



**Группировка
материала
статистических
наблюдений**

1. Адекватность математических методов.

2. Группировка материала статистических наблюдений

2.1. Простая группировка

2.2. Перекрестная группировка

Распределение изучаемой совокупности на однородные группы по существенным для нее признакам (характеристикам) называется ***статистической группировкой***.

Основное назначение группировки:

1. установление численности каждой отдельно взятой части совокупности, расчленённой в соответствии со значениями определенного признака (или нескольких признаков),
2. изучение влияния причин и зависимости явлений.

Ряды распределения

- Результат группировки единиц наблюдения по какому-либо признаку называется **статистическим рядом**.

Отдельные значения признака (x)	5	6	7	8	9	10
Частота (n)	1	0	1	2	0	4
Объем совокупности	8					

Сгруппированные данные

Образование (классы)	5—7	8	9—10
Частота	2	2	4

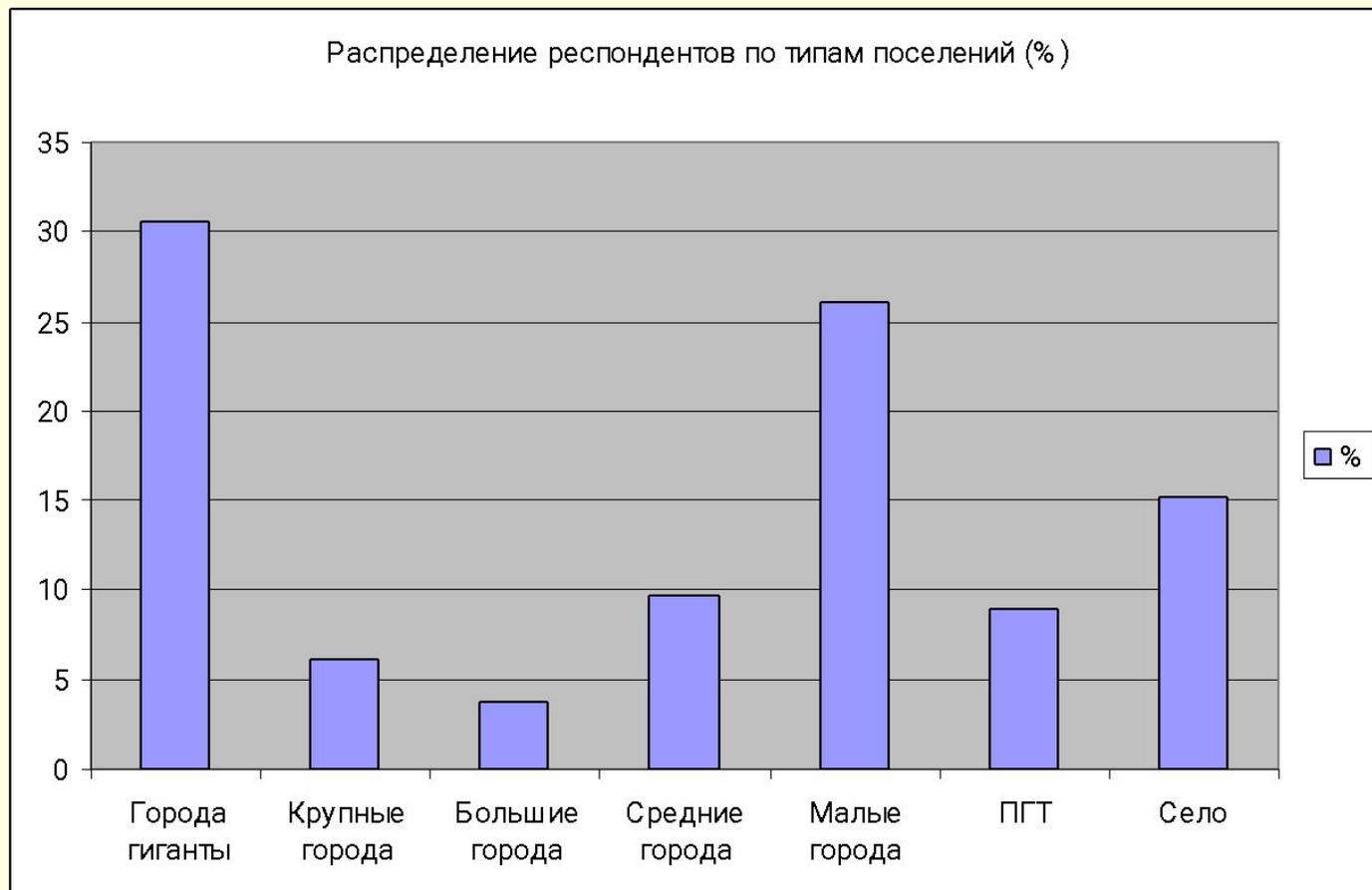
Простая группировка

- это классификация или упорядочение данных по одному признаку.

N	Значение:	F:	%отв.
1	Города гиганты	452	30,50
2	Крупные города	90	6,07
3	Большие города	55	3,71
4	Средние города	143	9,65
5	Малые города	386	26,05
6	ПГТ	132	8,91
7	Село	224	15,11
	Итого:	1482	100,00

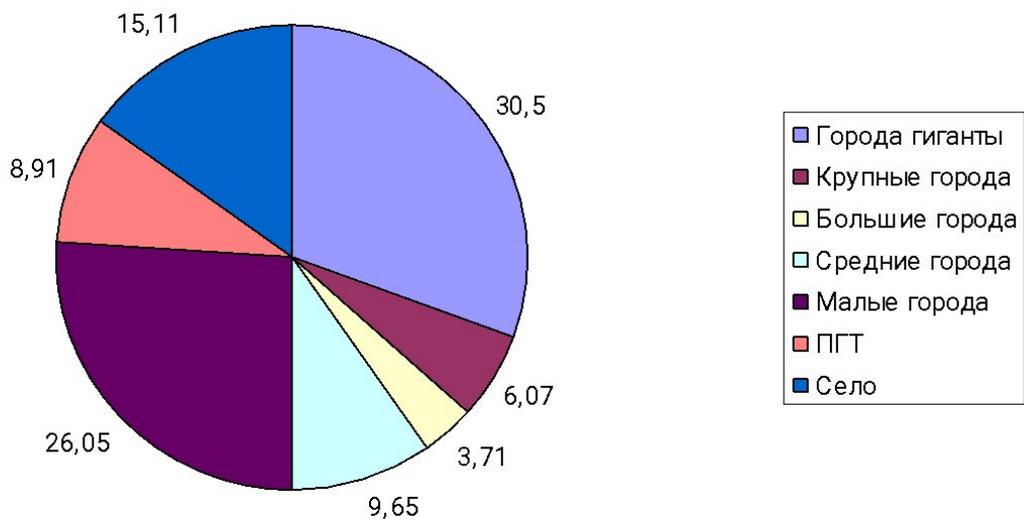
Гистограмма

- — это графическое изображение интервального ряда

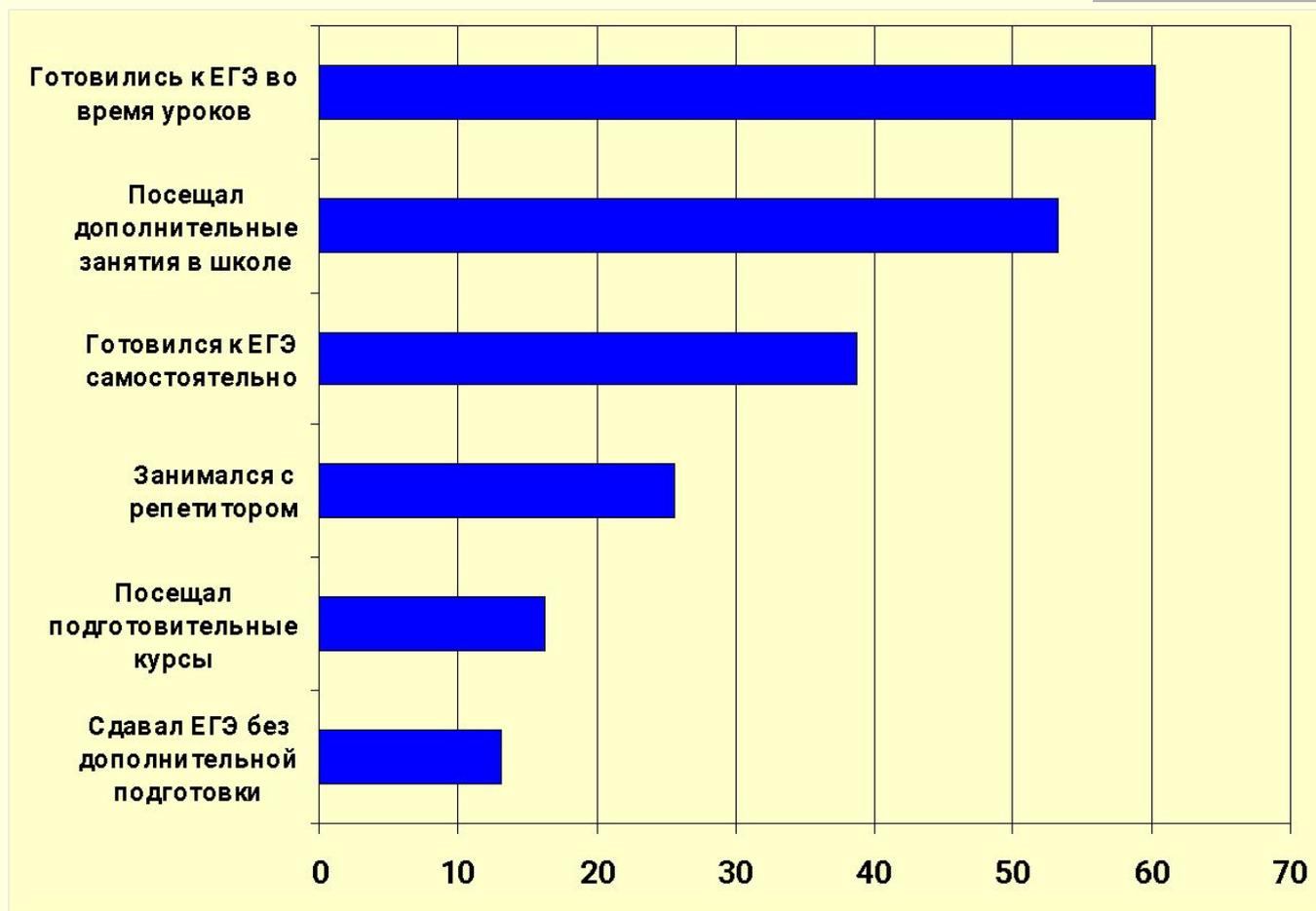


Круговая диаграмма

Распределение респондентов по типам поселения

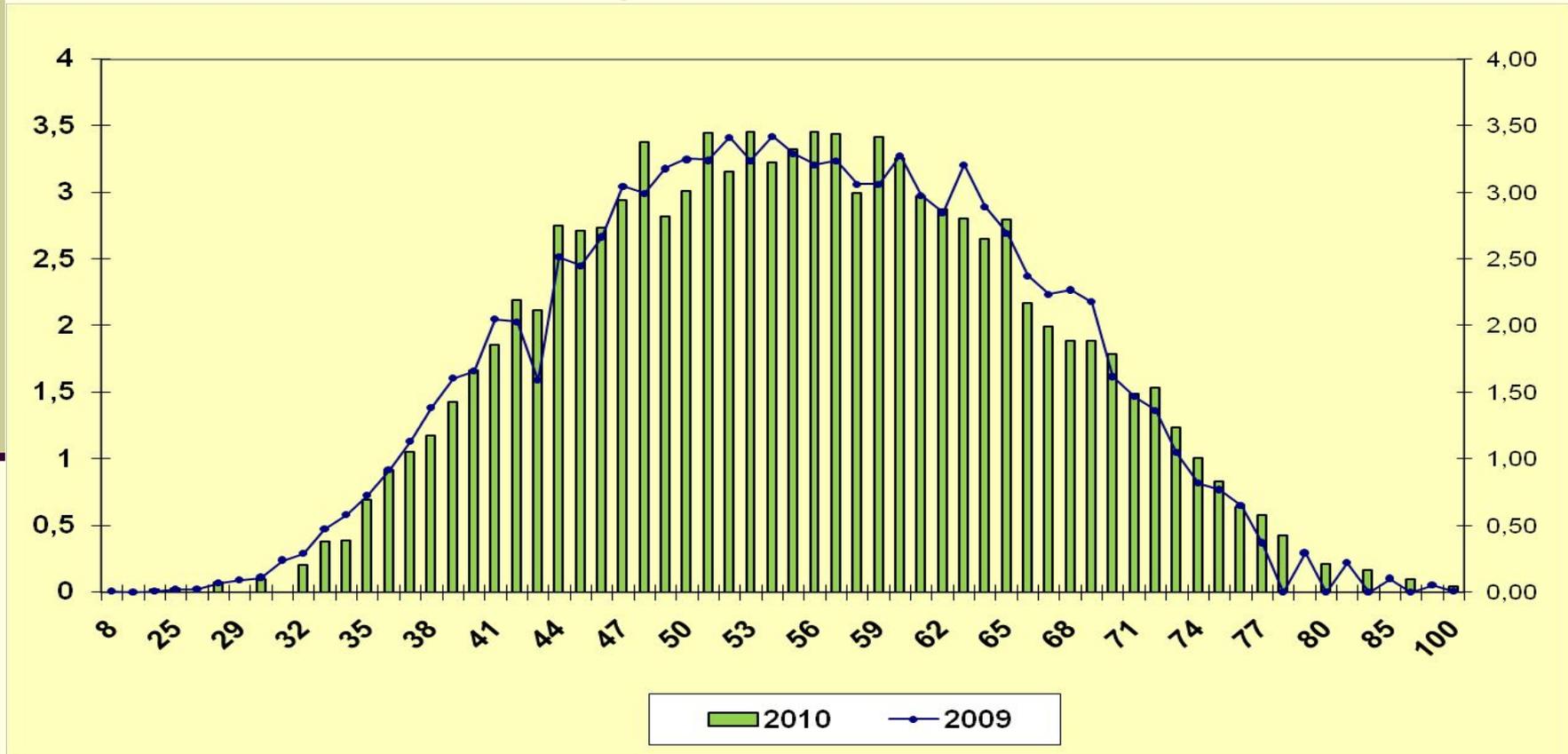


Поливариантные вопросы:



Полигон распределения

- это график распределения частот для порядковых и количественных переменных

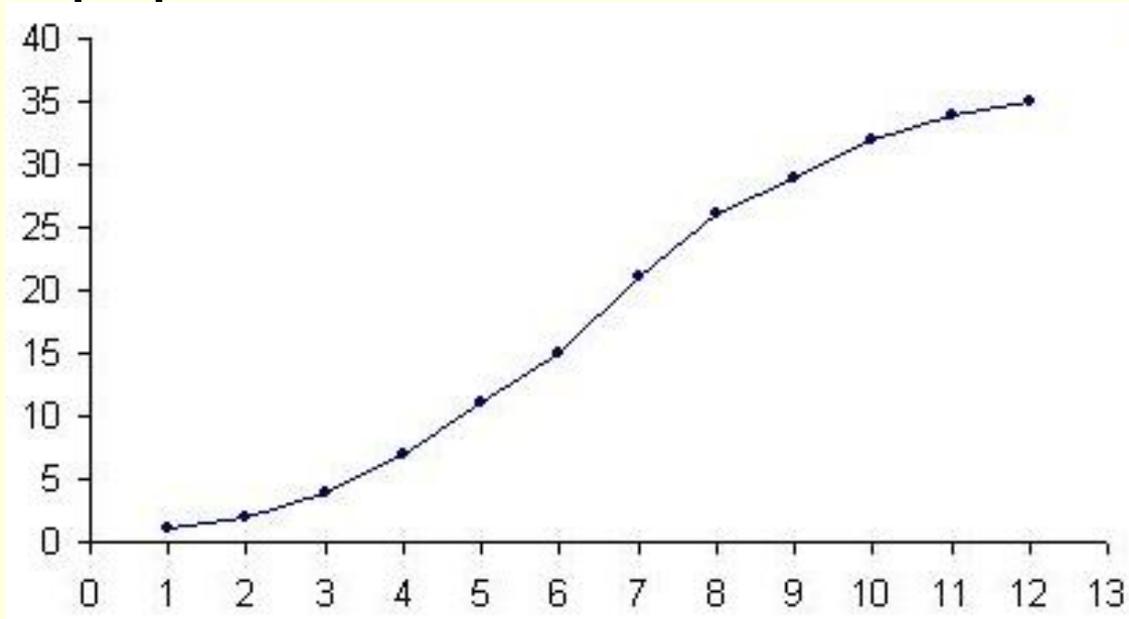


Полигон распределения



Кумулята

- графическое изображение статистического ряда накопленных данных, полученной информации.



Перекрестная группировка

— это связывание данных предварительно упорядоченных по двум признакам (свойствам, показателям).

Задачи:

- обнаружить какие-то взаимозависимости,
- осуществить взаимоконтроль показателей
- сформировать новый составной показатель (индекс) на основе совмещения двух свойств или состояний объекта, определить (об этом ниже) направление связей влияния одного явления (характеристики, свойства) на другое.

1. Поиск устойчивых связей

Возраст мужа /лет	Возраст жены (лет)								Итого
	15-19	20-24	35-29	30-34	35-39	10-44	45-19	50-»	
15-19	42	10	3						55
20-24	158	504	51	10	1				719
25-29	52	271	184	22	7	2			538
30-34	5	52	87	69	13	5			231
35-39	1	12	27	29	21	2	3		95
40-44		1	9	18	17	8	2	1	50
48-49	1		3	в	IS	16	7	1	49
50			1	4	11	15	21	43	95
Итого	259	850	365	188	8в	47	33	45	1838

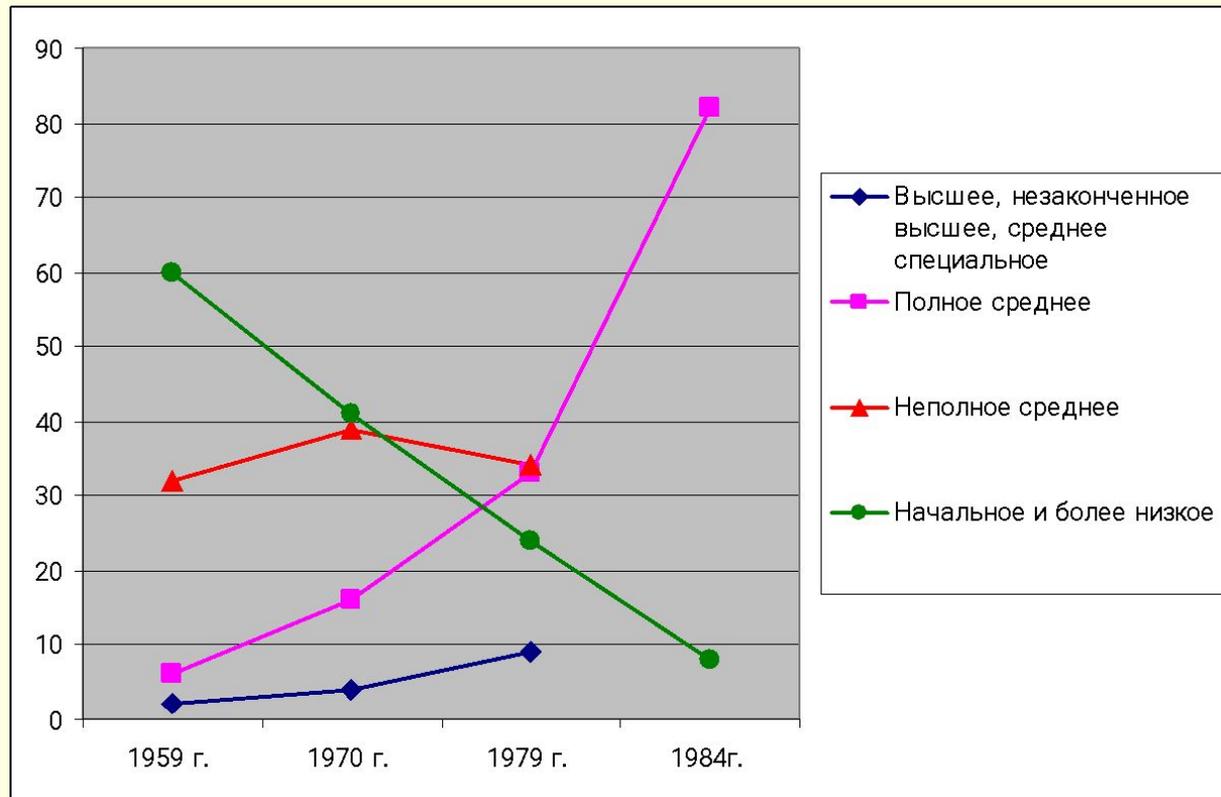
2. Установление зависимости

Наука	Время "полужизни" (лет)	Наука	Время "полужизни" (лет)
Экономика	33,0	Биология	5,0
География	16,0	Вычислительная математика	4,95
Геология	11,8		
Ботаника	10,0-11,8	Физика	3,9-4,6
Химия Математика	8,1 5,42-10,5	Биомедицина	3,0

3. Поиск тенденции, динамики процесса

Уровень образования	Доля рабочих, имеющих данный уровень образования. (%)			
	1959 г.	1970 г.	1979 г.	1984г.
Высшее, незаконченное высшее, среднее специальное	2	4	9	
Полное среднее	6	16	33	82
Неполное среднее	32	39	34	
Начальное и более низкое	60	41	24	8

Поиск тенденции, динамики процесса



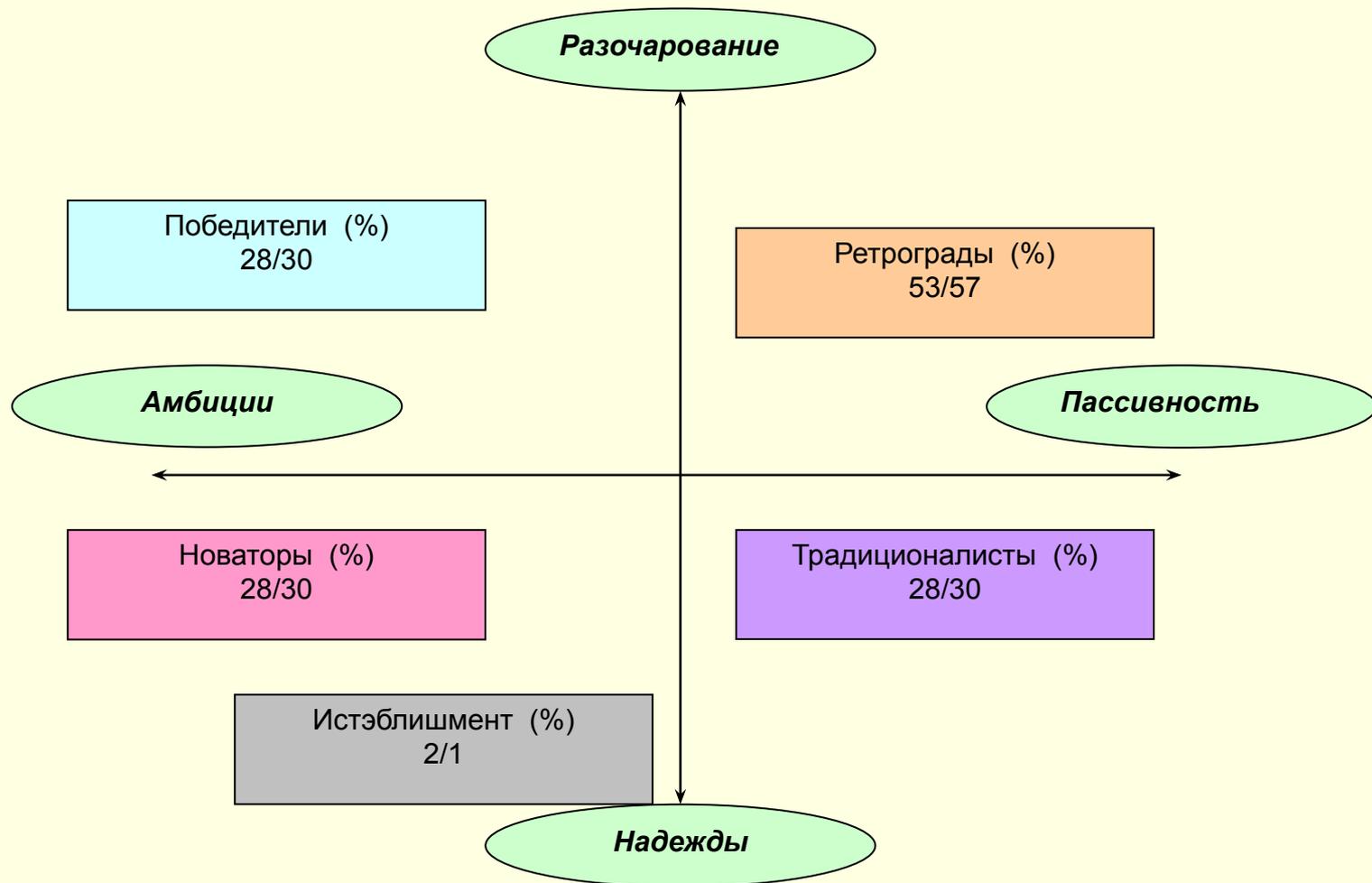
Эмпирическая типологизация

- поиск устойчивых сочетаний свойств социальных объектов (или явлений), рассматриваемых в соответствии с описательными гипотезами в нескольких измерениях одновременно.

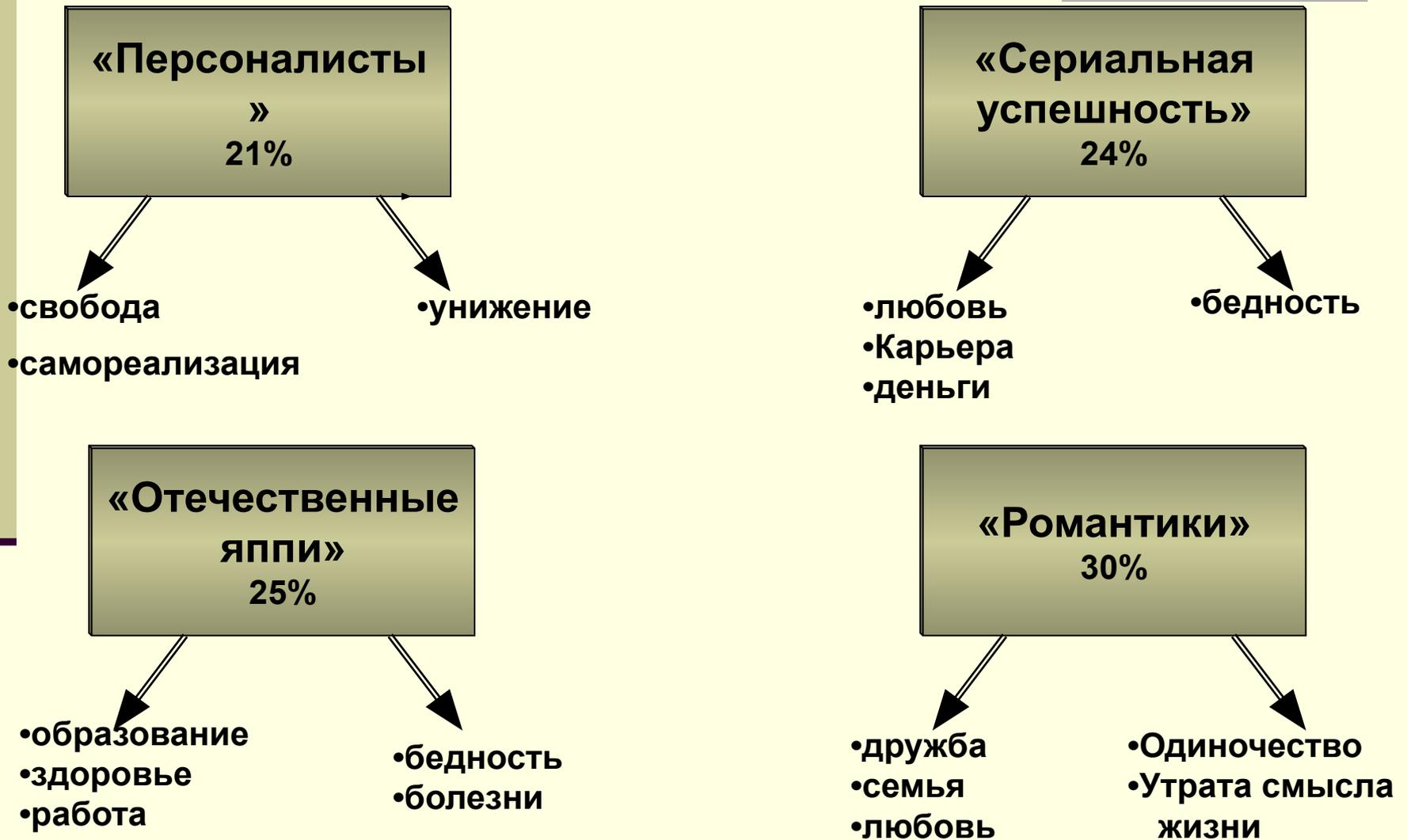
Таксономический анализ

Направление миграции	Типологические группы (в % к численности группы)					
	I	II	III	IV	V	VI
Из крупного или среднего города в село	7,3	0,8	0,0	1,2	0,0	3,0
Из малого города в село	8,0	6,6	3,2	10,0	0,0	21,2
Из села в село	47,7	14,6	13,8	41,7	22,8	15,2
Из села в малый город	16,2	47,0	35,0	4,1	31,0	22,7
Из села в средний или крупный город	20,8	31,0	48,0	42,4	46,2	37,9

Кластерный анализ



Кластерный анализ



Математические методы работы с различными типами шкал

Для **номинальной шкалы**, возможен расчет только следующих характеристик: частоты, проценты, мода, индекс качественной вариации, индекс энтропии, хи-квадрат, коэффициенты Юла, F, Крамера, Лямбда

Порядковые шкалы предполагают больший набор показателей: помимо показателей для номинальных шкал, порядковые позволяют рассчитывать медиану, квартили, относительное квартильное отклонение, процентиля, коэффициент гамма

Для **количественных шкал** возможен расчет любых аналитических величин, включая среднее, стандартное отклонение, коэффициент Пирсона и т.д.

Задачи

1. **Определить характер шкал в анкете.**
2. **Учебные достижения учащихся некоторого класса по математике характеризуются данными, представленными в таблице. Построить полигон частот.**

Количество баллов x	1	2	3	4	5	6	7	8	9	10	11	12
Число учащихся n	1	1	2	3	4	4	6	5	3	3	2	1

Задачи

3. По результатам тестирования по математике учащихся 7-го класса получены данные о доступности заданий теста (отношение числа учащихся, правильно выполнивших задания, к числу тестировавшихся учащихся), представленные ниже, в таблице. Тест содержал 25 заданий. Построить гистограмму.

Доступность задания x , %	25-35	35-45	45-55	55-65	65-75	75-85	85-95
Количество задач n	1	1	5	7	7	3	1

Задачи

4. По данным таблицы составить кумулятивный вариационный ряд, для которого построить кумуляту.

Количество баллов x	1	2	3	4	5	6	7	8	9	10	11	12
Число учащихся n	1	1	2	3	4	4	6	5	3	3	2	1