

Группировки в историческом исследовании

Одним из наиболее распространённых приёмов представления совокупности разрозненных данных в удобной для восприятия форме выступает группировка.

Метод группировки заключается в разбиении исходной совокупности данных на группы, каждая из которых объединена общими показателями.

Сгруппированные данные представляются в виде таблиц или графиков.

- Имеющийся в распоряжении исследователя набор чисел называется статистической совокупностью.
- Количественные показатели, характеризующие рассматриваемый признак и принимающие различные значения – вариантами или переменными.
- Сведения источника, систематизированные в возрастающем или убывающем порядке и оформленные в виде таблицы называются ранжированным рядом.

Правила при составлении (построении) таблиц

- 1. Каждая таблица должна иметь свой подзаголовок. При минимальном количестве слов он должен полностью отражать внутреннюю структуру таблицы.
- 2. В одной таблице не должно быть много признаков. Важно помнить, что чем меньше признаков, характеристик сведено в одну таблицу, тем выше её наглядность, проще анализ, представленных данных.
- 3. Не строить громоздких таблиц. Нет необходимости каждой variante признака выделять отдельную графу таблицы. Целесообразно объединить несколько граф в одну под названием «прочие», при том, что эта графа не будет охватывать более 0,1 от общего числа наблюдений.

- 4. Не путать употребление «итого» и «всего». «Итого» выступает итогом для определенной части совокупности, а «всего» является итогом для совокупности в целом.
- 5. Громоздкие числа принято округлять.
ВАЖНО! *Арифметическое правило: если округляется цифра больше 5, то округление идет в сторону увеличения числа. Если округляется цифра меньше 5, то в сторону уменьшения. Если округляется 5, то округление идет к чётной цифре.*

6. Каждая клеточка таблицы должна соответствовать определенному числу. Следует избегать включения в таблицу простых дробей.

- **ВАЖНО!**
 - *Если в распоряжении исследователя нет сведений по какому-то параметру, то рекомендуется ставить или прочерк (-), или троеточие (...)*
 - *Если сведения есть, но выражены крайне малой величиной, то в таблицу вносится 0,0.*
 - *Если какое-либо значение получено исследователем, автором таблицы в результате приближенных, условных вычислений, то оно должно быть заключено в круглые скобки (X).*
 - *Если исследователь сомневается в достоверности значения того или иного параметра, взятого из источника, то рядом с сомнительным показателем ставится вопросительный знак (X?).*

- 7. Таблицы сопровождаются сносками и примечаниями. Сноски относятся к части таблицы – строке, столбцу, клетке – и указывают на ограниченные обстоятельства, которые надо иметь в виду при чтении отмеченных фрагментов таблицы. Примечания относятся к таблице в целом. Чаще всего в них указывается источник информации. Если таблица авторская, следует указывать «Составлено по данным:...» Если таблица взята в готовом виде, то указывается источник информации.

- Признаки, положенные в основу составления таблицы, могут быть *дискретными*, т.е. принимающими только целые значения и *непрерывными*, если отдельные их значения могут отличаться друг от друга на сколь угодно малую величину.
- Для того чтобы не потерять информацию и в то же время составить компактную таблицу используют *интервальные ряды*.
- Выбор интервалов зависит от свойств изучаемого процесса или явления и от цели работы, вопрос этот решается содержательным, качественным анализом и зависит от профессиональных навыков историка.

Формула Г.Стерджесса

- $K = \frac{X_{max} - X_{min}}{1 + 3,2 * \lg n}$,

- Где K – величина интервала;
- X_{max} – наибольшее значение признака;
- X_{min} – наименьшее значение признака;
- N – число элементов совокупности

- Границы интервалов для дискретных признаков устанавливаются без совпадений крайних показателей смежных интервалов.
- Противоположное правило применяется для дробных (непрерывных) признаков – обязательное совпадение смежных границ интервалов.
- Статистика различает закрытые и открытые интервалы. В первом случае указывается верхняя и нижняя границы интервала, во втором – определена только верхняя или нижняя граница (например, «до 1 года» или «20 десятин и более»)

3 ОСНОВНЫХ ВИДА группировок:

- Типологические – расчленяют качественно – разнородную совокупность на однородные группы, на типы. В основу группировки закладывается качественный признак. (Напр.: распределение промышленности периода НЭПа по социальным секторам – государственная, кооперативная, частная)
- Структурные – представляют качественно-однородную совокупность в виде количественных групп. В основу этих группировок закладывается количественный метод. (Напр.: распределение менеджеров по стажу; - по размерам заработной платы; - по возрасту.

- Аналитические группировки позволяют установить и на определенном уровне изучить взаимосвязь между признаками. (в статистической литературе наз. факторными, при этом один из группированных признаков рассматривается как результат, а другой – как фактор). Изучение и интерпретация данных аналитической группировки должны начинаться с предварительного выяснения принципиальной возможности существования связи между признаками.

Группировка рабочих по полу и месту рождения (в %)

Пол	Место рождения	
	Город	Село
Мужской	7,5	92,5
женский	4,2	95,8

Группировка рабочих по стажу на данном предприятии и количеству детей

Стаж	Количество детей		
	0	1	2 и более
До 1 года	39	8	14
1-5	19	2	6
5-10	3	15	9

Группировка рабочих по стажу на данном предприятии и общему стажу работы по найму

Общий стаж	Стаж на данном предприятии					
	До 1	1-5	5-10	10-15	15-20	20 и более
До 1 г.	49	-	-	-	-	-
1-8	16	9	2	-	-	-
8-10	1	7	4	1	1	-
10-20	-	5	13	10	4	1
20-30	-	3	15	12	9	7
30-40	-	-	11	14	10	6
40 и более	-	-	7	5	12	5

- Метод группировки позволяет примерно определить роль того или иного признака в изучаемом явлении, силу его влияния на другие характеристики. Для этого используют безусловные распределения (структурные группировки значений интересующего признака)
- С помощью безусловного распределения можно выявить группу факторов, определяющих изучаемое событие и выделить признаки, имеющие место, но не играющие заметной роли.

Распределение служащих N-ского управления по возрасту

Возраст	20-25	25-30	30-35	35-40	40-45	45-50	50 и более
Количество служ. в %	0,3	0,7	1,9	3,7	86,9	4,0	2,5

Динамические ряды.

- - бывают **моментные**, в которых время задано в виде конкретных дат (моментов времени), и **интервальные**, где время задано в виде промежутков – лет, месяцев, суток.
- Показатели временного ряда называют **уровнями**.

Исследование динамических рядов начинается с доказательства нижеперечисленных требований предъявляемых к их построению

- 1. Однородность явления относительно каждой динамической группы, т.е. в один временной промежуток должны включаться одни и те же явления.
- 2. Неизменность территории, к которой относятся показатели. (Это особенно важно проверить, т.к. история знает примеры когда юридический статус территории остался прежним, а административно-территориальные границы изменились)
- 3. единство методологии учета показателей. Уровни динамического ряда могут быть заданы либо абсолютными, либо относительными, либо средними величинами.

- 4. «Временной показатель, положенный в основу динамического ряда, в случае его интервальной разбивки должен иметь сопоставимые временные промежутки».
- Показатели динамического ряда считаются сравнимыми при выполнении всех четырех условий его построения.