

ГРУППИРОВКА СТАТИСТИЧЕСКИХ ДАННЫХ

2.1. Сущность, задачи и виды
группировок. Группировочные признаки.

Группировка – это процесс выделения из статистической совокупности качественно однородных групп по одному или нескольким существенным признакам

Задача и типы группировок

Задача группировки	Тип группировки
Выделение социально-экономических типов явлений	Типологическая
Выявление состава и структуры совокупности	Структурная
Установление причинно-следственных связей между причинами явлений	Факторная (аналитическая)

Основанием группировки являются группировочные признаки

Группировка считается *простой*, если в ее основе лежит один признак.

Комбинационная группировка– это группировка, выполненная по двум и более признакам.

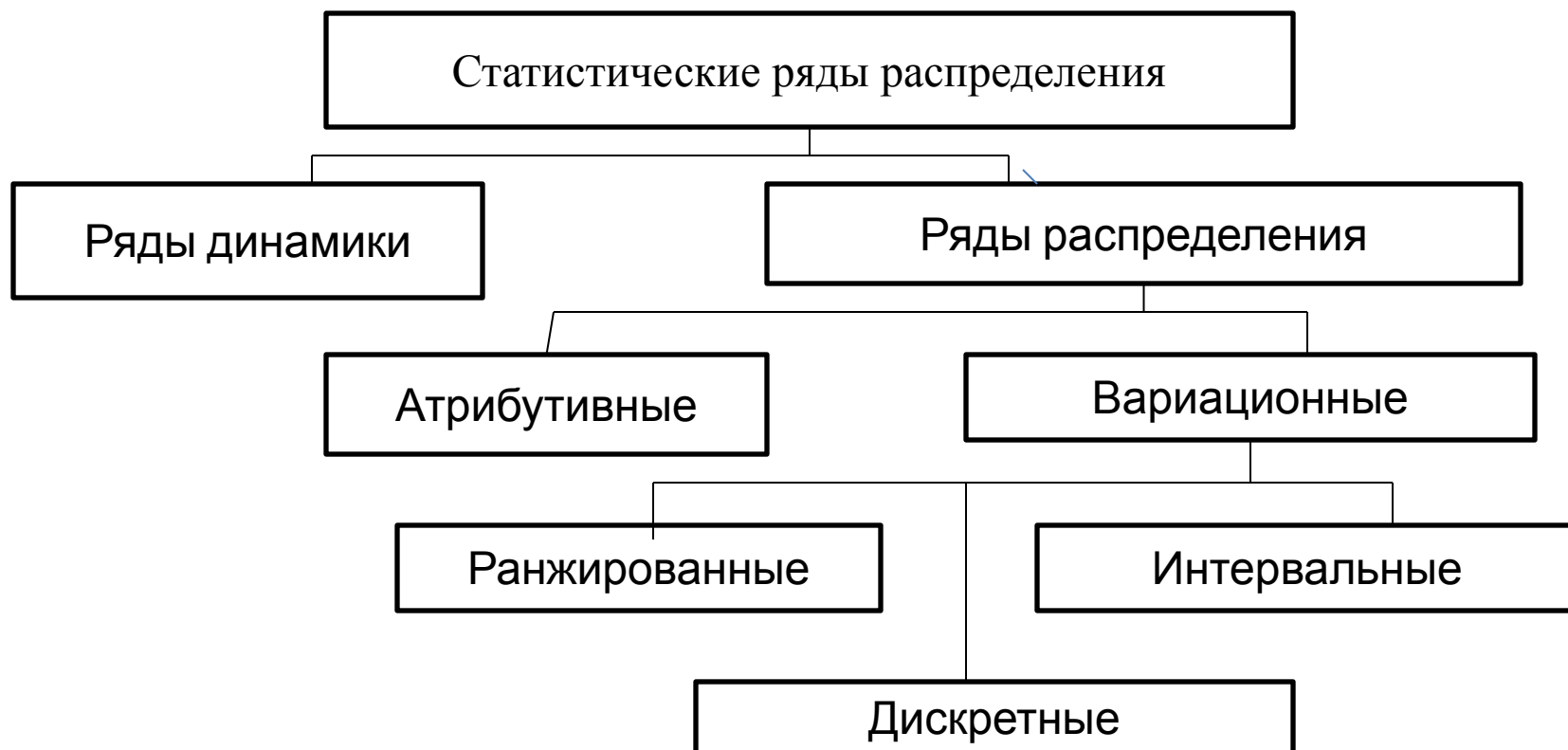
Первичной называют группировку, образованную на основе первичных данных, собранных в процессе статистического наблюдения.

Вторичная группировка выполняется по данным первичной, если есть необходимость получить групп меньше, но более крупных, или привести в сопоставимый вид данные, сгруппированные в разные по размеру интервалы с целью их возможного сравнения.

Результаты сводки и группировки данных статистического наблюдения представляют в виде статистических рядов распределения и таблиц.

2.2. Статистические ряды распределения

- Статистические ряды распределения характеризуют структуру изучаемого явления, позволяют судить об однородности (неоднородности) совокупности, границах ее изменения, закономерностях развития



Статистические ряды подразделяются на **ряды динамики** и **ряды распределения**.

Ряды динамики отражают изменения процесса, *явления во времени*.

Год	2001	2002	2003	2004	2005	2006	2007
Урожайность, ц/га	20,8	20,4	22,5	18,2	13,9	16,5	16,4

Вид ряда– ряд динамики. В основе группировки лежит *существенный, количественный признак*.

Группировка простая, т.к. выполнена по одному признаку (урожайность овса).

Ряды распределения показывают состояние исследуемого явления, его состав или структуру

Группы по возрасту, лет	Всего по региону	В том числе мужчин
До 20	9,7	8,8
20-30	13,2	12,3
30-40	29,5	29,0
40-50	27,7	27,4
50 и старше	19,9	22,5
Итого	100	100

Вид ряда –структурный ряд распределения; в основе группировки лежит существенный, количественный, непрерывный признак.

Группировка простая, поскольку выполнена по одному признаку (возраст работников)

Атрибутивные– ряды, построенные на основе качественных признаков.

Вариационные– ряды, построенные на основе количественных признаков. В зависимости от прерывности вариации признака различают *ранжированные, дискретные (прерывные) и интервальные (непрерывные) вариационные ряды.*

Ранжированным считается ряд, в котором единицы совокупности расположены в порядке возрастания (убывания) изучаемого признака.

Название озера	Максим. глубина, м	Название озера	Максим. Глубина, м
Дрисвяты	12,0	Мядель	24,6
Снуды	16,5	Нарочь	24,8
Селява	17,6	Лепельское	33,7
Лосвидо	20,2	Рычу	51,9
Струсто	23,0		

Вариационный ряд имеет два элемента:

– **варианту**– отдельное значение группировочного признака в вариационном ряду распределения (x_i);

– **частоту** – число повторений отдельных вариантов (f_i)

Частоты, выраженные в долях единицы или процентах к итогу, называют **частостями**.

Если в основе группировки лежит количественный дискретный признак, т.е. признак, который принимает определенные (например целые) значения, то число выделяемых групп будет равно числу вариант значений признака. Пример дискретного вариационного ряда представлен в таб

Успеваемость в баллах (x_i)	6	7	8	9	10
Число учеников (f_i)	2	3	6	4	5

Интервальные вариационные ряды значение варианты имеют в виде интервалов. **Интервал**– это разность между максимальным и минимальным значениями признака в каждой группе (нижней и верхней границами интервала). Интервалы могут быть открытые и закрытые, равные и неравные)

Предприятия с численностью промышленно-производственного персонала, чел	Число предприятий, %
До 100	40,7
101-200	25,7
201-500	18,5
501-1000	7,2
1001 и более	7,9

На основе этой формулы выявлено соотношение между численностью совокупности и оптимальным числом групп

2.3. Построение интервальных рядов распределения

При построении интервальных рядов распределения возникает вопрос о численности групп. Число образуемых групп (n) может быть представлено программой исследования или получено на основе формулы Стерджесса:

$$n = 1 + 3.322 \lg N,$$

N	15-24	25-44	45-89	90-179	180-359	360-719
n	5	6	7	8	9	10

Величину интервала (i) можно определить по формуле

$$i = \frac{X_{max} - X_{min}}{n}$$

Где x_{max} и x_{min} – максимальное и минимальное значение группировочного показателя в совокупности. Нижней границей первой группы принимается минимальное значение признака, а верхней – $x_{min} + i$

Верхняя граница первой группы является нижней границей второй группы, верхняя граница больше нижней на величину интервала. По этому принципу строятся границы всех групп. Верхняя граница последней группы равна x_{max} .

2.4 Статистические таблицы

Статистическая таблица– это способ представления результатов сводки и группировки статистических данных в упорядоченном виде.

На основании обобщения информации и представления ее в виде сводных статистических таблиц можно анализировать состав, структуру и динамику изучаемых явлений, рассчитывать итоговые, средние и другие производные показатели.

Для таблицы первоначально разрабатывают ее макет.

Общий заголовок таблицы

Подлежащее	Сказуемое						
	Внутренний заголовок, отражающий содержание всех строк	Верхние внутренние заголовки					
	А	1	2	3	4	5	6
Боковые заголовки (наименование строк)							

Подлежащее статистической таблицы есть перечень групп или единиц, составляющих исследуемую совокупность единиц наблюдения.

Сказуемое статистической таблицы - это цифровые показатели, с помощью которых дается характеристика выделенных в подлежащем групп и единиц.

Различают *простые, групповые и комбинационные таблицы.*

Простой считается таблица, подлежащее которой представляет:

- упорядоченный перечень единиц совокупности (пример таб. про озера)
- административно-территориальные единицы мира, государства, области и т.д.
- различные периоды (год, квартал или хронологические даты (таб. про овес))

Групповая таблица характеризуется подлежащим, в котором группировка изучаемой совокупности дана по одному признаку.

Комбинационная таблица имеет подлежащее, данные в котором сгруппированы по двум и более признакам

При составлении таблиц необходимо соблюдать общие правила:

- таблица должна быть легко обозримой;
- общий заголовок должен кратко выражать основное содержание;
- наличие строк «общих итогов»;
- наличие нумерации строк, которые заполняются данными;
- соблюдение правила округления чисел.