

Практическая работа №6

Microsoft Excel. Графическое представление данных и элементы статистической обработки

Откройте рабочую книгу EXCEL.

Перейдите на второй лист. Назовите Лист 2 – Работа 6

Задание №6.1.

1. Выбрать из таблицы 6.1 пять стран, начиная со своего номера варианта.
2. Создать таблицу с этими данными.
3. Дополнить таблицу данными по 2020 год. Необходимые данные можно найти на сайтах мировой статистики. (например, https://lirt.hse.ru/static_data, <https://ru.tradingeconomics.com>) В отчете необходимо указать источник данных.
4. Подсчитать общие значения по каждой стране и по каждому году
5. Оформить таблицу.

Таблица 6.1.

Численность занятых в экономике стран мира (миллионов человек)

Вариант	Страна	2000	2002	2003	2004	2005	2006
1	Россия	65,3	66,3	67,2	67,1	68,6	69,2
2	Австрия	3,8	3,8	3,8	3,7	3,8	3,9
3	Беларусь	4,4	4,4	4,3	4,3	4,3	4,4
4	Бельгия	4,1	4,1	4,1	4,1	4,2	4,3
5	Венгрия	3,8	3,9	3,9	3,9	3,9	3,9
6	Германия	36,6	36,5	36,2	35,7	36,6	37,3
7	Греция	4,1	4,2	4,3	4,3	4,4	4,5
8	Испания	15,5	16,6	17,3	18	19	19,7
9	Италия	21,2	21,9	22,1	22,4	22,6	23
10	Литва	1,4	1,4	1,4	1,4	1,5	1,5
11	Норвегия	2,3	2,3	2,3	2,3	2,3	2,4
12	Польша	14,5	13,8	13,6	13,8	14,1	14,6

13	Португалия	5	5,1	5,1	5,1	5,1	5,2
14	Австралия	9	9,2	9,5	9,6	10	10,2
15	Румыния	10,8	9,2	9,2	9,2	9,1	9,3
16	Словакия	2,1	2,1	2,2	2,2	2,2	2,3
17	Словения	0,9	0,9	0,9	0,9	0,9	1
18	Египет	17,2	17,9	18,1	18,7	19,3	19,4
19	Украина	20,2	20,1	20,2	20,3	20,7	20,7
20	Финляндия	2,4	2,4	2,4	2,4	2,4	2,5
21	Канада	14,8	15,3	15,7	15,9	16,2	16,5
22	Чешская Республика	4,7	4,8	4,7	4,7	4,8	4,8
23	Швейцария	3,9	4	4	4	4	4,1
24	Швеция	4,2	4,2	4,2	4,2	4,3	4,3
25	Эстония	0,6	0,6	0,6	0,6	0,6	0,6
26	Азербайджан	3,7	3,7	3,7	3,8	3,9	4
27	Израиль	2,2	2,3	2,3	2,4	2,5	2,6
28	Индонезия	89,8	91,6	90,8	93,7	94,9	95,2
29	Казахстан	6,2	6,7	7	7,2	7,3	7,4
30	Пакистан	36,8	38,9	39,9	42	42,9	47
	Всего	411,5	418,2	421	427,9	436,4	445,8

6. Используя условное форматирование:

выделить голубой заливкой значения, которые составляют от итогового значения за год более 40% и желтой заливкой, если меньше 20%

** Выделить красной заливкой название страны зеленой заливкой, если ее данные показывают рост численности занятых в экономике стране, и красной, если ситуация стабильна

7. Построить три вида диаграмм (столбчатую, цилиндрическую и круговую) доли страны в общей численности занятых в 2005 году.

Задание 6.2

1. Построить график изменения численности занятых в экономике страны для своего варианта.
2. Построить линию тренда таким образом, чтобы значение коэффициента достоверности результата R^2 был наиболее близок к единице. Создайте небольшую таблицу, показывающую тип линии тренда, коэффициент достоверности и соответствующее уравнение.
3. Сделать прогноз на один период вперед

Задание 6.3

Методом наименьших квадратов найти уравнение оптимальной прямой, наименее удаленной от заданных экспериментальных точек, и вычислить величину достоверности аппроксимации R^2 .

Для расчета взять $n=12$ точек. Точки берутся из таблицы подряд, начиная с номера S студента по журналу.

Варианты задания

N	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
x_i	2	3	3	5	6	7	13	13	11	10	9	8	2	2	4	5	6	7	8	8	3	9	11
y_i	1	2	3	4	7	7	15	17	11.5	10	8	6.5	1	3	4	5.5	6	6.5	7	9	3	8	10

N	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43
x_i	13	14	14	8	5	7	12	2	1	8	15	12	12	7	5	9	6	7	7	5
y_i	12	13	14	9	6	7	11	1	1	7	15	13	12	6.5	5	8	6	6.5	8	4

Задание следует выполнить тремя способами:

1. используя формулы для вычисления коэффициентов оптимальной прямой и величины достоверности аппроксимации R^2 ;
2. используя функции "ЛИНЕЙН" и "КОРРЕЛ";
3. отобразив экспериментальные точки на точечной диаграмме, добавив на диаграмму линию тренда и поместив на диаграмму величину достоверности аппроксимации R^2 .

Способ 1

Уравнение оптимальной прямой имеет вид

$$y = ax + b ,$$

где коэффициенты a и b определяются из системы уравнений

$$\begin{cases} (\sum x_i^2) \cdot a + (\sum x_i) \cdot b = \sum x_i y_i \\ (\sum x_i) \cdot a + n \cdot b = \sum y_i \end{cases} .$$

В матричном виде эта система уравнений имеет вид $\mathbf{Ax} = \mathbf{B}$,

где

$$\mathbf{A} = \begin{bmatrix} (\sum x_i^2) & (\sum x_i) \\ (\sum x_i) & n \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} a \\ b \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} \sum x_i y_i \\ \sum y_i \end{bmatrix} .$$

Решение системы уравнений в матричном виде $\mathbf{x} = \mathbf{A}^{-1}\mathbf{B}$,

где \mathbf{A}^{-1} — матрица, обратная к матрице \mathbf{A} .

Для вычисления обратной матрицы следует воспользоваться функцией "МОБР", а для умножения матрицы \mathbf{A}^{-1} на матрицу \mathbf{B} — функцией "МУМНОЖ". Для получения результатов с использованием этих функций необходимо предварительно выделить массив нужного размера. Для запуска этих функций следует пользоваться комбинацией клавиш {Ctrl+Shift+Enter}.

Величина достоверности аппроксимации R^2 , равная квадрату коэффициента корреляции, вычисляется по формуле

$$R^2 = 1 - \frac{E}{T} ,$$

$$\text{где } E = \sum (y_i - y_{i, \text{теор}})^2, \quad y_{i, \text{теор}} = ax_i + b, \quad T = \sum y_i^2 - \frac{(\sum y_i)^2}{n} .$$

Величина R^2 располагается в интервале от нуля до единицы. Чем R^2 ближе к единице, тем аппроксимация более достоверна.

Способ 2

Для вычисления коэффициентов a и b следует воспользоваться функцией "ЛИНЕЙН". Предварительно необходимо выделить две ячейки в строке. Для запуска этой функции следует пользоваться комбинацией клавиш {Ctrl+Shift+Enter}.

Вычисление коэффициента корреляции R производим при помощи функции "КОРРЕЛ". Исходными являются массивы y_i и $y_{i,теор}$. Для определения величины достоверности аппроксимации значение R возводим в квадрат.

Способ 3

Отображаем экспериментальные точки на точечной диаграмме, позволяющей сравнить пары значений.

В третьем окне Мастера диаграмм устанавливаем:

Во вкладке *Заголовки* Название диаграммы: "Аппроксимация экспериментальных данных", Ось X (категорий): "x", Ось Y (значений): "y".

Во вкладке *Линии сетки* — сетки для осей X и Y.

Во вкладке *Легенда* — добавить легенду.

В четвертом окне Мастера диаграмм устанавливаем, что диаграмму необходимо поместить на имеющемся листе.

Для дополнения экспериментальных данных линией регрессии (линией тренда), которая аппроксимирует точки данных, следует выделить диаграмму и выбрать команду *Добавить линию тренда* меню *Диаграмма*. В открывшемся окне диалога во вкладке *Тип* выбираем "Линейная", во вкладке *Параметры* отмечаем "показывать уравнение на диаграмме" и "поместить на диаграмму величину достоверности аппроксимации (R^2)".

Результаты, полученные каждым из трех способов, должны совпадать!

Пример оформления показан на рисунке 1.

Задание 6.4 Написать отчет:

- ✓ Алгоритм построения графика.
- ✓ Оформление графика
- ✓ МНК
- ✓ Линия тренда.
- ✓ Величина достоверности аппроксимации R^2

	A	B	C	D	E	F	G	H	I	J	K
1	Аппроксимация экспериментальных данных методом наименьших квадратов										
2											
3			Способ 1								
4	i		x_i	y_i		x_i²	x_i·y_i	y_i²	y_{i,теор}	(y_i-y_{i,теор})²	
5	1		1	1		1	1	1	0,82	0,03	
6	2		8	8		64	64	64	7,75	0,06	
7	3		15	14		225	210	196	14,68	0,46	
8	4		12	12		144	144	144	11,71	0,08	
9	5		12	13		144	156	169	11,71	1,66	
10	6		7	7		49	49	49	6,76	0,06	
11	7		5	6		25	30	36	4,78	1,49	
12	8		9	8		81	72	64	8,74	0,55	
13	9		6	5		36	30	25	5,77	0,59	
14	10		7	6		49	42	36	6,76	0,57	
15	11		7	7		49	49	49	6,76	0,06	
16	12		5	4		25	20	16	4,78	0,60	
17											
18	Σ		94	91		892	867	849	91,00	6,24	
19											
20	[A]=	892	94	[B]=	867						
21		94	12		91						
22											
23	[A⁻¹]=	0,0064	-0,0503		[x]=	a=	0,9904				
24		-0,0503	0,4775			b=	-0,1745				
25											
26						E=	6,24				
27						T=	158,92				
28						R²=	0,9608				
29											
30			Способ 2								
31	a=	b=		R=	R²=						
32	0,9904	-0,1745		0,9802	0,9608						
33											
34			Способ 3								
35											
36											
37											
38											
39											
40											
41											
42											
43											
44											
45											
46											
47											
48											
49											
50											
51											
52											

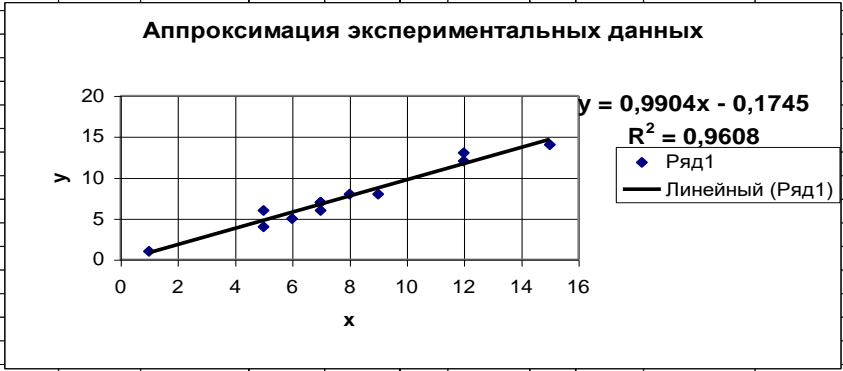


Рисунок 1 Пример оформления задания 6.3