

# Статистическая сводка и группировка данных

**Сводка** – это научная обработка первичных данных с целью получения обобщенных характеристик изучаемого социально-экономического явления по ряду существенных для него признаков.

## Статистическая сводка

По глубине и точности обработки информации:

**Простая** – подсчёт общих итогов по совокупности

**Сложная** – комплекс действий по разделению совокупности на группы, составлению системы показателей, подсчету числа единиц и итогов по группам, оформлению таблиц.

По форме обработки информации:

**Централизованная** – все данные поступают и обрабатываются в одной организации

**Децентрализованная** – отчеты предприятий поступают и сводятся в стат-органах субъектов РФ, затем итоги поступают в Федеральную службу и сводятся в целом по стране

**Группировка** – разбиение общей совокупности единиц объекта наблюдения по одному или нескольким существенным признакам на более однородные группы

# Виды группировок:

- Типологическая
- Структурная
- Аналитическая(факторная)

**Типологическая – разбиение  
разнородной статистической  
совокупности на отдельные качественно  
однородные группы и выявление на их  
основе устойчивых социально-  
экономических типов явлений**

**Структурная** – группировка, в которой происходит разделение однородной совокупности на группы, характеризующие ее структуру по какому-то варьирующему признаку. (группировка населения по уровню дохода).

**Аналитическая (факторная)** – позволяет выявить взаимосвязи между изучаемыми явлениями и их признаками. (группировка банков по сумме уставного капитала, величине активов и балансовой прибыли)

# По способу построения:

- 1) **Простая** группировка – группы образованы по одному признаку
- 2) **Комбинационная** – в которой образование группы производится по двум и более признакам, взятым в сочетании (комбинации) (сначала по качественному , а затем по количественному признаку)



# Принципы построения статистических группировок:

## 1. Определение цели группировки и группировочного признака

### Принципы отбора группировочных признаков:

- а) в качестве *основания* группировки следует брать типичные, существенные признаки изучаемого явления в соответствии с целями статистической работы;
- б) при выборе группировочных признаков должны быть приняты во внимание конкретные условия места и времени;
- в) при изучении сложных явлений группировку следует проводить *по ряду* существенных, характерных (комбинированных) признаков

## 2. Определение числа групп

### Число групп зависит от:

- вида признака
- степени его вариации
- объема изучаемой совокупности

## 3. Определение границ группы

## 4. Выбор признаков, которые будут характеризовать выделенные группы

### *В) Отбор атрибутивных признаков*

Отбираются признаки, которые характеризуют свойства данного явления без количественного выражения; при этом статистическая совокупность разделяется на столько групп, сколько разновидностей имеет признак; атрибутивные признаки могут отбираться и по альтернативному варианту

### *Г) Отбор количественных признаков*

Осуществляется по величине изучаемой совокупности, что даёт возможность разделить её на группы в соответствии с индивидуальными значениями или уровнями группировочного признака

**Группировочный признак** (*основание группировки*) – количественный или качественный признак, по которому производится разбиение единиц совокупности на группы.

Для **качественных** признаков число групп равно числу вариантов признака (пол сотрудников, цвет автомобилей жителей города)

Для **количественных** признаков при определении числа групп учитывают *размах вариации* группировочного признака (R):

$$R = X_{max} - X_{min}$$

Число групп можно определить математически или экспертным

путём:

Математический способ (формула Стерджесса):

$$m=1+3,322*\lg(n)=\log_2(n)+1$$

Применима при  $n > 30$  (при большом объёме совокупности)

**Интервал** – это значения признака, лежащие в определенных границах.

**Нижняя граница** – наименьшее, **верхняя граница** – наибольшее значение признака в интервале.

**Ширина интервала** – разность между его верхней и нижней границей.

Интервалы бывают **равные** и **неравные**, **закрытые** и **открытые**.

Ширина равного интервала:

$$i=h=R/m$$

Правила округления ширины интервала:

- 1) Если  $h$  имеет один знак до запятой (0,67; 3,82), полученные значения округляют до десятых -0,7 и 3,8.
- 2) Если  $h$  имеет две значащие цифры до запятой (14, 875), это значение округляют до целого числа - 15
- 3) Если  $h$  трех-, четыре- и более значимое число – его округляют до ближайшего числа кратного 100 или 50

**Вторичная группировка** – операция по образованию новых групп на основе ранее осуществленной группировки

**Во вторичной группировке применяются два способа образования новых групп:**

- Первый способ состоит в укрупнении первоначальных интервалов
- Второй способ называется *методом долевого перегруппировки* и состоит в том, что за каждой группой закрепляется определенная доля единиц совокупности.

**Ряд распределения – простейшая группировка, в которой каждая выделенная группа характеризуется только количеством входящих в нее единиц совокупностей**



Ряды распределения		
Атрибутивн ые	Вариационные	
	Интервальные	Дискретн ые

**Ранжированный ряд** - это распределение отдельных единиц совокупности в порядке возрастания или убывания исследуемого признака.

**Варианты** – отдельные значения признака  $X_i$

**Частоты** – количества отдельных вариантов или *численности* каждой группы вариационного ряда  $f_i$

**Частоты** – частоты, выраженные в долях единицы или процентах к итогу  $w_i$

**Плотность распределения** – частота рассчитанная на единицу ширины интервала

**Дискретный ряд** - это такой вариационный ряд, в основу построения которого положены признаки с прерывным изменением (дискретные признаки).

**Интервальный вариационный ряд распределения** – ряд распределения, в котором группировочный признак задан интервалами значений

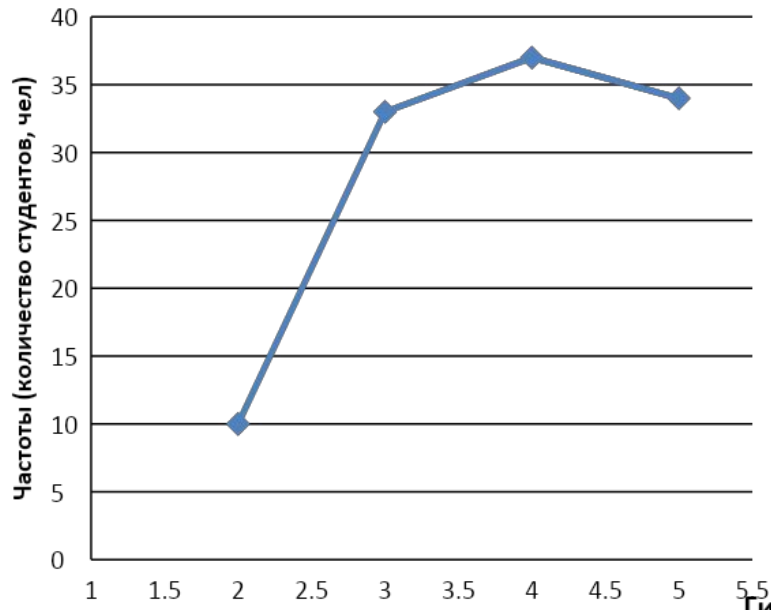
# Графическое изображение рядов распределения

Наглядно ряды распределения представляются при помощи графических изображений.

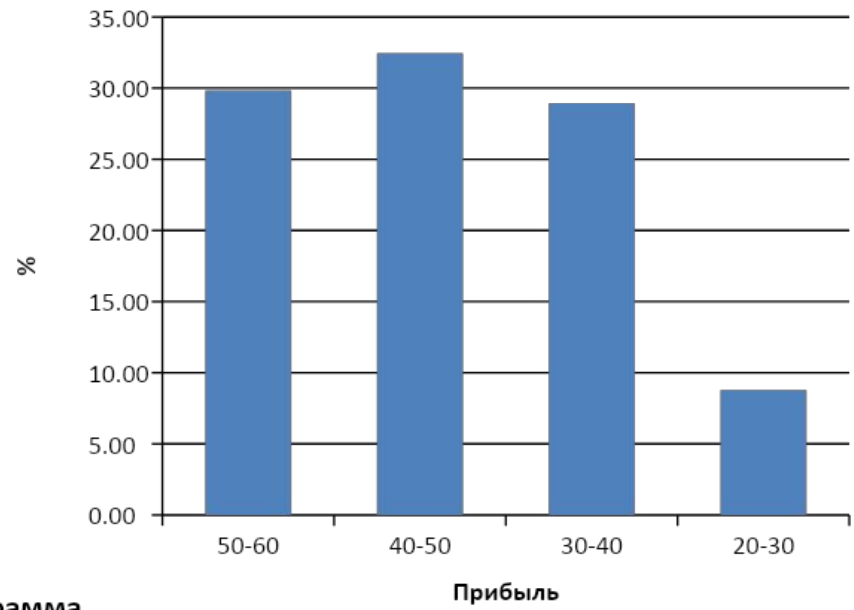
**Ряды распределения изображаются в виде:**

- Полигона (для дискретных вариационных рядов)
- Гистограммы
- Кумуляты
- Огивы

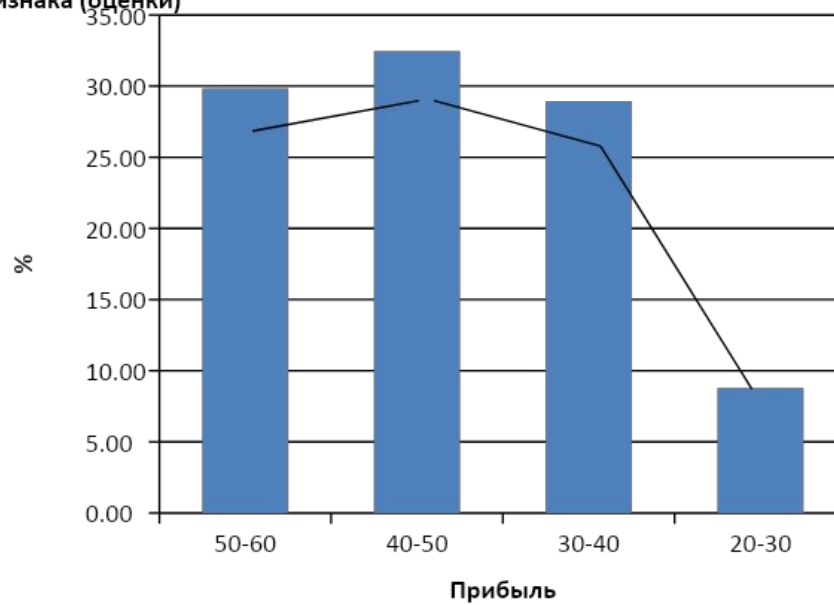
Полигон распределения



Гистограмма



Значения признака (оценки)



**Кумулята** или кумулятивная кривая строится по накопленным частотам или частостям. При этом на оси абсцисс помещают значения признака, а на оси ординат — накопленные частоты или частости

При построении кумуляты накопленная частота (частость) соответствующего интервала присваивается его верхней границе



**Огива** строится аналогично кумуляте с той лишь разницей, что накопленные частоты помещают на оси абсцисс, а значения признака — на оси ординат.

Разновидностью кумуляты является **кривая концентрации** или **график Лоренца**. Для построения кривой концентрации на обе оси прямоугольной системы координат наносится масштабная шкала в процентах от 0 до 100. При этом на оси абсцисс указывают накопленные частоты, а на оси ординат — накопленные значения доли (в процентах) по объему признака.

Равномерному распределению признака соответствует на графике диагональ квадрата. При неравномерном распределении график представляет собой вогнутую кривую в зависимости от уровня концентрации признака.

