

# **СВОДКА, КЛАССИФИКАЦИИ И ГРУППИРОВКА СТАТИСТИЧЕСКИХ ДАННЫХ**

# Классификации в статистической практике

- Признак, имеющий качественно отличную характеристику и служащий основой для разграничения элементов в совокупности на группы называется *группировочным признаком*
- Группировочным признаком может быть любой признак (атрибутивный или количественный). В зависимости от сложности явления и цели исследования группировочных признаков может быть несколько

- Если разграничение элементов совокупности на группы осуществляется по *атрибутивным (качественным) признакам*, то такой вид группировки называют *классификацией или номенклатурой*.
- *Классификации или номенклатуры* разрабатываются международными и национальными статистическими органами и рекомендуются как статистический стандарт.

- *Классификатор*, где каждому значению группировочного признака присвоен код, то есть условное цифровое обозначение.

Российская Федерация использует такие основные виды классификаторов.

■ *Общероссийский классификатор видов экономической деятельности, продукции и услуг (ОКДП)*, который включает 55 тысяч видов продукции и услуг по 6 уровням классификации (раздел, подраздел, группа, подгруппа, класс, подкласс);

■ *Общероссийский классификатор продукции (ОКП)*, предназначенный для автоматизированной обработки информации на национальном, региональном и отраслевом уровнях; содержит около 50 тыс. видов продукции по такой иерархии; класс, подкласс, группа, подгруппа, вид.

■ *Общероссийский классификатор форм собственности (ОКФС)*, дающий возможность установить распределение предприятий по формам собственности (российская собственность, иностранная собственность, совместная российская и иностранная собственность);

■ *Общероссийский классификатор организационно-правовых форм (ОКОПФ)*, который включает разделы: юридические лица для коммерческих и некоммерческих организаций; организации без права юридического лица; индивидуальные лица; индивидуальные предприниматели.

# Группировка статистических данных

- Если разграничение элементов совокупности на группы осуществляется по *количественным признакам*, то такой вид группировки называют *статистической группировкой*

# Типы группировок

В зависимости от цели и заданий исследования различают такие виды группировки:

- типологические;
- структурные;
- аналитические.

- *Типологическими* группировками, называют такие группировки, которые приводят к выделению в составе массовых явлений их социально-экономических типов (то есть однородных частей по качеству и условиям развития, в которых действуют одни и те же закономерности).

- *Структурные* группировки характеризуют состав однородной совокупности по любым признакам.

С помощью таких группировок анализируют структуру совокупности и структурные сдвиги в развитии социально-экономических явлений и процессов. К ним принадлежат группировки населения по полу, возрасту, а на производстве — группировка рабочих по производственному стажу, уровню квалификации и др.

- Группировки, которые направлены на выявление связи между отдельными признаками изучаемого явления, называются *аналитическими*.

Примерами таких группировок могут быть группировки, в которых изучаются взаимосвязи между себестоимостью и ее факторами, производительностью труда и ее факторами и т. п.

По количеству группировочных признаков, положенных в основу группировки, различают:

- простые группировки;
- комбинированные группировки.

- *простой* называют группировку, которая проводится по одному признаку.
- *Комбинированной* называют группировку при сочетании двух и больше группировочных признаков.

При использовании метода группировки решают такие задачи:

- а) выбор группировочного признака;
- б) определение количества групп и величины интервала;
- в) установление перечня показателей, которыми должны характеризоваться выделенные группы относительно конкретной группировки;
- г) составление макетов таблиц, где будут представлены результаты группировки;
- д) вычисление абсолютных, относительных и средних показателей;
- ж) табличное и графическое оформление результатов группировки.

- Принципиальное значение при построении группировок имеет выбор *группировочного признака*, на основе которого выделяют разные типы, группы и подгруппы.
- В качестве группировочного принимают наиболее существенные признаки.
- Группировочным признаком может быть атрибутивный (качественный) или количественный признак.

- Если группировка осуществляется по атрибутивному признаку, то выделяют столько групп, сколько имеется наименований признака.

# Пример группировки по атрибутивному признаку

- Рассмотрим распределение городских земельных площадей по виду застройки (атрибутивным признакам)

Вид городской застройки	Площадь, га	% от итога
Промышленная площадь	670	22,1
Жилищная	1920	63,5
Общественная	315	10,4
Другая	120	4,0
Всего	3025	100

- Такой вид группировки относится к структурной, которая в этом примере характеризует масштабы и весомость отдельных видов застройки на территории города.

- При составлении группировок на основе *количественных признаков* (дискретных или непрерывных) определяют количество групп и интервалы группировки.
- В массовой совокупности оптимальное количество групп с равными интервалами приблизительно можно определить по формуле американского ученого Стерджеса:

$$m=1+3,322\lg n,$$

где  $m$  - количество групп;  $n$  - объем совокупности.

- Формула Стерджеса может быть использована при условии, что распределение единиц совокупности за данным признаком приближается к нормальному закону распределения.
- *Интервалы*, то есть промежуток между значениями признака в группе единиц, бывают равные, неравные, открытые и закрытые. Выбор вида интервала зависит от характера распределения единиц исследуемой совокупности.

- *Равные интервалы* используют в случаях, когда значения варьируемого признака  $x$  изменяются плавно, постепенно, равномерно. Ширина интервала  $h$  определяется по формуле:

$$h = \frac{x_{\max} - x_{\min}}{m}$$

где  $x_{\max}$ ,  $x_{\min}$  - наибольшее и наименьшее значение признака в совокупности.

# Пример

- Прибыльность активов коммерческих банков колеблется в пределах от 5 к 45%.
- При принятии количества групп  $m=4$  ширина интервала - 10. Тогда границы интервалов составляют соответственно:  
5-15, 15-25, 25-35, 35-45.

- Поскольку границы интервалов совпадают (15 - в первой и второй группе, 25 - во второй и третьей, 35 - в третьей и четвертой), то для исключения неопределенности отнесения граничных значений признака к той или другой группе используют следующее правило:

# Правило

- левое одинаковое число не включает в себя значения признака, правое - включает.

- Тогда, например, число 15 должно быть отнесено ко второй группе, а не к первой. Все интервалы в данном примере называются *закрытыми*.

- Приведенное распределение прибыльности активов банков может быть представлено в другом виде:  
до 15. 15-25, 25-35, 35 и более.
- Первый и последний интервалы имеют лишь одну границу и называются *открытыми*

- Неравные интервалы используются в случае, когда диапазон значений признака слишком широкий и распределение совокупности по этому признаку неравномерно.

## Пример

- Рассмотрим распределение поселков городского типа по количеству жителей (тыс. чел.):  
до 3; 3-4,9; 5-9,9; 10-49,9.

# Ряды распределения

- Особый вид группировок в статистике представляют ряды распределения, которые являются самым простым способом обобщения статистических данных.
- *Рядом распределения* называют группировку, характеризующую состав (структуру) явления в данный период времени.

В зависимости от того, какой признак (качественный или количественный) положен в основу группировки различают:

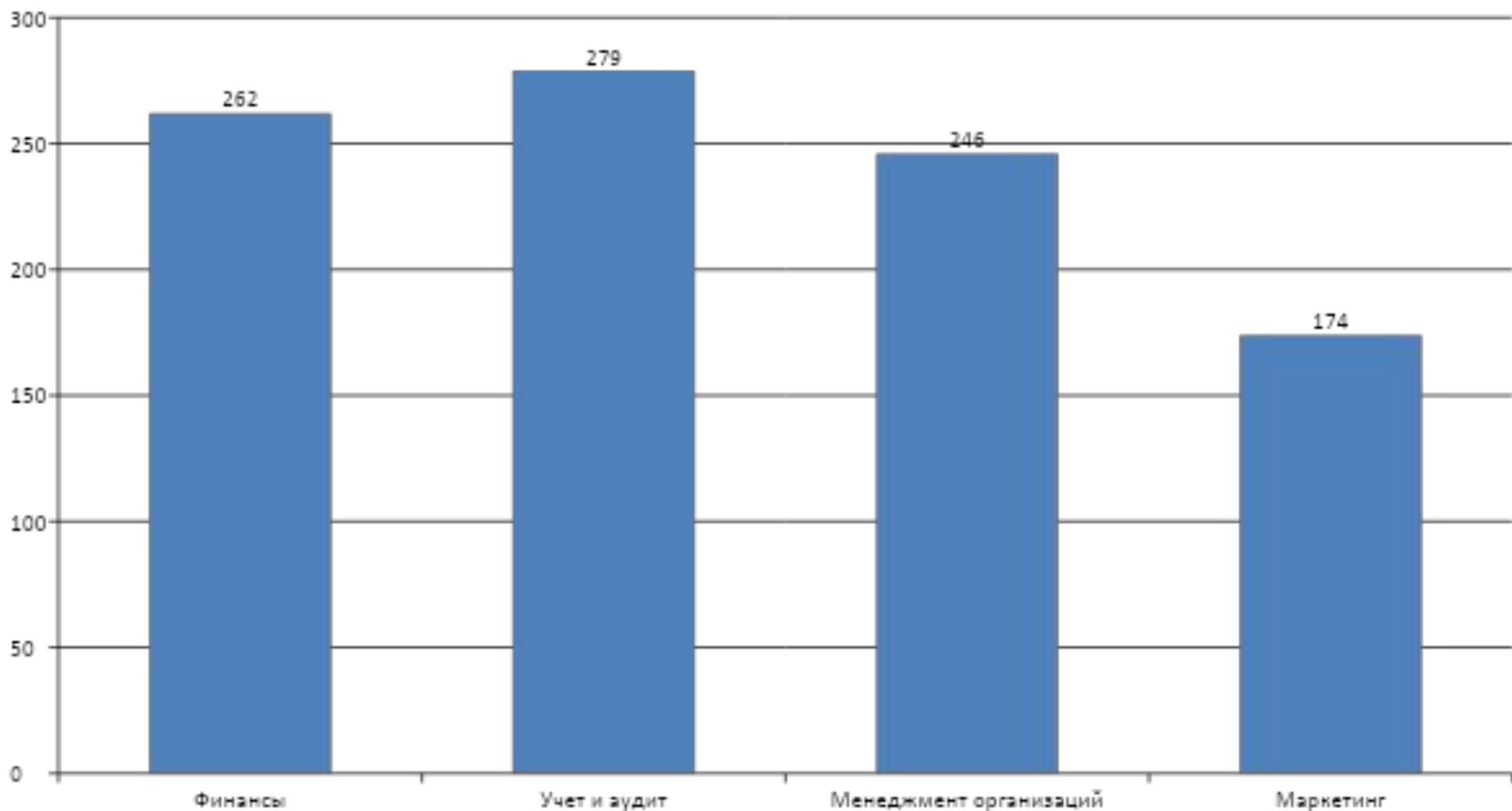
- атрибутивные (качественными) ряды распределения
- вариационные (количественными) ряды распределения .

# Пример

- Рассмотрим распределения студентов университета по экономическими специальностями

<sup><math>\varphi_i</math></sup> Наименование специальности	Число студентов, чел., $f_i$	% от общего Количества,
Финансы	262	27,3
Учет и аудит	279	29
Менеджмент организаций	246	25,6
Управление персоналом	174	18,1
Всего	961	100

# Атрибутивный ряд распределения



*Элементами* (характеристиками) этого ряда распределения являются:

*значения атрибутивного признака* (первая графа таблицы);

*частоты  $f_i$*  - численные характеристики отдельных значений признака, то есть числа, которые показывают, как часто встречается то или другое значение признака в ряду (вторая графа);

- *частоты*  $\varphi_i$  - это частоты, выраженные в относительных величинах (коэффициентах или процентах), что приведено в третьей графе таблицы.

*Вариационные ряды* распределения бывают

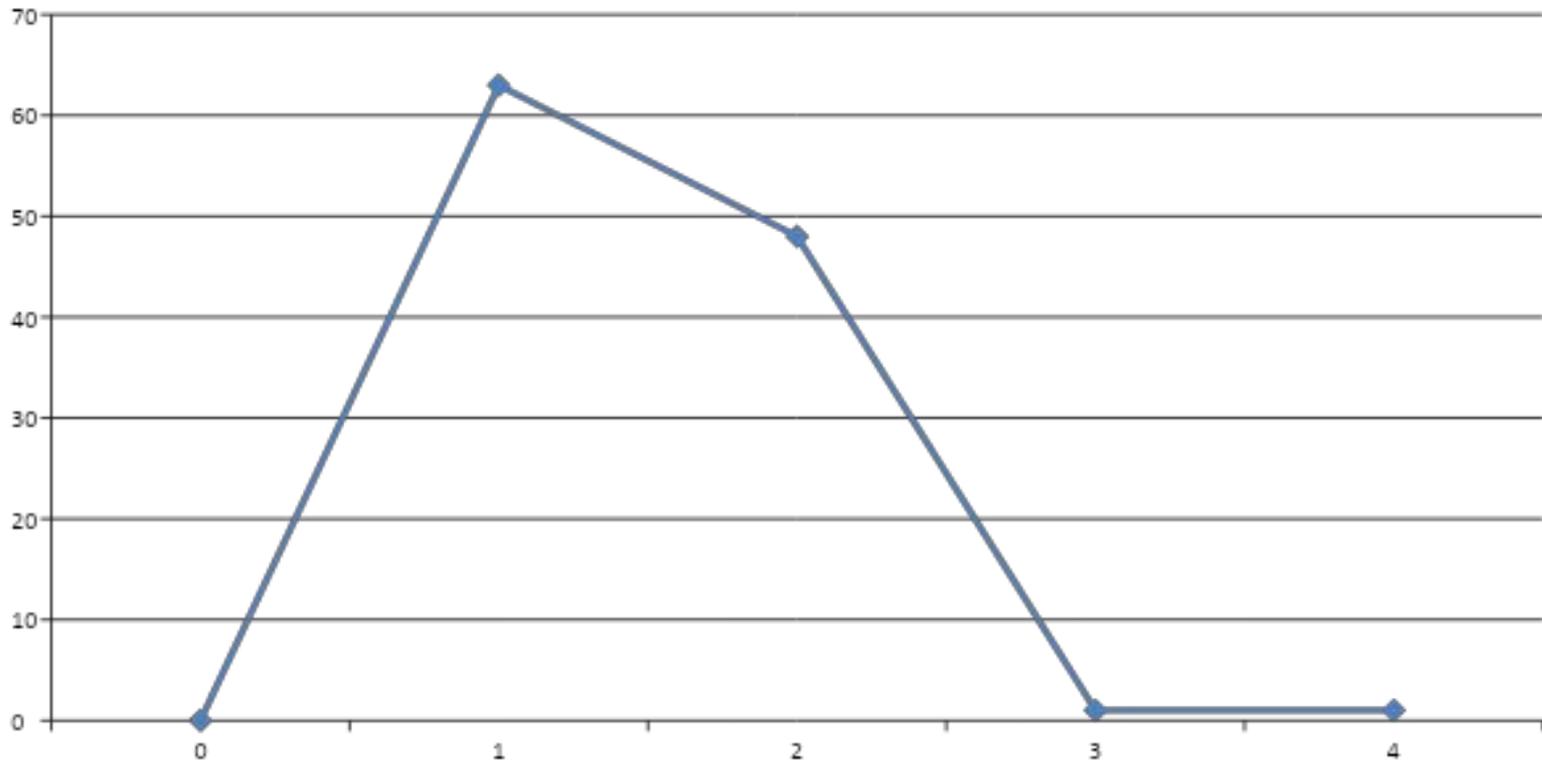
- дискретными
- интервальными.
  
- *Дискретные* вариационные ряды основаны на величинах признаков, которые имеют целые значения (например, тарифный разряд рабочих, количество марок автомобилей и т. п.).

# Пример

Рассмотрим дискретный ряд  
распределения семей в населенном  
пункте по количеству детей

Количество детей, $x$ ;	Количество семей, $f_i$	В % от общего количества семей,
1	63	56,2
2	48	42,8
3	1	1,0
Всего	112	100

# Дискретный ряд распределения полигон распределения



К основным элементам вариационных рядов распределения относятся:

*варианты*  $x$ . — численные значения количественного признака в группировке (первая графа таблиц), которые могут быть положительными, отрицательными, абсолютными, относительными;

*частоты*  $f_i$  — значение отдельных вариантов (вторая графа);

*частости*  $\varphi_i$  (третья графа).

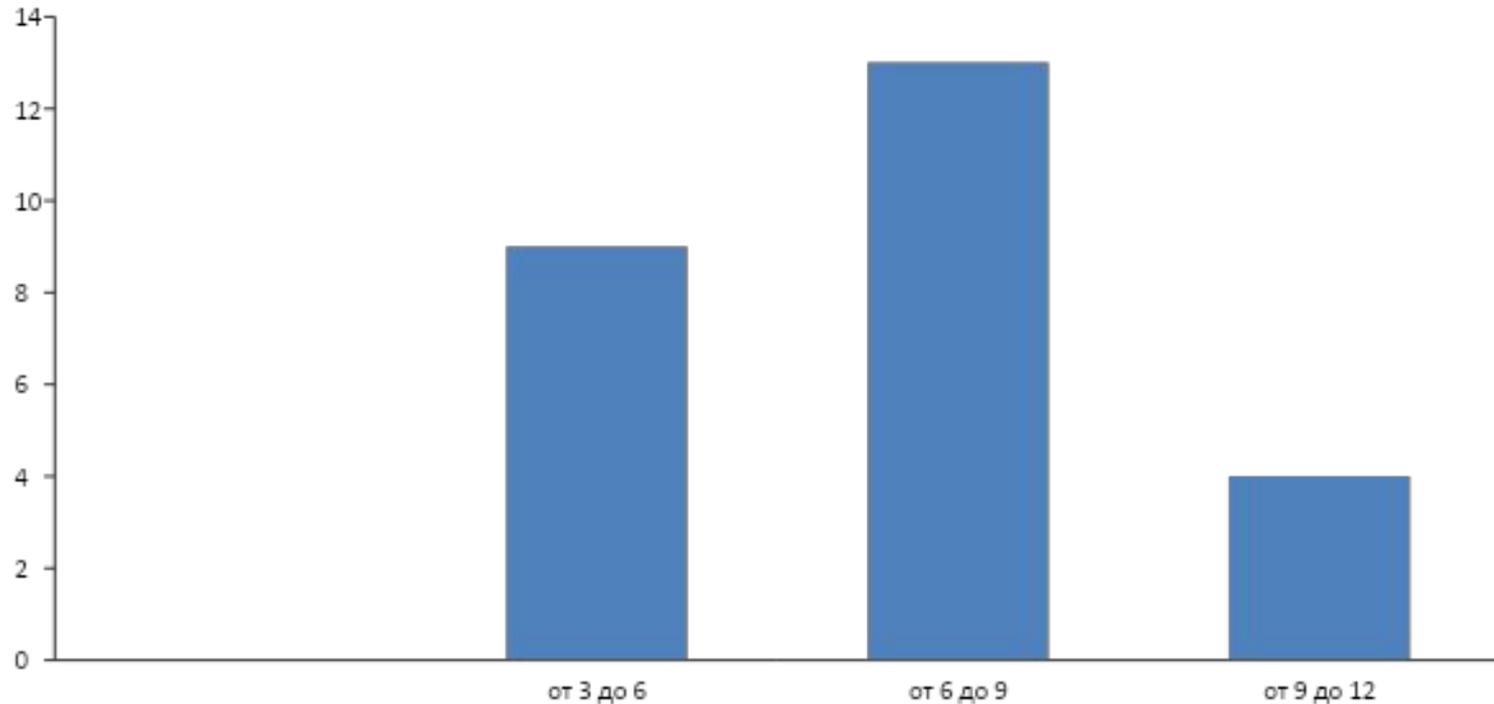
- В *интервальных* вариационных рядах группировочный признак может принимать любое значение (целое, дробное) в пределах каждого интервала (например, распределение заработной платы работающих в организации, распределение основных фондов предприятия и т. п.).

# Пример

- Рассмотрим ряд распределения коммерческих банков в населенном пункте по величине прибыли

Размер прибыли, млн д. е., $x_i$	Число банков	В % к общей прибыли, (/),
3-6	9	34,6
6-9	13	50.0
9-12	4	15,4
Всего	26	100

# Интервальный вариационный ряд распределения. Гистограмма



Кроме отмеченных элементов  
вариационных рядов к их  
характеристикам могут быть отнесены  
также такие показатели-

$\frac{f_i}{h}$

- **плотность частоты**

где  $h$  - ширина интервала;

$$F_i = \sum_{i=1}^n \varphi_i$$

накопленная частота .

Такие показатели используются при  
анализе построенных рядов  
распределения.

# Пример

Рассмотрим использование характеристик вариационного интервального ряда на примере распределения сельских хозяйств области.

Количество голов, $x_t$	Количество хозяйств $f_i, \%$	Плотность распределения я * 100%	Накопленая частотность $F = \sum_{i=1}^n \varphi_i$
До 300	4,4	1,47	4,4
300-599	11,3	3,77	15,7
600-999	24,5	6,12	40,2
1000-3000	53,7	2,69	93,9
Более 3000	6,1	0,303	100
Всего	100	X	

# Статистические таблицы

- Таблица по своему логическому содержанию рассматривается как «статистическое предложение», которое имеет свое подлежащее и сказуемое.
- *Подлежащее таблицы* характеризует объект исследования.
- *Сказуемое* — это система показателей, которые отображают подлежащее как объект.

В зависимости от структуры подлежащего статистические таблицы разделяют на

- простые,
- групповые,
- комбинированные.
- Подлежащее *простой таблицы* представляет собой перечень единиц явления или процесса. В сказуемом простой таблицы отсутствует группировка статистических данных. Такие таблицы могут быть перечневыми, хронологическими, территориальными.

# Пример простой перечневой таблицы

## Наличие строительных машин в строительных управлениях региона по состоянию на в 01.01.2013 года

Вид машин	Количество машин данного вида, тыс. шт.
Экскаваторы	32,6
Скреперы	8,7
Бульдозеры	31,6
Краны передвижные	40,3
Всего	113,2

- Подлежащим таблицы является вид машин, сказуемым - количество разных их видов.

- В *групповых статистических* таблицах подлежащее группируется по одному признаку, а у *комбинированных* - по двум и больше признакам

# Пример групповой таблицы

## Группировка магазинов по уровню производительности труда работников

Уровень производительности труда магазинов, тыс. д. е.	Количество магазинов	Фондоотдача на 1 д. е. активной части основных фондов, д. е.	Рентабельность активной части основных фондов %
А	1	2	3
До 60	4	40,42	2,3
60-70	4	43,1	2,8
70-80	7	75,8	4,7
80-90	7	65,9	4,0
90-100	3	93,1	5,1
Более 100	7	109,3	6,4
Всего	32	X	X
В среднем	X	75	4,4

- В приведенной групповой таблице подлежащим являются магазины города, которые разделены на группы по уровню производительности труда, сказуемым - показатели этих организаций (количество магазинов, фондоотдача, рентабельность).

# Пример комбинированной таблицы

Группировка продовольственных магазинов города по части площади торгового зала и длительности рабочего дня за отчетный период

Группы и подгруппы магазинов по части площади торгового зала (%) и длительности рабочего дня (в часах)	Количество магазинов	Фондоотдача на 1 д.е. активной части Основных фондов, д. е.	Рентабельность Активной части Основных фондов %
А	1	2	3
К 35%	13	48,5	3,10
В том числе: 8-10 часов	4	41,2	2,20
более 10 часов	9	57,5	4,02
35-45%	21	69,8	5,20
В том числе: 8-10 часов	6	54,6	3,08
более 10 часов	15	77,4	7,10
45-55%	18	90,6	6,40
В том числе: 8-10 часов	5	68,9	4,17
более 10 часов	13	108,7	7,98
Всего	52	X	X
В среднем	X	73,5	4,70

- **Подлежащим** данной таблицы являются магазины, распределенные на группы и подгруппы по части площади торгового зала и длительности рабочего дня; в сказуемом приведены показатели, которые наиболее полно характеризуют эффективность работы магазинов.

- *Макет таблицы* - это комбинация горизонтальных строк и вертикальных граф (столбцов), на пересечении которых образуются клетки для записи соответствующей информации. Левые боковые и верхние клетки предназначены для словесных заглавий подлежащего и системы показателей **сказуемого**. Другие клетки таблицы предусмотрены для численных статистических данных. Основное содержание таблицы указывается в ее названии.

# Технические правила составления таблиц

1. Название таблицы, заглавия строк и граф должны быть четкими, лаконичными, без сокращений, лишней и второстепенной информации.
2. В названии таблицы указывается ее порядковый номер, объект исследования, его временной и географический признак. Если названия отдельных граф или строк повторяются, то их целесообразно объединить общим заглавием.

3. В верхних и боковых заглавиях указываются единицы измерения с использованием общепринятых сокращений (руб., т, м и т. п.). Если единица измерения общая для всех данных таблицы, ее указывают в названии таблицы.

4. Для сложных в построении таблиц графы целесообразно нумеровать: графу с названием подлежащего помечают литерой **алфавита**, а другие графы — цифрами.

5. Обобщенная информация граф таблицы содержится в итоговой строке с отметкой «Итого» (промежуточный итог) «Всего» (окончательный итог), «В среднем».
6. Значение показателей в клетках таблицы следует округлять в границах одной строки или графы с одинаковой степенью точности (до целых; 0,1; 0,01 и т. д.).

7. Если нет сведений о любом показателе таблицы, ставится три точки (...). Отсутствие любого признака в таблице отражается тире. В тех случаях, когда клетка таблицы не подлежит заполнению по логическому смыслу содержания, ставится знак X.

8. К таблице в случае необходимости добавляются примечания, в которых указывают источники данных, дают подробное толкование содержания отдельных показателей и другие объяснения.