

## § 4

## Инструменты анализа данных

## 4.1. Диаграммы

Как правило, электронные таблицы содержат большое количество числовых данных, которые требуется сравнивать, оценивать их изменение с течением времени, определять соотношение между ними т. д. Проводить подобный анализ большого количества числовых данных значительно легче, если изобразить их графически (визуализировать). Для графического представления числовых данных используются диаграммы.

**Диаграмма** — это графическое представление числовых данных, позволяющее быстро оценить соотношение нескольких величин.

Табличные процессоры позволяют строить диаграммы следующих типов:

- гистограмма;
- линейчатая диаграмма;
- круговая диаграмма;
- график;
- диаграмма с областями;
- поверхностная диаграмма;
- лепестковая диаграмма и др.

Чтобы в Microsoft Excel просмотреть все доступные типы диаграмм, изучите группу **Диаграммы** на вкладке **Вставка** (рис. 1.8).

Выясните, какие типы диаграмм можно создавать в табличном процессоре, имеющемся в вашем распоряжении.



Рис. 1.8. Типы диаграмм в Microsoft Excel

В диаграмме любого типа можно выделить следующие объекты (рис. 1.9):

- 1 — область диаграммы (в ней размещаются все объекты диаграммы);
- 2 — название диаграммы, чётко описывающее то, что представлено на диаграмме;

- 3 — область построения диаграммы (непосредственно в ней располагается сама диаграмма);
- 4 — ось значений (вертикальная, ось  $Y$ ). На ней находится шкала с определённым шагом, устанавливаемым автоматически, в зависимости от наименьшего и наибольшего значений данных, изображённых на диаграмме. Именно по этой шкале можно оценить данные, представленные на диаграмме;
- 5 — ряды данных — наборы числовых данных, некоторым образом связанных между собой и размещённых в электронной таблице в одной строке или столбце. На диаграмме ряд данных изображается геометрическими фигурами одного вида и цвета;
- 6 — ось категорий (горизонтальная, ось  $X$ ). На ней отображаются значения определённого свойства данных;
- 7 — легенда, поясняющая соответствие между названиями рядов и используемыми на диаграмме цветами. По умолчанию названия рядов являются названиями строк (или столбцов) диапазона данных, по которым построена диаграмма;
- 8 — названия осей.

Воспроизведите в табличном процессоре диаграмму, представленную на рисунке 1.9. С помощью контекстного меню исследуйте свойства каждого объекта этой диаграммы.



Рис. 1.9. Основные элементы диаграммы

На диаграммах разных типов числовые данные могут быть представлены точками, отрезками, прямоугольниками, секторами круга, прямоугольными параллелепипедами, цилиндрами, конусами и другими геометрическими фигурами. При этом размеры геометрических фигур или расстояния от них до осей пропорциональны числовым данным, которые они отображают.

Диаграммы, создаваемые в электронных таблицах, динамические — при редактировании данных в таблице размеры или количество фигур, обозначающих эти данные, автоматически изменяются.



Вспомните основные приёмы построения диаграмм, известные вам из курса информатики основной школы.

Рассмотрим самые распространённые типы диаграмм.

**Гистограммы** целесообразно создавать тогда, когда нужно сравнить значения нескольких наборов данных, графически изобразить отличия значений одних наборов данных от других, показать изменения данных с течением времени.

Различают следующие виды гистограмм:

- гистограмма с группировкой;
- гистограмма с накоплением;
- нормированная гистограмма с накоплением;
- объёмная гистограмма.

В гистограмме с группировкой прямоугольники, которые являются графическими изображениями числовых данных из разных наборов, располагаются рядом друг с другом (см. рис. 1.9). В гистограмме с накоплением прямоугольники, изображающие числовые данные, располагаются друг над другом (рис. 1.10). Это даёт возможность оценить суммарные данные и вклад каждой составляющей в общую сумму.

В нормированной гистограмме с накоплением вертикальная ось имеет шкалу в процентах. Это даёт возможность оценить долю (процентную часть) данных в общей сумме (рис. 1.11).

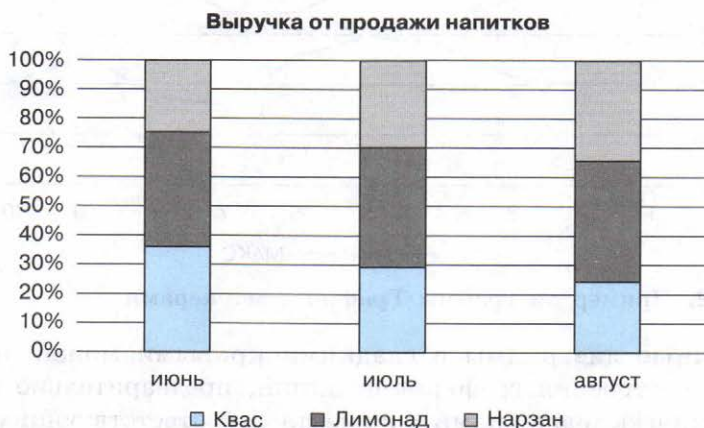


Подумайте, по какой из трёх диаграмм проще всего определить:

- 1) продажа каких напитков неуклонно возрастала;
- 2) продажа каких напитков принесла наибольшую прибыль в июле;
- 3) динамику изменений суммарной выручки от продажи всех трёх напитков;
- 4) вклад от продажи каждого напитка в общую выручку.



**Рис. 1.10.** Пример гистограммы с накоплением



**Рис. 1.11.** Нормированная гистограмма с накоплением

**Линейчатые** диаграммы аналогичны гистограммам и отличаются от них лишь горизонтальным расположением геометрических фигур.

К типу диаграмм **Круговая** относятся плоские и объёмные круговые диаграммы. Их целесообразно использовать тогда, когда нужно отобразить части одного целого, сравнить соотношение частей между собой и отношение частей к целому.

Круговые диаграммы позволяют отобразить только один ряд данных. Они теряют наглядность, если содержат много элементов данных. Несколько круговых диаграмм можно заменить, например, одной нормированной гистограммой с накоплением.



Подумайте, сколько разных круговых диаграмм можно построить по информации, содержащейся в диаграмме, представленной на рисунке 1.9.

Сколько круговых диаграмм потребуется для того, чтобы изобразить информацию, представленную на гистограмме с накоплением (см. рис. 1.10)?

Диаграммы типа **График** целесообразно использовать, если количество данных в наборе достаточно большое, если нужно отобразить динамику изменения данных во времени, сравнить изменения нескольких рядов данных (рис. 1.12).



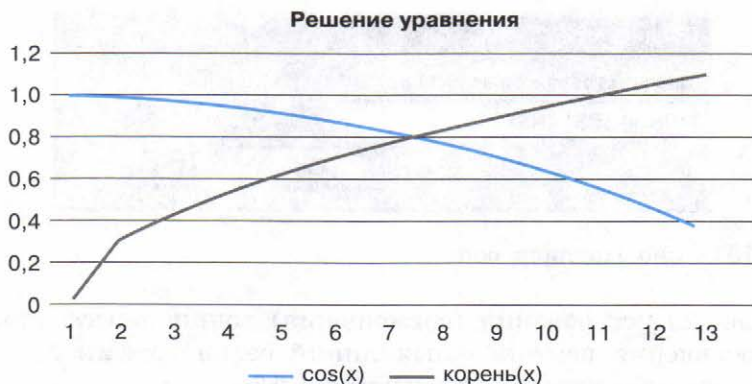
**Рис. 1.12.** Пример диаграммы **График с маркерами**

**Точечные** диаграммы с гладкими кривыми можно использовать для построения графиков функций, предварительно заполнив диапазон ячеек значениями аргумента и соответствующими значениями функции. Можно построить на одной диаграмме графики двух функций и использовать их для приближённого решения уравнения.

**Пример.** Найдём на отрезке  $[0; 1,2]$  корень уравнения  $\cos(x) = \sqrt{x}$ , построив в табличном процессоре графики функций, соответствующих левой и правой частям равенства. Для этого:

- используя стандартные функции **COS** и **КОРЕНЬ**, построим таблицу значений функций для  $x$ , изменяющегося с шагом  $0,1$ :

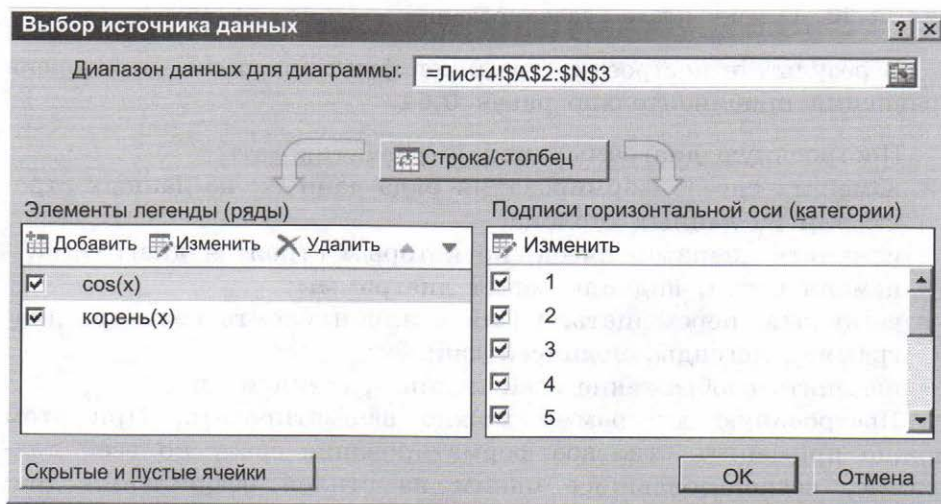
	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	x	0	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1	1,1	1,2
2	cos(x)	1	1	0,98	0,96	0,92	0,88	0,83	0,76	0,7	0,62	0,54	0,45	0,36
3	корень(x)	0	0,32	0,45	0,55	0,63	0,71	0,77	0,84	0,89	0,95	1	1,05	1,1



**Рис. 1.13.** Графики функций

- 2) по значениям диапазона B2:N3 построим графики функций  $\text{COS}(x)$  и  $\text{КОРЕНЬ}(x)$  (рис. 1.13);
- 3) заменим номера точек, проставленные по горизонтальной оси, на значения аргумента  $x$  рассматриваемых функций. Для этого вызовем контекстное меню горизонтальной оси и выберем пункт **Выбрать данные**. Появится окно **Выбор источника данных** (рис. 1.14).

В открывшемся окне нажмём на кнопку изменения подписей горизонтальной оси и выберем диапазон со значениями аргумента (рис. 1.15).



**Рис. 1.14.** Окно **Выбор источника данных**

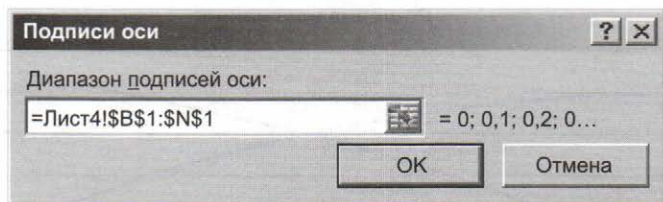


Рис. 1.15. Окно Подписи оси

После редактирования (совмещения) точки пересечения осей и добавления вертикальных линий сетки график приобретёт вид, представленный на рисунке 1.16.

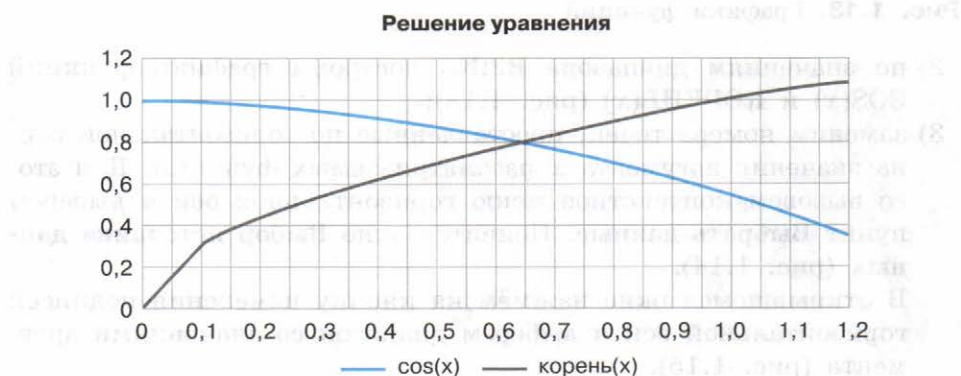


Рис. 1.16. График после редактирования

В результате построения графиков функций видно, что корень уравнения приблизительно равен 0,64.

Построенную диаграмму можно редактировать:

- изменять способ формирования ряда данных: из данных строки или из данных столбца;
- изменять диапазон ячеек, по которым строится диаграмма;
- изменять тип, вид или макет диаграммы;
- вставлять, перемещать, удалять или изменять название диаграммы, легенды, подписей данных;
- изменять отображение осей и линий сетки и др.

Построенную диаграмму можно форматировать. При этом можно применить стилевое форматирование сразу ко всей диаграмме, воспользовавшись одним из стилей оформления диаграмм. Кроме того, можно форматировать отдельные объекты диаграммы, которые предварительно надо выделить.

Некоторые объекты диаграммы, например ряд, состоят из нескольких частей. Чтобы выделить только одну часть, например отдельную точку ряда, необходимо сначала выделить весь объект, а затем выбрать нужную его часть.

## 4.2. Сортировка данных

Данные в электронной таблице можно сортировать, т. е. изменять порядок их расположения в строках или столбцах. В отсортированных данных легче найти необходимые значения, осуществить их анализ, выявить имеющиеся закономерности и др.

---

**Сортировка** — это упорядочение данных в таблице.

---

Сортировка данных может проводиться по возрастанию (от наименьшего к наибольшему) или по убыванию (от наибольшего к наименьшему). В Microsoft Excel соответствующие инструменты размещены на вкладке **Данные** в группе **Сортировка и фильтр**.

В Microsoft Excel сортировка данных по возрастанию заключается в следующем:

- символы упорядочиваются в порядке размещения их кодов в кодовой таблице Unicode;
- числа и даты упорядочиваются от наименьшего значения к наибольшему и располагаются перед текстовыми данными, причём сначала располагаются числа;
- текстовые данные сначала упорядочиваются по их первым символам; если первые символы в текстах совпали, то они упорядочиваются по вторым символам; тексты, в которых совпали первые два символа, упорядочиваются по их третьим символам и т. д.;
- логическое значение **ЛОЖЬ** размещается перед значением **ИСТИНА**;
- пустые ячейки всегда располагаются последними.

При сортировке данных по убыванию порядок расположения будет обратный, за исключением пустых ячеек, которые всегда располагаются последними.

Далее представлены введённые данные (диапазон A1:A15) и результаты их сортировки. Определите, где представлены данные, отсортированные по возрастанию, где — по убыванию. Объясните полученные результаты.





	А	В	С
1		-11	ИСТИНА
2	Петрозаводск	11	ЛОЖЬ
3		27.03.2017	Петрозаводск
4	11	25.02.2020	Петербург
5	11А	11А	Москва
6	Москва	Best	Best
7		Москва	11А
8	Петербург	Петербург	25.02.2020
9		Петрозаводск	27.03.2017
10	-11	ЛОЖЬ	11
11	ИСТИНА	ИСТИНА	-11
12	ЛОЖЬ		
13	25.02.2020		
14	27.03.2017		
15	Best		

Сортировка в выделенном связном диапазоне ячеек из нескольких столбцов выполняется по данным первого из выделенных столбцов. Иначе говоря, данные во всех других столбцах выделенного диапазона ячеек не сортируются, а расставляются по строкам электронной таблицы в соответствии с перестановкой данных первого столбца.

Если установить курсор в одну из ячеек связного диапазона и воспользоваться инструментом сортировки, то данные всего связного диапазона будут отсортированы в выбранном порядке по данным текущего столбца этого диапазона.

Если перед сортировкой данных выделить только часть связного диапазона, то при вызове инструмента сортировки откроется окно **Обнаружены данные вне указанного диапазона**. В этом окне при необходимости можно расширить выбранный диапазон ячеек до всего связного диапазона или подтвердить то, что сортировать данные следует только в пределах выделения.

В произвольном выделенном диапазоне ячеек отсортировать данные можно по значениям не одного, а нескольких столбцов (по нескольким уровням сортировки). Сортировка данных по значениям нескольких столбцов выполняется так:

- сначала данные сортируются по значениям первого из выбранных столбцов;
- сортировка данных по значениям каждого следующего из выбранных столбцов происходит лишь для тех строк элек-

тронной таблицы, в которых значения во всех предыдущих выбранных для сортировки столбцах совпадают.

Уровни такой сортировки и её порядок задаются в окне **Сортировка (Данные → Сортировка и фильтр → Сортировка)**. На рисунке 1.17 приведены введённые данные и результат их сортировки по значениям двух столбцов, а соответствующие настройки окна **Сортировка** приведены на рисунке 1.18.

В окне **Сортировка** можно:

- указать требуемые уровни сортировки;
- удалить любой уровень сортировки из списка для сортировки;
- переместить любой уровень сортировки выше или ниже в списке для сортировки;
- задать требуемые режимы сортировки.

В этом же окне можно установить параметры сортировки (кнопка **Параметры...**): сортировать по столбцам (по умолчанию) или по строкам; сортировать без учёта регистра (по умолчанию) или с учётом регистра.

	A	B	C	D	E	F	G
1	Фамилия	Имя	Класс		Фамилия	Имя	Класс
2	Васечкин	Дима	10		Васечкин	Юра	5
3	Иванов	Саша	9		Иванов	Петя	5
4	Петров	Коля	11		Сидоров	Денис	5
5	Сидоров	Денис	10		Петров	Руслан	6
6	Васечкин	Андрей	8		Васечкин	Андрей	8
7	Иванов	Андрей	11		Иванов	Саша	9
8	Петров	Руслан	6		Васечкин	Дима	10
9	Сидоров	Денис	5		Иванова	Ирина	10
10	Васечкин	Юра	5		Сидоров	Денис	10
11	Иванов	Петя	5		Иванов	Андрей	11
12	Иванова	Ирина	10		Петров	Коля	11

Рис. 1.17. Пример сортировки по значениям двух столбцов

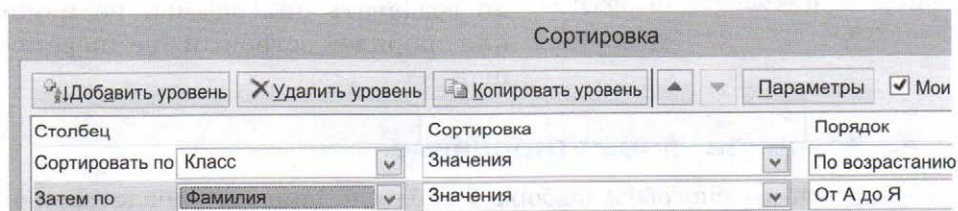


Рис. 1.18. Фрагмент окна **Сортировка**

### 4.3. Фильтрация данных

Ещё одним инструментом эффективного анализа данных в электронных таблицах является фильтрация.

**Фильтрация** — это выбор в электронной таблице данных, соответствующих определённым условиям.

Операция фильтрации, в отличие от операции сортировки, не меняет порядок строк. В отфильтрованном списке отображаются только строки, отвечающие условиям отбора данных, а остальные строки временно скрываются.

Если установить табличный курсор в произвольную ячейку заполненного данными диапазона (причём некоторые ячейки этого диапазона могут быть пустыми) и вызвать инструмент **Фильтр (Данные → Сортировка и фильтр → Фильтр)**, то около правой границы каждой ячейки первой строки этого диапазона появятся кнопки открытия списков, в которых находятся:

- команды сортировки данных по значениям данного столбца;
- команда **Фильтр по цвету**;
- команда **Снять фильтр с**;
- команда открытия меню команд для установки условий фильтрации:
  - числовые фильтры (если в столбце числовые данные);
  - текстовые фильтры (если в столбце текстовые данные);
  - фильтры по дате (если в столбце даты).

	А	В	С
1	Фамил ▾	Имя ▾	Класс ▾
4	Петров	Коля	11
7	Иванов	Андрей	11

**Рис. 1.19.** Пример фильтрации с помощью числового фильтра

Результаты выбора с помощью фильтрации из списка (см. рис. 1.17) всех учеников 11 класса представлены на рисунке 1.19.

Отфильтрованную таблицу можно редактировать, форматировать, выводить на печать. Для неё можно создавать диаграммы, не изменяя порядок строк и не перемещая их.

### 4.4. Условное форматирование

Ещё одним способом выбора в таблице данных, удовлетворяющих определённым условиям, является так называемое условное форматирование.

Условное форматирование автоматически изменяет формат ячейки на заданный, если для значения в данной ячейке выполняется определённое условие.

Для установки условного форматирования необходимо:

- 1) выделить нужный диапазон ячеек;
- 2) вызвать инструмент **Условное форматирование** (Главная → Стили → Условное форматирование);
- 3) выбрать в списке кнопки **Условное форматирование** необходимый тип правил (Правила выделения ячеек, Правила отбора первых и последних значений, Гистограммы, Цветовые шкалы, Наборы значков);
- 4) в списке правил указанного типа выбрать нужное правило;
- 5) в открывшемся окне задать условие и выбрать из списка форматов тот, который будет установлен;
- 6) завершить операцию щелчком по кнопке **ОК**.

Результат применения операции условного форматирования к диапазону C2:C12 списка учеников (см. рис. 1.17) представлен на рисунке 1.20 (тип правил — **Правила выделения ячеек**; условие выбора — **Больше 8**; формат — **Зелёная заливка и тёмно-зелёный текст** (рис. 1.21)).

	A	B	C
1	Фамилия	Имя	Класс
2	Васечкин	Дима	10
3	Иванов	Саша	9
4	Петров	Коля	11
5	Сидоров	Денис	10
6	Васечкин	Андрей	8
7	Иванов	Андрей	11
8	Петров	Руслан	6
9	Сидоров	Денис	5
10	Васечкин	Юра	5
11	Иванов	Петя	5
12	Иванова	Ирина	10

Рис. 1.20. Пример условного форматирования



**Рис. 1.21.** Пример условного форматирования

В отличие от фильтрации условное форматирование не скрывает ячейки со значениями, не удовлетворяющими определённому условию, а лишь выделяет заданным образом те ячейки, для значений которых условие выполняется.

## 4.5. Подбор параметра

Если известны параметры (исходные данные) и формула, по которой они должны быть преобразованы, то пользователь вводит их в ячейки электронной таблицы и получает некоторый результат. В электронных таблицах есть и обратная возможность: подобрать такие параметры, которые при подстановке их в известную формулу будут приводить к желаемому заранее известному результату. В Microsoft Excel это можно сделать с помощью одной из функций специального инструмента **Анализ «что-если»**.

Рассмотрим эту возможность на примере решения квадратного уравнения  $x^2 - 5x + 6 = 0$ .

Нам известна формула для вычислений ( $x^2 - 5x + 6$ ) и желаемый результат (0). Внесём эту информацию в ячейки таблицы:

	A	B
1	x	0
2	F(x)	=B1^2-5*B1+6

При подборе параметра в Microsoft Excel используется итерационный (циклический) процесс. Количество итераций и точность вычислений можно установить в окне **Параметры Excel (Файл → Параметры → Формулы → Параметры вычислений)**.

Вызовем окно подбора параметра (**Данные → Анализ «что-если» → Подбор параметра**) и заполним в нём поля ввода (рис. 1.22).

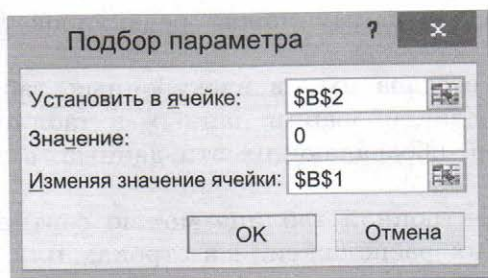


Рис. 1.22. Окно Подбор параметра

Ниже представлен результат подбора параметра:

	A	B
1	x	2
2	F(x)	0

Получен один из двух корней квадратного уравнения. Инструмент подбора параметра устроен так, что он возвращает одно решение, причём то, которое ближе к начальному значению. Напомним, что мы в качестве начального значения параметра приняли  $x = 0$ .

Самостоятельно поэкспериментируйте с другими начальными значениями параметра и найдите второй корень квадратного уравнения.

## САМОЕ ГЛАВНОЕ

Проводить анализ большого количества числовых данных значительно легче, если изобразить их графически (визуализировать). Для графического представления числовых данных используются диаграммы.

Табличные процессоры позволяют строить гистограммы, линейчатые диаграммы, круговые диаграммы, графики, диаграммы с областями, поверхностные диаграммы, лепестковые диаграммы и др.

В диаграмме любого типа можно выделить следующие объекты: область диаграммы, название диаграммы, область построения диаграммы, ось категорий, ось значений, названия осей, ряды данных, легенду.

Построенную диаграмму можно редактировать и форматировать.

Диаграммы, создаваемые в электронных таблицах, динамические — при редактировании данных в таблице размеры или количество фигур, обозначающих эти данные, автоматически изменяются.

Данные в электронной таблице можно сортировать, т. е. изменять порядок их расположения в строках или столбцах. В отсортированных данных легче найти необходимые значения, осуществить их анализ, выявить имеющиеся закономерности и др.

Ещё одним инструментом эффективного анализа данных в электронных таблицах является фильтрация, позволяющая из многочисленных данных отобрать только те, которые соответствуют заданным условиям. Операция фильтрации, в отличие от операции сортировки, не меняет порядок строк. В отфильтрованном списке отображаются только строки, отвечающие условиям отбора данных, а остальные строки временно скрываются.

Условное форматирование автоматически изменяет формат ячейки на заданный, если для значения в данной ячейке выполняется определённое условие. В отличие от фильтрации условное форматирование не скрывает ячейки со значениями, не удовлетворяющими определённому условию, а лишь выделяет заданным образом те ячейки, для значений которых условие выполняется.

Для анализа данных может быть полезна имеющаяся в электронных таблицах возможность подобрать такие параметры, которые при подстановке их в известную формулу будут приводить к желаемому заранее известному результату.



### Вопросы и задания

1. Для чего предназначены диаграммы? Какой анализ числовых данных можно выполнить с их помощью?
2. Назовите основные типы диаграмм, которые могут быть построены в электронных таблицах.
3. Назовите основные объекты диаграмм и их свойства.
4. Опишите виды гистограмм. Для чего предназначен каждый из этих видов?
5. Для чего предназначены круговые диаграммы?
6. Для чего предназначены графики?
7. Перечислите основные операции редактирования диаграмм.

8. Перечислите основные операции форматирования диаграмм.
9. По представленной ниже информации составьте таблицу распределения суши и воды на поверхности земного шара.

Площадь поверхности Земли — 510 072 тыс. кв. км, в том числе площадь суши — 148 940 тыс. кв. км (29,2%), площадь водной поверхности — 361 132 тыс. кв. км (70,8%). При этом суша большей частью лежит в Северном полушарии, а водная поверхность — наоборот, в Южном. В Северном полушарии водная поверхность занимает 61%, а поверхность суши — 39%; для Южного полушария эти соотношения таковы: 81% воды и 19% суши.

По данным полученной таблицы постройте следующие диаграммы:

- 1) гистограмму с группировкой;
  - 2) гистограмму с накоплением;
  - 3) нормированную гистограмму с накоплением;
  - 4) объёмную гистограмму с накоплением;
  - 5) круговую;
  - 6) линейчатую с группировкой.
10. Дан фрагмент электронной таблицы:

	A	B	C
1	2	1	
2	=C1-B1*4	=(B1+C1)/A1	=C1-4



Какое целое число должно быть записано в ячейке C1, чтобы после выполнения вычислений диаграмма, построенная по значениям диапазона ячеек A2:C2, соответствовала рисунку?

11. Можно ли построить круговые диаграммы для данных, содержащих отрицательные числа? Подкрепите свой ответ примерами.
12. В табличном процессоре постройте график функции  $y = \frac{1}{x^2 + 1}$  на отрезке  $[-2; 2]$  с шагом 0,2.
13. В табличном процессоре на одной диаграмме постройте графики трёх функций  $y = \sin x$ ,  $y = 2\sin x$ ,  $y = \sin 2x$  на отрезке  $[-2\pi; 2\pi]$  с шагом  $\frac{\pi}{8}$ .





14. На интервале  $[-1; 1]$  с шагом  $0,1$  решите графически систему уравнений:

$$\begin{cases} y = 2x + 7; \\ y = 2x^2 + 9. \end{cases}$$

15. Что называют сортировкой? Для чего она используется?
16. Сформулируйте правила, определяющие порядок сортировки данных разных типов по убыванию.
17. Какой порядок сортировки можно задать для числовых данных? Для текстовых данных?
18. Что называют фильтрацией? Для чего она используется?
19. Сравните операции сортировки и фильтрации. Что у них общего? Чем они различаются?
20. Используя возможность подбора параметра, решите квадратное уравнение  $x^2 + 2x - 15 = 0$ .



Дополнительные материалы к главе смотрите в авторской мастерской.