



Курс лекций по дисциплине «Социология»

Токарев В.В.

Лекция 04

Методы получения и обработки
информации в прикладном
исследовании

План лекции и логика изложения

- **Цель лекции:** Дать представление об особенностях применения выборочного метода в социологии. Познакомиться с основными подходами к получению, хранению и обработке информации в рамках прикладного исследования.
- **План лекции:**
 - Выборочный метод в социологическом исследовании.
 - ✓ Основные виды выборки, области применения и особенности. Погрешности в выборочном исследовании.
 - Структура и общие принципы разработки инструментария исследования.
 - ✓ Типы вопросов и типы измерительных шкал. Особенности различных видов шкал и области их применения.
 - Методы обработки социальной информации и программные средства для анализа данных.
 - ✓ Основные группы методов обработки первичной социальной информации. Программные средства анализа данных – SPSS, Statistica, Vortex, MS Excel.

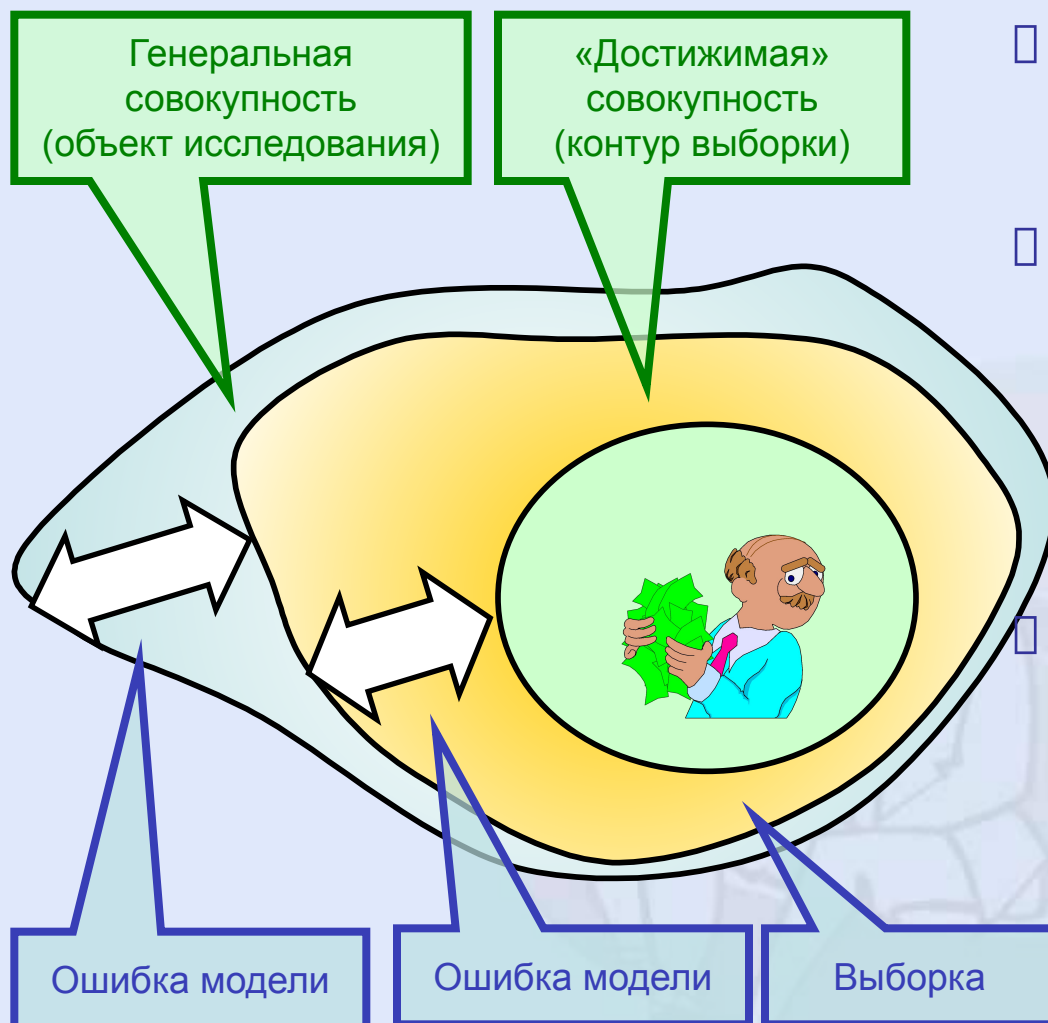
Программа прикладного исследования



Понятия выборочного исследования

- Генеральная совокупность – вся социальная группа, про которую необходимо собрать информацию.
 - В большинстве случаев «генеральная совокупность» и «объект исследования» - синонимы (состоят из одних и тех же элементов).
- Выборка (выборочная совокупность).
 - Часть генеральной совокупности, которая непосредственно подвергается исследованию.
- Единица выборки (единицы отбора).
 - Элементы генеральной совокупности, из которых формируется выборочная совокупность.
- Ошибка выборки.
 - Степень рассогласования (невязки) между оценками важных с точки зрения исследования признаков, полученными на выборке и теми значениями, которые могли бы быть получены при проведении сплошного исследования (на всей генеральной совокупности).

Ошибки в выборочном исследовании



- ❑ Социологическое исследование – это измерение.
- ❑ Измерение (всегда!!!) содержит ошибки:
 - ❑ Ошибки методики измерения.
 - ❑ Ошибки инструмента (прибора).
- ❑ Ошибка включает в себя два компонента:
 - ❑ Систематическая ошибка.
 - ❑ Случайная ошибка (статистическая).

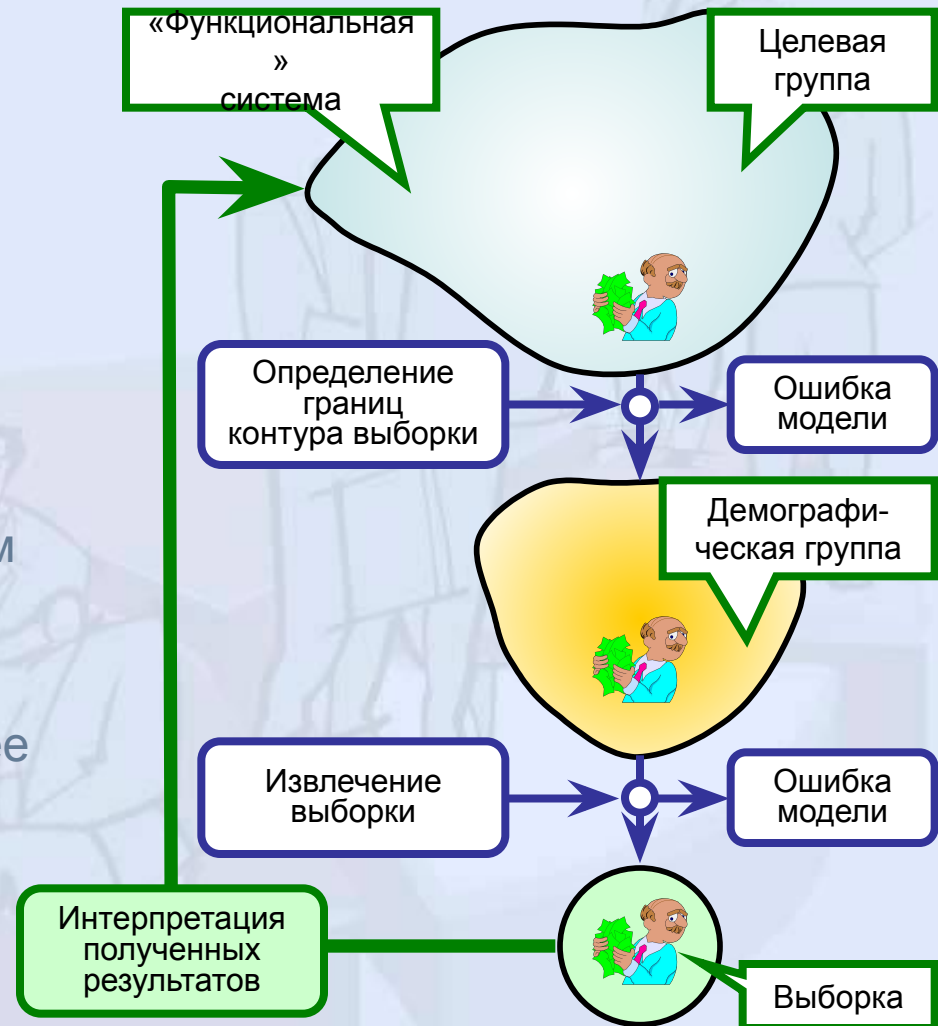
Систематическая и случайная ошибка

□ Систематическая ошибка.

- Источники ошибки – методика измерений и инструментарий.
- Может быть (хотя бы теоретически) компенсирована в процессе обработки.
- Не снижается с увеличением объема выборки.

□ Статистическая погрешность.

- Зависит от типа выборки и ее объема.
- Снижается (в некоторых случаях) с увеличением объема выборки.



Репрезентативность выборки

- Репрезентативность выборки.
 - Свойство выборки достаточно полно и точно представлять наиболее важные для исследователя и коррелирующие с ними признаки генеральной совокупности.
 - Результаты, полученные на нерепрезентативных выборка всегда содержат систематическую ошибку.
- Ограничения понятия «репрезентативность».
 - Не бывает «репрезентативности вообще». Репрезентативность возможна только по каким-либо определенным (важным для исследователя) признакам.
 - Репрезентативность не обязательно обеспечивает «надежность» или «точность» получаемых данных.
 - Для контроля репрезентативности необходимо привлечение «внешних» источников информации.
 - Репрезентативными могут быть (а могут и не быть) как «собственно случайные», так и направленные выборки.

Обеспечение репрезентативности

Статистические данные для расчета распределения выборки по типам поселений и демографическим группам, тыс. человек											
			11	12	13	14	21	22	23	24	
Код района	Район	Всего в возрасте 16 лет и старше	Мужчины 16-29	Мужчины 30-44	Мужчины 45-59	Мужчины 60 лет и старше	Женщины 16-29	Женщины 30-44	Женщины 45-59	Женщины 60 лет и старше	FIN
1	(1) Тракторозаводский	116.78	14.26	14.67	13.79	9.31	14.62	14.98	17.34	17.81	FIN
2	(2) Краснооктябрьский	133.43	15.72	15.39	15.55	11.14	16.62	17.07	20.25	21.69	FIN
3	(3) Центральный район	77.37	11.38	8.88	7.57	6.32	12.12	9.19	9.78	12.13	FIN

Распределение населения по районам и демографическим группам, %

Код района	Район	Всего в возрасте 16 лет и старше	Мужчины 16-29	Мужчины 30-44	Мужчины 45-59	Мужчины 60 лет и старше	Женщины 16-29	Женщины 30-44	Женщины 45-59	Женщины 60 лет и старше	FIN
1	(1) Тракторозаводский	9.5%	1.2%	1.2%	1.1%	0.8%	1.2%	1.2%	1.4%	1.5%	FIN
2	(2) Краснооктябрьский	10.9%	1.3%	1.3%	1.3%	0.9%	1.4%	1.4%	1.7%	1.8%	FIN
3	(3) Центральный район	6.3%	0.9%	0.7%	0.6%	0.5%	1.0%	0.8%	0.8%	1.0%	FIN

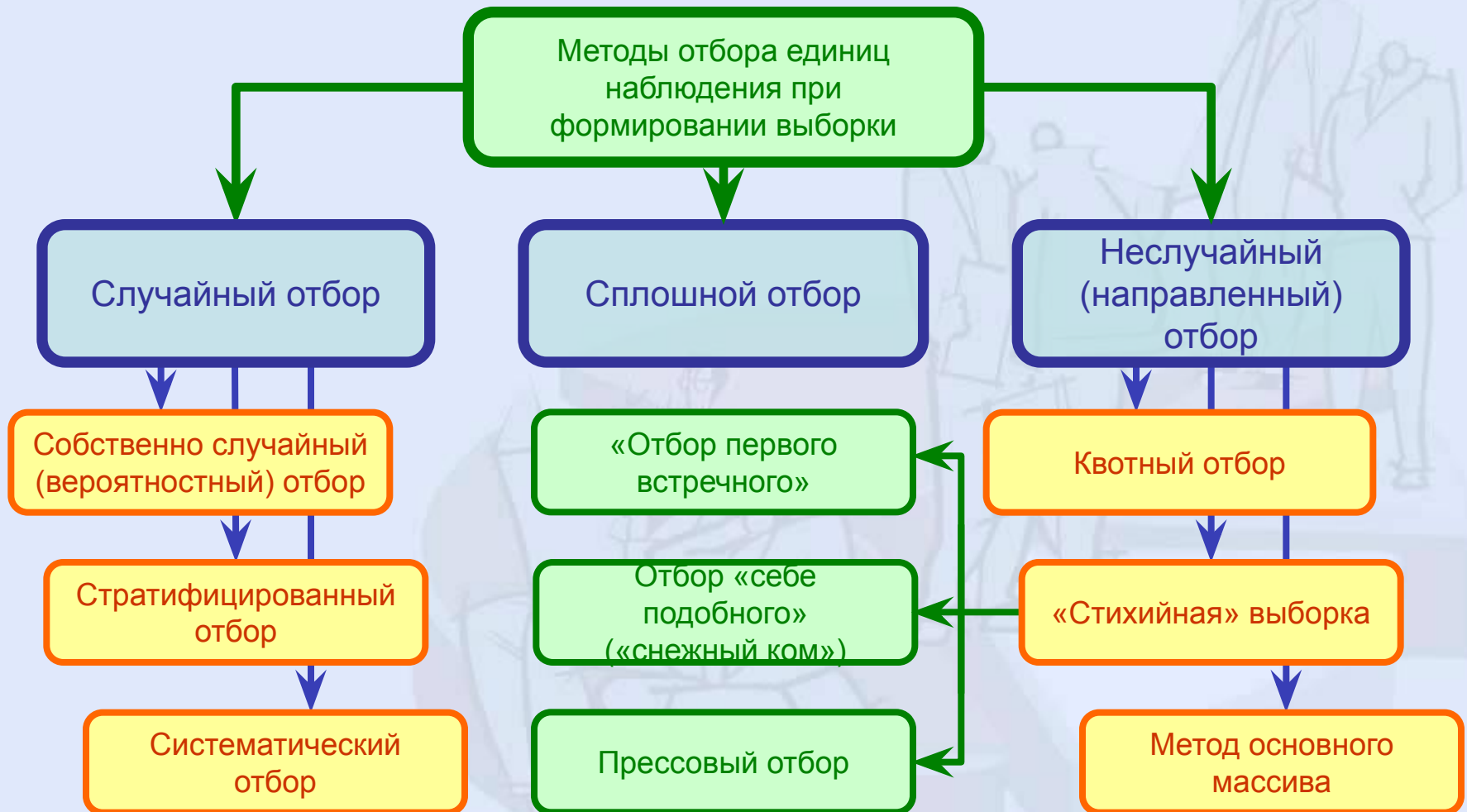
Структура выборки общим объемом в 1000 респондентов

1000

Код района	Район	Всего в возрасте 16 лет и старше	Мужчины 16-29	Мужчины 30-44	Мужчины 45-59	Мужчины 60 лет и старше	Женщины 16-29	Женщины 30-44	Женщины 45-59	Женщины 60 лет и старше	FIN
							12	12	14	15	FIN
							14	14	17	18	FIN
							10	8	8	10	FIN
							16	16	18	16	FIN
							7	7	8	9	FIN
							9	7	8	9	FIN
							9	9	10	11	FIN
							15	16	17	18	FIN
							29	29	34	31	FIN
							10	10	12	13	FIN
							FIN	FIN	FIN	FIN	FIN
Всего респондентов, человек							131	128	146	150	FIN

Для репрезентативной выборки выборочное распределение должно соответствовать генеральной совокупности по основным контролируемым признакам (в данном случае – район проживания, пол и возраст)

Методы формирования выборки



Вероятностный отбор

- Вероятностный отбор - это способ формирования выборки, при котором каждая единица генеральной совокупности имеет равную вероятность быть включенной в выборку.
 - При вероятностном отборе имеется возможность количественной оценки статистической погрешности выборки.
 - Является единственным приемлемым методом формирования выборки в тех случаях, когда отсутствуют или недостаточны предварительные данные о распределении объектов, входящих в генеральную совокупность, по интересующим исследователя характеристикам.

Вероятностные выборки строятся с использованием таблиц или генераторов случайных чисел.

10	09	73	25	3	09	37	67	07	15
37	54	20	48	0	80	15	73	61	47
08	42	26	89	5					
99	01	90	25	29					
12	80	79	99	70					
66	06	57	47	17	34	07	27	68	50
31	06	01	08	05	45	57	18	24	06
85	26	97	76	02	02	05	16	56	92
63	57	33	21	35	05	32	54	70	48
73	79	64	57	53	03	52	96	47	78

Фрагмент таблицы равномерно распределенных случайных чисел

Вероятностный отбор

4791	A	B	C	D	E	F	G	H
4792	ЗАВОДОУКОВСКИЙ РАЙОН							
4793								
4794		(101) Заводоуковский район						
4795		2						
4796								
4797		Регион	Город	Код города	Станция	Число номеров	Минимальный	Максимальный
4798		(72) Тюменская	(101) Заводоуковский	34542	2	17	1064	6110
4799								
4800		Число инт.	5		Диапазон	5046		
4801		Кратность	50		Число номеров	250		
4802								
4803								
4804		##	Полный номер	Краткий номер	FIN			
4805		1	83454224480	24480	FIN			
4806		2	83454223073	23073	FIN			
4807		3	83454223561	23561	FIN			
4808		4	83454220283	20283	FIN			
4809		5	83454224451	24451	FIN			
4810		6	83454222733	22733	FIN			
4811		7	83454221451	21451	FIN			
4812		8	83454221939	21939	FIN			
4813		9	83454224316	24316	FIN			
4814		10	83454220844	20844	FIN			
4815		11	83454220432	20432	FIN			
4816		12	83454223236	23236	FIN			

Номер телефонной станции (в пределах города)

Диапазон разрешенных номеров

Телефонные номера, сгенерированные случайным образом

Телефонные номера с учетом кода города

Вероятностный отбор практически всегда используется в телефонных опросах, когда стоимость контакта практически не зависит от местонахождения респондента

Систематический отбор

- Систематический отбор - способ формирования выборочной совокупности, при котором выбор первого объекта репрезентации (как правило, случайный) полностью определяет набор объектов репрезентации, включаемых в выборку.
 - Отбор единиц осуществляется через один и тот же интервал (шаг) в исходном списке. Номер первого объекта в списке, включаемого в выборку, определяется случайным образом или принимается равным половине шага.

МАРШРУТНОЕ ЗАДАНИЕ

Организатор опроса выдает Вам маршрутное задание, в котором указаны адреса отобранных к опросу домов. Выполнять маршрутное задание можно в любом направлении. Не так важно, в каком порядке Вы проходите дома, важно, чтобы дом был указан в маршрутном задании.

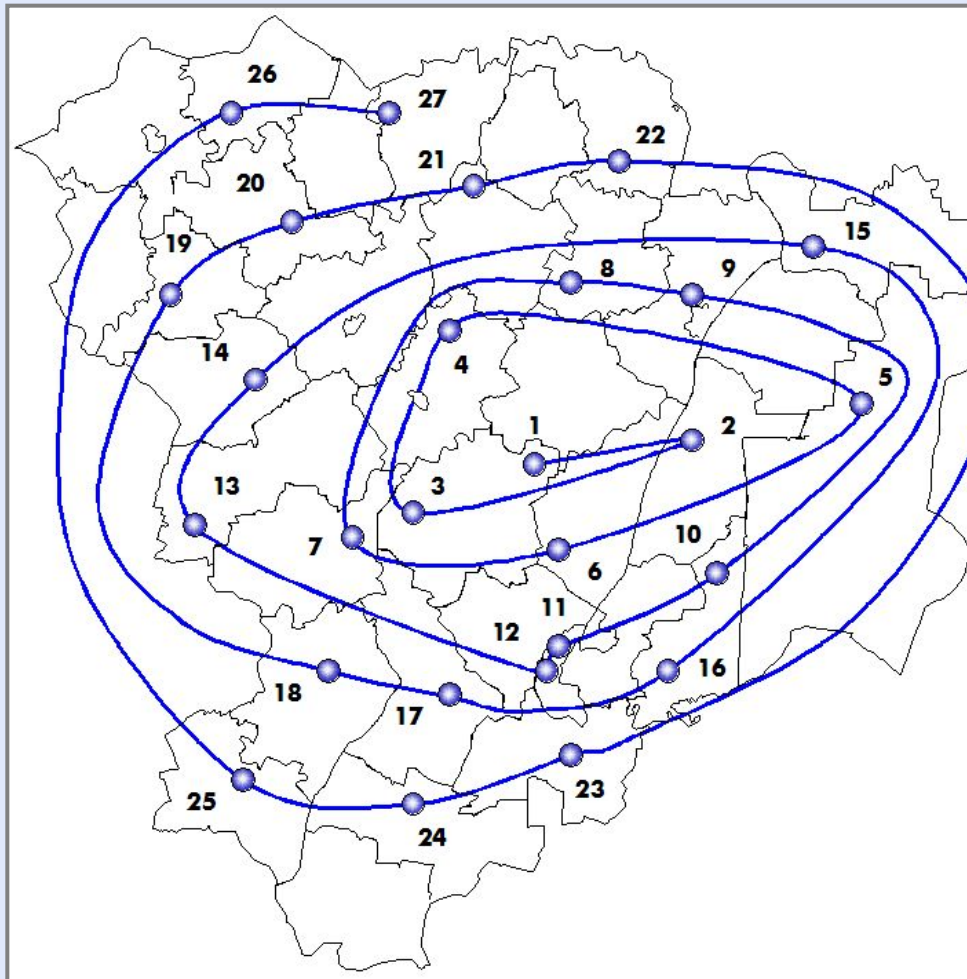
ОТБОР КВАРТИРЫ

В ПЕРВОМ ВЫБРАННОМ многоквартирном доме или общежитии интервьюер обращается в квартиру №5.

Требование строгой случайности нарушается. Выбор первого объекта и шага отбора полностью определяет выборку

ЕСЛИ В КВАРТИРЕ ИНТЕРВЬЮ БЫЛО ВЗЯТО, то следующая квартира, куда надо обращаться, вычисляется согласно описываемому ниже шагу. Так, в одно, двух и трехэтажных домах следующая квартира вычисляется прибавлением цифры 3 (№5, №8, №11...), в четырех и пятиэтажных домах шаг равен 5 (№5, №10, №15...) и в домах свыше 5 этажей шаг равен 7 (т.е. №5, №12, №19...). Когда дом закончится, следующий дом надо рассматривать как продолжение предыдущего.

Систематический отбор



На территории области расположено «N» объектов.
Объекты упорядочиваются в списке «раскручивающейся спиралью», вычерченной на карте.

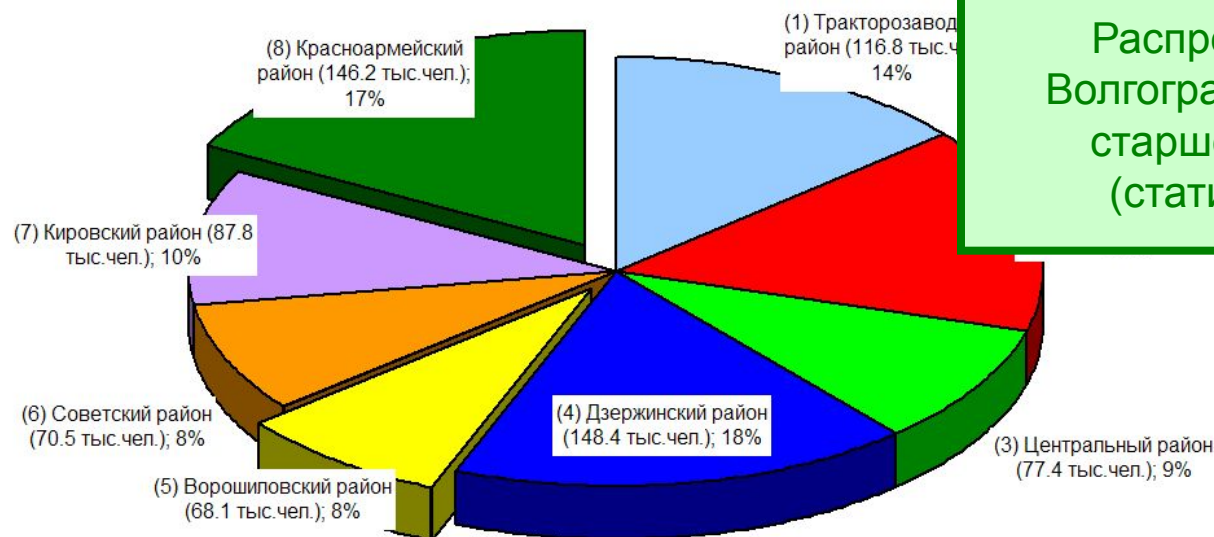
В выборку включается «n» объектов. Объем выборки определяется целями исследования, его бюджетом и сроками проведения

Объекты, включаемые в выборку, отбираются с шагом $\text{Step} = \text{ОКРУГЛВНИЗ}(N/n)$. Номер первого объекта, включаемого в выборку, определяется случайным образом.

Стратифицированный отбор

- Процедура районированного (стратифицированного, расслоенного) отбора.
 - До извлечения выборки производится разделение генеральной совокупности на «страты», различающиеся по контролируемым (важным с точки зрения цели исследования) признакам.
 - Выборка (случайная или направленная) извлекается независимо из отдельных страт.
 - Объем выборки из каждой страты определяется исследователем исходя из специфики исследования (как правило, из соображений минимизации статистической погрешности выборки).
 - На этапе обработки результатов исследования проводится перевзвешивание полученных данных (различным стратам присваиваются различные «весовые коэффициенты»).
- Стратифицированный отбор – единственный вид отбора, который при определенных условиях может обеспечить ошибку выборки ниже, чем при вероятностном отборе.

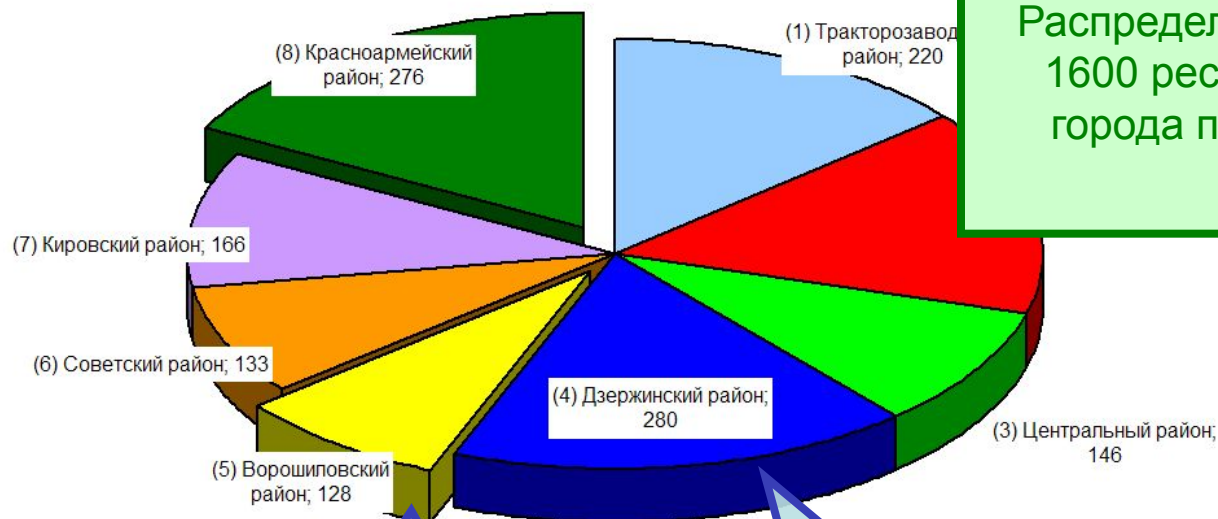
Пример стратифицированного отбора



Район	Всего в возрасте 16 лет и старше, тыс. человек	Доля в общей совокупности, %	Объем выборки при пропорциональном отборе	Максимальная статистическая погрешность	Объем выборки при стратифицированном отборе	Максимальная статистическая погрешность выборки
(1) Тракторозаводский район	116.8	14%	220	6.6%	200	6.9%
(2) Краснооктябрьский район	133.4	16%	252	6.2%	200	6.9%
(3) Центральный район	77.4	9%	146	8.1%	200	6.9%
(4) Дзержинский район	148.4	17%	280	5.9%	200	6.9%
(5) Ворошиловский район	68.1	8%	128	8.7%	200	6.9%
(6) Советский район	70.5	8%	133	8.5%	200	6.9%
(7) Кировский район	87.8	10%	166	7.6%	200	6.9%
(8) Красноармейский район	146.2	17%	276	5.9%	200	6.9%
Общий итог	848.6	100%	1600	2.5%	1600	2.5%

Пример стратифицированного отбора

Распределение выборки объемом 1600 респондентов по районам города при пропорциональном отборе



Район	Всего в возрасте 16 лет и старше, тыс. человек	Доля в общей совокупности, %	Объем выборки при пропорциональном отборе	Максимальная статистическая погрешность	Объем выборки при стратифицированном отборе	Максимальная статистическая погрешность выборки
(1) Тракторозаводский район	116.8	14%	220	6.9%	200	6.9%
(2) Красноармейский район	138.4	16%	276	6.9%	200	6.9%
(3) Центральный район	77.4	9%	146	8.1%	200	6.9%
(4) Дзержинский район	148.4	17%	280	5.9%	200	6.9%
(5) Ворошиловский район	68.1	8%	128	8.7%	200	6.9%
(6) Советский район	70.5	8%	133	8.5%	200	6.9%
(7) Кировский район	87.8	10%	166	7.6%	200	6.9%
(8) Красноармейский район	146.2	17%	276	5.9%	200	6.9%
Общий итог	848.6	100%	1600	2.5%	1600	2.5%

Максимальная статистическая погрешность – 8.7%

Максимальная статистическая погрешность – 5.9%

Пример стратифицированного отбора

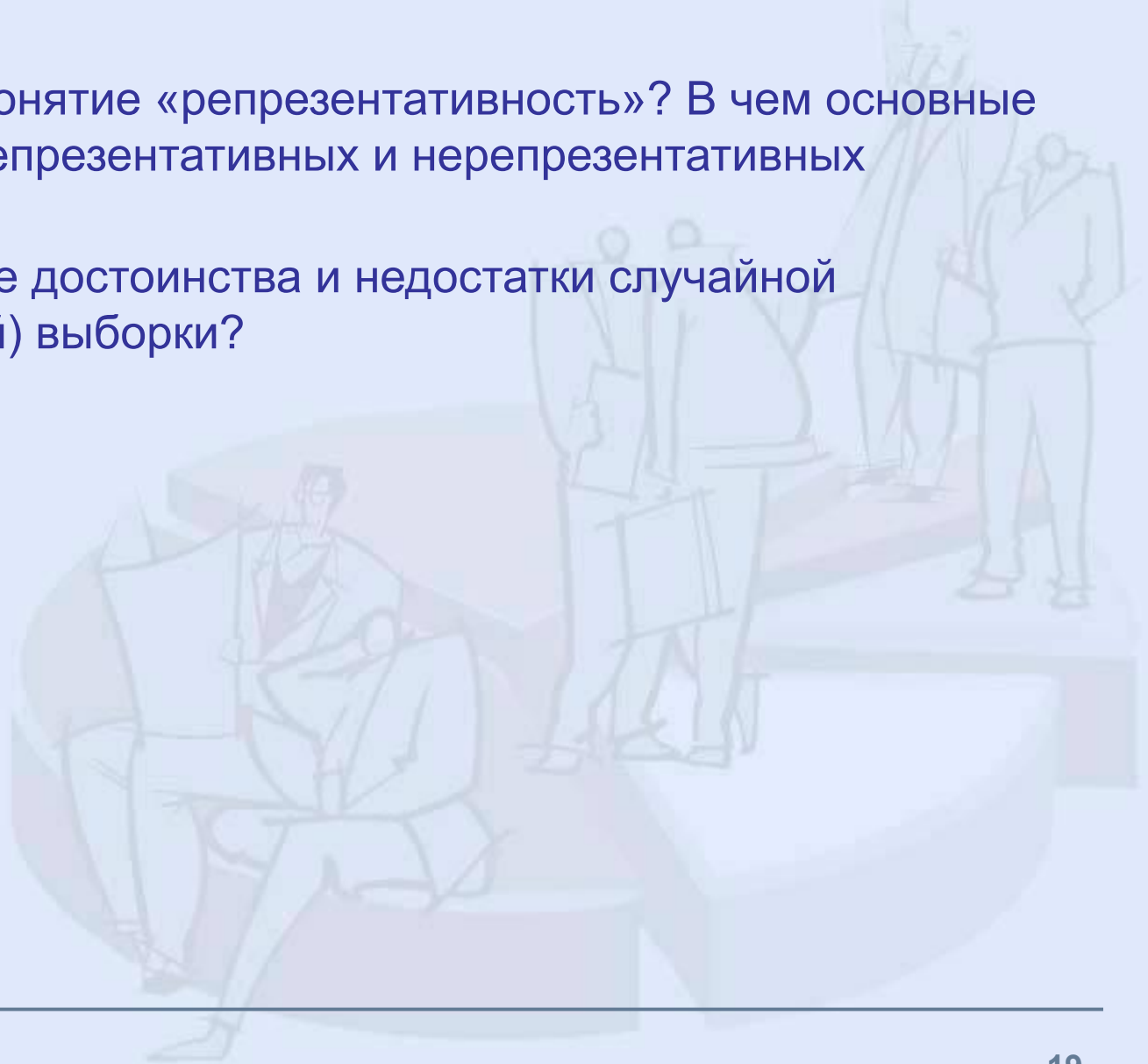


Направленный отбор

- «Направленным отбором» называют такой способ отбора, при котором при формировании выборки не соблюдаются условия случайности.
 - Случайный отбор возможен лишь в том случае, если имеется «основа выборки» - полный список единиц отбора, входящих в генеральную совокупность. В противном случае использование направленного отбора неизбежно.
 - Часто в исследованиях используется «функциональная выборка» («целевая выборка»), которая формируется из объектов, удовлетворяющих некоторым критериям («молодежь», «студенчество»). Такой вид отбора также является направленным.
- Оценка погрешности при направленном отборе.
 - Для направленных выборок принципиально невозможно оценить случайную погрешность оценки. На практике для оценки погрешности используются те же формулы, что и для одноступенчатой вероятностной выборки.

Контрольные вопросы

- Что означает понятие «репрезентативность»? В чем основные особенности репрезентативных и нерепрезентативных выборок?
- В чем основные достоинства и недостатки случайной (вероятностной) выборки?



Литература по теме занятия

- Белановский С.А. Метод фокус–групп. — М.: Издательство Магистр, 1996. - 272 с.
- Дулина Н. В., Токарев В. В. Элементы прикладной социологии. Часть I: Учебное пособие / ВолгГТУ - Волгоград, 1996. — 142 с.
- Дулина, Н. В. Основы прикладной социологии: учеб. пособие / Н. В. Дулина, И. А. Небыков, В. В. Токарев / ВолгГТУ. — Волгоград, 2006. —155 с.
- Дулина Н. В., Небыков И. А., Ситникова О. И. Практикум по социологии: Учебное пособие / ВолгГТУ. — Волгоград, 2004. — 132 с.
- Мельникова О.Т. Фокус-группа в маркетинговом исследовании. Методология и техники качественного исследования в социальной психологии: учеб. пособие. - М.: ИЦ «Академия», 2003 - 272 с.
- Рабочая книга социолога / отв. ред. Г. В. Осипов. — М.: КомКнига, 2006. — 480 с.
- Российская социологическая энциклопедия. Под общей редакцией Г. В. Осипова. — М.: Издательская группа НОРМА ИНФРА М, 1999. — 672 с.
- Ядов, В. А. Социологическое исследование: методология, программа, методы / В. А. Ядов; изд. 2-е, перераб и доп. — М.: Наука, 1987. — 248 с.
- Ядов, В. А. Стратегия социологического исследования. Описание, объяснение, понимание социальной реальности / В. А. Ядов — М.: «Добросвет», Книжный дом «Университет», 1998. — 596 с.