

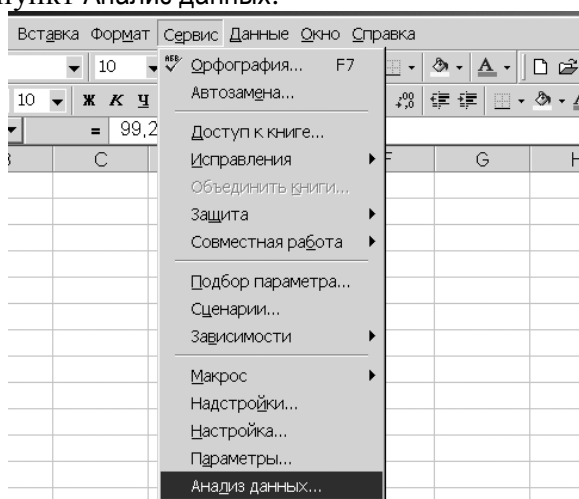
РЕШЕНИЕ ЗАДАЧ ОПИСАТЕЛЬНОЙ СТАТИСТИКИ С ПОМОЩЬЮ ПАКЕТА АНАЛИЗА MS EXCEL

Простейшие задачи описательной статистики могут решаться с использованием табличных процессоров. Далее все примеры приводятся для пакета Microsoft Excel 2000.

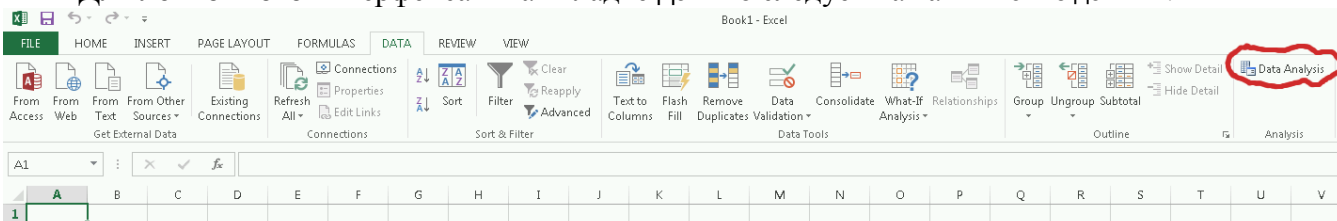
Пусть выборка, содержащая 1000 вариантов, расположена в первом столбце первого рабочего листа текущей рабочей книги:

	A	B	C	D
992	99,210			
993	101,259			
994	95,157			
995	101,878			
996	98,946			
997	98,187			
998	97,747			
999	96,665			
1000	101,426			
1001	106,734			
1002	99,113			
1003	97,963			
1004	94,625			
1005	99,220			
1006	93,999			
1007	103,472			
1008				
1009				
1010				

Для нахождения точечных оценок числовых параметров неизвестного распределения, которому подчинена генеральная совокупность, следует: для **ортодоксального** интерфейса – из меню Сервис выбрать пункт Анализ данных:

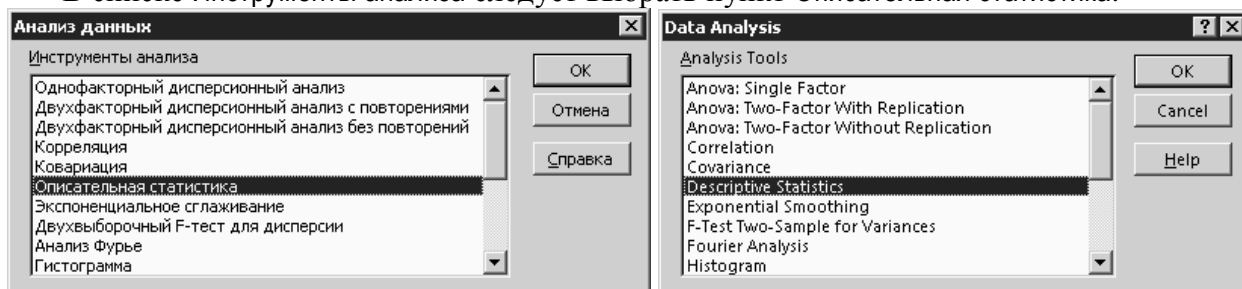


Для **ленточного** интерфейса – на вкладке Данные следует нажать Анализ данных:

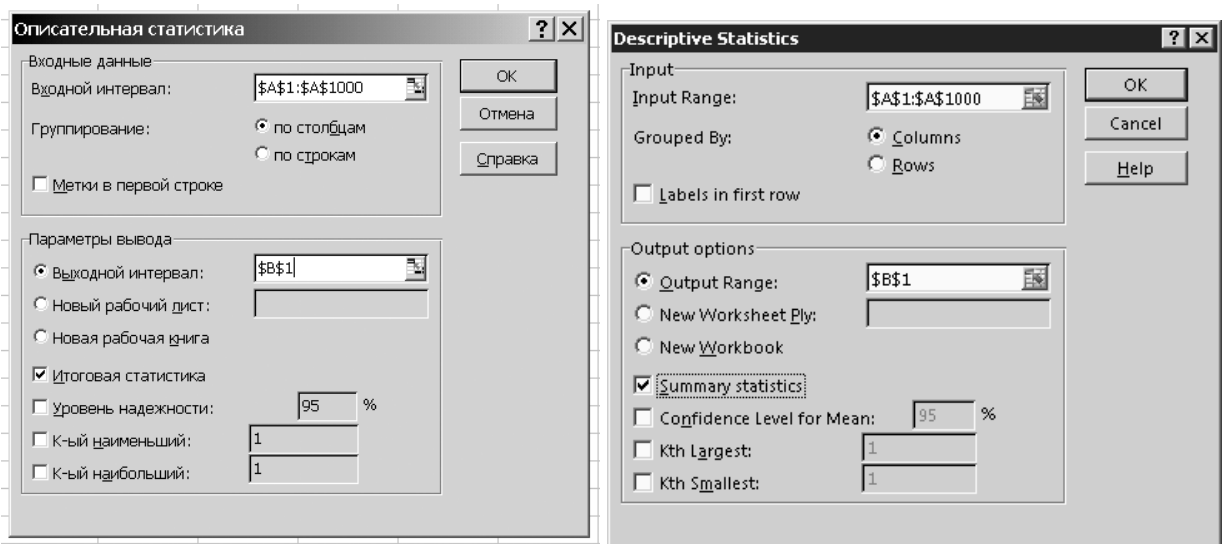


Если указанные действия недоступны, то необходимо установить пакет анализа (для **ортодоксального** интерфейса: выбор Сервис – Надстройки; в диалоговом окне установить флажок Пакет анализа; для **ленточного** интерфейса способ подключения пакета анализа усложняется, см. <http://www.google.com/search?q=ms+excel+load+analysis+toolpak>).

В списке Инструменты анализа следует выбрать пункт Описательная статистика:



В диалоговом окне Описательная статистика



необходимо указать диапазон рабочего листа, содержащий выборку; в данном примере – \$A\$1:\$A\$1000. В качестве выходного интервала достаточно указать первую ячейку второго столбца – \$B\$1. Дополнительно следует установить флажок Итоговая статистика, после чего нажать Enter. Результаты анализа будут помещены во второй столбец:

	A	B	C
1	99,210	Столбец1	
2	101,259		
3	95,157	Среднее	99,91778
4	101,878	Стандартная ошибка	0,093233
5	98,946	Медиана	99,85605
6	98,187	Мода	97,38525
7	97,747	Стандартное отклонение	2,948278
8	100,798	Дисперсия выборки	8,692346
9	99,088	Экссесс	-0,17834
10	104,529	Асимметричность	0,074057
11	98,301	Интервал	17,02596
12	101,805	Минимум	91,45482
13	102,784	Максимум	108,4808
14	95,570	Сумма	99917,78
15	102,903	Счет	1000
16	97,288		

Указать диапазон, содержащий выборку, можно следующим образом: после перевода фокуса ввода на поле Входной интервал щелкнуть на первой ячейке диапазона (\$A\$1), затем удерживая клавиши Shift и Control нажать PageDown; при этом диапазон будет расширен до последней заполненной ячейки (\$A\$1000).

Пакет анализа MS Excel содержит встроенные средства построения непрерывного вариационного ряда и гистограммы, однако эти средства функционируют не совсем корректно. Поэтому часть данных, необходимых для построения гистограммы, следует подготовить отдельно.

Найдем границы разрядов. Интервал изменения вариантов (в данном примере – от 91,5 до 108,5) уже известен. В качестве левой границы первого разряда выберем 90, в качестве правой границы последнего – 110.

Так как

$$l = 1 + 3,31 \lg 1000 = 10,9,$$

то число разрядов можно взять равным 10, тогда длина каждого разряда будет равна

$$\frac{110 - 90}{10} = 2.$$

Вычисление границ удобно выполнять с использованием автозаполнения. После двойного щелчка на ячейке D1 вводим 90; нажатие на Enter переводит на ячейку D2. В эту ячейку следует ввести число 92. Затем следует выделить ячейки D1 и D2 (щелчок на D1, нажать и

удерживать Shift, щелчок на D2), подвести курсор к маркеру автозаполнения (черный квадрат в правом нижнем углу ячейки D2):

	A	B	C	D
1	99,210	Столбец1		90
2	101,259			92

и, удерживая левую клавишу мыши, перевести маркер до ячейки D11:

	A	B	C	D
1	99,210	Столбец1		90
2	101,259			92
3	95,157	Среднее	99,91778	94
4	101,878	Стандартная ошибка	0,093233	96
5	98,946	Медиана	99,85605	98
6	98,187	Мода	97,38525	100
7	97,747	Стандартное отклонение	2,948278	102
8	100,798	Дисперсия выборки	8,692346	104
9	99,088	Эксцесс	-0,17834	106
10	104,529	Асимметричность	0,074057	108
11	98,301	Интервал	17,02596	110
12	101,805	Минимум	91,45482	

После этого из меню Сервис вновь следует выбрать Анализ данных, и в списке инструментов анализа выбрать пункт Гистограмма. Как и ранее, входным интервалом будет диапазон \$A\$1:\$A\$1000. Интервал, содержащий границы разрядов, указывается в поле Интервал карманов (в данном примере – \$D\$1:\$D\$11). В качестве выходного интервала достаточно указать первую ячейку пятого столбца – \$E\$1:

	A	B	C	D	E	F	G	H	I	J
1	99,210	Столбец1		90						
2	101,259			92						
3	95,157	Среднее	99,91778	94						
4	101,878	Стандартная ошибка	0,093233	96						
5	98,946	Медиана	99,85605	98						
6	98,187	Мода	97,38525	100						
7	97,747	Стандартное отклонение	2,948278	102						
8	100,798	Дисперсия выборки	8,692346	104						
9	99,088	Эксцесс	-0,17834	106						
10	104,529	Асимметричность	0,074057	108						
11	98,301	Интервал	17,02596	110						
12	101,805	Минимум	91,45482							
13	102,784	Максимум	108,4808							
14	95,570	Сумма	99917,78							
15	102,903	Счет	1000							
16	97,288									
17	98,109									
18	106,141									
19	99,306									
20	100,290									
21	105,748									
22	99,836									
23	97,502									
24	102,663									
25	100,805									

Гистограмма [?] [X]

Входные данные

Входной интервал: \$A\$1:\$A\$1000 [...]

Интервал карманов: \$D\$1:\$D\$11 [...]

Метки

Параметры вывода

Выходной интервал: \$E\$1 [...]

Новый рабочий лист: [...]

Новая рабочая книга

Парето (отсортированная гистограмма)

Интегральный процент

Вывод графика

OK Отмена Справка

Частоты, соответствующие каждому разряду, будут помещены в ячейки F3:F12:

	D	E	F
	90	Карман	Частота
	92	90	0
78	94	92	1
33	96	94	20
05	98	96	69
25	100	98	160
78	102	100	272
46	104	102	229
34	106	104	165
57	108	106	61
96	110	108	21
32		110	2
08		Еще	0

Перед построением гистограммы следует:

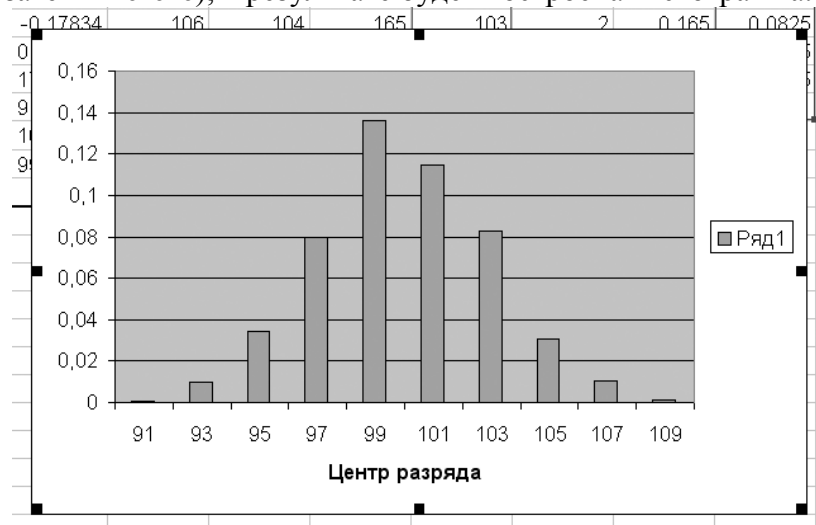
– вычислить значения, соответствующие центру каждого разряда (в ячейку G3 вводится «=(E3+E2)/2», нажатие Enter, маркер автозаполнения переводится от ячейки G3 до ячейки G12);

– при необходимости – вычислить длины каждого разряда (в ячейку H3 вводится «=E3-E2», нажатие Enter, маркер автозаполнения переводится от H3 до H12); если все длины одинаковы, можно заполнить диапазон H3:H12 соответствующими значениями (выделить диапазон H3:H12, ввести 2, нажать Ctrl+Shift+Enter).

– вычислить относительные частоты (в ячейку I3 вводится «=F3/1000», нажатие Enter, маркер автозаполнения переводится от I3 до I12; в данном примере число 1000 – это объем выборки);

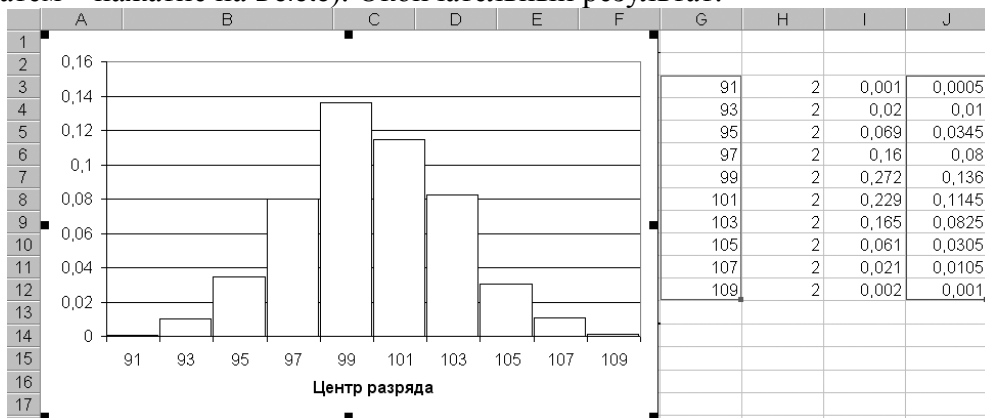
– вычислить высоту каждого прямоугольника гистограммы (в ячейку J3 вводится «=I3/H3», нажатие Enter, маркер автозаполнения переводится от ячейки J3 до ячейки J12).

Для визуализации результатов в случае использования **ортодоксального** интерфейса из меню Вставка следует выбрать пункт Диаграмма. На вкладке Стандартные выбирается Гистограмма. После перехода к следующему диалоговому окну (нажатие на Далее) на вкладке Диапазон данных в поле Диапазон указывается интервал ячеек, содержащий высоты прямоугольников (в данном примере – «=Лист1!\$J\$3:\$J\$12»). В этом же диалоговом окне на вкладке Ряд в поле Подписи оси X указывается интервал ячеек со значениями, соответствующими центру каждого разряда (в данном примере – «=Лист1!\$G\$3:\$G\$12»). В следующем диалоговом окне на вкладке Заголовки в поле Ось X (категорий) можно ввести строку «Центр разряда». В последнее диалоговое окно мастера диаграмм никакой информации вводить не нужно (выбирается Далее, затем – Готово); в результате будет построена гистограмма:

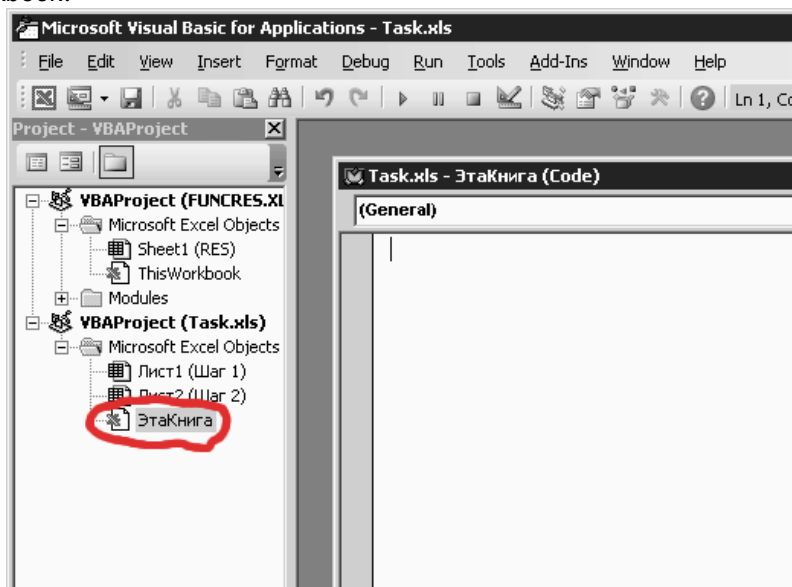


После этого можно изменить ширину каждого прямоугольника (двойной щелчок на любом из них, в диалоговом окне Формат ряда данных на вкладке Параметры установить зна-

чение в поле Ширина зазора равным 0 или 1) и удалить заголовок ряда (щелчок на заголовке «Ряд 1», затем – нажатие на Delete). Окончательный результат:



Для ленточного интерфейса операции, которые нужно выполнить посредством графического интерфейса пользователя для построения гистограммы, будут отличаться от версии к версии, и запоминать их бессмысленно. Рациональнее использовать переносимую между различными версиями ВАСИК-программу. Нажать Alt+F11, в древовидном списке проектов развернуть ветвь, соответствующую редактируемому файлу и сделать двойной щелчок на элементе ThisWorkbook:



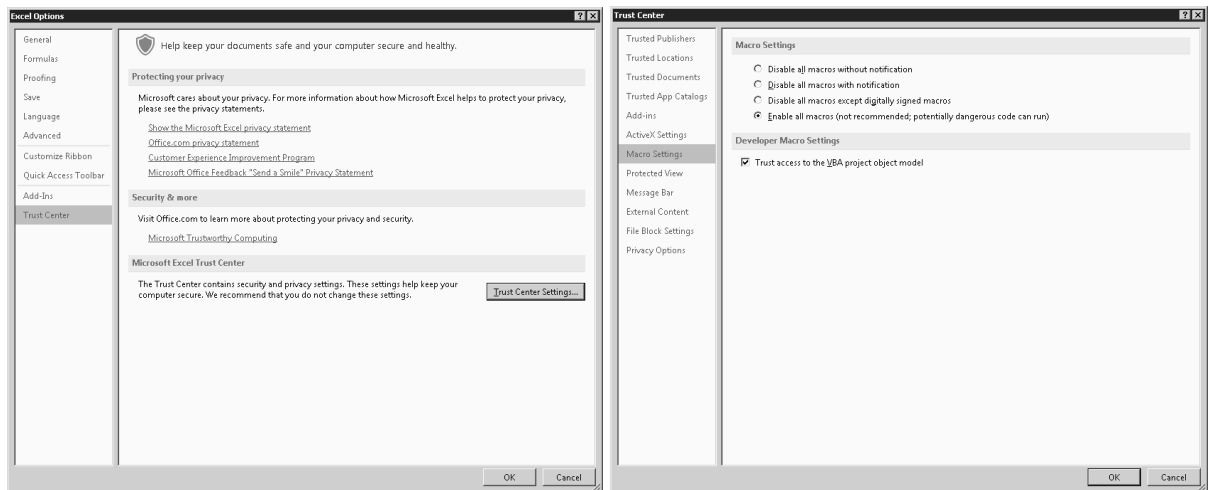
В открывшемся окне текстового редактора ввести:

```
Option Explicit
```

```
Sub PlotHistogram()
    Dim plotOut As Chart
    Set plotOut = ActiveSheet.Shapes.AddChart(xlColumnClustered).Chart
    plotOut.HasLegend = False
    plotOut.ChartGroups(1).GapWidth = 0
    With plotOut.SeriesCollection.NewSeries
        .XValues = Range("G3:G12")
        .Values = Range("J3:J12")
    End With
End Sub
```

После ввода текста следует установить каретку в любое место внутри функции PlotHistogram() и нажать F5. При этом MS Excel, скорее всего, выдаст сообщение о том, что выполнение ВАСИК-программ заблокировано. Для того, чтобы разблокировать их вы-

полнение, в ленточных червях^w версиях следует обратиться к результатам поиска <http://www.google.com/search?q=ms+excel+enable+macros>. Ключевые диалоговые окна:



Результат выполнения ВАСИК-программы в MS Excel 2013:

