

# Семинар 2

Методы группировки в  
исторических исследованиях  
(продолжение)

# Виды признаков в таблице

- Дискретные:
- Принимают только целые значения
- ПРИМЕР: количество детей в семье
- Непрерывные
- Отдельные их значения могут отличаться друг от друга на сколь угодно малую величину
- ПРИМЕР: стаж работы человека

# Интервальная разбивка признака

- В исторических исследованиях используют чаще всего таблицы с интервальной разбивкой признака, поскольку признаки обладают большим количеством вариантов

# Интервальная разбивка признака

- Определение границ интервалов
- Цель: найти оптимальное число групп
- Оптимальное количество интервалов признака
- Установить границы интервалов
- Решение этой задачи зависит от степени однородности рассматриваемой совокупности

# Интервальная разбивка признака. Формальные способы

- Использование формулы Г. Стерджесса:

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

*K* – величина интервала

*X max* – наибольшее значение признака

*X min* – наименьшее значение признака

*n* – число элементов совокупности

# Интервальная разбивка признака. Формальные способы

- ПРИМЕР:
- Дано: 100 рабочих со стажем от 1 до 42 лет. Определить оптимальную величину интервала для группировки рассматриваемой совокупности по стажу
- $X_{\max} = 42$
- $X_{\min} = 1$
- $n = 100$

# Интервальная разбивка признака. Формальные способы

$$\frac{42-1}{10}$$

$$K = 1 + 3,2 \lg 100 = 5,5$$

Таким образом, оптимальной величиной интервала является 5,5 и группировка примет следующий вид:

# Интервальная разбивка признака. Формальные способы

Стаж	1-6,5 лет	6,5 -12	12 -17,5
------	-----------	---------	----------

# Границы интервалов для дискретных признаков

- Устанавливаются без совпадения крайних показателей смежных интервалов
- Это важно для того, чтобы один и тот же показатель не был зафиксирован в двух интервальных группах
- **Пример: группировка количества студентов в группе должна выглядеть примерно так:**
- ***9-15 ;16-22; 23-28...***

# Границы интервалов для непрерывных признаков

- Противоположная тенденция:  
ОБЯЗАТЕЛЬНОЕ СОВПАДЕНИЕ СМЕЖНЫХ ГРАНИЦ ИНТЕРВАЛОВ
- *Пример: стаж работы по найму – варианты от 1 месяца до 42 лет*
- *Характер признака - непрерывный*
- *По признаку совпадения границ интервалов распределение примет следующий вид:*
- *До 1 г.; 1-8; 8-20; 20-30; 30-40; 40 и более*

# Границы интервалов для непрерывных признаков

- Такое распределение предполагает наличие предлогов: «от» и «до»
- *В полном виде пример выглядит так:*
- *Стаж работы до 1 г.; от 1 до 8; от 8 до 20; от 20 до 30 и т.д.*
- *Следовательно, стаж 8 лет входит в третью группу.*

# Виды интервалов

- **Закрытые интервалы:**
- Указывают верхнюю и нижнюю границы интервала
- *например: от 8 до 15 лет*
- **Открытые интервалы:**
- Определяется только верхняя или нижняя граница
- *Например: «до 1 года»; «20 десятин и более»*

# Виды группировок

- **Типологические:** расчленяют качественно разнородную совокупность на однородные группы (типы)
- *Пример: В.И. Ленин: при анализе крестьянских хозяйств: бедняцкие, середняцкие, кулацкие. Затем происходит подсчет внутри каждой совокупности*

# Виды группировок

- **Структурные группировки:** представляют качественно-однородную совокупность в виде количественных групп. В основе закладывается количественный признак.
- *Пример: распределение рабочих по стажу.*
- Проблема: выбор группировочного признака

# Виды группировок

- **Аналитические группировки:** позволяют установить и на определенном уровне изучить взаимосвязь между признаками
- Пример:

Пол	Место рождения: город	Место рождения : село
Мужской	7,5	92,5
женский	4,2	95,3

# Виды группировок

- Группировка рабочих по месту рождения. ВАЖНО: выяснить – зависит ли пол человека от места рождения или определяется ли место рождения рабочего его полом

# Графические методы группировки

- Впервые о технике составления статистических графиков упоминается в работе английского экономиста У. Плейфейра “Коммерческий и политический атлас”, опубликованной в 1786 году и положившей начало развитию приемов графического изображения статистических данных.
- Трактовка графического метода представления статистических данных как особой знаковой системы - искусственного знакового языка - связана с развитием семиотики, науки о знаках и знаковых системах.
- Знак в семиотике служит символическим выражением некоторых явлений, свойств или отношений.

# Графические методы группировки

- Существующие в семиотике знаковые системы принято разделять на языковые и неязыковые.
- Неязыковые знаковые системы дают представления о явлениях окружающего нас мира (например, шкала измерительного прибора, высота столбика ртути в термометре и т.д.).

# Графические методы группировки

- С точки зрения семиотики человеческая речь, выраженная знаками-буквами, составляет естественный язык.
- Искусственные языковые системы используются в различных областях жизни и техники. К ним относятся системы математических, химических знаков, алгоритмические языки, графики и др.
- Не исключая естественного языка, искусственные, или символические языки упрощают изложение специальных вопросов определенной области знаний.

# Графические методы группировки

- Статистический график - это чертеж, на котором статистические совокупности, характеризуемые определенными показателями, описываются с помощью условных геометрических образов или знаков.

# Графические методы группировки

- Представление данных таблицы в виде графика производит более сильное впечатление, чем цифры, позволяет лучше осмыслить результаты статистического наблюдения, правильно их истолковать, значительно облегчает понимание статистического материала, делает его наглядным и доступным.
- Это, однако, вовсе не означает, что графики имеют лишь иллюстративное значение. Они дают новое знание о предмете исследования, являясь методом обобщения исходной информации.

# Формы графического изображения.

## Познавательная ценность

Графический метод группировки – продолжение табличного метода группировки:

Наглядность

Доступность

Ясность

Интерес

Возможность иллюстрации

Анализ явлений

# Ограниченность графического метода группировки данных

- График не может включить столько данных, сколько может таблица
- На графике показываются всегда приблизительные, округленные значения
- Фиксируется только общая ситуация. Пропадает детальность
- Построение графиков требуют аккуратности от исследователя

# Требования к оформлению графических изображений

- Изображение именуется рисунком
- Всегда после слова «рисунок»  
сопровождаются заголовком
- Допускается размещение заголовка над  
рисунком и под ним.

# Требования к оформлению графических изображений

- Заголовок должен содержать **информацию о характере изображенного показателя, единицах его измерения, территории и времени его определения**
- В заголовке (названий) графика **определяется задача, которая решается при помощи графика, дается характеристика места и времени, к которому относится график.**

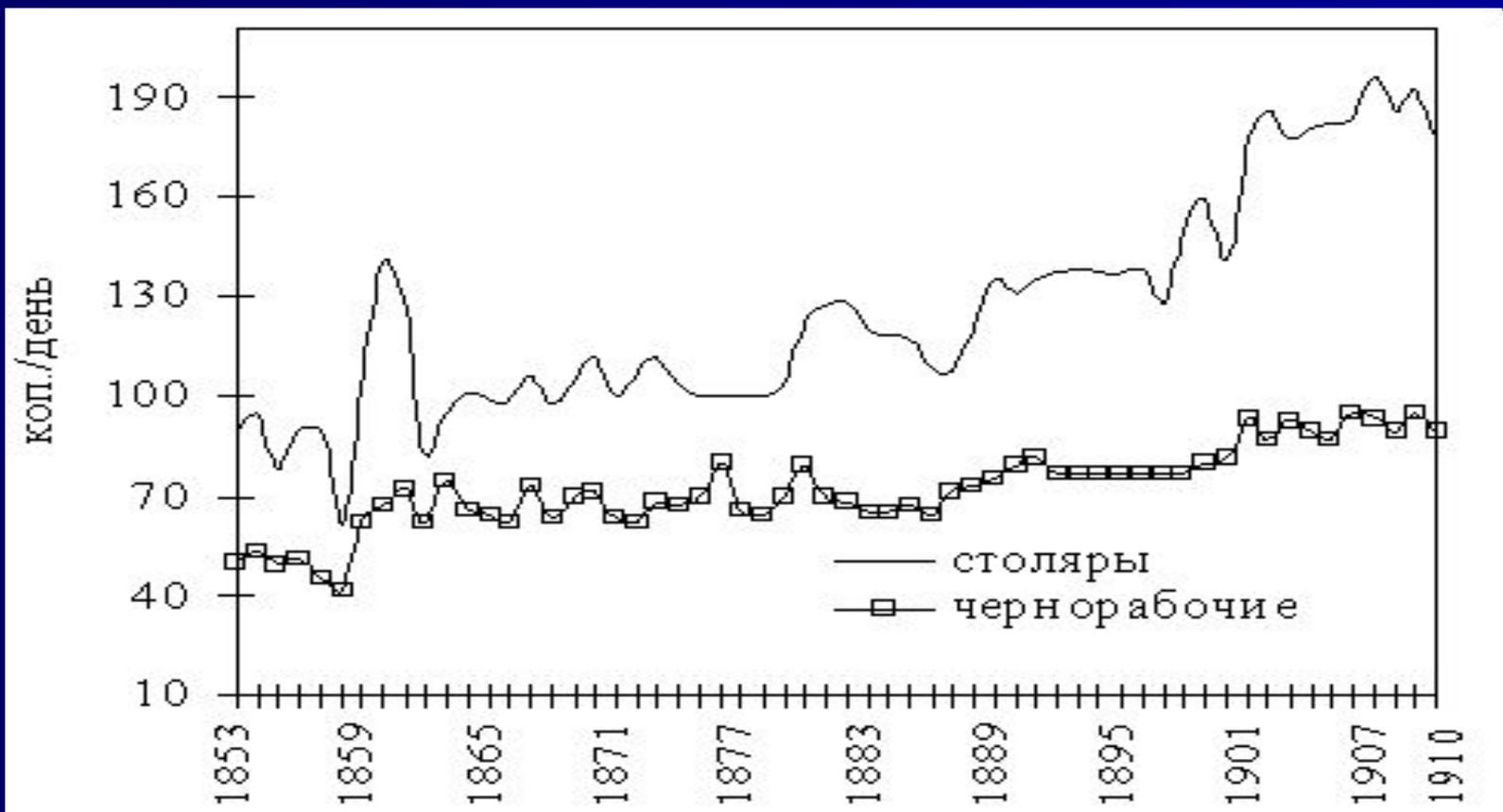
# Требования к оформлению графических изображений

- Прежде всего график должен быть достаточно наглядным, так как весь смысл графического изображения как метода анализа в том и состоит, чтобы наглядно изобразить статистические показатели.
- График должен быть выразительным, доходчивым и понятным.
- График состоит из графического образа и вспомогательных элементов.

# Требования к оформлению графических изображений

- Графический образ - это совокупность линий, фигур, точек, которыми изображены статистические данные. Диаметрические знаки, рисунки или образы, применяемые в статистических графиках, многообразны.
- Это точки, отрезки прямых линий, знаки в виде фигур различной формы, штриховки или окраски (круги, квадраты, прямоугольники и др.).
- Эти знаки применяются для сравнения статистических величин, изображающих абсолютные и относительные размеры сравниваемых совокупностей.

**Пример. Рис 2. Динамика дневной зарплаты рабочих-строителей Санкт-Петербурга (столяры и чернорабочие), коп. в день. 1853-1910 гг.**



# Требования к оформлению графических изображений

- Сравнение на графике производится по некоторым измерениям: площади или длине одной из сторон фигуры, местонахождению точек, их густоте, густоте штриховке, интенсивности или цвету окраски.
- Вспомогательные элементы включают общий заголовок, условные обозначения, оси координат, шкалы с масштабами и числовую сетку.
- Словесные пояснения (экспликация графика) помещенных на графике геометрических образов, различных по их конфигурации, штриховке или цвету, позволяют мысленно перейти от геометрических образов к явлениям и процессам, изображенным на графике.

# Требования к оформлению графических изображений

- В статистических графиках чаще всего применяется система прямоугольных координат, но есть и графики, построенные по принципу полярных координат (круговые графики).
- Когда график строится в прямоугольных координатах, на горизонтальной оси абсцисс и вертикальной оси ординат в определенном порядке располагаются характеристики статистических признаков изображаемых явлений или процессов, а в поле графика размещаются геометрические знаки, составляющие сам график. Поле графика - это пространство, в котором располагаются геометрические знаки, образующие график.
- Признаки, располагаемые на осях координат, могут быть качественными и количественными.

# Построение графиков

- Графические формы представления количественного материала называются рисунками, графиками, диаграммами

# Оформление графиков

- Каждый рисунок должен иметь свой номер
- Заголовок
- Горизонтальная и вертикальная оси специально обозначаются и шкалируются
- В случае нанесения нескольких линий под рисунком следует поместить подробное описание

# Требования к оформлению графических изображений

- В статистических графиках чаще всего применяется система прямоугольных координат, но есть и графики, построенные по принципу полярных координат (круговые графики).
- Когда график строится в прямоугольных координатах, на горизонтальной оси абсцисс и вертикальной оси ординат в определенном порядке располагаются характеристики статистических признаков изображаемых явлений или процессов, а в поле графика размещаются геометрические знаки, составляющие сам график. Поле графика - это пространство, в котором располагаются геометрические знаки, образующие график.
- Признаки, располагаемые на осях координат, могут быть качественными и количественными.

# Виды графических изображений

- По способу построения статистические графики делятся **на диаграммы и статистические карты.**
- **Диаграммы** - наиболее распространенный способ графических изображений. Это графики количественных отношений. Виды и способы их построения разнообразны.
- Диаграммы применяются для наглядного сопоставления в различных аспектах (пространственном, временном и др.) независимых друг от друга величин: территорий, населения и т. д.
- При построении точечных диаграмм в качестве графических образов применяются совокупности точек; при построении линейных - линии.
- Основной принцип построения всех плоскостных диаграмм сводится к тому, что статистические величины изображаются в виде геометрических фигур и, в свою очередь, подразделяются на столбиковые, полосовые, круговые, квадратные и фигурные

# Виды графиков.

## Линейная диаграмма

- Соединяет все значения одной переменной непрерывной линией
- Дает возможность сравнения значений разных переменных путем нанесения нескольких аналогичных линий, часто различающихся цветом или манерой изображения
- Полезны при изображении динамики развития ситуации как средство обобщения разнородных данных

# Оформление линейных диаграмм

Применяют систему прямоугольных координат.

Обычно по оси абсцисс откладывается время (годы, месяцы и т. д.) , а по оси ординат - размеры изображаемых явлений или процессов.

На оси ординат наносят масштабы. Особое внимание следует обратить на их выбор, так как от этого зависит общий вид графика.

Если масштаб для шкалы на оси абсцисс очень растянут по сравнению с масштабом на оси ординат, то колебания в динамике явлений мало выделяются, и, наоборот, при увеличении масштаба по оси ординат по сравнению с масштабами на оси абсцисс дает резкие колебания.

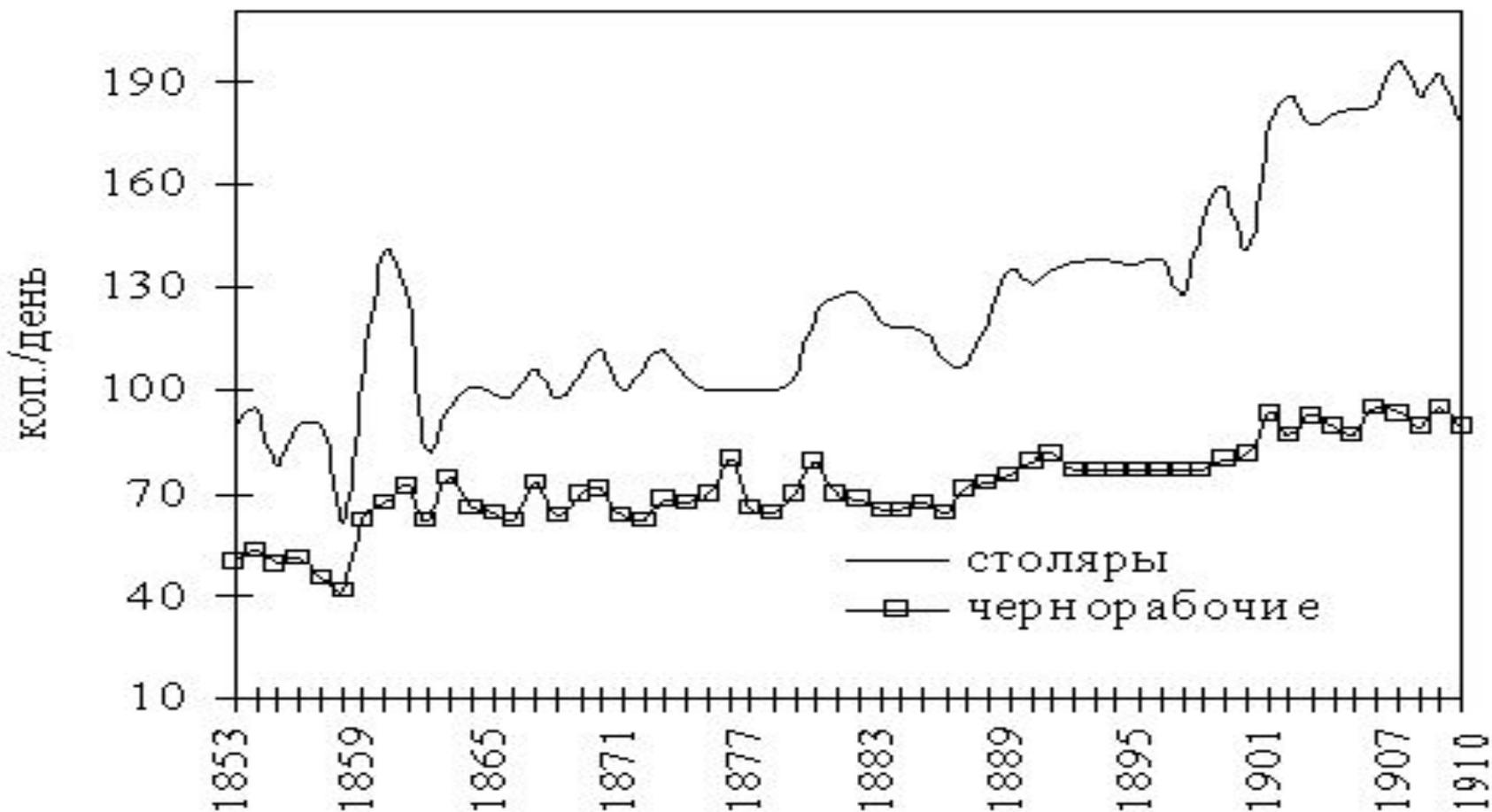
Равным периодам времени и размерам уровня должны соответствовать равные отрезки масштабной шкалы.

# Виды графиков.

## Линейная диаграмма

- Соединяет все значения одной переменной непрерывной линией
- Дает возможность сравнения значений разных переменных путем нанесения нескольких аналогичных линий, часто различающихся цветом или манерой изображения
- Полезны при изображении динамики развития ситуации как средство обобщения разнородных данных

# Пример графического изображения



# Пример графического изображения

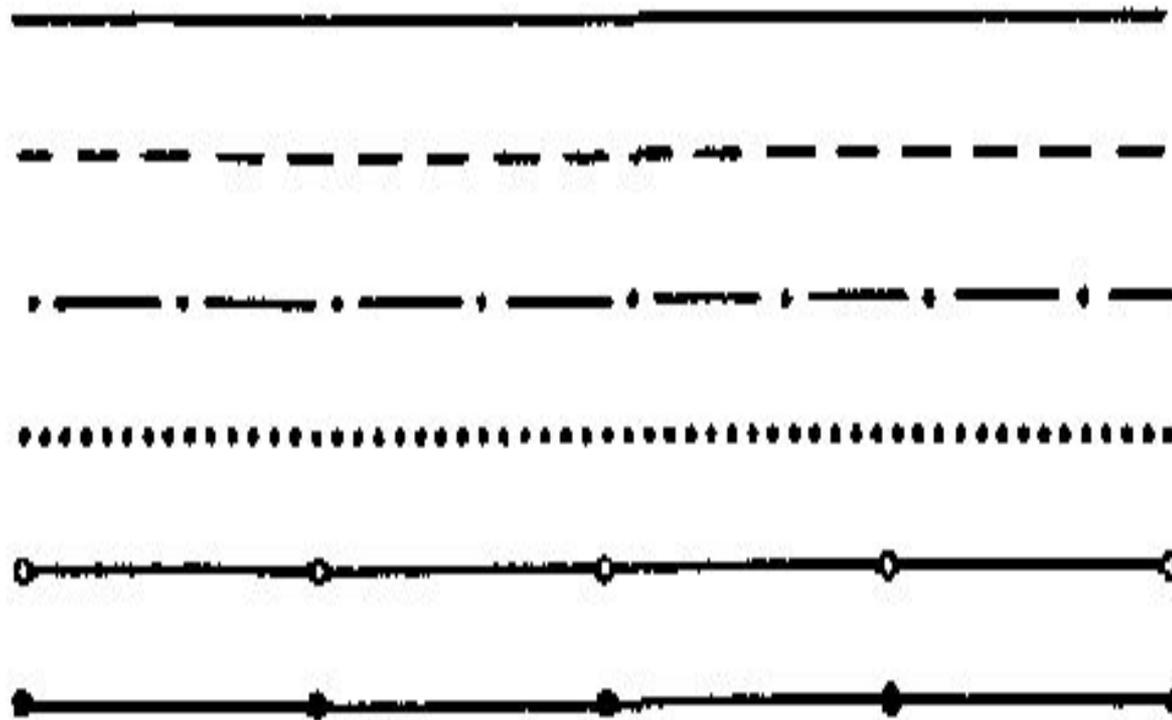


# Оформление линейных диаграмм

- Нередко на одном линейном графике приводится несколько кривых, которые дают сравнительную характеристику динамики различных показателей или одного и того же показателя.
- Однако на одном графике не следует помещать более трех-четырех кривых, так как большое их количество неизбежно осложняет чертеж и линейная диаграмма теряет наглядность.

# Диаграммы

Легче всего читать диаграмму, если количество линий сведено к минимуму. Однако иногда бывает полезно поместить данные о нескольких переменных на одной диаграмме. Тогда диаграмму следует сделать как можно больше и, если какие-нибудь из линий пересекаются, придать максимум разнообразия изображенным линиям. Наиболее часто употребляемые способы изображения линий таковы:



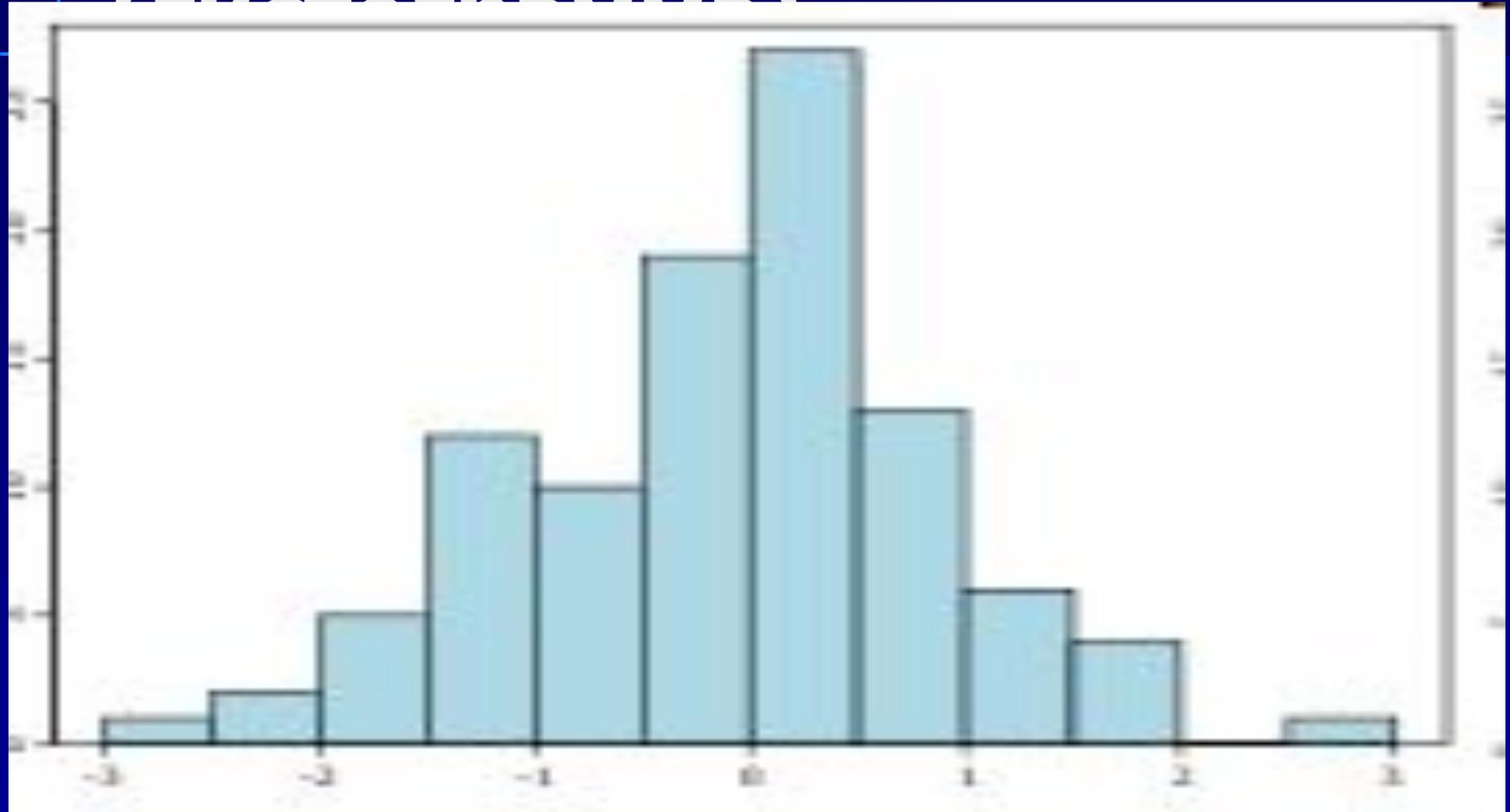
# Виды графиков.

## Столбцевая диаграмма (гистограмма)

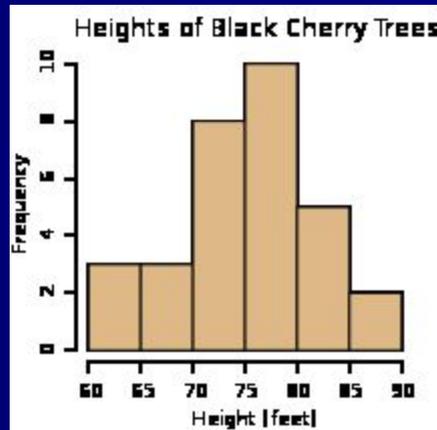
- Графическое изображение, в котором высота и иногда ширина серии столбцов иллюстрируют наблюдения по одной или нескольким переменным

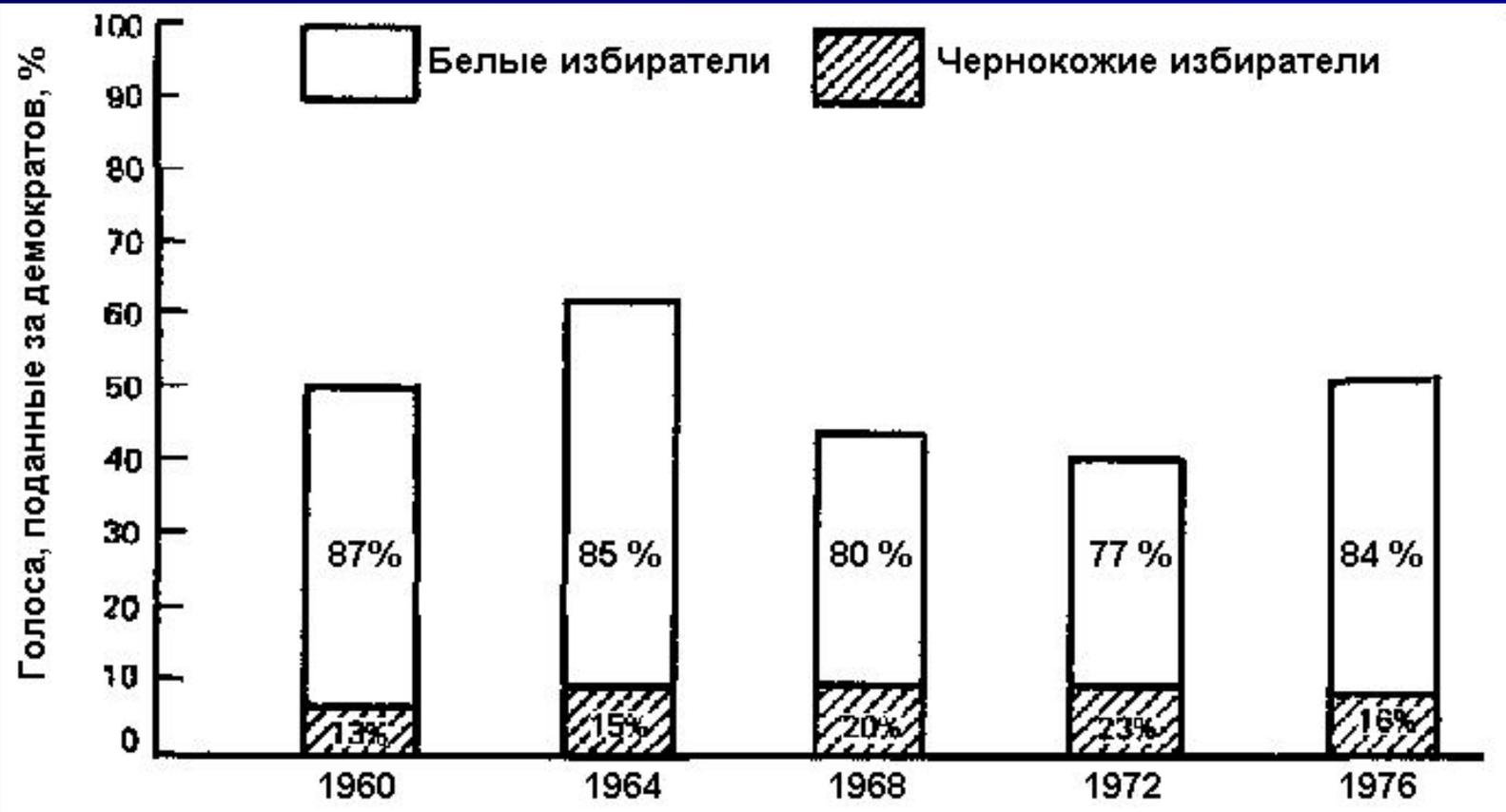
# Виды графиков.

## Столбцевая диаграмма (гистограмма)

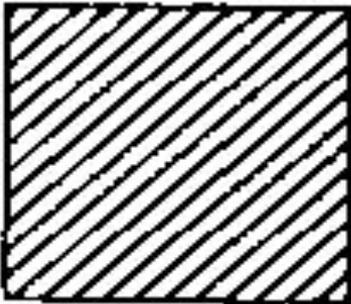
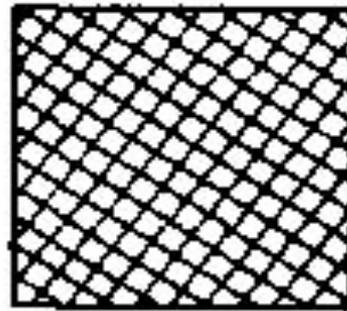


# Виды графиков. Столбцевая диаграмма (гистограмма)





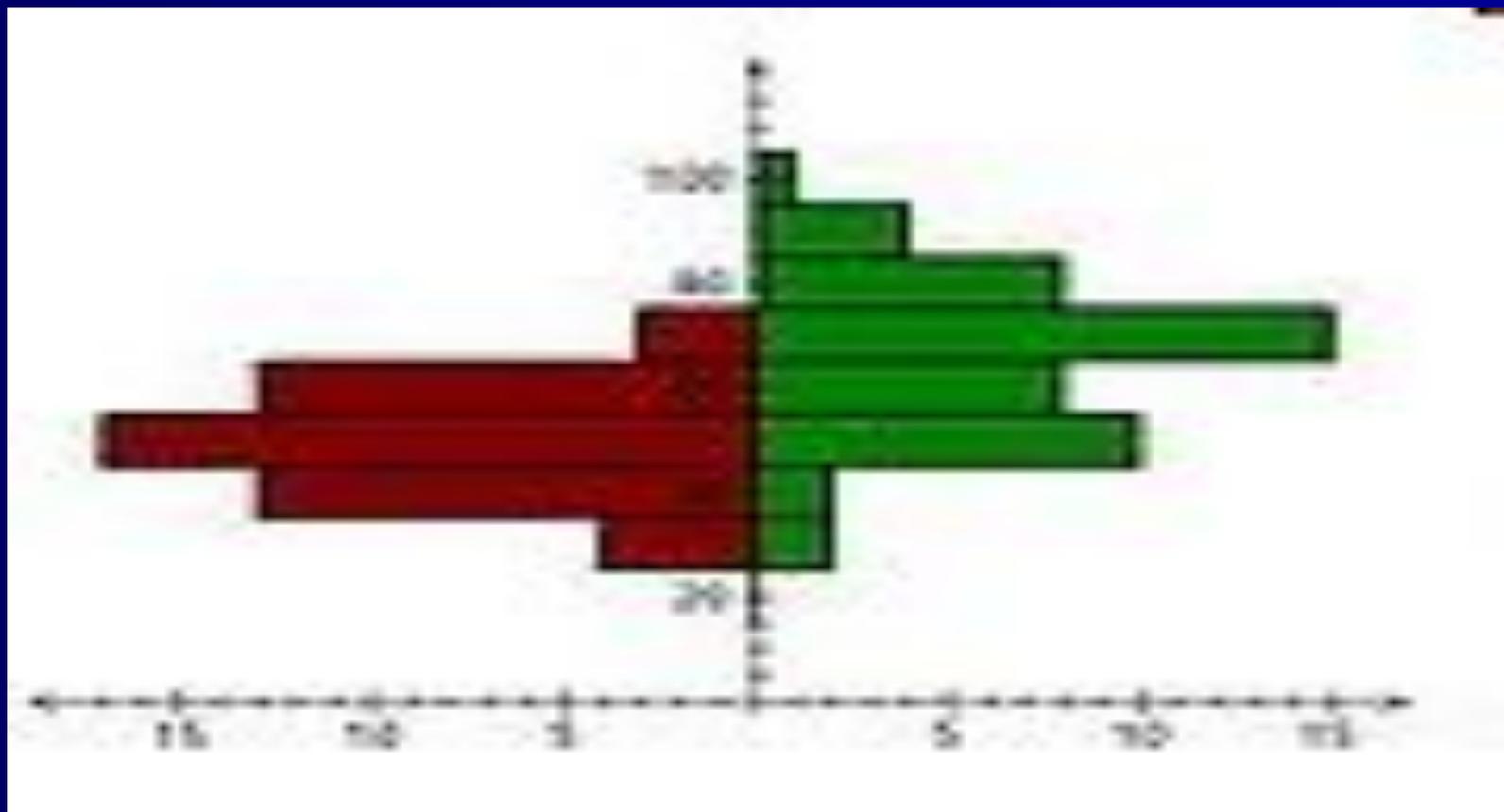
**Некоторые из наиболее распространенных способов штриховки, используемых для секторных диаграмм и секторных гистограмм, показаны на рисунке.**



# Двусторонняя диаграмма

- Расположенный в двух направлениях график, используемый для того, чтобы показать колебания сверх или менее нормы, представленной центральной линией
- Ориентирован на углубление общих представлений о распределении количественных показателей для отдельных переменных

# Двусторонняя диаграмма



# Двусторонняя диаграмма



Рис. 13.5. Двусторонняя гистограмма: колебания расовой составляющей на президентских выборах (1960—1976 гг.).

# Правила оформления столбцевых диаграмм

- Одномерность изображаемых статистических показателей и их одномасштабность для различных столбиков и полос требуют выполнения единственного положения: соблюдения соразмерности (столбиков - по высоте, полос - по длине) и пропорциональности изображаемым величинам.
- Для выполнения этого требования необходимо: во-первых, чтобы шкала, по которой устанавливается размер столбика (полосы), начиналась с нуля;
- во-вторых, эта шкала должна быть непрерывной, т.е. охватывать все числа данного статистического ряда;
- разрыв шкалы и соответственно столбиков (полос) не допускается.
- Невыполнение указанных правил приводит к искаженному графическому представлению анализируемого статистического материала.

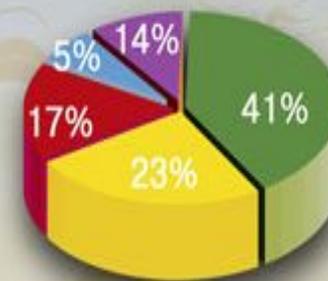
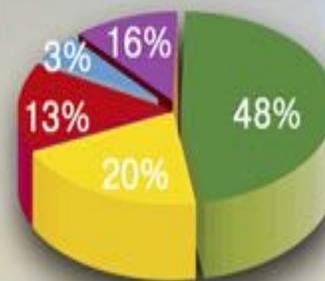
### Сравнительный график обращений граждан в Администрацию города Реутова



### Тематика обращений граждан за 9 месяцев

2006г.

2005г.

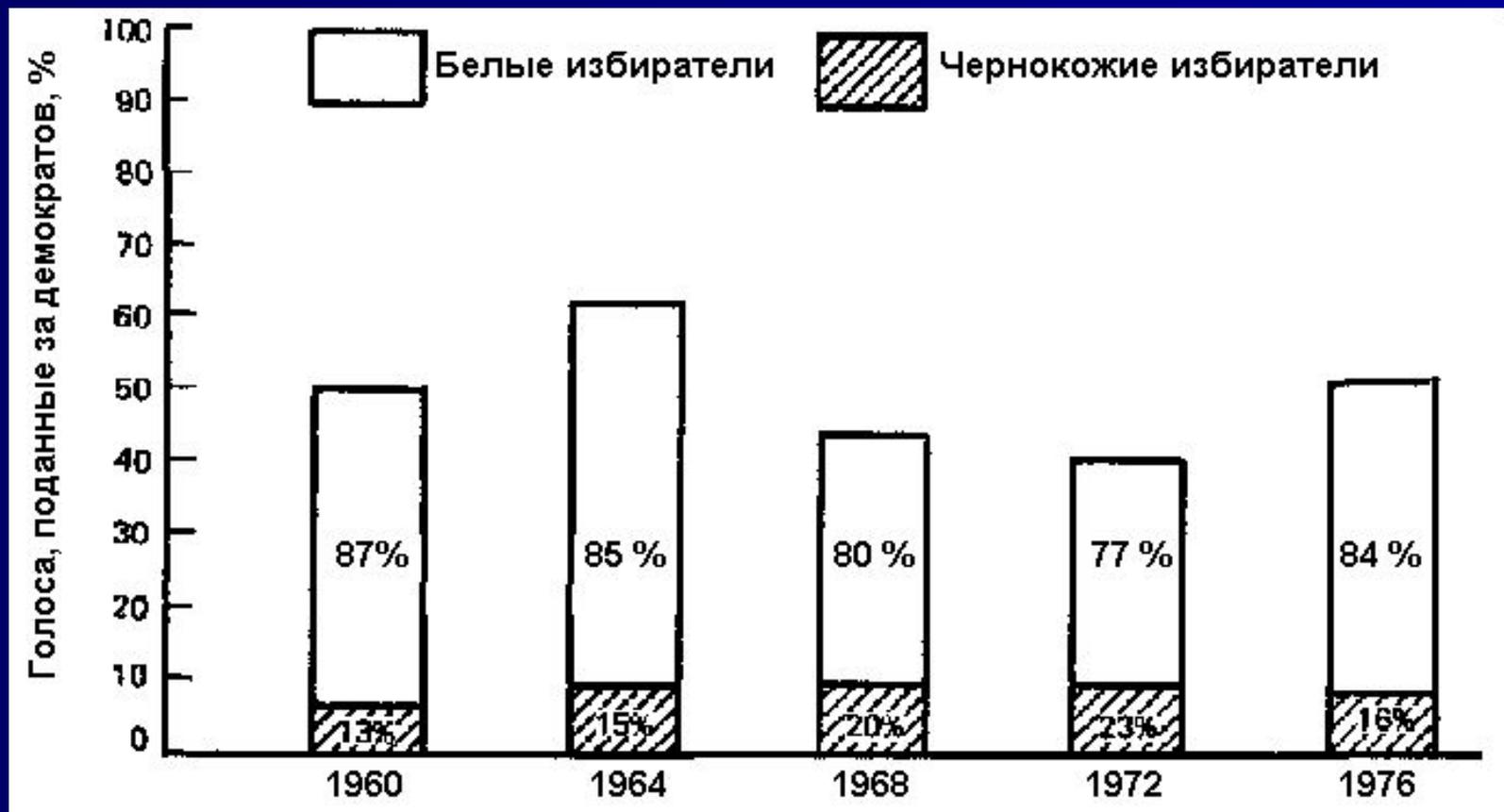


- Вопросы социальной сферы
- Вопросы коммунального и дорожного хозяйства
- Жилищные вопросы

- Вопросы связи
- Вопросы иных сфер деятельности

РОСТ КОЛИЧЕСТВА ОБРАЩЕНИЙ ГРАЖДАН ВЫЗВАН ПРОВЕДЕНИЕМ РЕФОРМ НА ГОСУДАРСТВЕННОМ И МЕСТНОМ УРОВНЕ - ЖИЛИЩНАЯ РЕФОРМА, РЕФОРМА ЖИЛИЩНО-КОММУНАЛЬНОГО ХОЗЯЙСТВА И Т.Д.

# Столбцевая диаграмма (гистограмма)



# Правила оформления столбцевых диаграмм

- Важнейший признак любой диаграммы - масштаб.
- Поэтому чтобы правильно построить фигурную диаграмму, необходимо определить единицу счета. В качестве последней принимается отдельная фигура (символ), которой условно присваивается конкретное численное значение.
- исследуемая статистическая величина изображается отдельным количеством одинаковых по размеру фигур, последовательно располагающихся на рисунке.
- Однако в большинстве случаев не удастся изобразить статистический показатель целым количеством фигур. Последнюю из них приходится делить на части, так как по масштабу один знак является слишком крупной единицей измерения.
- Обычно эта часть определяется на глаз. Сложность точного ее определения является недостатком диаграмм. Однако большая точность представления статистических данных не преследуется, и результаты получаются вполне удовлетворительными.

## НЕКОТОРЫЕ ПРЕДОСТЕРЕЖЕНИЯ

- Во-первых, они должны быть и наглядны и конструктивны. Как часть самого исследовательского процесса, они могут быть чрезвычайно полезны для наиболее глубокой разработки нашей концепции и для твердого понимания того, что же наши данные говорят нам. Гибкость и способность к новым формам анализа могут значительно помочь в углублении наших знаний о политических событиях, а методики, даже такие простые, как эти, окажут действие в оформлении выводов.

# НЕКОТОРЫЕ ПРЕДОСТЕРЕЖЕНИЯ

- Во-вторых, таблицы и графики нужно правильно применять. Даже из этого короткого обзора ясно, что совершенно не исключена возможность ошибочного представления результатов из-за небольшого искажения методики, и точно так же не исключена возможность, что кто-то будет одурачен этим бездумным злоупотреблением.
- У исследователей существуют моральные обязательства по отношению к другим – излагать свои выводы не только точно, но и правдиво – и интеллектуальные обязательства по отношению к себе – тщательно их проверять. Эти обязательства составляют краеугольный камень исследования. Мы не должны забывать о них.

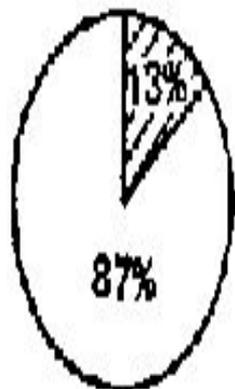
# НЕКОТОРЫЕ ПРЕДОСТЕРЕЖЕНИЯ

- Наконец, ваша работа не должны быть перегружена таблицами и графиками. Обилие подобного материала забивает текст и ухудшает его восприятие.
- Авторское решение включить таблицу или диаграмму в текст воспринимается читателем как сигнал к тому, что автор придает особое значение заключенной в них информации. Автор обязан делать выбор, а не просто предлагать читателю "шведский стол" информации.
- Такая осмотрительность не только повышает ценность работы, но и вынуждает его или ее подумать и решить, что важнее, и, таким образом, вносит дополнительный вклад в работу.

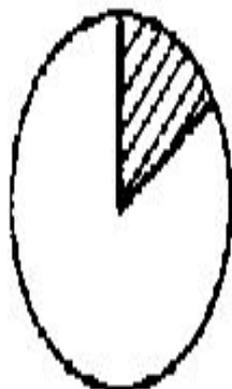
# Секторная диаграмма

- Рисунок, на котором круг представляет собой определенную совокупность и разделен на сегменты, показывающие относительную долю каждой части
- Применяется для интерпретации изменений относительных величин

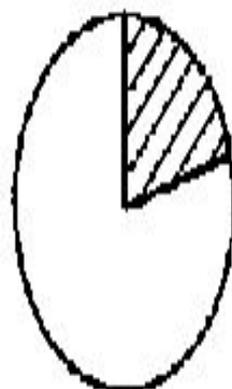
# Секторная диаграмма: расовая составляющая демократической коалиции, 1960-1976 гг



(а) 1960



(б) 1964



(в) 1968



(г) 1972



(д) 1976



Процент белых избирателей



Процент чернокожих избирателей

# Секторная диаграмма

- Рисунок, на котором круг представляет собой определенную совокупность и разделен на сегменты, показывающие относительную долю каждой части
- Применяется для интерпретации изменений относительных величин

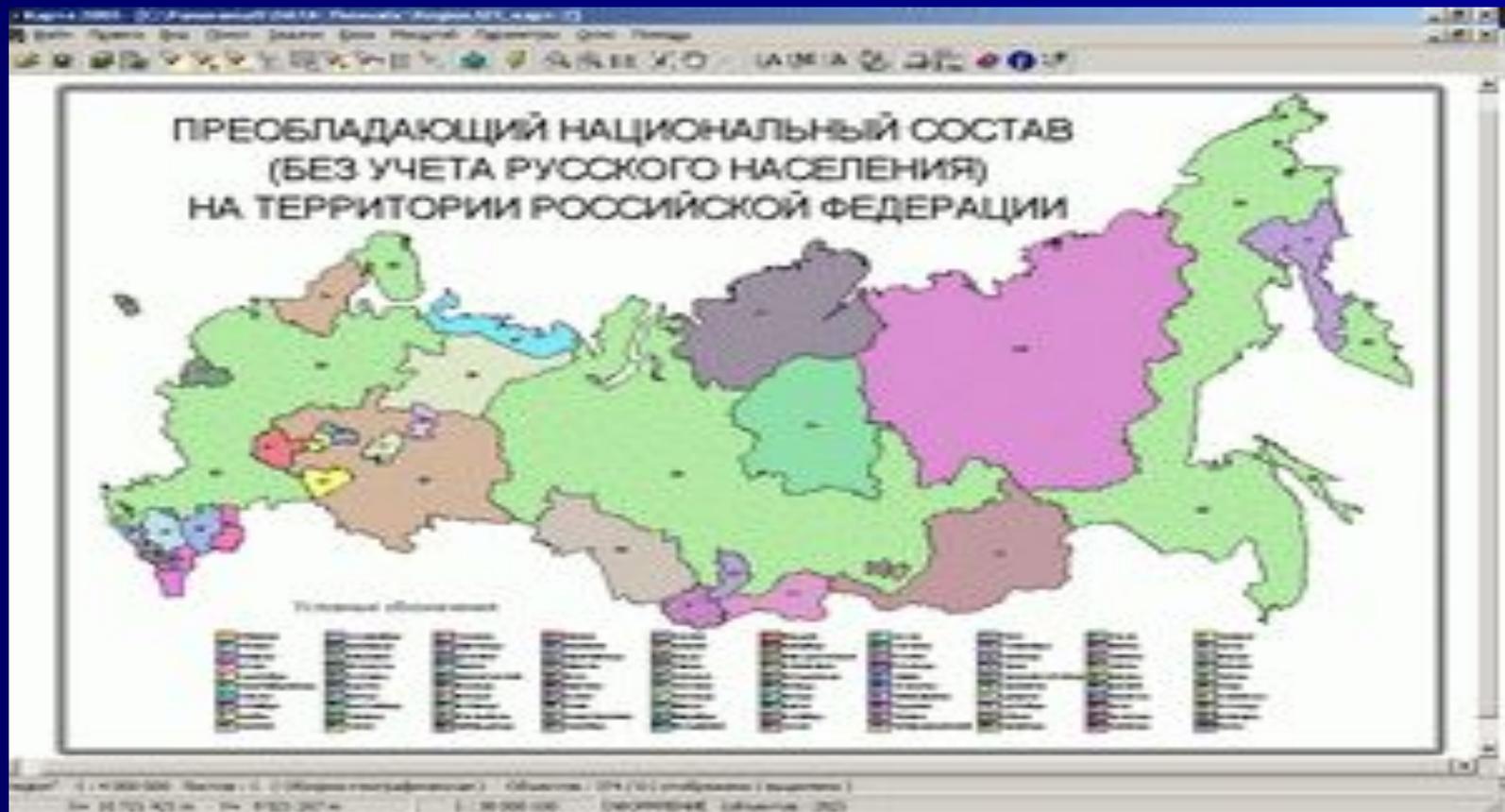
# Картограммы и картодиаграммы

- Статистические карты представляют собой вид графических изображений статистических данных на схематичной географической карте, характеризующих уровень или степень распространения того или иного явления на определенной территории.
- Средствами изображения территориального размещения являются штриховка, фоновая раскраска или геометрические фигуры. Различают картограммы и картодиаграммы.
- Картограммы - это схематическая географическая карта, на которой штриховкой различной густоты, точками или окраской определенной степени насыщенности показывается сравнительная интенсивность какого-либо показателя в пределах каждой единицы нанесенного на карту территориального деления (например, плотность населения по областям или республикам распределения районов по урожайности зерновых культур и т. п.).
- Картограммы делятся на фоновые и точечные.

# Картограммы

- Статистические карты - графики количественного распределения по поверхности. По своей основной цели они близко примыкают к диаграммам и специфичны лишь в том отношении, что представляют собой условные изображения статистических данных на контурной географической карте, т. е. показывают пространственное размещение или пространственную распространенность статистических данных. Геометрические знаки - это либо точки, либо линии или плоскости, либо геометрические тела. В соответствии с этим различают графики точечные, линейные, плоскостные и пространственные (объемные).

# Картограмма пространственная





# картограммы и картодиаграммы.

- Основное их различие заключается в том, что картограмма показывает распределение одной характеристики между указанными объектами карты, а по картодиаграммам можно оценить соотношение нескольких характеристик для каждого из объектов.
- Например, на размещенной слева иллюстрации, численность населения в регионах отображена штриховой картограммой, а процентное соотношение национальностей, населяющих каждый из регионов, показано в виде секторных картодиаграмм.

# картограммы и картодиаграммы.

- Картограммы могут быть цветными (заливка полигонов различными цветами), штриховыми (нанесение на полигоны различного вида штриховок) и точечными (заполнение полигонов массивами точек различного размера и степени сгущения).
- Цветные картограммы могут быть градиентными и мозаичными. Цвета полигонов на цветной градиентной картограмме плавно изменяются от одного до другого цвета в зависимости от значения анализируемой характеристики. Для построения градиентной заливки может быть использовано до 7 основных цветов). В случае, если градиент формируется по одному цвету, полигоны будут раскрашены в оттенки одного цвета (от белого до черного с переходом через указанный цвет).

# Резюме о графических изображениях

- Во-первых, они должны быть и наглядны и конструктивны.
- Как часть самого исследовательского процесса, они могут быть чрезвычайно полезны для наиболее глубокой разработки нашей концепции и для твердого понимания того, что же наши данные говорят нам.
- Гибкость и способность к новым формам анализа могут значительно помочь в углублении наших знаний о политических событиях, а методики, даже такие простые, как эти, окажут действие в оформлении выводов.

# Резюме о графических изображениях

- Во-первых, они должны быть и наглядны и конструктивны.
- Как часть самого исследовательского процесса, они могут быть чрезвычайно полезны для наиболее глубокой разработки нашей концепции и для твердого понимания того, что же наши данные говорят нам.
- Гибкость и способность к новым формам анализа могут значительно помочь в углублении наших знаний о политических событиях, а методики, даже такие простые, как эти, окажут действие в оформлении выводов.

# Резюме о графических изображениях

- Во-вторых, таблицы и графики нужно правильно применять.
- совершенно не исключена возможность ошибочного представления результатов из-за небольшого искажения методики, и точно так же не исключена возможность, что кто-то будет одурачен этим бездумным злоупотреблением.
- У исследователей существуют моральные обязательства по отношению к другим – излагать свои выводы не только точно, но и правдиво – и интеллектуальные обязательства по отношению к себе – тщательно их проверять. Эти обязательства составляют краеугольный камень исследования. Мы не должны забывать о НИХ.

# Резюме о графических изображениях

- Наконец, ваша работа не должны быть перегружена таблицами и графиками.
- Обилие подобного материала забивает текст и ухудшает его восприятие.
- Авторское решение включить таблицу или диаграмму в текст воспринимается читателем как сигнал к тому, что автор придает особое значение заключенной в них информации.
- Автор обязан делать выбор, а не просто предлагать читателю “шведский стол” информации.
- Такая осмотрительность не только повышает ценность работы, но и вынуждает его или ее подумать и решить, что важнее, и, таким образом, вносит дополнительный вклад в работу

# Резюме

- При построении таблиц и графиков в автоматическом режиме применяются программы WORD и EXCELL

**Спасибо!**



