

Обработка статистических данных с помощью Excel

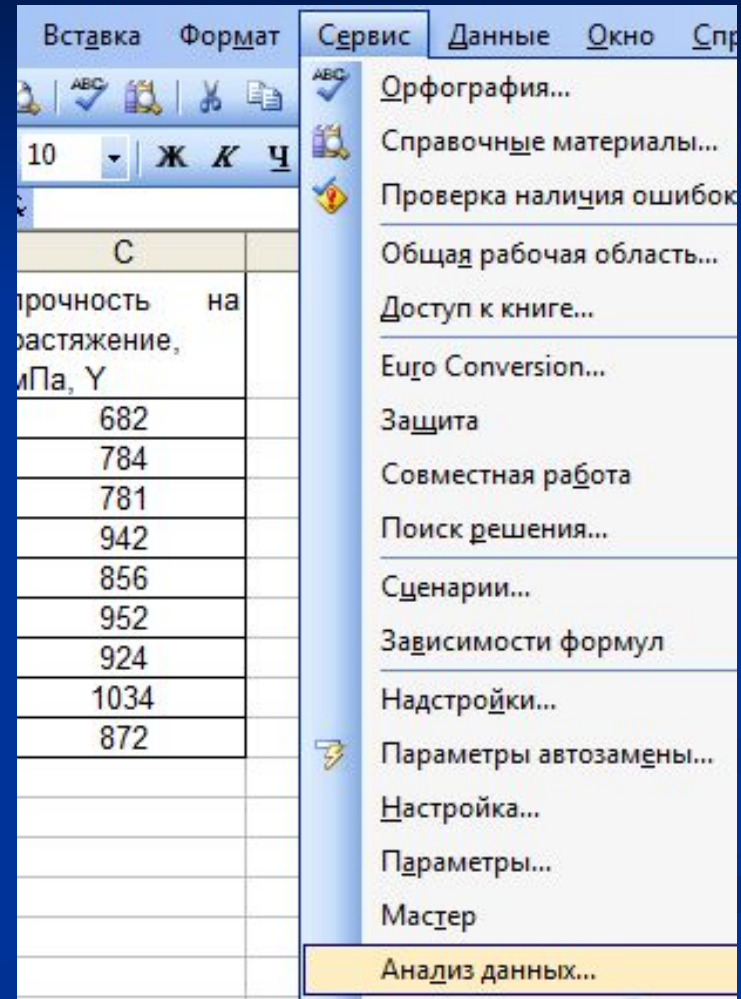
Исходные данные

	A	B	C
1	V, X1	Al, X2	прочность на растяжение, мПа, Y
2	1,78	0,82	682
3	2,05	2,88	784
4	2,07	0,95	781
5	2,02	2,81	942
6	5,8	0,95	856
7	5,68	2,7	952
8	5,89	0,88	924
9	5,54	2,78	1034
10	3,6	1,83	872

Пакет анализа данных

Набор средств анализа данных, называемый **Пакет анализа**, предназначен для решения сложных статистических и инженерных задач. Для проведения анализа данных с помощью этих инструментов следует указать входные данные и выбрать параметры; анализ будет проведен с помощью подходящей статистической или инженерной макрофункции, а результат будет помещен в выходной диапазон. Чтобы вывести список доступных инструментов анализа, выберите команду **Анализ данных** в меню **Сервис**. Если она отсутствует, необходимо выполнить следующие действия: выполнить команду **Надстройки**. На экране появится окно диалога **Надстройки**; выбрать пункт **Пакет анализа**, а затем нажать кнопку ОК. Начнет загружаться пакет **Анализ данных**.

После окончания загрузки в списке опций пункта **Сервис** основного меню появится строка **Анализ данных**. При выборе этой строки появляется окно диалога **Анализ данных** (рис. 1).



Перечень операций

- 1. Генерация случайных чисел.
- Выборка.
- Гистограмма.
- Описательная статистика.
- Скользящее среднее.
- Экспоненциальное сглаживание.
- Ковариационный анализ.
- Корреляционный анализ.
- Двухвыборочный F-тест для дисперсий.
- Двухвыборочный Z-тест для средних.
- Парный двухвыборочный t-тест для средних.
- Двухвыборочный t-тест с одинаковыми дисперсиями.
- Двухвыборочный t-тест с разными дисперсиями.
- Однофакторный дисперсионный анализ.
- Двухфакторный дисперсионный анализ с повторениями.
- Двухфакторный дисперсионный анализ без повторений.
- Регрессия.
- Ранг и перцентиль.
- Анализ Фурье.
- Во многих инструментах статистического анализа есть *одинаковые параметры*. К ним относятся следующие. Категория **Входные данные**.

Регрессия

- Линейный регрессионный анализ заключается в подборе графика для набора наблюдений с помощью метода наименьших квадратов. Регрессия используется для анализа воздействия на отдельную зависимую переменную значений одной или более независимых переменных.
- Например, на спортивные качества атлета влияют несколько факторов, включая возраст, рост и вес. Регрессия пропорционально распределяет меру качества по этим трем факторам на основе его спортивных результатов. Результаты регрессии впоследствии могут быть использованы для предсказания качеств нового, непроверенного атлета.
- Регрессия использует функцию ЛИНЕЙН.

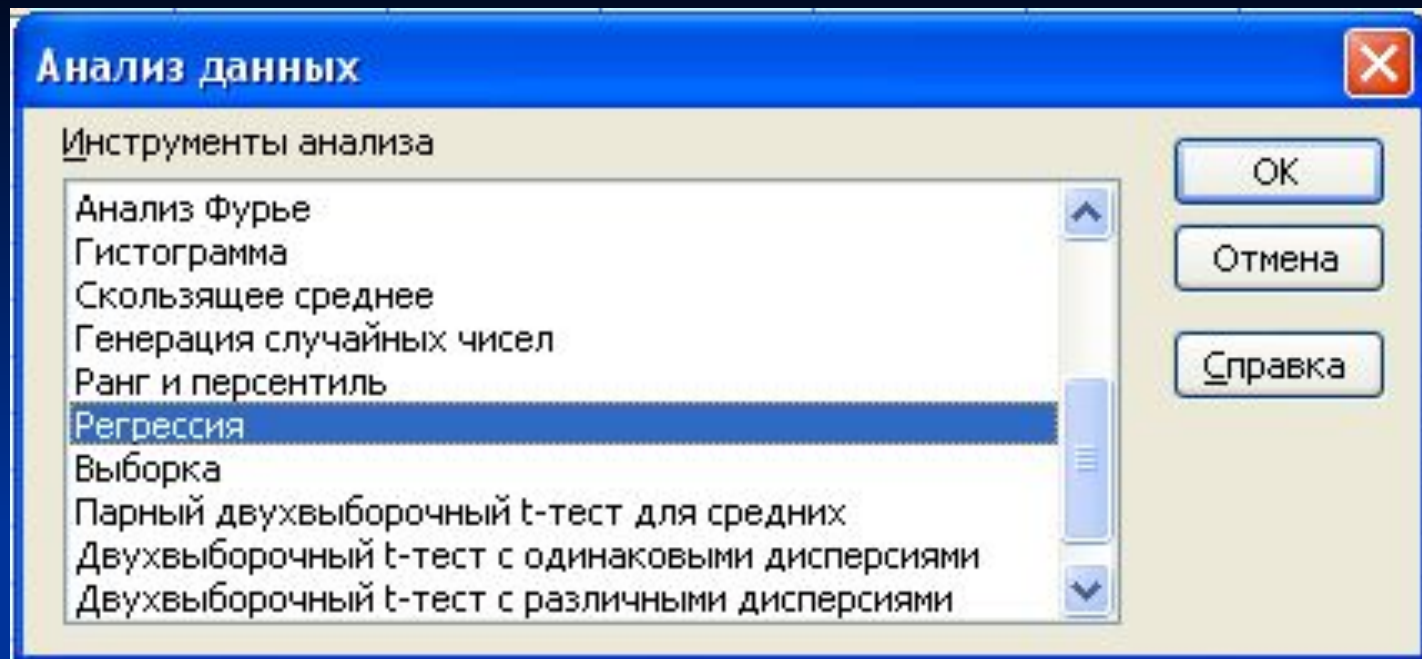


Рис. 1. Окно диалога Анализ данных

В окне диалога **Анализ данных** отображается список инструментов, приведенный ниже (перечень дан в соответствии с перечислением по дальнейшему тексту книги).

	A	B	C
			прочность на
1	V, X1	AI, X2	растяжение, мПа, Y
2	1,78	0,82	682
3	2,05	2,88	784
4	2,07	0,95	781
5	2,02	2,81	942
6	5,8	0,95	856
7	5,68	2,7	952
8	5,89	0,88	924
9	5,54	2,78	1034
10	3,6	1,83	872
11			
12			
13			

Регрессия

Входные данные

Входной интервал Y:

Входной интервал X:

☐ Метки

☐ Константа - ноль

☐ Уровень надежности: %

Параметры вывода

☒ Выходной интервал:

☐ Новый рабочий лист:

☐ Новая рабочая книга

В окно входной интервал Y вставляем значения столбца Y. В окно входной интервал X вставляем значения столбца X1.

Получаемые значения при использовании регрессии

11	ВЫВОД ИТОГОВ								
12									
13	Регрессионная статистика								
14	Множеств	0,667675							
15	R-квадрат	0,44579							
16	Нормиров	0,366618							
17	Стандартн	85,61436							
18	Наблюден	9							
19									
20	Дисперсионный анализ								
21		<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>значимость F</i>			
22	Регрессия	1	41271,26753	41271,27	5,630599	0,049396			
23	Остаток	7	51308,73247	7329,819					
24	Итого	8	92580						
25									
26		<i>Коэффициент</i>	<i>стандартная ошибка</i>	<i>статистика t</i>	<i>P-Значение</i>	<i>нижние 95%</i>	<i>верхние 95%</i>	<i>нижние 95,0%</i>	<i>верхние 95,0%</i>
27	Y-пересеч	723,4935	67,89074144	10,65673	1,4E-05	562,9574	884,0296	562,9574	884,0296
28	Переменн	38,20966	16,10259382	2,372888	0,049396	0,133071	76,28624	0,133071	76,28624

Составление уравнения по полученным данным

$$Y=723,5+38,21 \cdot X_1$$

Высчитываем значения функции с использованием полученного уравнения

При этом значения X задаем с интервалом 0,3. Значения Y получаем протягиваем по столбцу B

B30		fx =723,49350787176+38,2096552179542*A30				
	A	B	C	D	E	F
30	1,5	780,808				
31	2	799,9128				
32	2,5	819,0176				
33	3	838,1225				
34	3,5	857,2273				
35	4	876,3321				
36	4,5	895,437				
37	5	914,5418				
38	5,5	933,6466				
39	6	952,7514				

Построение по данным диаграммы

Для этого на панели управления нажимаем

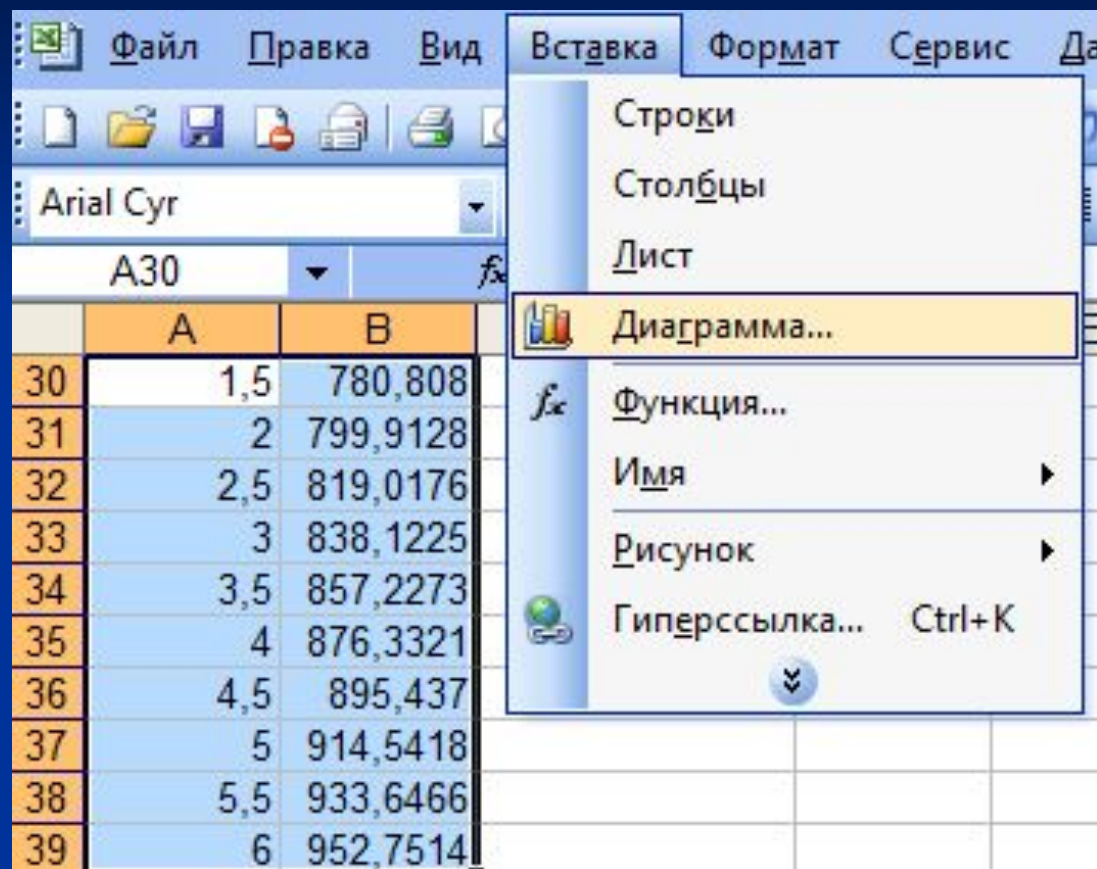
1. Вставка

2. Диаграмма

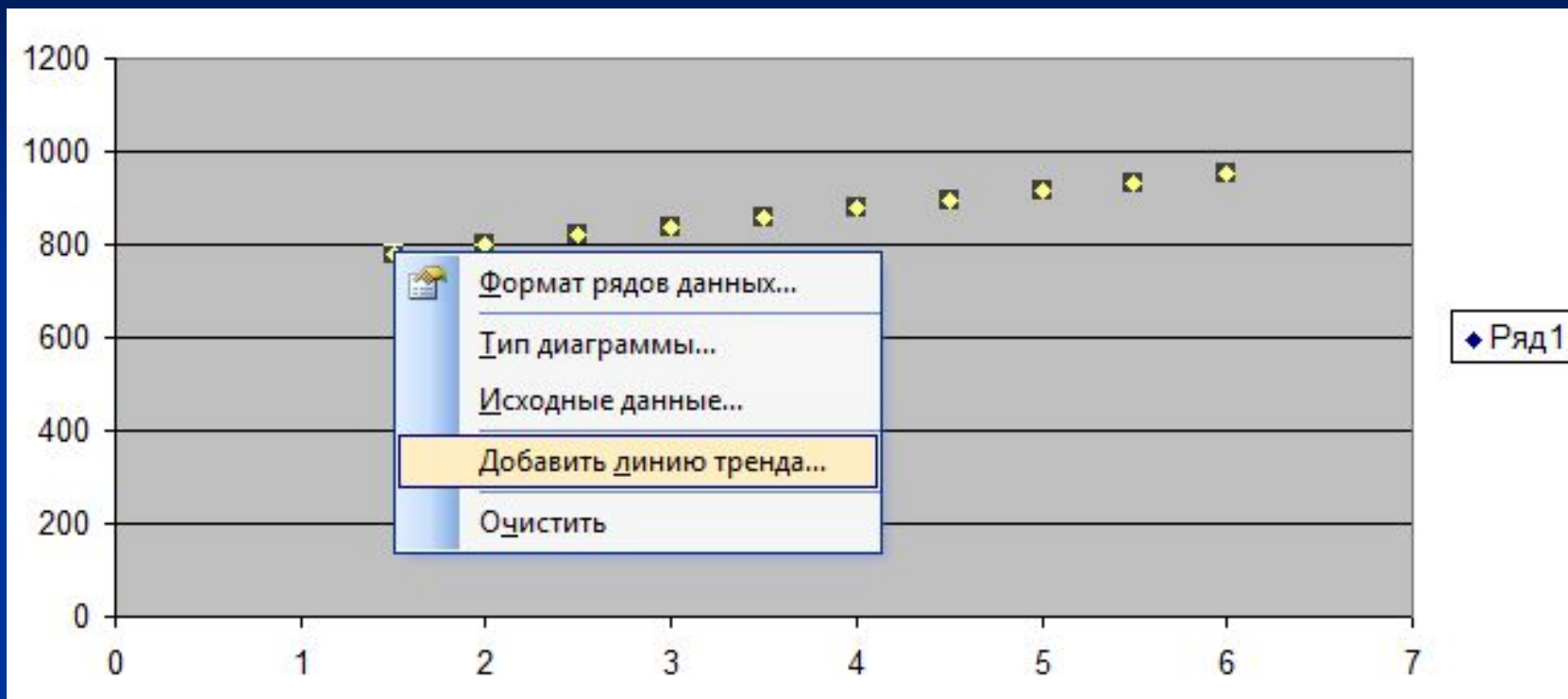
3. Выбор типа

диаграммы (точечная)

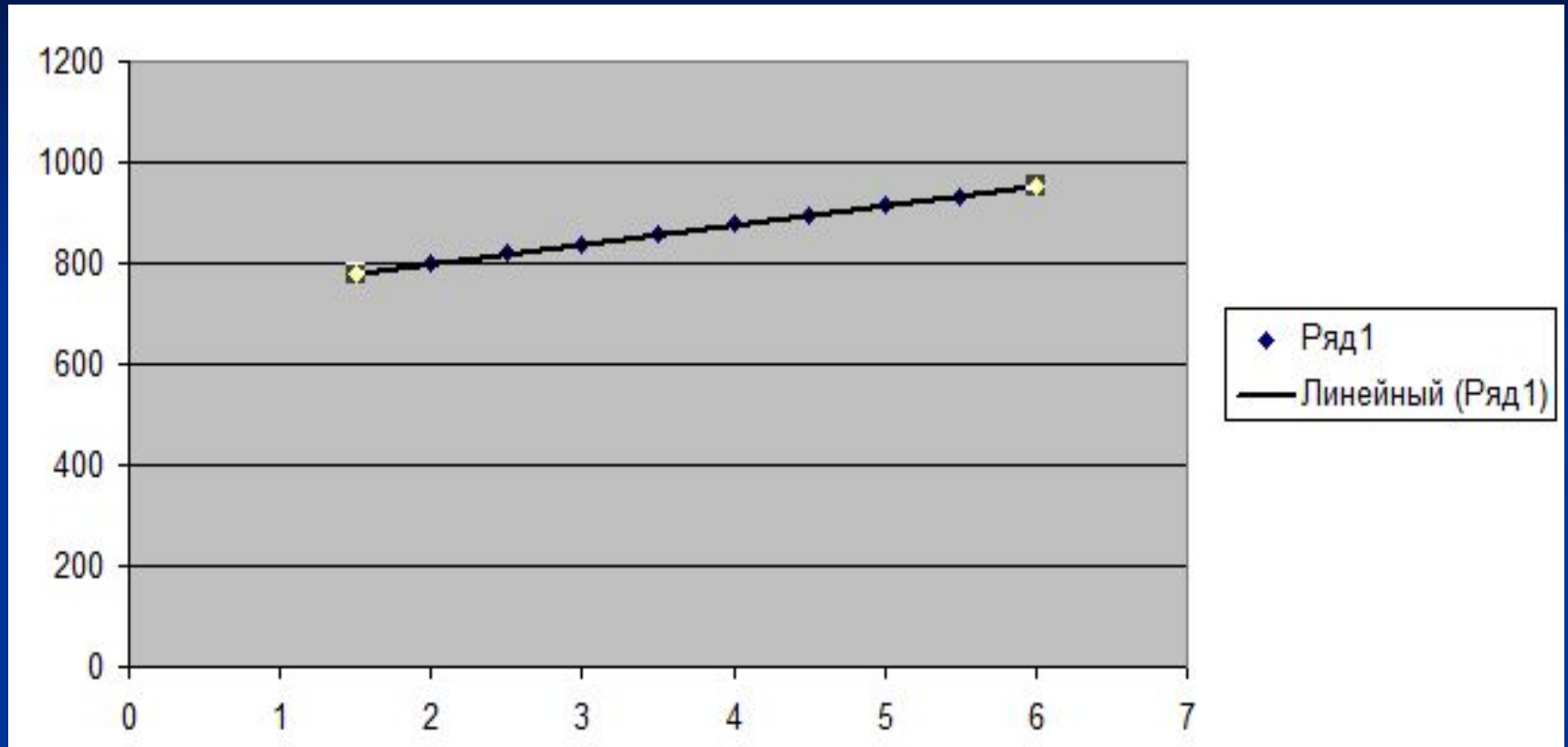
4. Далее действуем по предоставленному шаблону



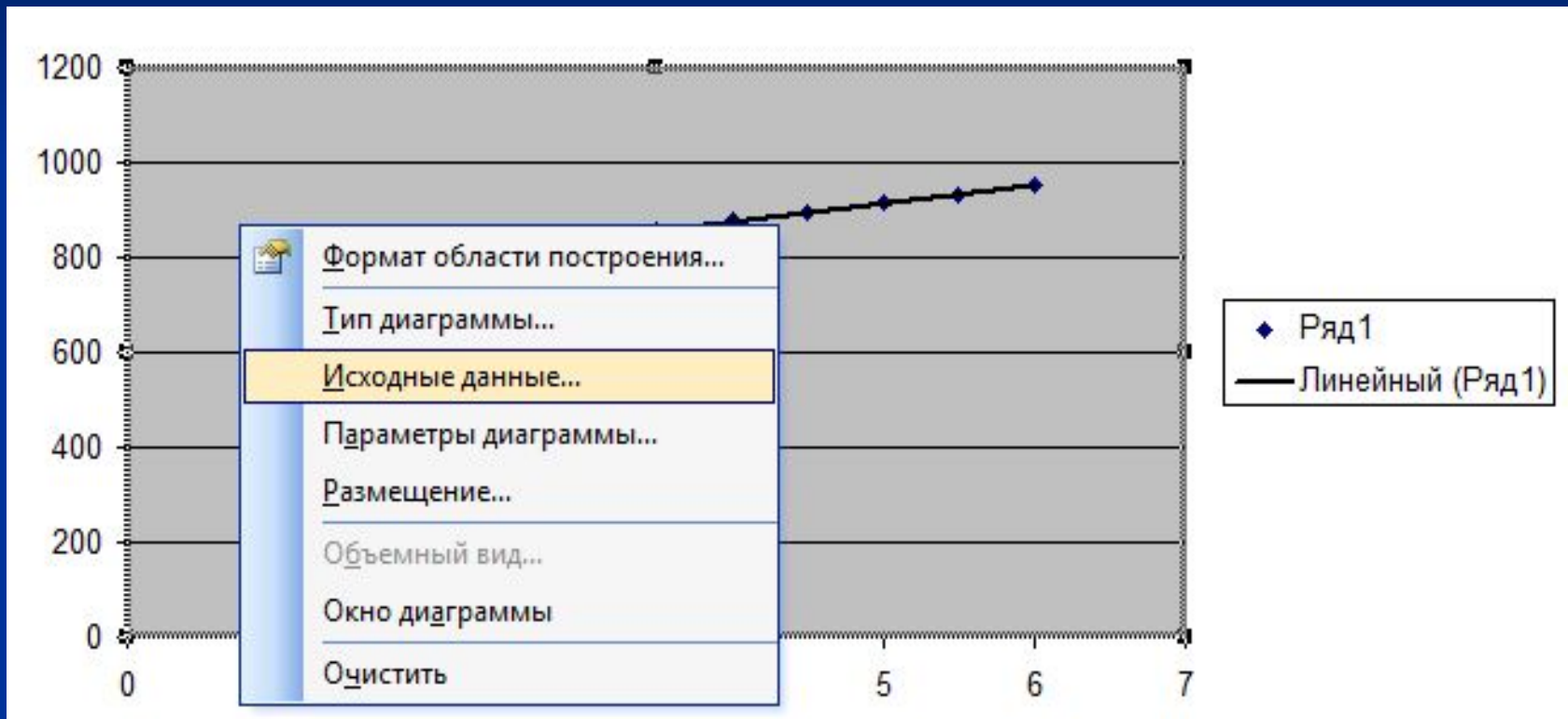
После построения диаграммы выделяем все точки
и соединяем их линией тренда



Построенная диаграмма

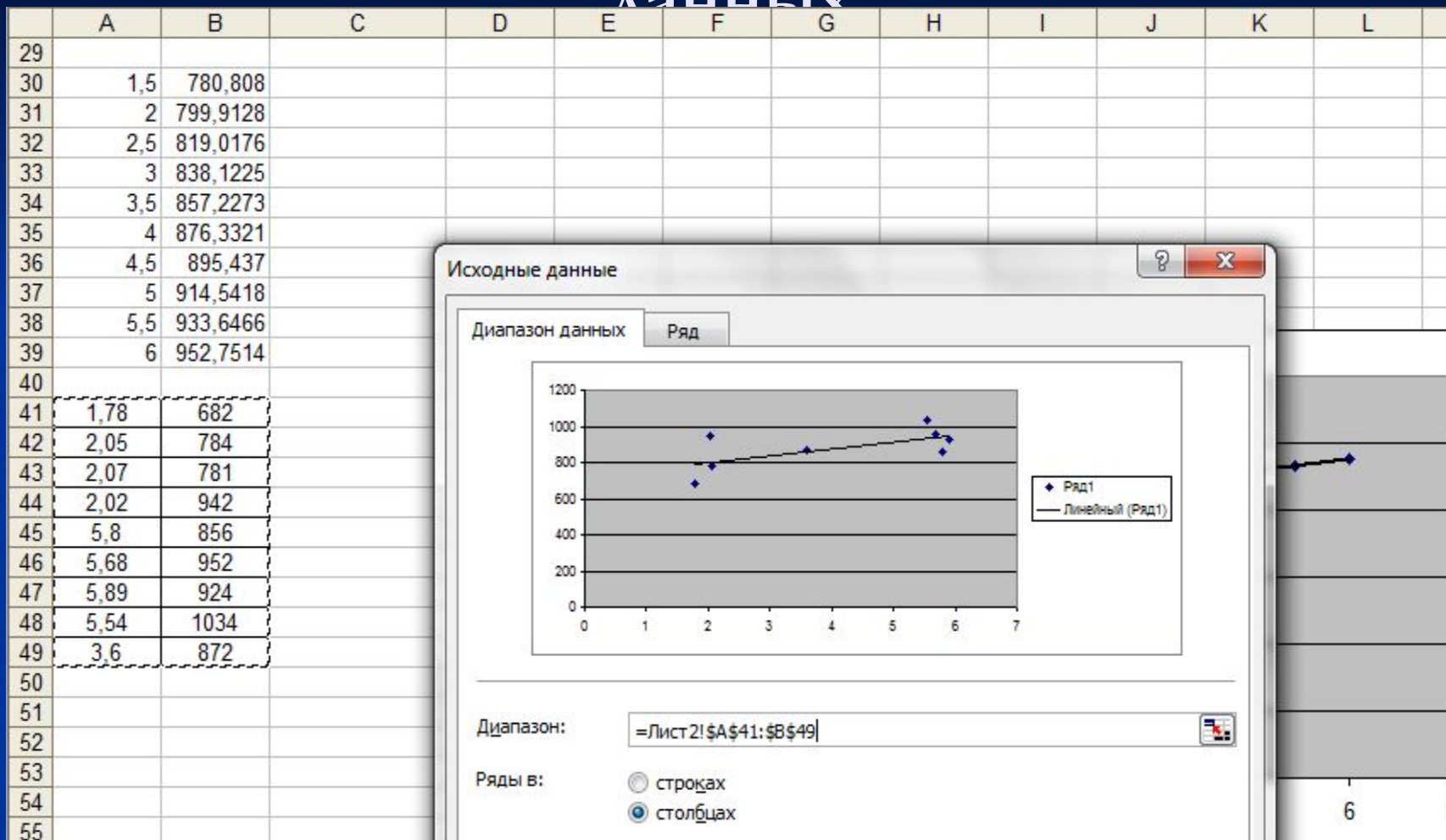


прямой, полученной по уравнению,
на диаграмме наносим точки исходных данных . Нажимаем на
диаграмму правой кнопкой мыши и выбираем исходные данные.
Далее добавляем данные из исходной таблицы

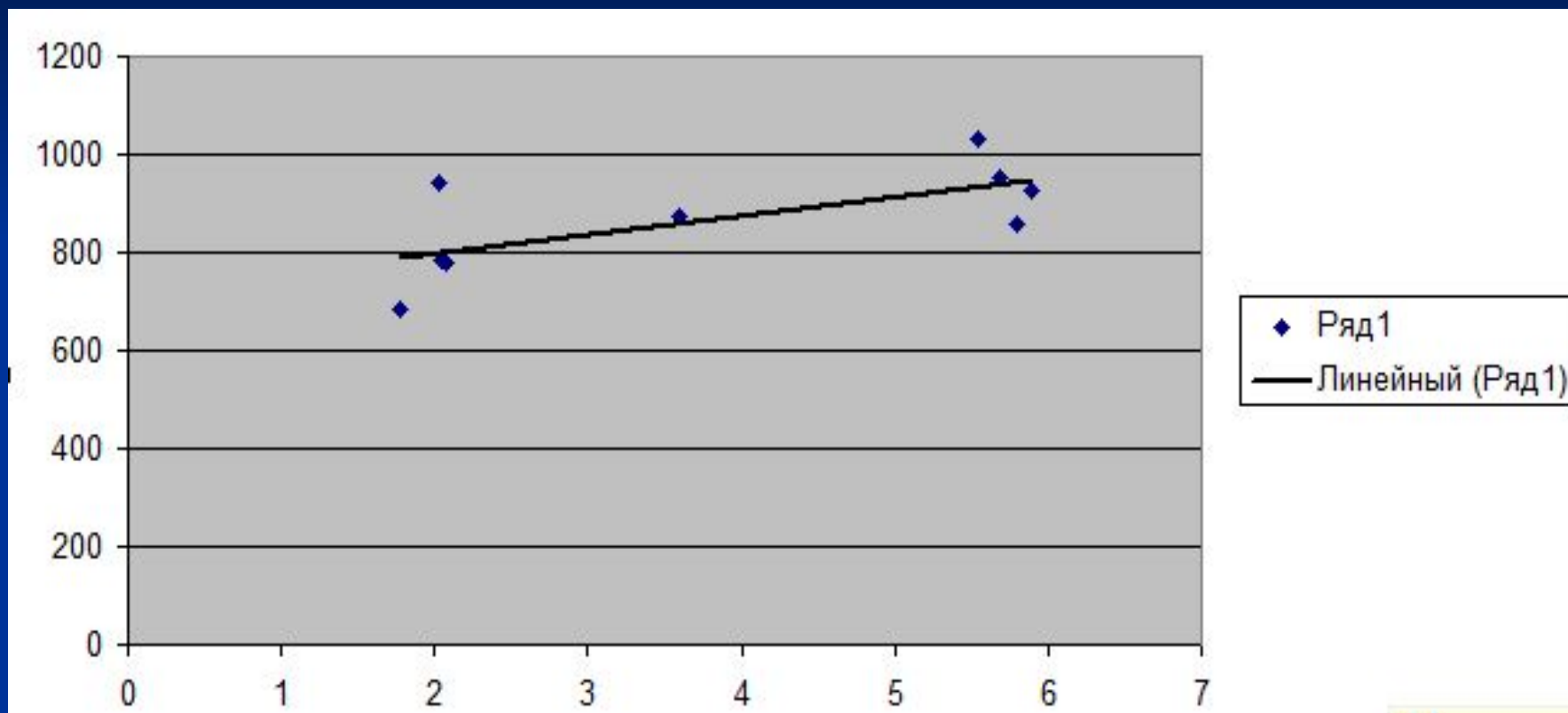


В возникающем окне изменяем диапазон

ДАННЫХ

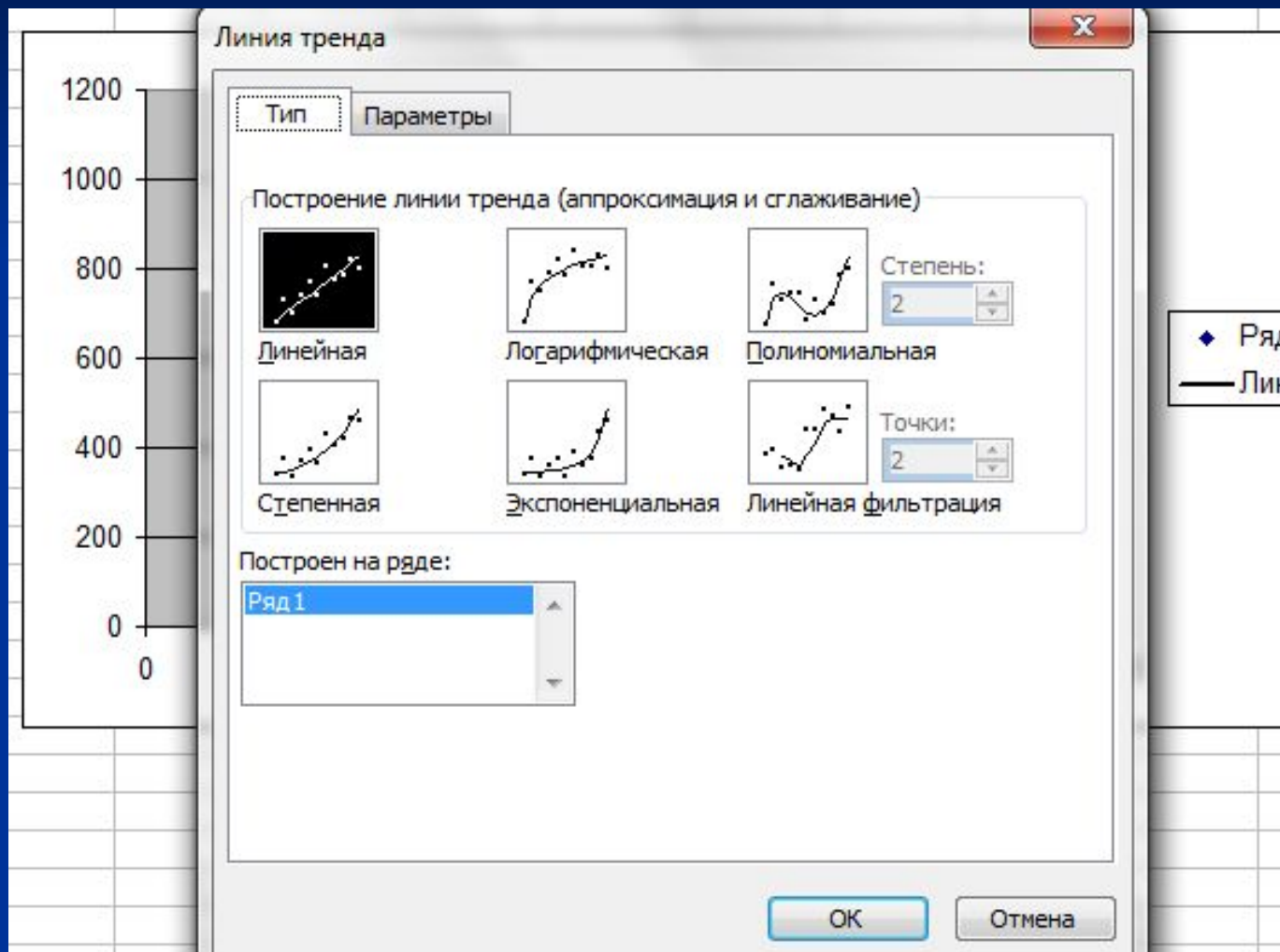


Конечная диаграмма



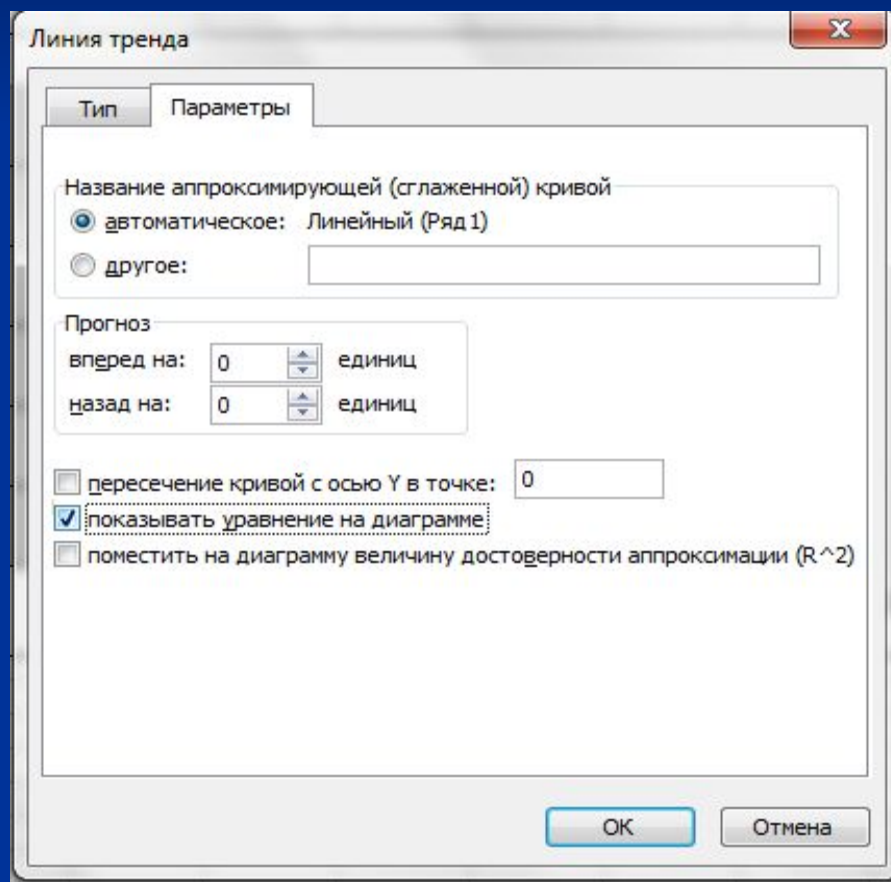
Подбор уравнения регрессии

1. Линейная зависимость

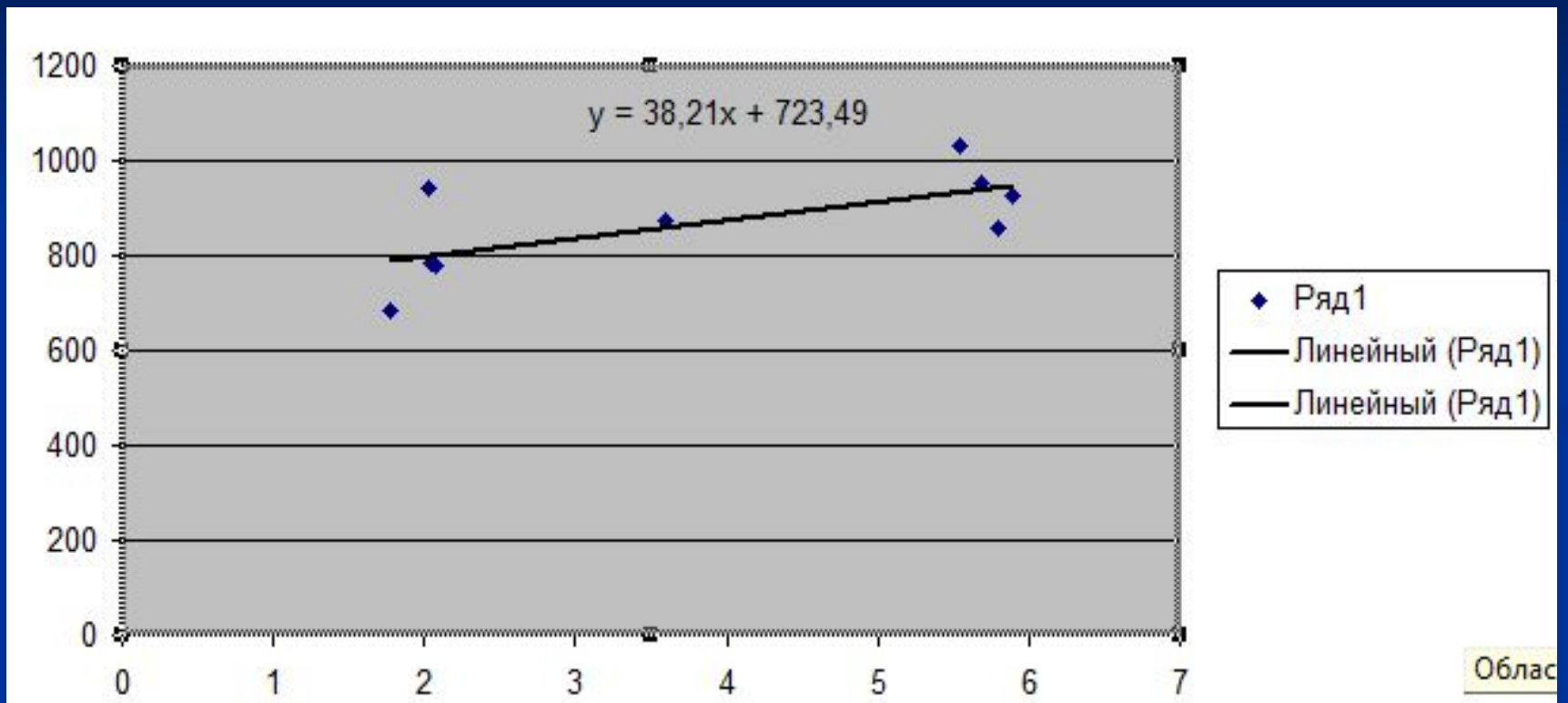


Подбор уравнения регрессии

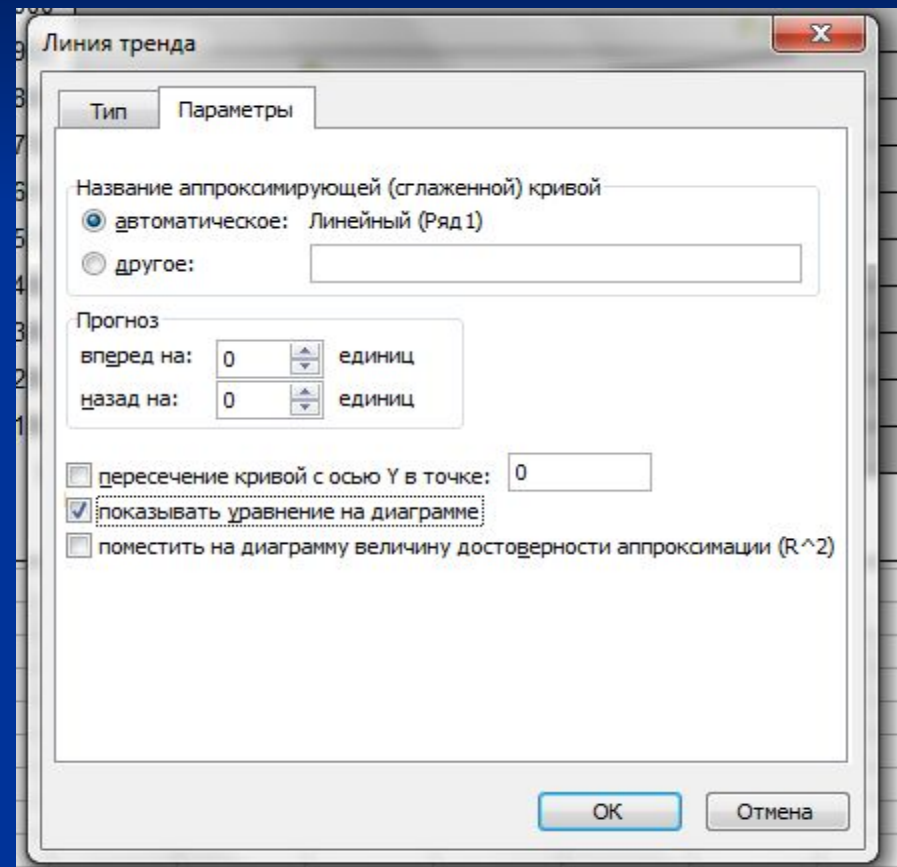
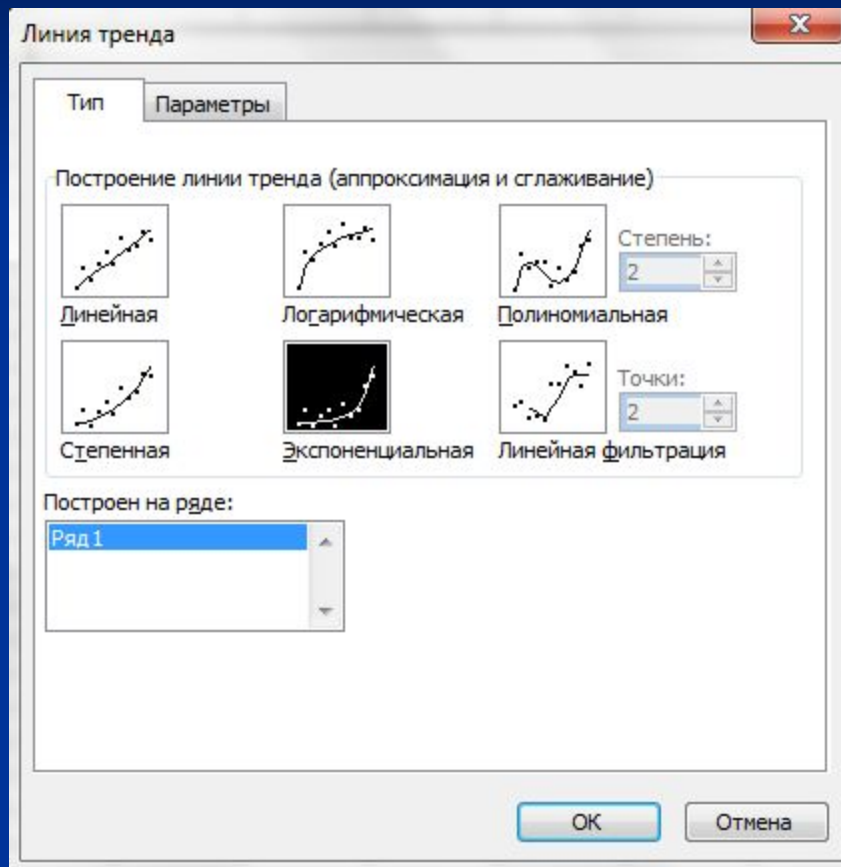
2. Отмечаем на вкладке «показать уравнение на диаграмме»



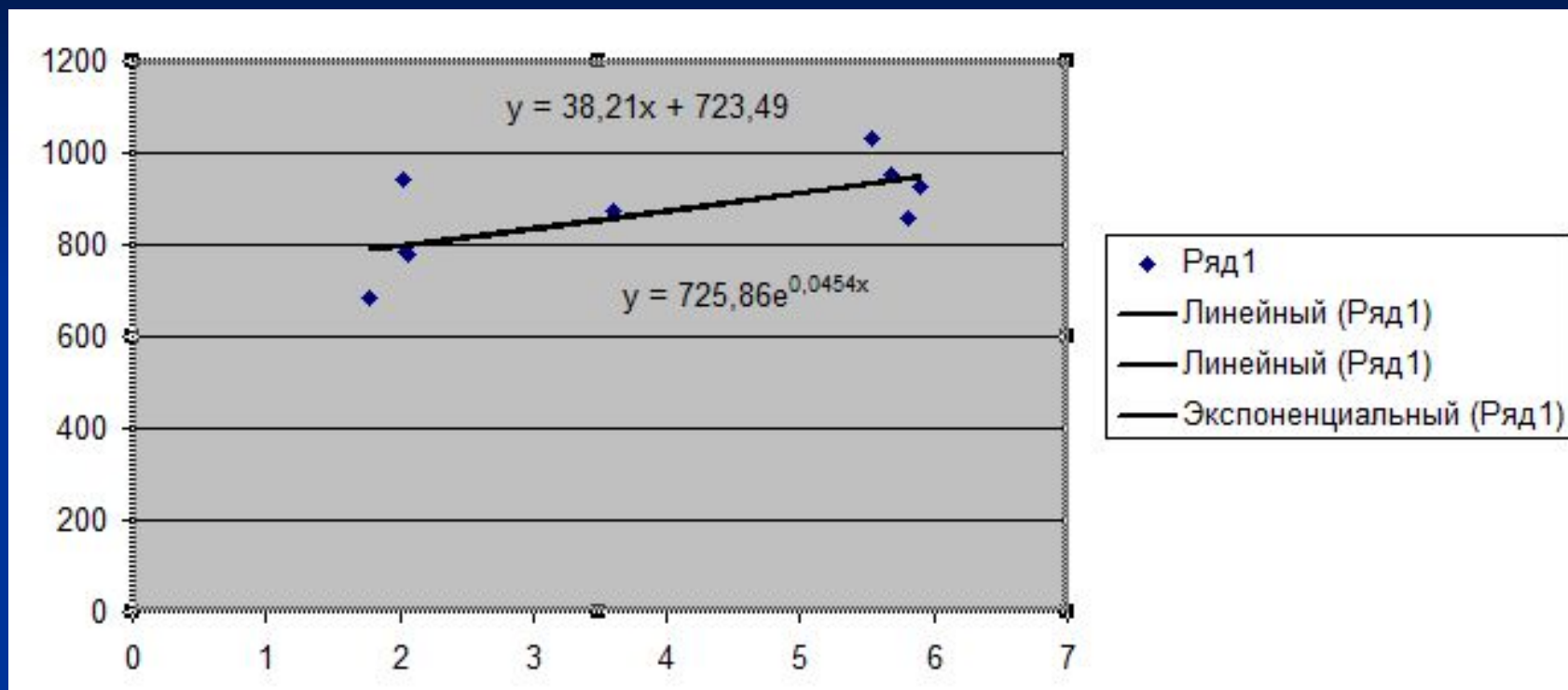
Получаем уравнение на диаграмме



Добавление экспоненциальной линии тренда



Полученная диаграмма



Построение поверхности

исходные данные

	А	В	С
			прочность на растяжение, мПа, Y
1	V, X1	Al, X2	
2	1,78	0,82	682
3	2,05	2,88	784
4	2,07	0,95	781
5	2,02	2,81	942
6	5,8	0,95	856
7	5,68	2,7	952
8	5,89	0,88	924
9	5,54	2,78	1034
10	3,6	1,83	872

Используем регрессию

Используя регрессию мы устанавливаем зависимость между X и Y .

Пользоваться регрессией так же как было описано ранее, за

исключением

Вводимого диапазона X (Выделяется

одновременно

2 столбца X_1 и X_2).

	A	B	C	D	E	F	G	H	I
			прочность на растяжение, мПа, Y						
1	V, X1	Al, X2							
2	1,78	0,82	682						
3	2,05	2,88	784						
4	2,07	0,95	781						
5	2,02	2,81	942						
6	5,8	0,95	856						
7	5,68	2,7	952						
8	5,89	0,88	924						
9	5,54	2,78	1034						
10	3,6	1,83	872						
11									
12									
13									
14									
15									

Регрессия

Входные данные

Входной интервал Y:

Входной интервал X:

☐ Метки ☐ Константа - ноль

☐ Уровень надежности: %

Параметры вывода

☒ Выходной интервал:

☐ Новый рабочий лист:

☐ Новая рабочая книга

Остатки

Получаемые значения при использовании регрессии

11	ВЫВОД ИТОГОВ								
12									
13	Регрессионная статистика								
14	Множеств	0,87009							
15	R-квадрат	0,757056							
16	Нормиров	0,676074							
17	Стандартн	61,22606							
18	Наблюден	9							
19									
20	Дисперсионный анализ								
21		df	SS	MS	F	Значимость F			
22	Регрессия	2	70088,22034	35044,11	9,348512	0,014339			
23	Остаток	6	22491,77966	3748,63					
24	Итого	8	92580						
25									
26	Коэффициенты стандартная ошибка статистики t-Значения нижние 95% верхние 95% нижние 95,0% верхние 95,0%								
27	Y-пересеч	602,5638	65,26540512	9,232514	9,12E-05	442,8651	762,2625	442,8651	762,2625
28	Переменн	39,28403	11,52208941	3,409454	0,014328	11,0905	67,47757	11,0905	67,47757
29	Переменн	63,33595	22,84349517	2,772603	0,032311	7,439932	119,232	7,439932	119,232

Получаемое уравнение регрессии

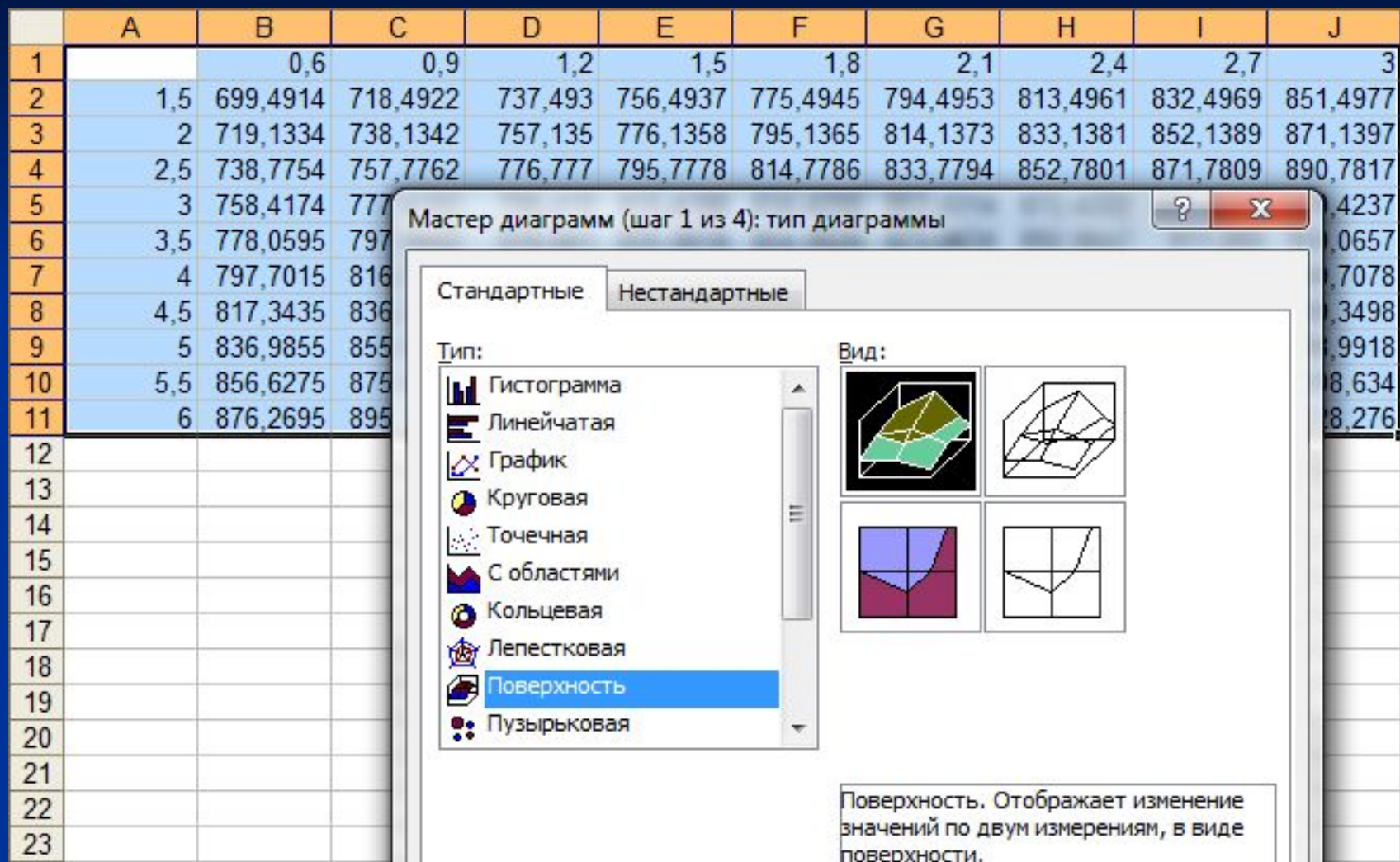
$$Y=602,6+39,3*X1+63,34*X2$$

Рекомендации по построению поверхности

Работа по построению поверхности предполагает использование следующей методики:

1. Подготовить диапазон изменения функции по двум координатам, расположив изменения координаты X_1 вдоль вертикального столбца вниз, а другой X_2 — вдоль прилегающей строки вправо (по оси X).
2. Ввести на пересечении координат (ячейка B2 (рис.2, слайд 29) полученное уравнение регрессии для построения поверхности и воспользоваться маркером автозаполнения для ее копирования на всю область построения поверхности (первоначально протянув по оси X , а затем с максимального значения диапазона данных по X_2 , вертикально вниз по соответствующему столбцу)
3. Выделить подготовленные данные и воспользоваться мастером построения диаграмм (тип диаграммы — **Поверхность**).
4. Отформатировать полученную поверхность.

Построение поверхности с использованием «Мастера диаграмм»



Построение поверхности

В ячейку **A2** введем текст "**Y \ X**" для определения строки таблицы, в которой будут записаны значения аргумента **X**, и столбца, в котором будут записаны значения аргумента **Y**.

Выполним формирование строки значений аргумента **X**. Для этого в ячейку **B2** запишем начальное значение аргумента **X** - "-1", в ячейку **C2** запишем значение "-0,75". Выделим ячейки **B2** и **C2**. Переместим **УМ** в правый нижний угол ячейки **C2**, превратив **УМ** в "маркер заполнения". **ФЛКМ** и протягиваем **УМ** до ячейки **J2**. Ячейки от **D2** до **J2** заполняются значениями аргумента **X** от -0,5 до 1.

Выполнив аналогичные действия, заполним ячейки столбца **A** от **A3** до **A7** значениями аргумента **Y**, изменяющегося от -1 до +1 с шагом 0,5.

Выполним формирование таблицы значений функции **Z**. Для этого в

ячейку **B3** введем формулу

$$= 53,065 + 5,4 * \$A2^2 + 2,36 * B\$1 \quad (1)$$

Рассмотрим более подробно последовательность действий при вводе этой формулы.

- **ЩЛК** выберем ячейку **B3** и введем в нее с клавиатуры символ "=".
- **УМ** переместим на ячейку **B2** и **ЩЛК**. При этом в строке формул после символа "=" появляется ссылка на ячейку **B2**, в которой находится первое значение аргумента **X**.
- Для того, чтобы сделать эту ссылку абсолютной по строке B, дважды нажмем функциональную клавишу **F4**. При этом ссылка последовательно преобразуется

$$B2 \rightarrow \$B\$2 \rightarrow B\$2$$

- Продолжаем ввод формулы: **"=B\$2^2+"**
- Далее по аналогии с формированием ссылки на ячейку **B\$2**, создаем абсолютную по столбцу ссылку **\$A3** и заканчиваем ввод формулы (1) нажатием на кнопку "Ввод" в строке формул.

Подводим **УМ** в правый нижний угол ячейки **B3**, образуя "маркер заполнения". **ФЛКМ** и, протаскивая его, заполняем ячейки таблицы от **B3**

до **J7**. Созданная таблица значений функции **Z** показана на приведенном в конце описания рисунке.

B2		fx =602,563768589013+39,2840341056241*\$A2^1+63,3359511109786*\$B\$1								
	A	B	C	D	E	F	G	H	I	J
1		0,6	0,9	1,2	1,5	1,8	2,1	2,4	2,7	3
2	1,5	699,4914	718,4922	737,493	756,4937	775,4945	794,4953	813,4961	832,4969	851,4977
3	2	719,1334	738,1342	757,135	776,1358	795,1365	814,1373	833,1381	852,1389	871,1397
4	2,5	738,7754	757,7762	776,777	795,7778	814,7786	833,7794	852,7801	871,7809	890,7817
5	3	758,4174	777,4182	796,419	815,4198	834,4206	853,4214	872,4222	891,4229	910,4237
6	3,5	778,0595	797,0602	816,061	835,0618	854,0626	873,0634	892,0642	911,065	930,0657
7	4	797,7015	816,7023	835,703	854,7038	873,7046	892,7054	911,7062	930,707	949,7078
8	4,5	817,3435	836,3443	855,3451	874,3458	893,3466	912,3474	931,3482	950,349	969,3498
9	5	836,9855	855,9863	874,9871	893,9879	912,9887	931,9894	950,9902	969,991	988,9918
10	5,5	856,6275	875,6283	894,6291	913,6299	932,6307	951,6315	970,6322	989,633	1008,634
11	6	876,2695	895,2703	914,2711	933,2719	952,2727	971,2735	990,2743	1009,275	1028,276

Рисунок 2.

Построение поверхности

Для построения поверхности функции **Z** выделим ячейки в диапазоне от **B3** до **J7**, содержащие таблицу значений функции.

На **ПШ** "Стандартная" нажмем кнопку "**Мастер диаграмм**". В открывающемся окне "**Мастер диаграмм (шаг 1 из 4): тип диаграммы**" на вкладке "**Стандартные**" в поле "**Тип:**" выбираем строку "**Поверхность**". В поле "**Вид:**" выбираем один из образцов и нажимаем кнопку "**Далее**".

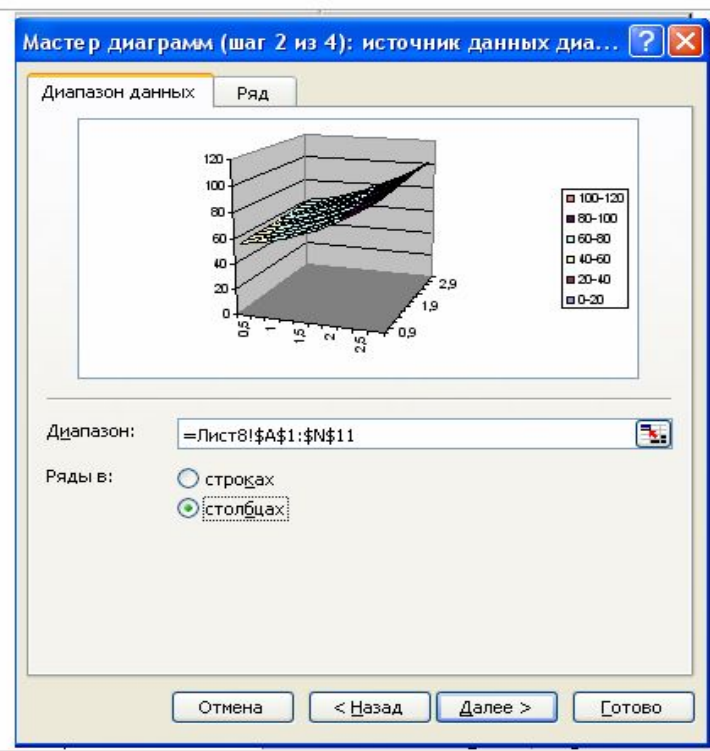
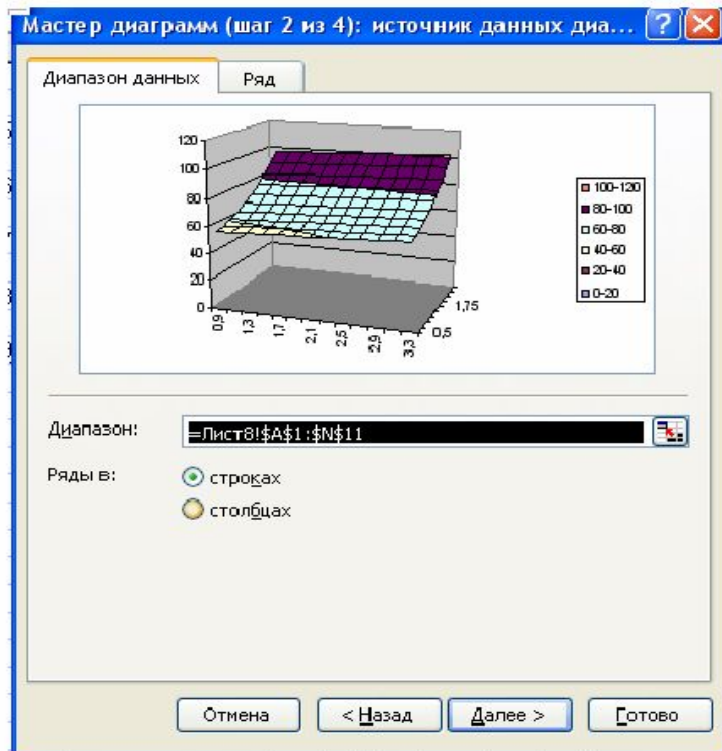
В окне "**Мастер диаграмм (шаг 2 из 4): источник данных диаграммы**" (Рис. 1) выполняется проверка правильности выделения данных, используемых для построения диаграммы.

- На вкладке "**Диапазон данных**" в поле ввода "**Диапазон:**" проверяем соответствие диапазона ячеек, выбранных для построения диаграммы, выделенному диапазону ячеек. При необходимости можно изменить диапазон выбранных ячеек.
- Из двух переключателей "**Ряды в: строках, столбцах**" можно выбрать вариант, позволяющий получить наиболее наглядный вариант диаграммы. (На рис. 1 показан вид диаграммы при выборе варианта "в столбцах", а на рис. 2 - "в строках".)
- На вкладке "**Ряд**" в одноименном поле выбираем последовательно строки Ряд1, Ряд2 и т.д. Каждой выбранной строке присваиваем

"Имя". Для этого **ЩЛК** по кнопке в правом конце поля ввода *"Имя:"* сворачивает окно диалога. Последующий **ЩЛК** по ячейке со значением аргумента, соответствующим выбранному ряду данных вносится в поле имени и изменяет соответствующую этой строке отметку на оси диаграммы.

- Для нанесения отметок по второй оси диаграммы используется поле ввода *"Подписи по оси X:"*. **УМ** на кнопку в правом конце поля ввода, а затем выделяем строку со значениями аргумента **X** (диапазон **B2:J2**).
- Выполнив эти действия, нажимаем кнопку *"Далее >"*.
- Открывается окно *"Мастер диаграмм (шаг 3 из 4): параметры диаграммы"*. В этом окне на вкладке *"Заголовки"* можно определить название диаграммы, метки осей диаграммы. На вкладке *"Линии сетки"* можно определить линии разметки, наносимые на плоскости

осевого триэдра. Нажимаем кнопку "*Далее >*" и переходим в окно следующего шага мастера диаграмм.



- В открывающемся окне "*Мастер диаграмм (шаг 4 из 4): размещение диаграммы*" определяется размещение создаваемой диаграммы на отдельном или имеющемся листе книги. Сделав выбор варианта размещения, нажимаем кнопку "*Готово*".
- Дальнейшее форматирование диаграммы выполняется с использованием меню "*Диаграмма*" и "*Формат*" для выбранных элементов диаграммы.

Построенная поверхность

