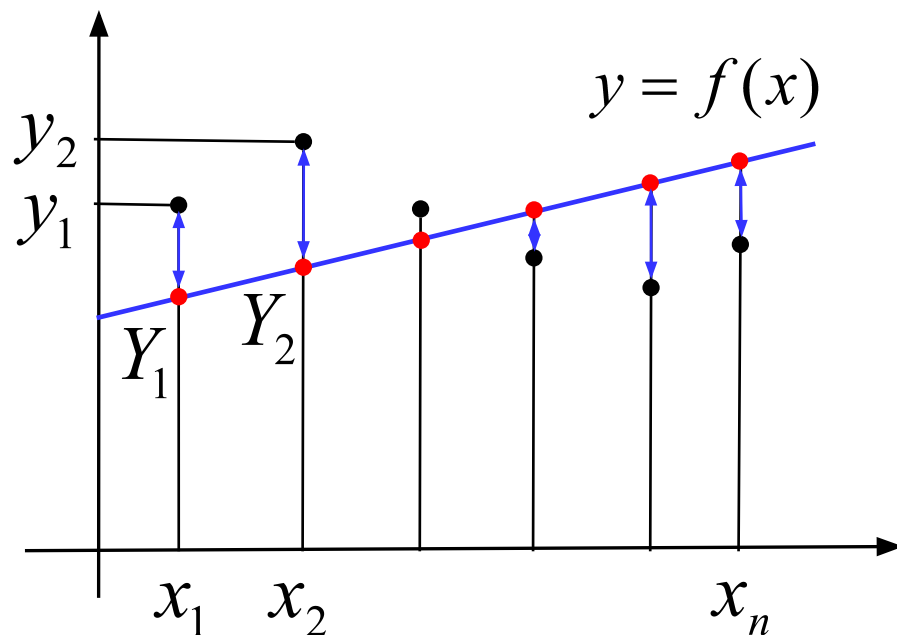
**График функции не обязательно проходит через заданные точки!**



**Как выбрать функцию?**

# Что значит «лучше всего соответствует»?

## Метод наименьших квадратов (МНК):



$(x_i, y_i)$  заданные пары значений

$$Y_i = f(x_i)$$

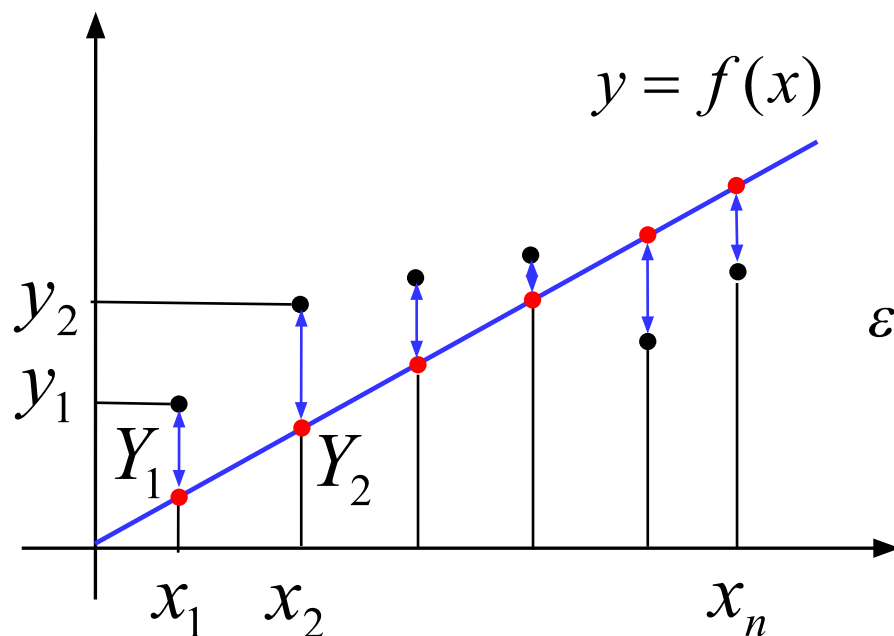
$$\varepsilon = \sum_{i=1}^n (y_i - Y_i)^2 \rightarrow \min$$



**Зачем возведение в квадрат?**

- 1) чтобы складывать положительные значения
- 2) решение сводится к системе линейных уравнений (просто решать!)

# МНК для линейной функции



НЕИЗВЕСТНО!

$$Y_i = k \cdot x_i$$

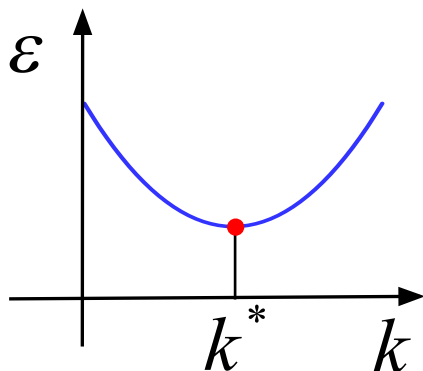
$$\begin{aligned} \varepsilon(k) &= \sum_{i=1}^n (y_i - Y_i)^2 = \sum_{i=1}^n (y_i - kx_i)^2 \\ &= k^2 \cdot \sum_{i=1}^n x_i^2 - k \cdot 2 \sum_{i=1}^n x_i \cdot y_i + \sum_{i=1}^n y_i^2 \end{aligned}$$

$a$

$-b$

$c$

$$\varepsilon(k) = ak^2 + bk + c \rightarrow \min$$



$$k^* = -\frac{b}{2a} = \frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n x_i^2}$$

# Коэффициент достоверности

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - Y_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

$(x_i, y_i)$  заданные пары значений

$$Y_i = f(x_i)$$

$\bar{y}$  – среднее значение  $y_i$

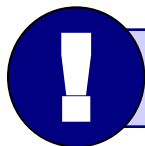
Крайние случаи:

- если график проходит через точки:

$$R^2 = 1$$

- если считаем, что  $y$  не меняется и  $Y_i = \bar{y}$ .

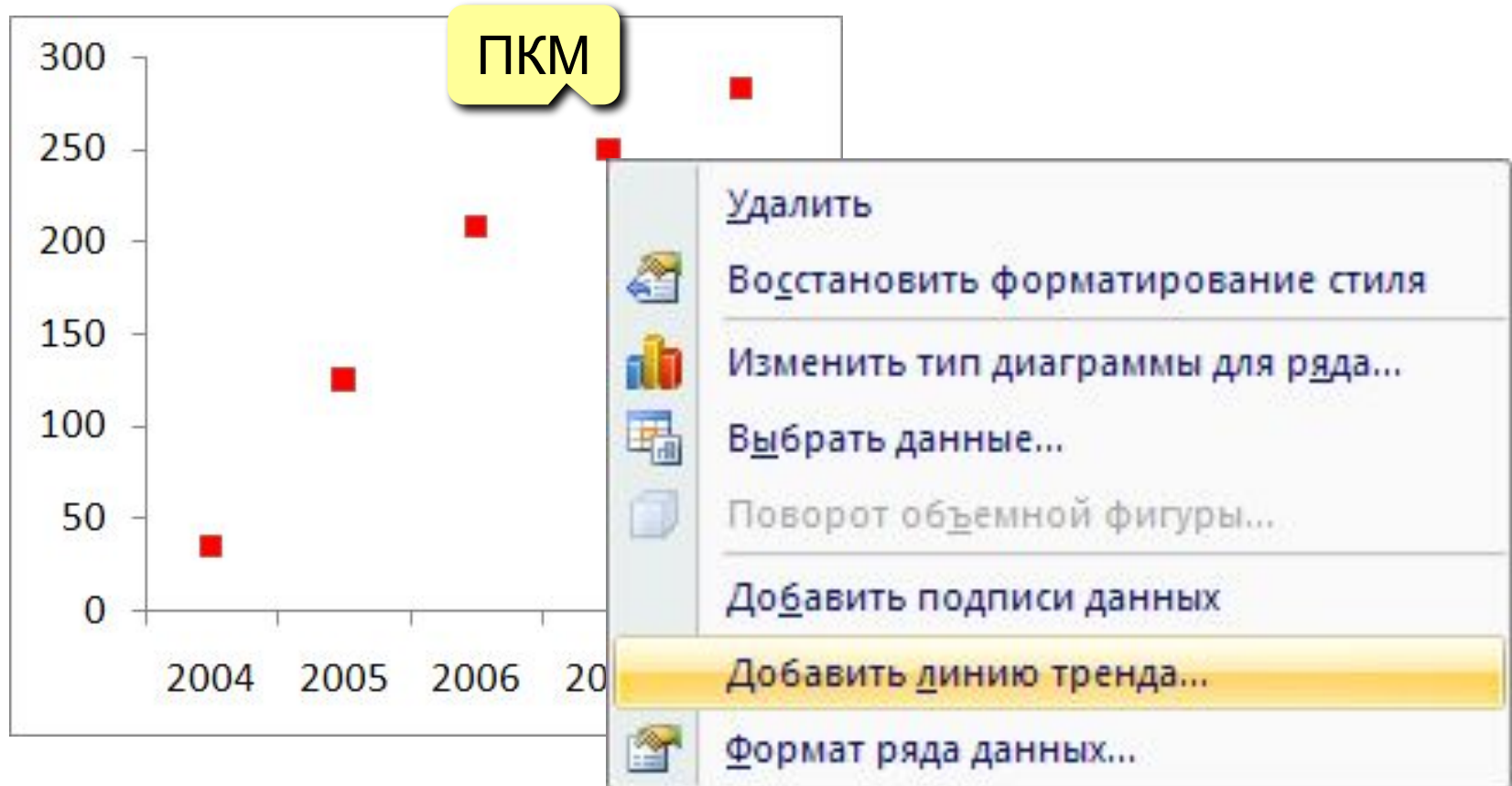
$$R^2 = 0$$

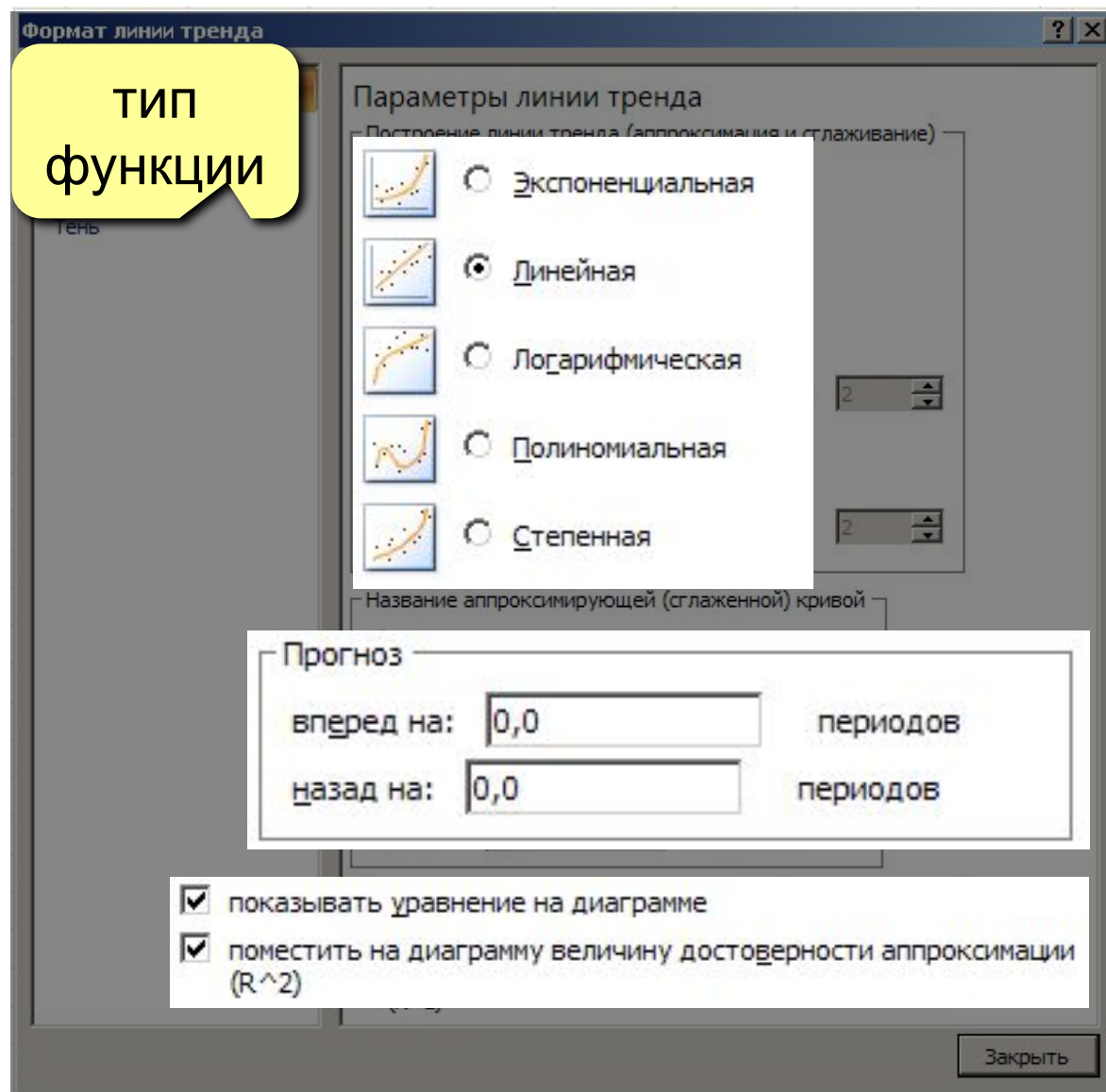


**Фактически – метод наименьших квадратов!**

# Восстановление зависимостей

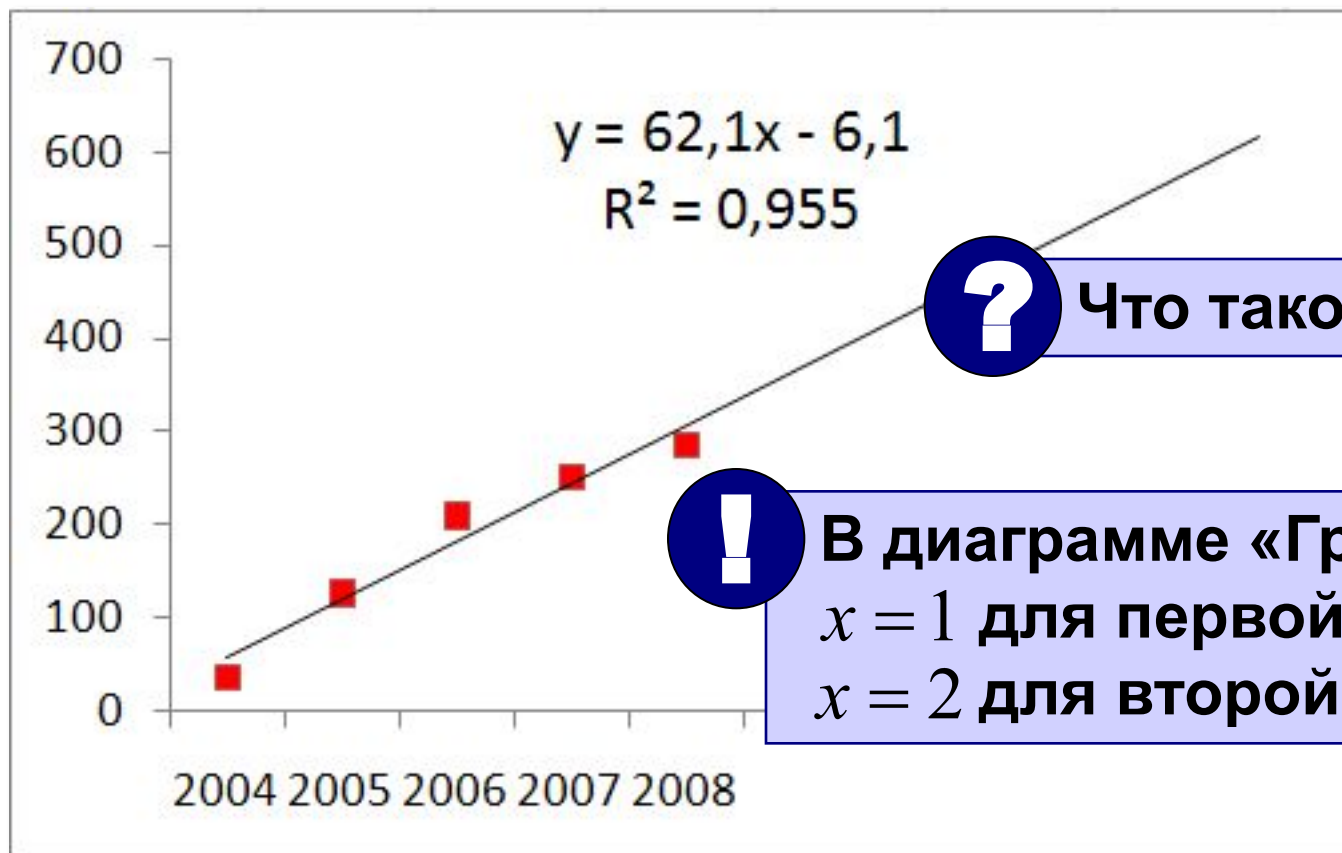
## Диаграмма «График»:







# Восстановление зависимостей



Что такое  $x$  ?



В диаграмме «График»  
 $x = 1$  для первой точки,  
 $x = 2$  для второй и т.д.



Насколько хорошо выбрана функция?

# Восстановление зависимостей

**Сложные случаи** (нестандартная функция):

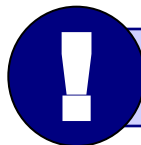
$$f(x) = a \cdot \sin kx + b$$



Что делать?

**Алгоритм:**

- 1) выделить ячейки для хранения  $a, k, b$
- 2) построить ряд  $Y_i = f(x_i)$  для тех же  $x_i$
- 3) построить на одной диаграмме ряды  $y_i$  и  $Y_i$
- 4) попытаться подобрать  $a, k, b$  так, чтобы два графика были близки
- 5) вычислить  $R^2$  в отдельной ячейке  
 функции: СУММКВРАЗН – сумма квадратов разностей рядов  
 ДИСПР – дисперсия
- 6) Поиск решения:  $R^2 \rightarrow \min$



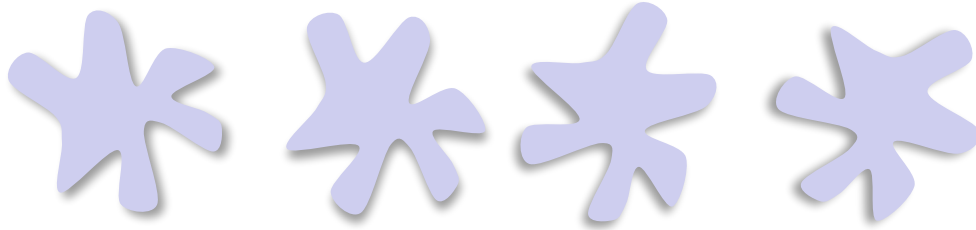
Это задача оптимизации!

# Работа в Excel 2007

## Тема 6. Моделирование

(по материалам учебника Н.В. Макаровой)

# Модель деления

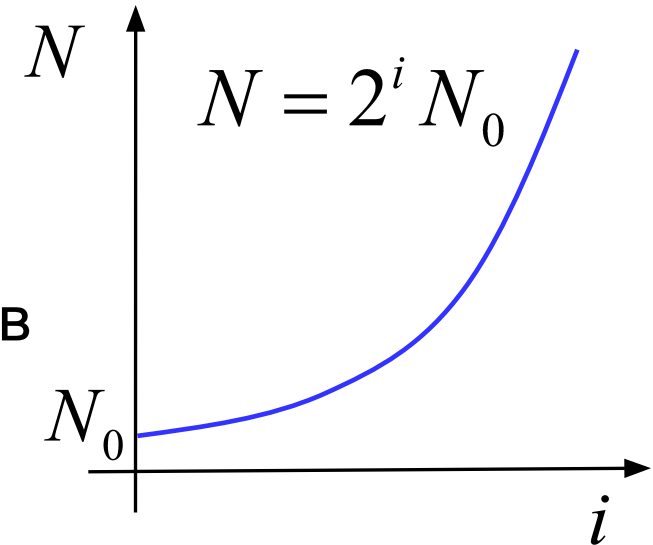


$N_0$  – начальная численность

$N_1 = 2N_0$  – после 1 цикла деления

$N_2 = 2N_1 = 4N_0$  – после 2-х циклов

$N_i = 2N_{i-1} = 2^i N_0$



## Особенности модели:

- 1) не учитывается смертность
- 2) не учитывается влияние внешней среды
- 3) не учитывается влияние других видов

# Рождаемость и смертность

$$N_i = N_{i-1} + K_p \cdot N_{i-1} - K_c \cdot N_{i-1}$$

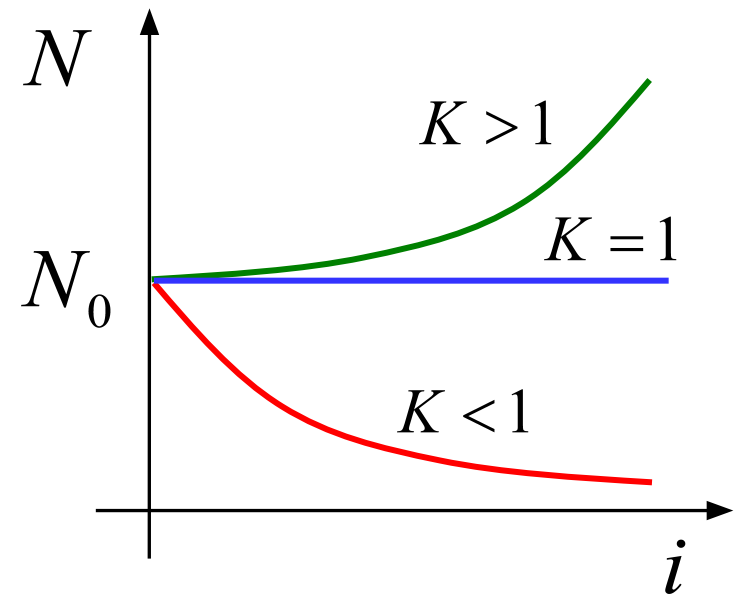
$K_p$  – коэффициент рождаемости

$K_c$  – коэффициент смертности

$$N_i = K \cdot N_{i-1}$$

Коэффициент изменения  
численности

$$K = 1 + K_p - K_c$$



## Особенности модели:

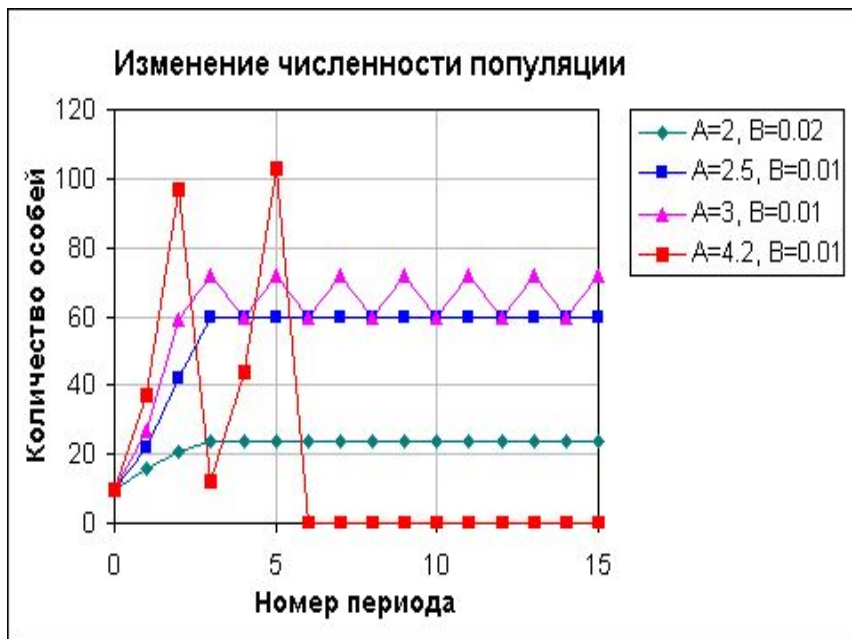
- 1) не учитывается влияние численности  $N$  и внешней среды на  $K$
- 2) не учитывается влияние других видов на  $K$

# Влияние численности и внешней среды

$$N_i = K \cdot N_{i-1} \quad K = A \cdot (1 - B \cdot N_{i-1})$$

$A$  – коэффициент устойчивости вида

$B$  – коэффициент среды обитания



## Варианты:

- устанавливается постоянная численность
- постоянно меняется (колебания)
- вымирание

# Влияние других видов

$N_i$  – численность белок,  $M_i$  – численность бурундуков

$$N_i = N_{i-1} (2 - K_1 \cdot N_{i-1} - K_2 \cdot M_{i-1})$$

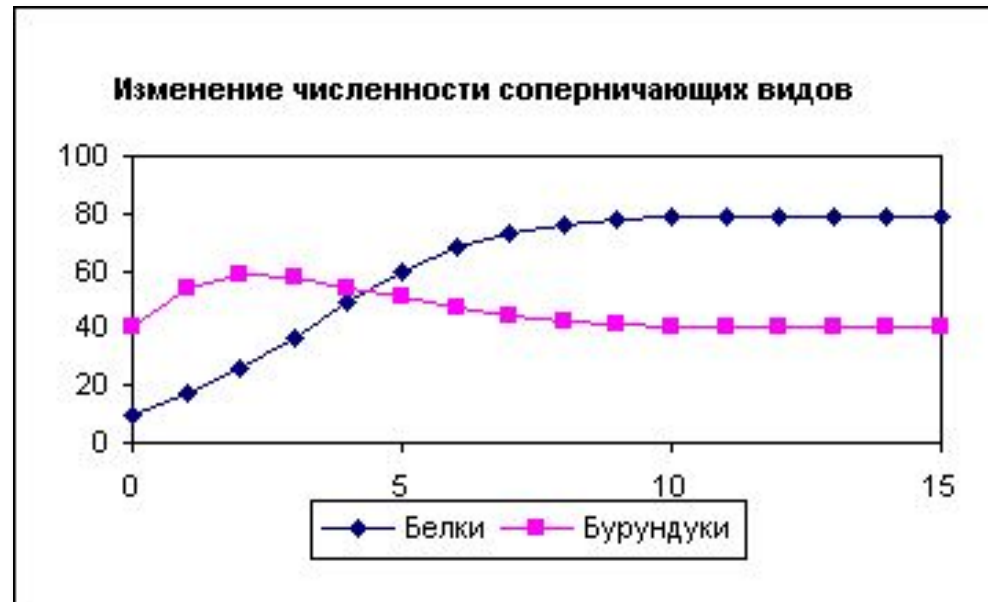
$$M_i = M_{i-1} (2 - K_3 \cdot M_{i-1} - K_4 \cdot N_{i-1})$$



Откуда видно влияние?

$K_2, K_4$  – взаимное влияние

если  $K_2 > K_1$  или  $K_4 > K_3$  – враждующие виды

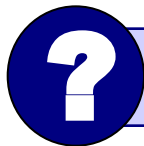


# Моделирование двух популяций

 $N_0$ 
 $M_0$ 

	A	B	C	D	E	F
1	$i$	$N$	$M$		$K1$	0,05
2	0	20	20		$K2$	0,01
3	1	$=B2*(2-F1*B2-F2*C2)$			$K3$	0,05
4	2				$K4$	0,01
5	3					

$$N_i = N_{i-1} (2 - K_1 \cdot N_{i-1} - K_2 \cdot M_{i-1})$$



Как скопировать формулы «вниз»?



