СТАТИСТИЧЕСКИЙ АНАЛИЗ В ЭКОНОМИКЕ

(теория статистики)

1) Предмет и метод статистической науки

2) Статистическое наблюдение

3) Сводка и группировка статистических данных

4) Понятие о статистической таблице

5) Графические изображения в статистике

6) Абсолютные и относительные величины
Средние величины, показатели вариации

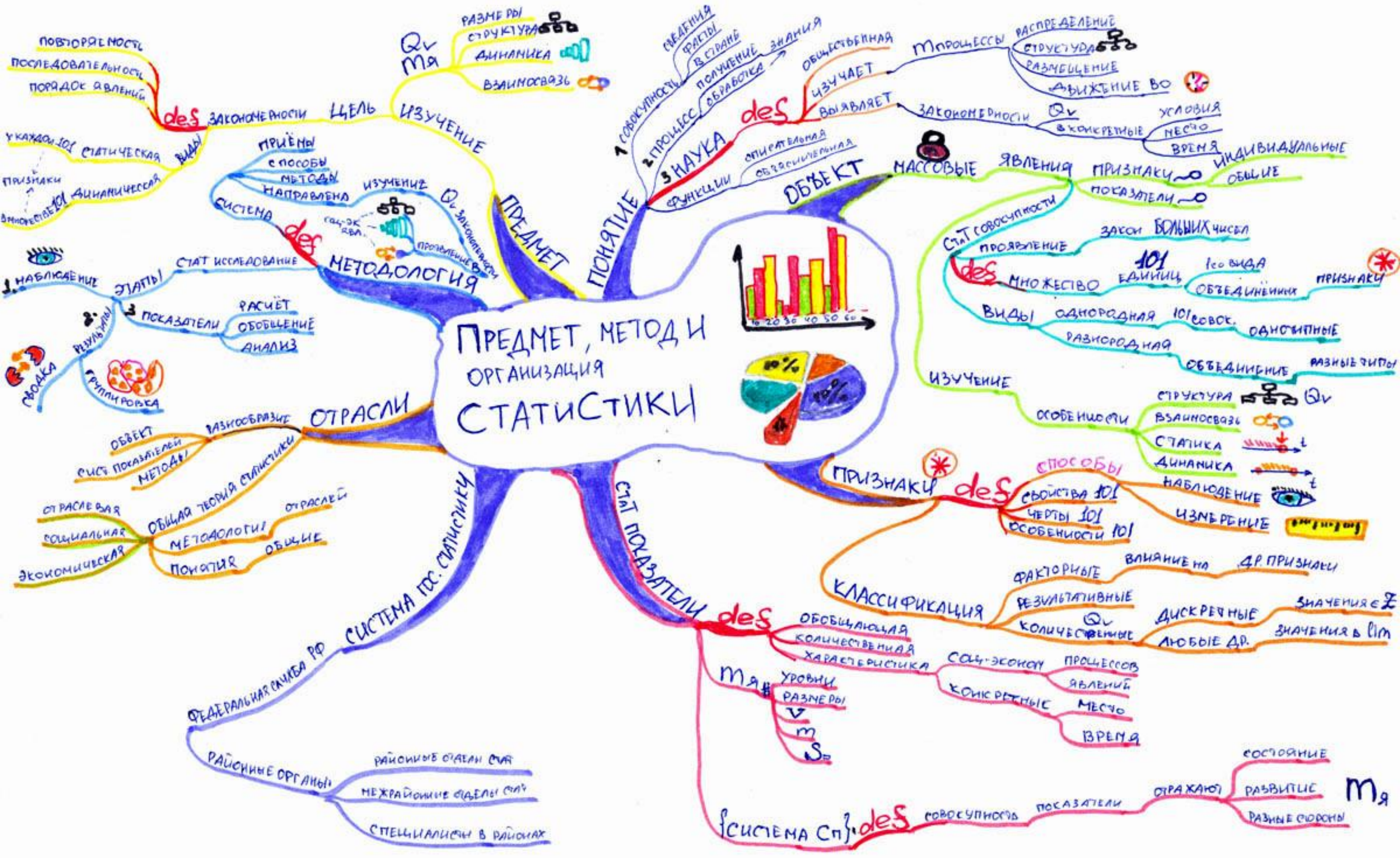
7) Индексы

8) Ряды динамики

9) Измерения связи

10) Основы выборочного наблюдения

ПРЕДМЕТ И МЕТОД СТАТИСТИЧЕСКОЙ НАУКИ



101 - единица
def - определение
-o - элемент
mя - массовые явления
Qv - количественность
Qy - качественность

Китай, более 2000 лет до н. э	Учет населения по полу, возрасту, учет сведений о промышленности и сельском хозяйстве
Древний Рим	Учет населения и имущественного положения граждан
1-я половина XVII века, Англия	Школа «Политических арифметиков» (демография), впервые проводится анализ данных
2-я половина XVII века, Германия	Школа государствоведения. Описание устройств государств, их быта, нравов населения
1749 г., г.Ахенваль, Германия	Введение нового термина СТАТИСТИКА
1-я половина XIX века, А.Кетле – Бельгия, Ф.Гальтон и К.Пирсон - Англия	Возникновение Статистико-Математического направления

Основоположники математической статистики

		
Вильям ГОССЕТ - "СТЬЮДЕНТ" (W.S.Hosset –"Student" 1876-1937) известный английский статистик	**Рональд Эйлмер ФИШЕР** (**R.A.Fisher** 1890-1962) английский математик и генетик, основоположник современной прикладной статистики	**Карл ПИРСОН** (**Carl Pearson,** .1857-1936) английский математик, биолог- евгеник, фипософ-позитивист

Основоположники математической статистики

		
Стефан Альфред ФОРБС (**S. A. Forbes,** 1844-1930) первым предложил рассматривать озеро как единую экосистему, микрокосм	**Раймонд Лаурел ЛИНДЕМАН** (**R.L. Lindeman,** 1915-1942) автор всего шести публикаций, определивших трофико-динамическое направление исследований в экологии	**Август ТИНЕМАН** (**A. Thienemann,** 1882-1960) гидробиолог и лимнолог, сформулировал «биотические принципы», ввел понятие «продукция»

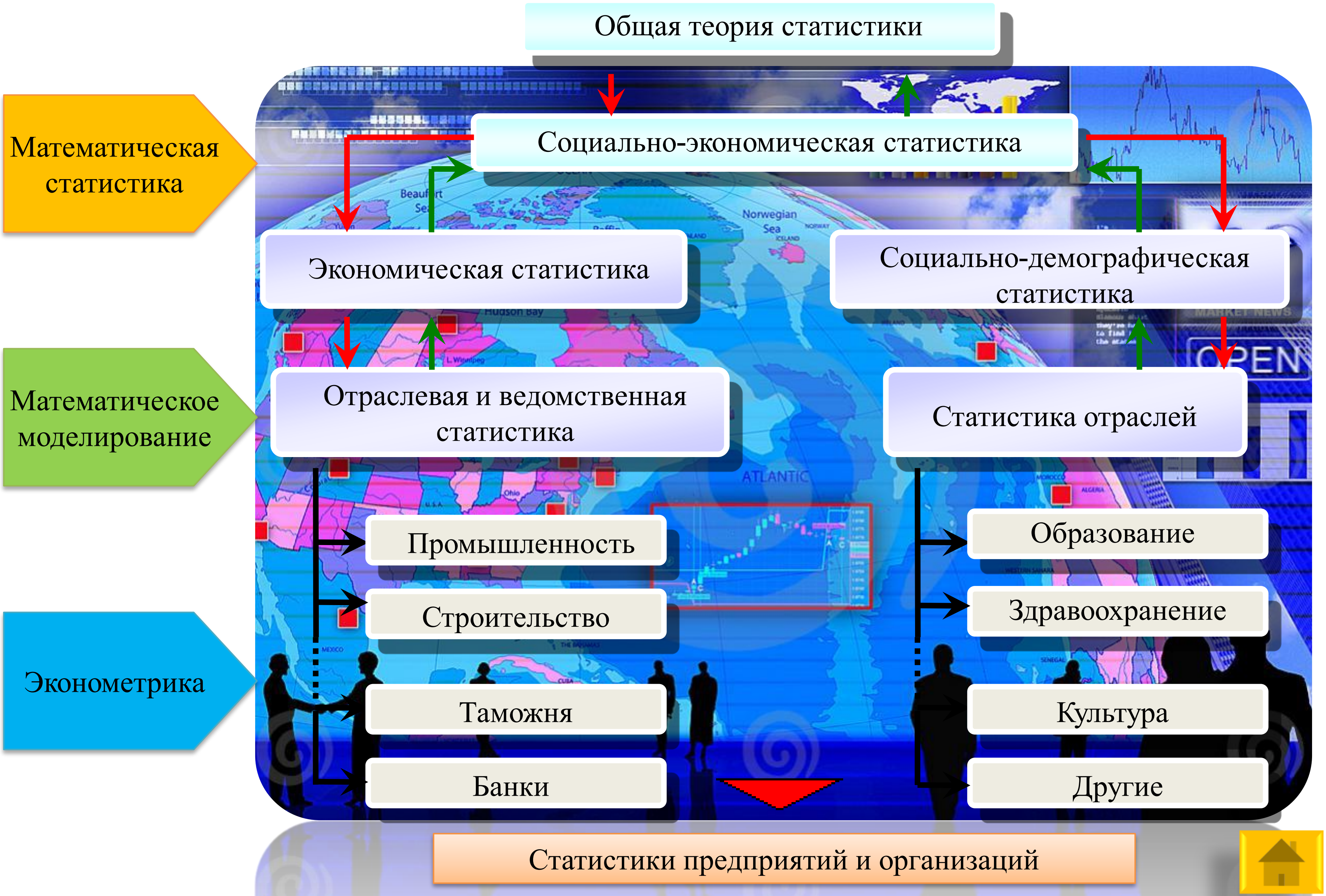
Основоположники математической статистики

		
Клод Элвуд ШЕННОН (**C.E. Shannon**, 1916 - 2001) "отец" математической теории информации	**КОЛМОГОРОВ Андрей Николаевич** (1903-1987) выдающийся отечественный математик	**НАЛИМОВ Василий Васильевич** (1910-1997), математик, крупный специалист по наукометрии и теории эксперимента

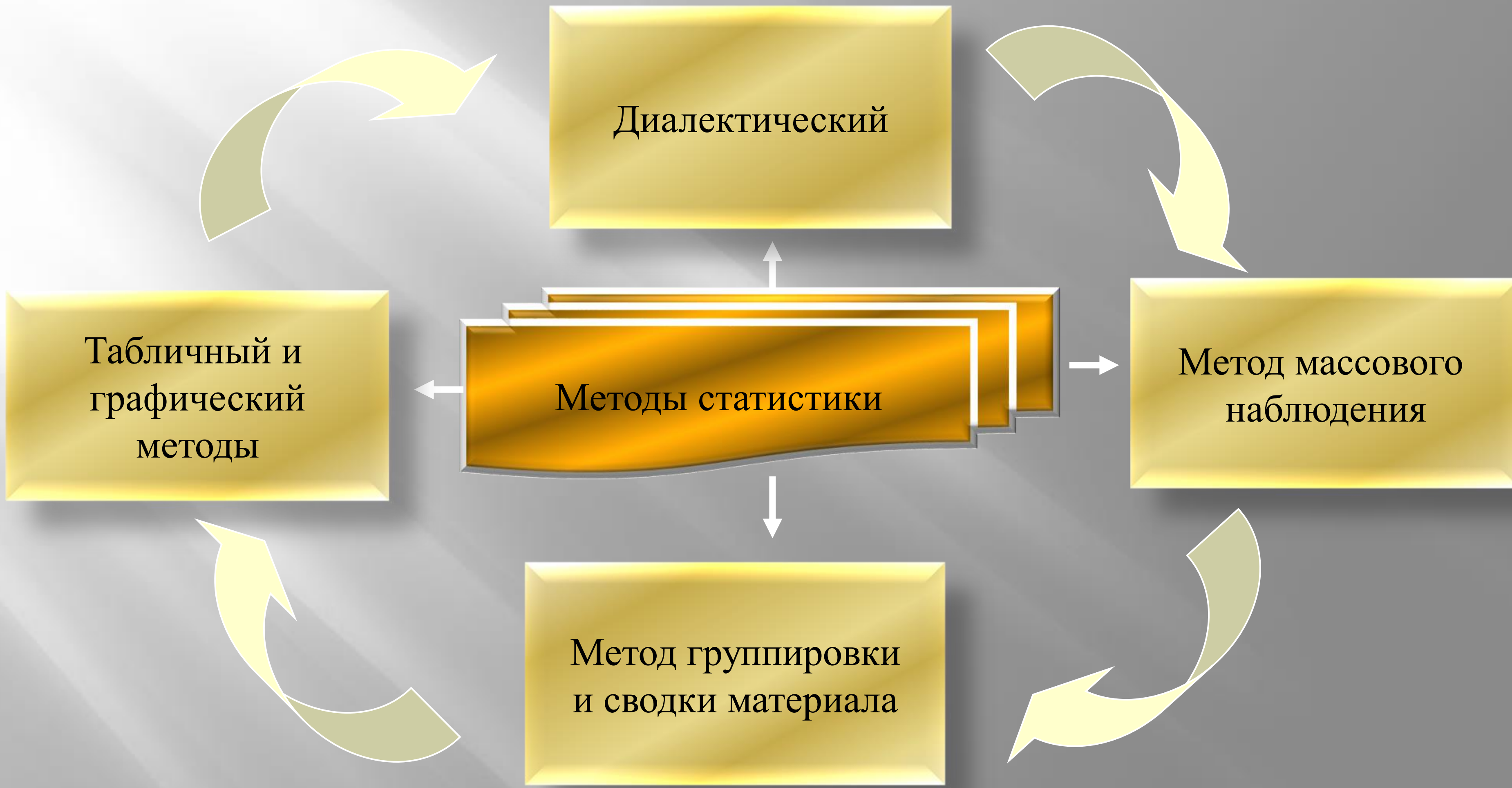
Основоположники математической статистики

		
Джордж Эвелин **ХАТЧИНСОН** (**G.E. Hutchinson**, 1903 - 1991) гидробиолог, лимнолог и эколог, развивший экосистемные представления; предложил современное понятие «экологической ниши»	**Виктор Сергеевич ИВЛЕВ** (1907-1964) ихтиолог, гидробиолог	**Георгий Георгиевич ВИНБЕРГ** (1905-1987) крупный специалист в области гидробиологии и продуктивности экосистем

Структура современной статистической науки



Методы статистики



Классификация признаков

Виды признаков

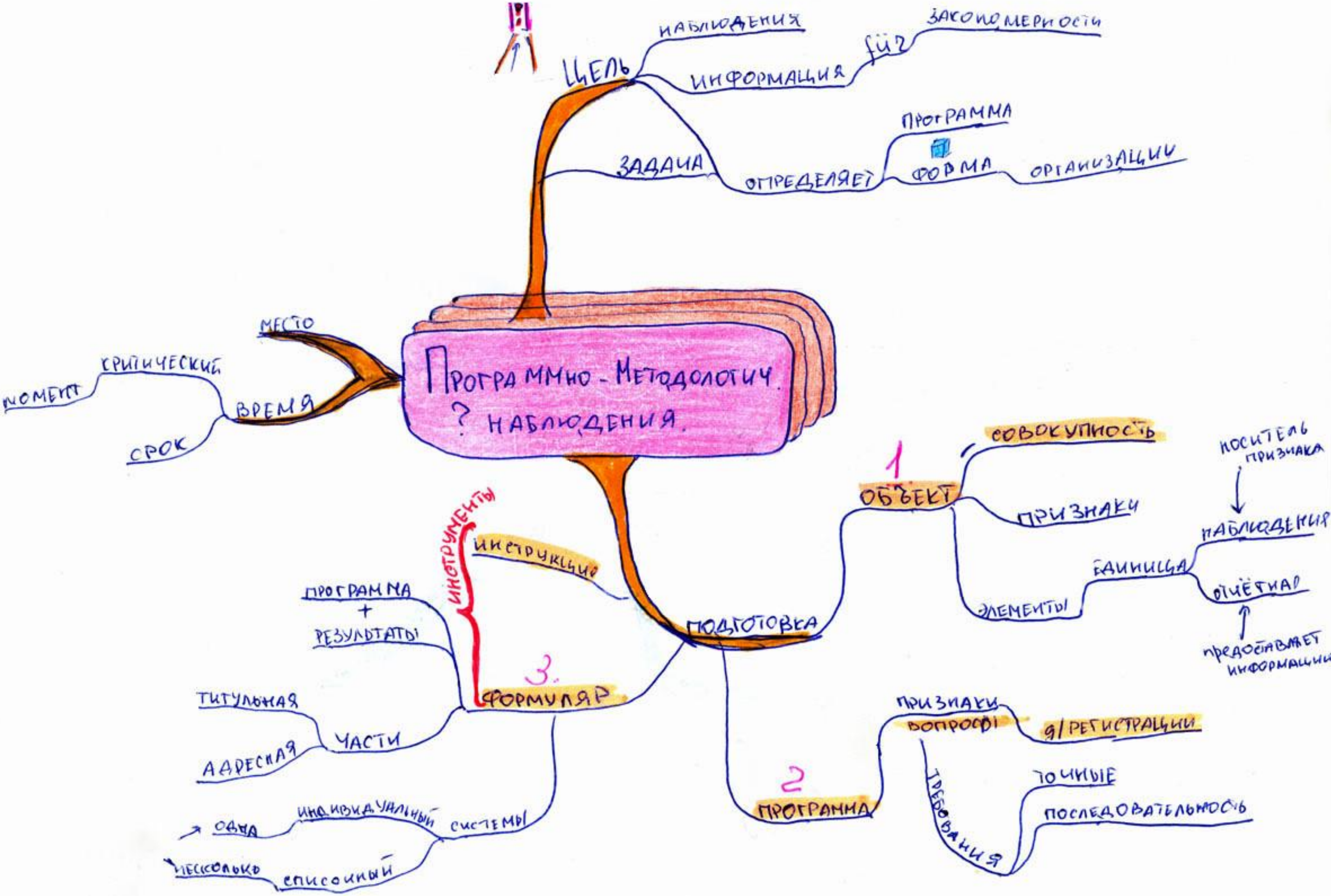
- Количественные
  - Дискретные
  - Непрерывные
- Качественные

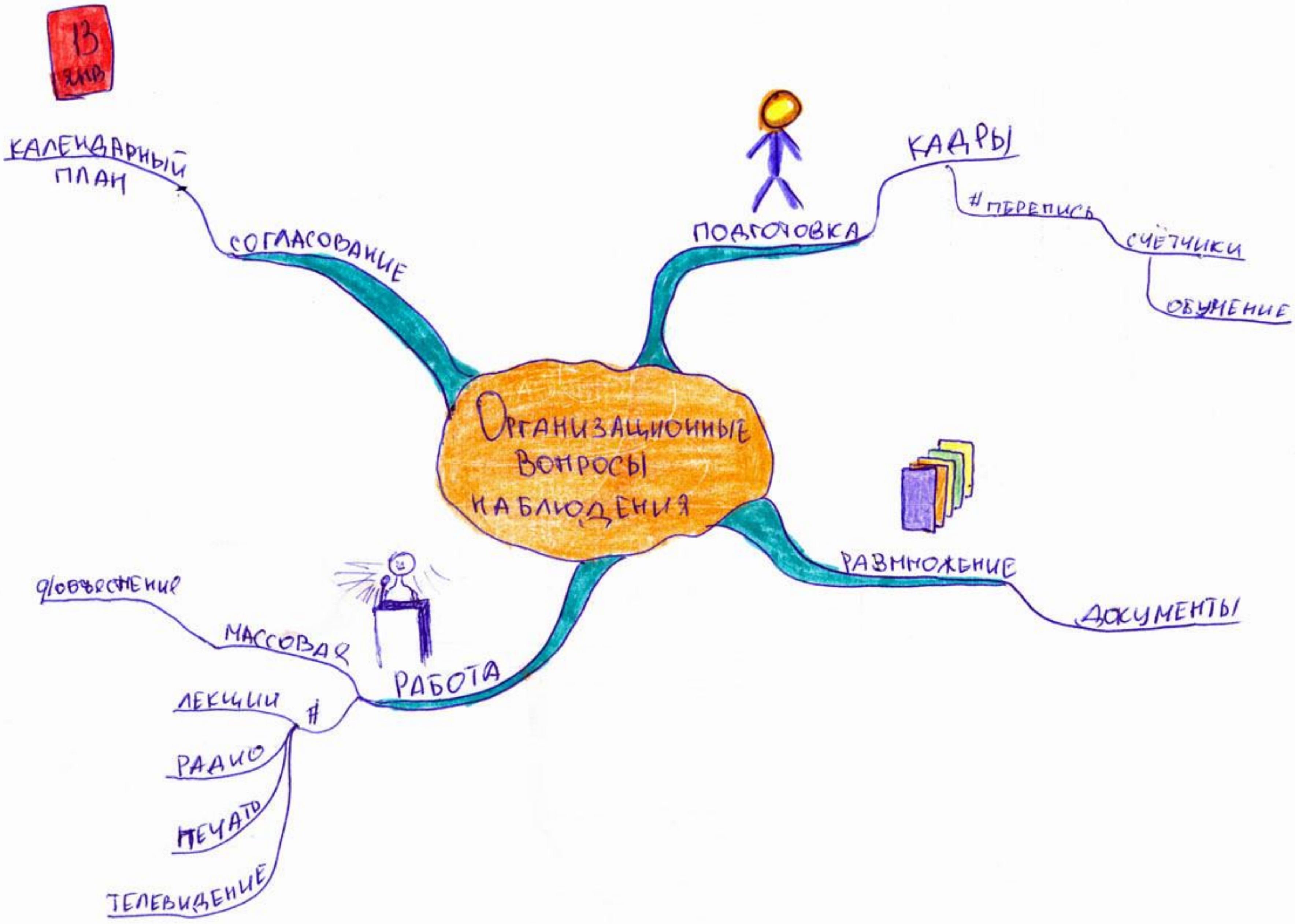
Виды статистических признаков

СТАТИСТИЧЕСКИЕ ПРИЗНАКИ

- По характеру выражения
  - Количественные
  - Качественные
- По целям исследования
  - Существенные
  - Несущественные
- По способу измерения
  - Учетные
  - Расчетные
- По способу взаимовлияния
  - Результативные
  - Факторные
- По характеру вариации
  - Дискретные
  - Непрерывные
  - Альтернативные
- По отношению ко времени
  - Моментные
  - Интервальные

СТАТИСТИЧЕСКОЕ НАБЛЮДЕНИЕ


Программно-Методологич. ? наблюдения.
ЦЕЛЬ
НАБЛЮДЕНИЯ
ИНФОРМАЦИЯ
für
ЗАКОНОМЕРНОСТИ
ЗАДАЧА
ОПРЕДЕЛЯЕТ
ПРОГРАММА
ФОРМА
ОРГАНИЗАЦИИ
МЕСТО
ВРЕМЯ
КРИТИЧЕСКИЙ
МОМЕНТ
СРОК
1
ОБЪЕКТ
СОВОКУПНОСТЬ
ПРИЗНАКИ
ЭЛЕМЕНТЫ
ЕДИНИЦА
НАБЛЮДЕНИЯ
НОСИТЕЛЬ ПРИЗНАКА
ОТЧЁТНАЯ
ПРЕДОСТАВЛЯЕТ ИНФОРМАЦИЮ
ПОДГОТОВКА
2
ПРОГРАММА
ПРИЗНАКИ
ВОПРОСЫ
Д/ РЕГИСТРАЦИИ
ТРЕБОВАНИЯ
ТОЧНЫЕ
ПОСЛЕДОВАТЕЛЬНОСТЬ
ИНСТРУМЕНТЫ
ИНСТРУКЦИИ
3.
ФОРМУЛЯР
ПРОГРАММА
+
РЕЗУЛЬТАТЫ
ЧАСТИ
ТИТУЛЬНАЯ
АДРЕСНАЯ
СИСТЕМЫ
ИНДИВИДУАЛЬНЫЙ
ОДНА
СПИСОЧНЫЙ
НЕСКОЛЬКО



13 янв
КАЛЕНДАРНЫЙ ПЛАН
СОГЛАСОВАНИЕ
ПОДГОТОВКА
КАДРЫ
ПЕРЕПИСЬ
СЧЁТЧИКИ
ОБУЧЕНИЕ
ОРГАНИЗАЦИОННЫЕ ВОПРОСЫ НАБЛЮДЕНИЯ
РАЗМНОЖЕНИЕ
ДОКУМЕНТЫ
РАБОТА
МАССОВАЯ
Д/ОБЪЯСНЕНИЕ
ЛЕКЦИИ
#
РАДИО
ПЕЧАТЬ
ТЕЛЕВИДЕНИЕ

Схема проведения статистического наблюдения

Подготовка наблюдения

↓

Проведение массового сбора данных, их фиксация на носителе информации

↓

Всесторонняя проверка и контроль полученных данных

↓

Подготовка данных к автоматизированной обработке

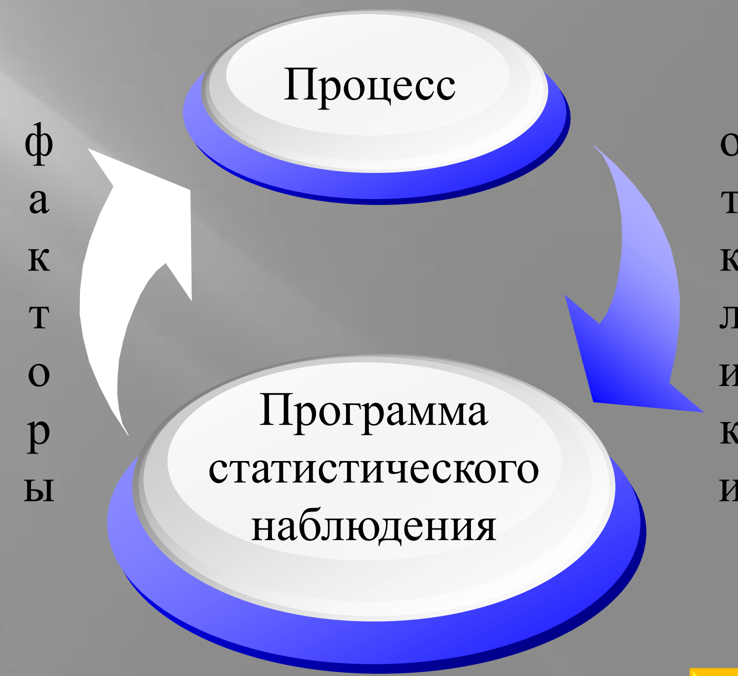
↓

Разработка предложений по совершенствованию статистического наблюдения

Схемы активного эксперимента и статистического наблюдения

Активный эксперимент

Статистическое наблюдение



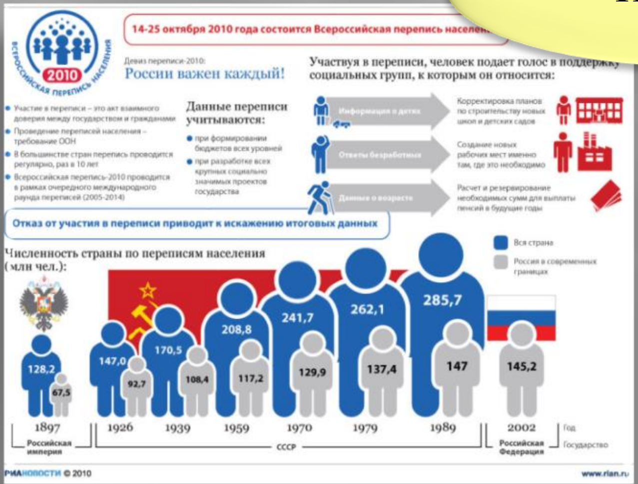
Классификация статистического наблюдения

Признак классификации	Виды наблюдений	Разновидности
1. По форме	1.1. Отчетность	По длительности (периодическая, годовая) По оперативности (срочная, почтовая) По использованию (внешняя, внутренняя)
	1.2. Специально организованное	–
2. По охвату единиц совокупности	2.1. Сплошное	–
	2.2. Несплошнное	Обследование основного массива Выборочное Монографическое
3. По фактору времени	3.1. Непрерывное	–
	3.2. Прерывное	Единовременное Периодическое
4. По способу учета факторов	4.1. Непосредственный учет	–
	4.2. Документальный учет	–
	4.3. Опрос	Экспедиционный Саморегистрация Корреспондентский

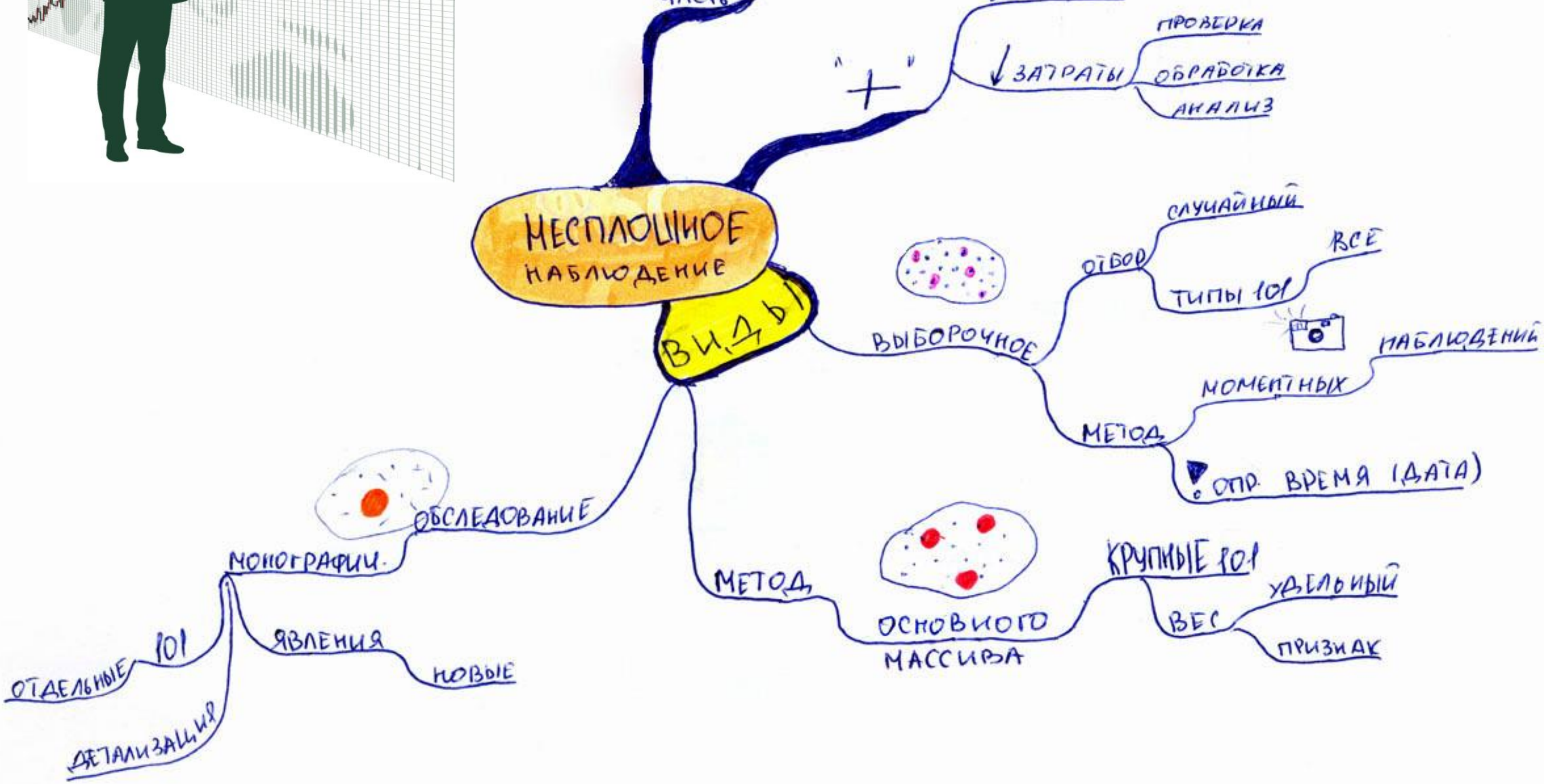
Формы статистического наблюдения

Формы организации статистического наблюдения

- Отчетность
  - типовая
  - специализированная
  - текущая
  - годовая
- Регистры
- Специально организованное статистическое наблюдение
  - Перепись

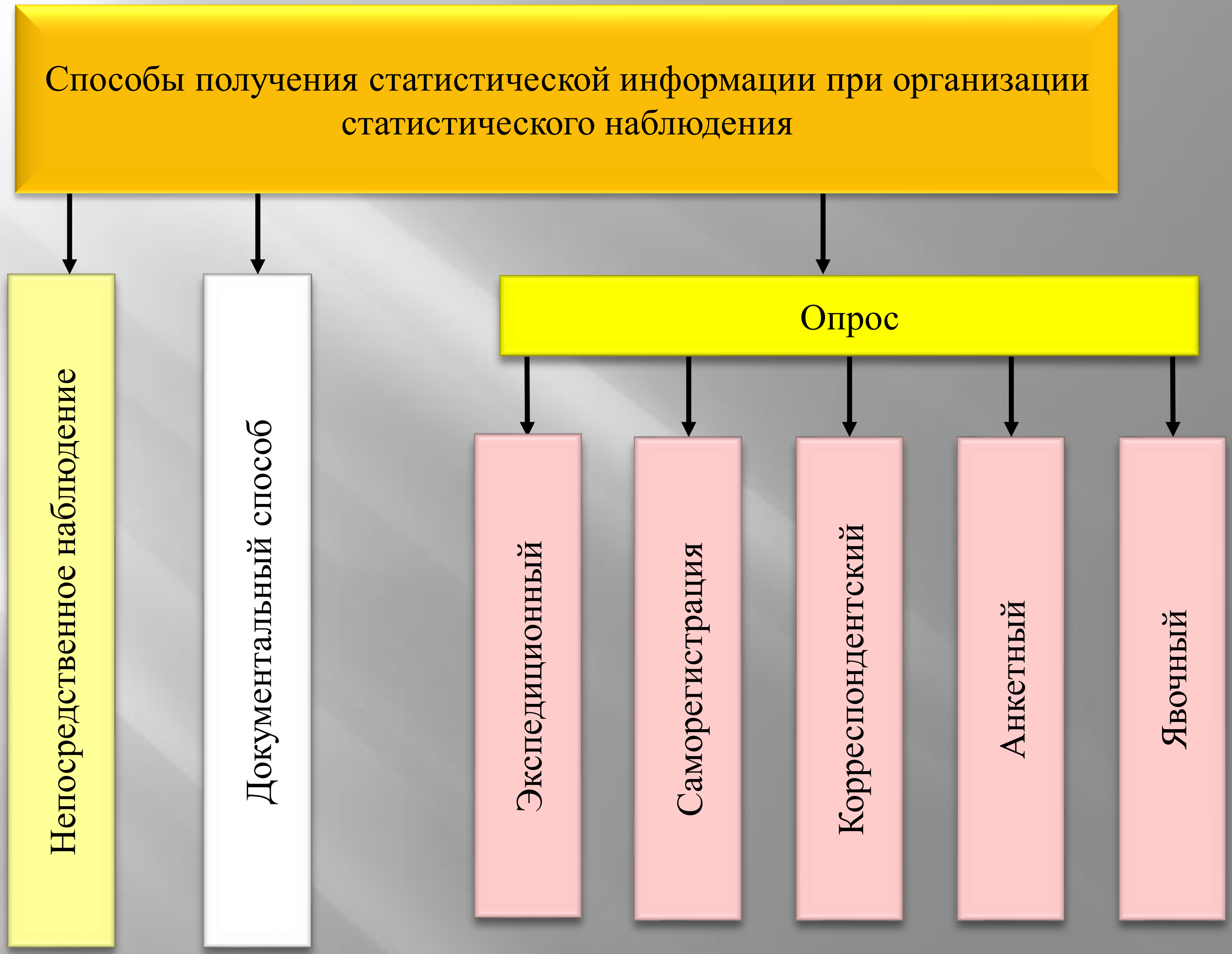



Виды статистического наблюдения
По времени регистрации фактов
По охвату единиц совокупности
Прерывное
Текущее или непрерывное
Несплошное
Сплошное
Периодическое
Единовременное
Выборочное
Монографическое
Основного массива



НЕСПЛОШНОЕ НАБЛЮДЕНИЕ
ВИДЫ
ЧАСТЬ
СОВОКУПНОСТИ
"+"
БЫСТРЕЕ
↓ЗАТРАТЫ
ПРОВЕРКА
ОБРАБОТКА
АНАЛИЗ
ВЫБОРОЧНОЕ
ОТБОР
СЛУЧАЙНЫЙ
ТИПЫ 101
ВСЕ
МЕТОД
МОМЕНТНЫХ
НАБЛЮДЕНИЙ
! ОПР. ВРЕМЯ (ДАТА)
МЕТОД
ОСНОВНОГО
МАССИВА
КРУПНЫЕ 101
ВЕС
УДЕЛЬНЫЙ
ПРИЗНАК
ОБСЛЕДОВАНИЕ
МОНОГРАФИИ.
ЯВЛЕНИЯ
НОВЫЕ
101
ОТДЕЛЬНЫЕ
ДЕТАЛИЗАЦИЯ

Способы получения информации про проведении статистического исследования



Виды и содержание контроля

Способы контроля

Синтаксический

-структура документа

-полнота документа

- полнота заполнения строк

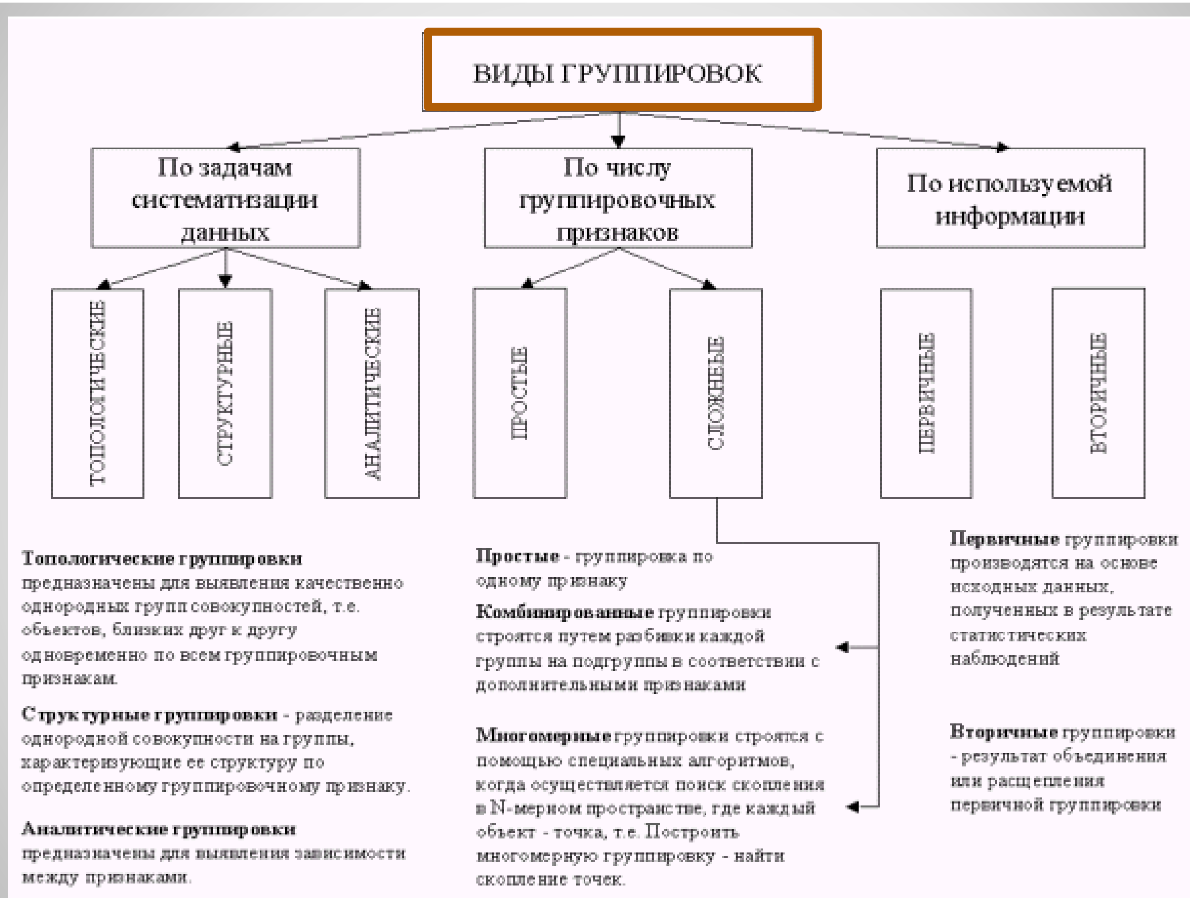
Логический

-соответствие кодов и наименований признаков
- наличие отклонений от заданных величин
- наличие логических связей между показателями

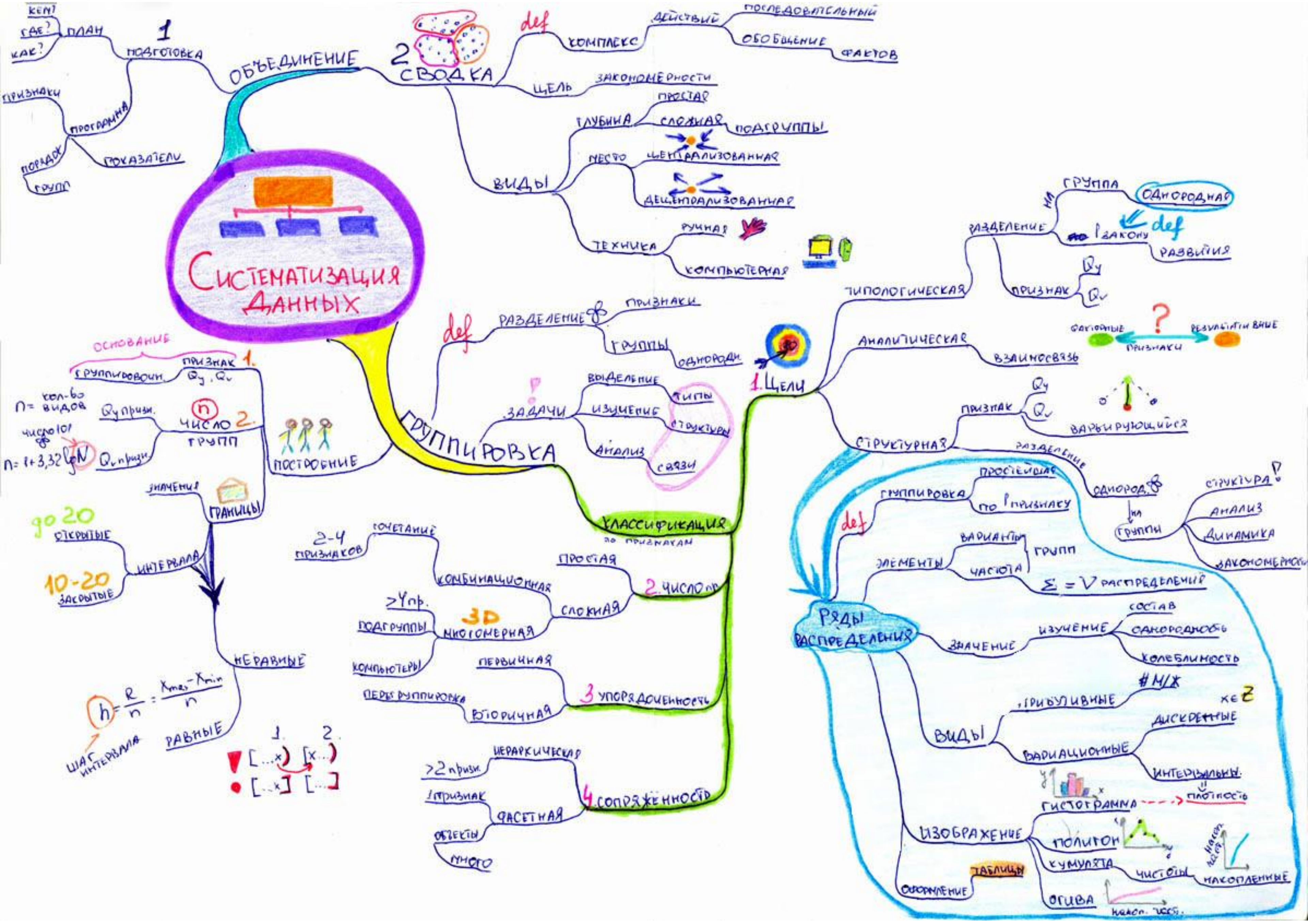
Арифметический

-соответствие построчных контрольных сумм документа и контрольных сумм ЭВМ
- соответствие пографных контрольных сумм документа и сумм ЭВМ

Предметы контроля



Сводка и группировка статистических данных



Систематизация данных
1 подготовка
план
кем?
где?
как?
программа
признаки
показатели
порядок групп
объединение
2 сводка
def
комплекс
действий
последовательный
обобщение
фактов
цель
закономерности
виды
глубина
простая
сложная
подгруппы
место
централизованная
децентрализованная
техника
ручная
компьютерная
группировка
def
разделение
признаки
группы
однородн.
задачи
выделение
типы
изучение
структуры
анализ
связи
1. цели
типологическая
разделение
на
группа
однородная
def
по закону
развития
признак
Q_y
Q_v
аналитическая
взаимосвязь
факторные
признаки
результативные
структурная
признак
Q_y
Q_v
варьирующийся
разделение
однород.
на
группы
структура
анализ
динамика
закономерность
классификация
по признакам
2. число пр.
простая
сложная
комбинационная
сочетание
2-4 признаков
многомерная
3D
>4 пр.
подгруппы
компьютеры
3 упорядоченность
первичная
вторичная
перегруппировка
4 сопряженность
иерархическая
>2 призн.
фасетная
1 признак
объекты
много
основание
1. признак
группировочн.
Q_y, Q_v
2. число групп
n
n = кол-во видов
Q_y призн.
Q_v призн.
число 10!
$n = 1 + 3{,}32 \lg N$
построение
границы
значения
интервалы
до 20 открытые
10-20 закрытые
неравные
равные
$h = \frac{R}{n} = \frac{x_{max} - x_{min}}{n}$
шаг интервала
1.
2.
[...x) [x...)
[...x] [...]
ряды распределения
def
группировка
простейшая
по 1 признаку
элементы
варианты
частота
групп
$\Sigma = V$ распределений
значение
изучение
состав
однородность
колеблимость
виды
атрибутивные
вариационные
дискретные
интервальные
$x \in Z$
изображение
гистограмма
плотность
полигон
кумулята
частоты
накопленные
огива
накоп. частоты
оформление
таблицы

Виды сводки

Виды сводки статистических данных

- По глубине обработки материала
  - Простая
  - Сложная
- По форме обработки материала
  - Децентрализованная
  - Централизованная
- По технике выполнения
  - Механизированная (компьютерная)
  - Ручная

Этапы программы статистической сводки

Выбор группировочных признаков

Определение порядка формирования групп

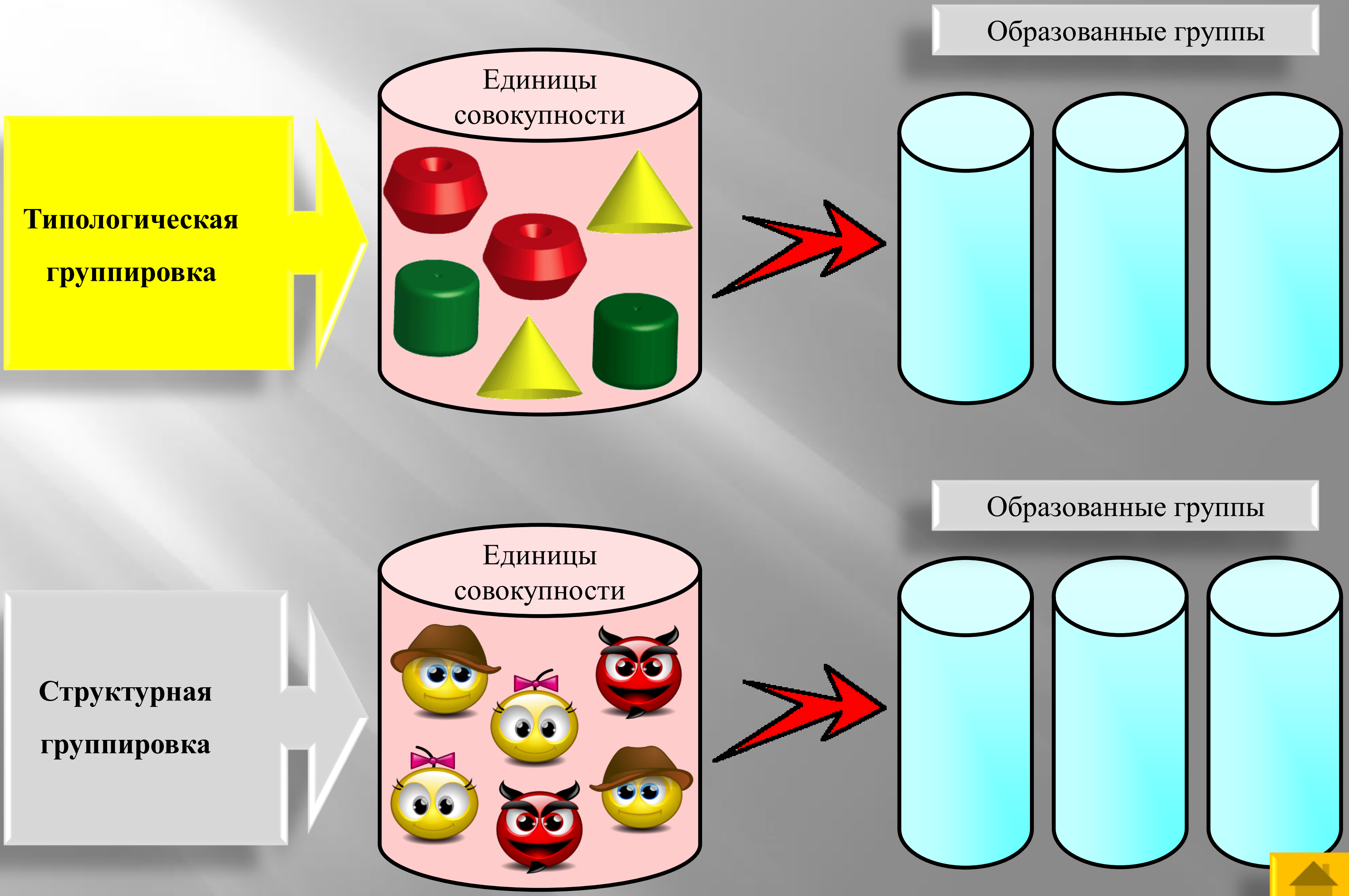
Разработка системы статистических показателей для характеристики групп и объекта вцелом

Разработка макетов статистических таблиц для предоставления результатов сводки

Виды статистических группировок

- По задачам систематизации данных
  - Изучение социально-экономических явлений
    - Типологическая
  - Анализ структуры явления и структурных сдвигов
    - Структурная
  - Выявление связей и зависимостей между отдельными признаками
    - Аналитическая (факторная)
- По числу группировочных признаков
  - Простые
  - Сложные
    - Комбинационная
    - Многомерная
- По используемой информации
  - По атрибутивным признакам
  - По количественным признакам
    - Групповая
    - Интервальная

Разделение разнородной совокупности на однородные группы



Типологические группировки

задача – выявление социально-экономических типов или однородных в существенном отношении групп.

№ п/п	Социально-экономические типы	Мужчины		Женщины	
		1980	1992	1980	1992
1.	Работники	–	–	–	–
2.	Крестьяне	–	–	–	–
3.	Служащие	–	–	–	–

Структурные группировки

задача – изучение состава отдельных типических групп при помощи объединения единиц совокупности, близких друг к другу по величине группировочного признака.

№ п/п	Количество посадочных мест	Количество столов	Число занятых	Товарооборот на 1 место
1.	до 25	–	–	–
2.	16 – 50	–	–	–
3.	51 – 70	–	–	–
4.	71 – 100	–	–	–

Аналитические группировки

задача – выявления влияния одних признаков на другие (выявить связь между социально-экономическими явлениями).

№ п/п	Группы магазинов по числу рабочих мест	Число магазинов	Товарооборот	
			на 1 работника	на 1 раб. место
1.	до 5	100	12,0	13,0
2.	6 – 10	50	14,0	16,0
3.	11 – 15	10	15,0	17,0
4.	16 – 20	4	30,0	39,0
5.	21 – 25	2	31,0	42,0

Комбинационные группировки

В них производится разделение совокупности на группы по двум или более признакам. При этом группы, образованные по одному признаку, разбиваются на подгруппы по другому признаку.

№ п/п	Группы предприятий по объему основных фондов	Оплата труда в рублях	Пол	Количество единиц
1.	до 200	100 – 120	М	–
			Ж	–
		120 – 140	М	–
			Ж	–
		140 – 160	М	–
			Ж	–
2.	200 – 400	100 – 120	М	–
			Ж	–
		120 – 140	М	–
			Ж	–
		140 – 160	М	–
			Ж	–

Группировка предприятий торговли в 2007 году
(цифры условные)

Пример типологической группировки

Группы предприятий	Количество предприятий	
	всего	% к итогу
Магазины	827	49,3
Предприятия мелкорозничной сети	366	21,8
Рынки	16	1
Предприятия оптовой торговли	150	9
Предприятия общественного питания	317	18,9
Всего	1676	100

Группировка экономически активного населения по уровню образования, %

Пример структурной группировки

Уровень образования	*Работающие по найму*	*Работающие не по найму*	*Безработные*	*Экономически активное население, всего*
А	1	2	3	4
Высшее	17,0	9,4	10,6	16,1
Неполное высшее	1,6	1,8	3,1	1,7
Среднеспециальное	32,1	22,6	27,3	31,1
Среднее общее	31,9	36,0	37,2	32,5
Неполное среднее	14,3	22,9	19,1	15,2
Начальное	3,1	7,3	2,7	3,4
Итого	100	100	100	100

Качество продукции и продолжительность договорных отношений поставщиков с торговыми предприятиями

Пример аналитической группировки

Продолжительность связей торговых предприятий с поставщиками, лет	Число поставщиков		Доля нестандартной и бракованной продукции, %
	Абсолютная величина	% к итогу	
А	***1***	***2***	***3***
До 1 года	8	16	11,2
1-5	18	36	9,1
5-10	14	28	5,8
Свыше 10	10	20	2,4
Итого	***50***	***100,0***	***7,17***

Распределение студентов 1-го курса кооперативного колледжа

Пример комбинационной группировки

Группы студентов по форме обучения	Группы студентов по факультету	Число студентов	В том числе	
			женщин	мужчин
Очное отделение	Товароведный	110	87	23
	Технологический	90	61	29
	Экономический	100	73	27
	Итого	***300***	***221***	***79***
Заочное отделение	Товароведный	31	27	4
	Технологический	35	23	12
	Экономический	26	19	7
	Итого	**92**	**69**	**23**
Вечернее отделение	Товароведный	50	41	9
	Технологический	40	33	7
	Экономический	60	39	21
	Итого	***150***	***113***	***37***
	Всего	*542*	*403*	*139*

Определение числа групп

Формула Стерджесса:

$$n = 1 + 3,322 \cdot \lg N$$

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

N – численность единиц совокупности;

n – число групп;

или

$$i = \frac{x_{\max} - x_{\min}}{1 + 3,322 \cdot \lg N}$$

Группировка торговых предприятий двух районов по торговой площади.

Первый район		Второй район	
Торговая площадь, $м^2$	Число торговых предприятий, % от их общего количества	Торговая площадь, $м^2$	Число торговых предприятий, % от их общего количества
10-50 50-100 100-200 200-600 600-1000	18 12 30 34 6	10-100 100-400 400-1000 свыше 1000	10 20 40 30
Итого	100	Итого	100

Вторичная группировка торговых предприятий по размеру торговой площади.

№ группы	Группы торговых предприятий по торговой площади, м^2	Удельный вес предприятий, % к итогу		Расчет
		Второй район	Первый район	
1	10-100	10	30	18 + 12 = 30
2	100-400	20	47	30 + 34 : 2 = 47
3	400-1000	40	23	34 : 2 + 6 = 23
4	Свыше 1000	30	-	—
	Итого	100	100	100

Пример вторичной группировки.
Произвести укрупнение интервалов

Группы магазинов по размеру товарооборота за IV квартал, тыс.руб.	Число магазинов	Товарооборот за IV квартал, тыс.руб.
До 10	15	93
10 — 15	8	112
15 — 20	13	200
20 — 30	3	68
30 — 50	9	378
50 — 60	7	385
60 — 70	3	180
70 — 100	8	600
100 — 200	22	2400
Свыше 200	12	3744
Итого	100	8160

Уплотним ряды распределения, образовав шесть групп

Группы магазинов по размеру товарооборота за IV квартал, тыс.руб.	Число магазинов	Товарооборот за IV квартал, тыс.руб.	Товарооборот в среднем на 1 магазин, тыс.руб.
До 10	15	93	6,2
10 — 20	21	312	14,8
20 — 50	12	446	37,1
50 — 100	18	1165	64,8
100 — 200	22	2400	109,0
Свыше 200	12	3744	312,0
Итого	100	8160	81,6

Основные показатели деятельности коммерческих банков одного из регионов (цифры условные), тыс. руб.

№ банка	Капитал	Работающие активы	Уставный капитал
1	31070	28101	3999
2	29918	47647	29699
3	13915	6141	4466
4	88889	104616	3572
5	36986	69624	39272
6	71584	236330	31765
7	36359	61435	8953
8	11678	14777	3788
9	57440	191513	11560
10	15419	24245	5925
11	53498	72019	23112
12	31058	50803	15256
13	12235	39998	3819
14	15328	21883	15409
15	35194	76128	6074
16	83 777	130651	12584
17	15521	51439	7254
18	24982	98 693	8708
19	23648	71457	16999
20 min	10135	26064	5056
21	33637	128275	5808
22	20426	54307	8125
23	14810	28193	8551
24	36034	65606	10389
25	34459	168557	10136
26 max	112615	298097	29273
27	84305	216888	34774
28	90985	244121	18192
29	22225	43795	5017
30	62276	306564	20558

Рассчитаем оптимальное количество групп в данном примере по формуле Стерджесса:

$$n = 1 + 3{,}322 \times \lg N = 1 + 3{,}322 \times \lg 30 \approx 5$$

$$h = \frac{\chi_{\max} - \chi_{\min}}{n} = \frac{112615 - 10135}{5} = 20496\ \textit{ò û ñ. ðóá.}$$

10135-30631 – 1-ая группа;
30631-51127 – 2-ая группа;
51127-71623 – 3-я группа;
71623-92119 – 4-ая группа;
92119-**112615** -5-ая группа.

Группировка малых и средних коммерческих банков одного из регионов по величине уставного капитала.

№ группы	Группы банков по величине уставного капитала, тыс. руб.	Число банков, ед.	Капитал, тыс. руб.	Работающие активы, тыс. руб.	Уставный капитал, тыс. руб.
1	10135-30631	13	230240	528639	122816
2	30631-51 127	8	274797	648529	99 887
3	51 127-71623	4	244798	806426	86995
4	71623-92119	4	347956	696276	69122
5	92119-112615	1	112615	298097	29273
Итого		30	1210406	2977967	408093

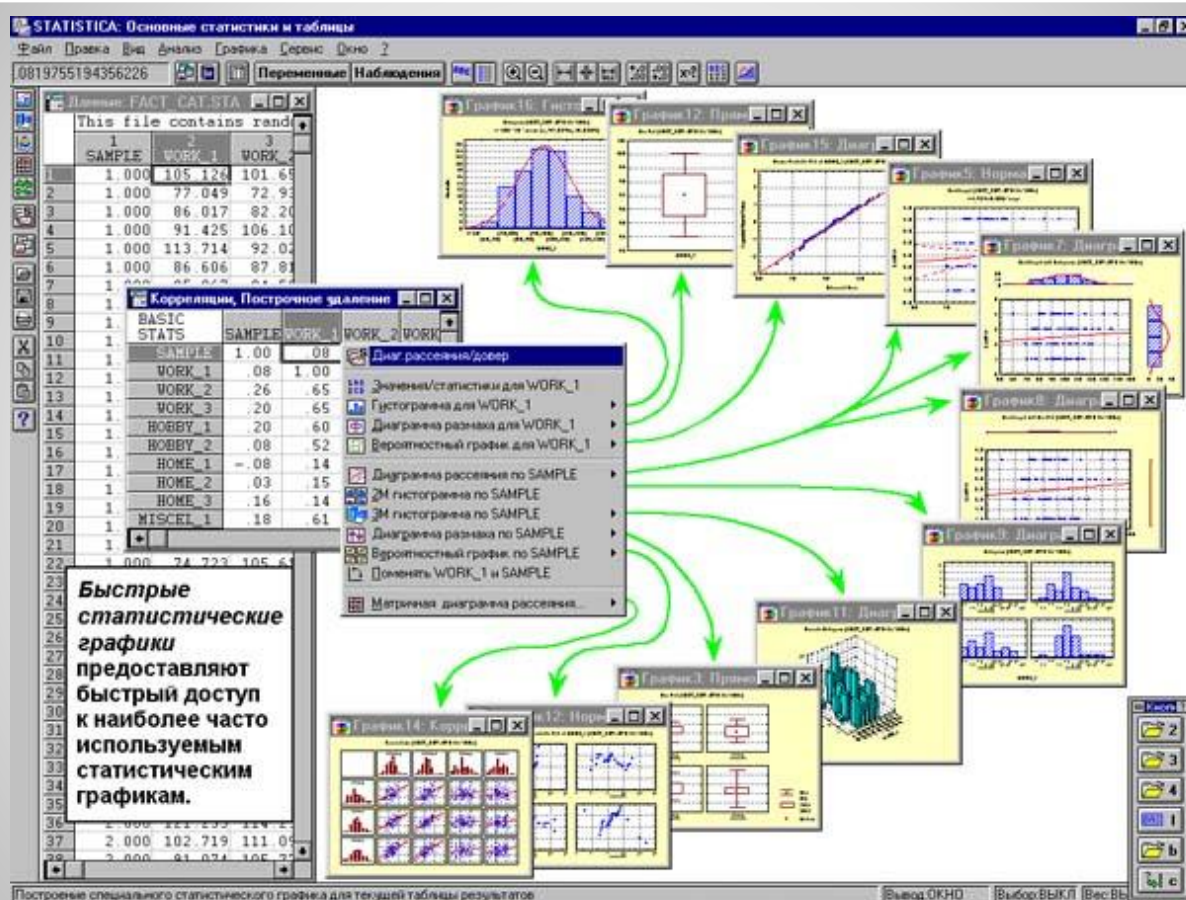
Структурная группировка банков

№ группы	Группы банков по величине уставного капитала, тыс. руб.	Число банков, %	Капитал, % к итогу	Работающие активы, % к итогу	Уставный капитал, % к итогу
1	10135-30631	43,4	19	17,8	30,1
2	30631-51 127	26,7	22,7	21,8	24,5
3	51 127-71623	13,3	20,2	27	21,3
4	71623-92119	13,3	28,8	23,4	16,9
5	92119-112615	3,3	9,3	10	7,2
Итого		100	100	100	100

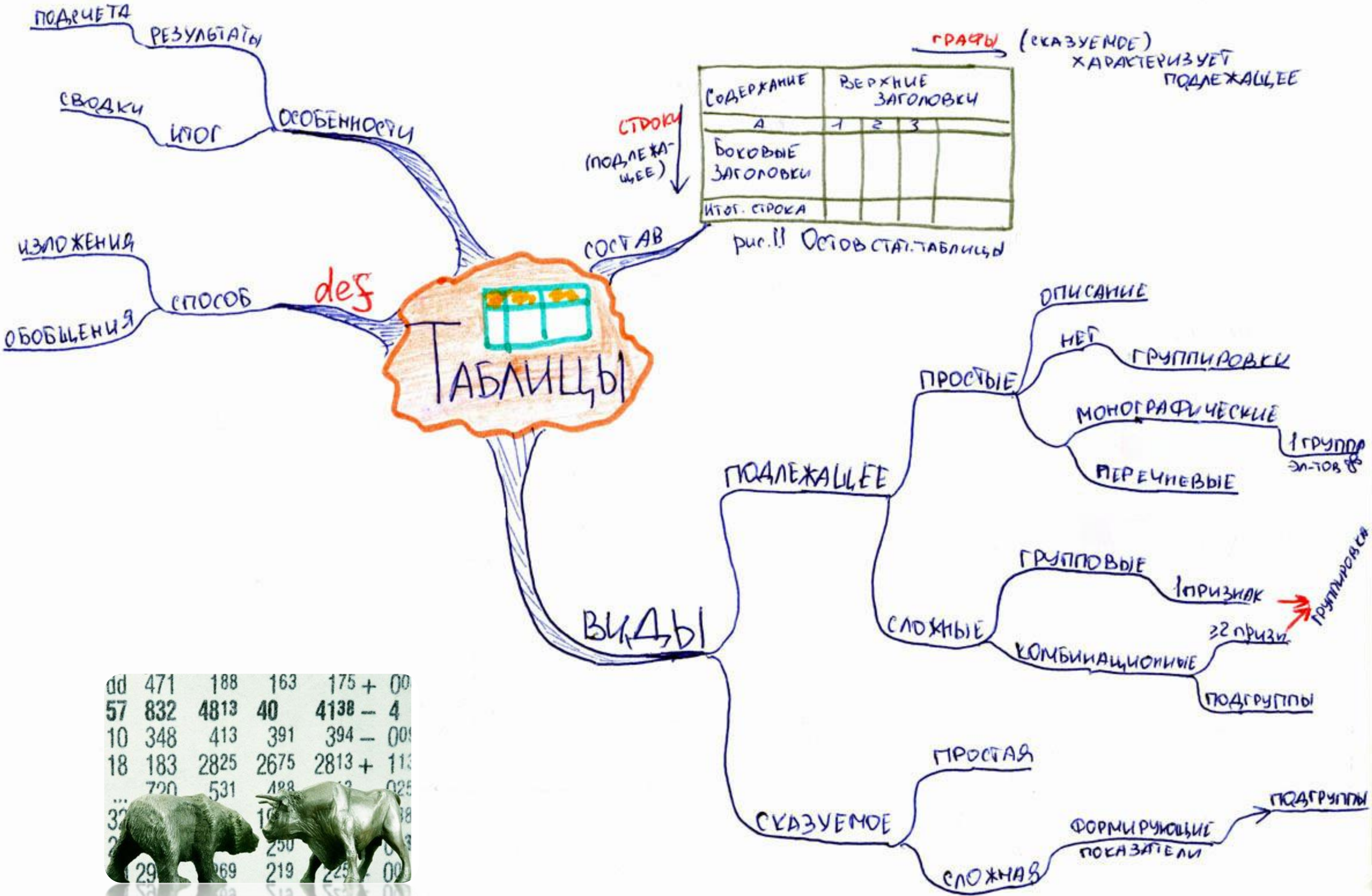
Группировка малых и средних коммерческих банков одного из регионов по величине уставного капитала.

Аналитическая группировка

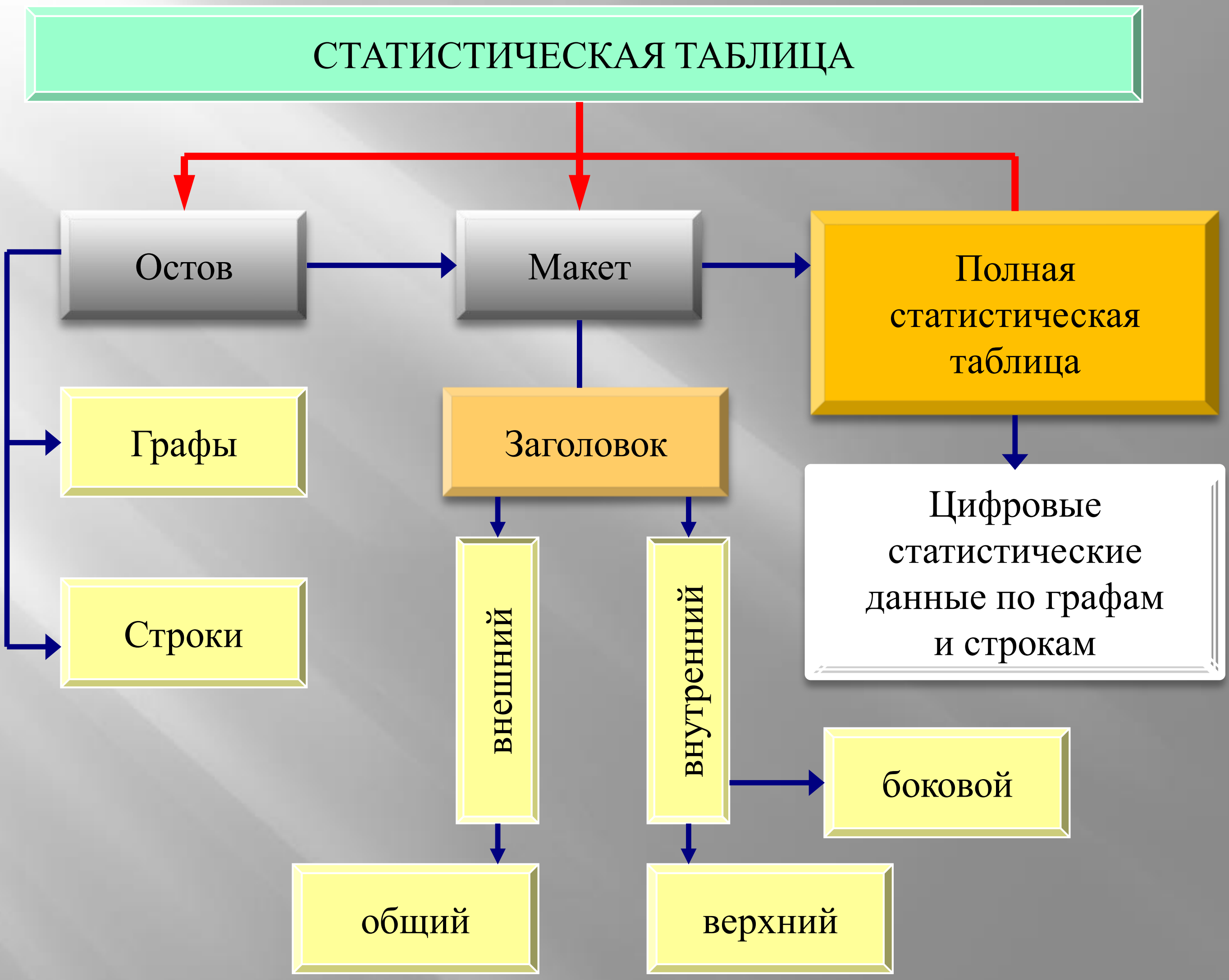
№ груп-пы	Группы банков по величине уставного капитала, тыс. руб.	Число банков, ед.	Капитал, тыс. руб.		Работающие активы, тыс. руб.	
			всего	в среднем на один банк	всего	в среднем на один банк
1	10135-30631	13	230240	17710,77	528639	40664,54
2	30631-51127	8	274797	34349,63	648529	81066,13
3	51 127-71623	4	244798	61 199,5	806426	201606,5
4	71623-92119	4	347956	86989	696276	174069
5	92119-112615	1	112615	112615	298097	298097
Итого		30	1210406	—	1210406	—
В среднем на один банк			—	40346,87	—	99265,57



ПОНЯТИЕ О СТАТИСТИЧЕСКОЙ ТАБЛИЦЕ

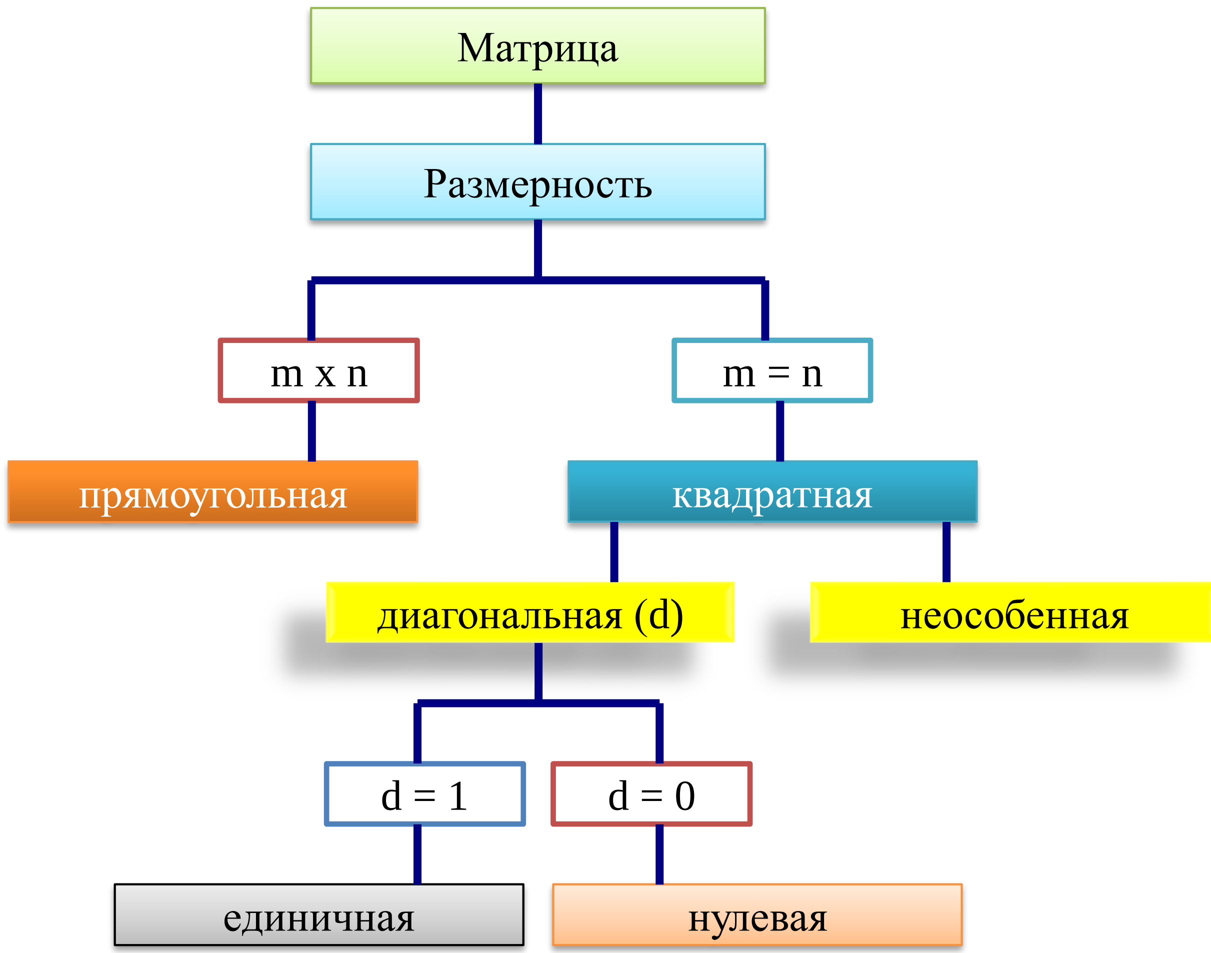

Таблицы
особенности
результаты
подсчета
итог
сводки
способ
изложения
обобщения
def
состав
строки
(подлежащее)
графы (сказуемое)
характеризует подлежащее
Содержание
Верхние заголовки
А
1
2
3
Боковые заголовки
Итог. строка
рис.11 Остов стат. таблицы
виды
подлежащее
простые
описание
нет группировки
монографические
1 группа эл-тов
перечневые
сложные
групповые
1 признак
≥2 призн.
группировка
комбинационные
подгруппы
сказуемое
простая
сложная
формирующие показатели
подгруппы

Основные составляющие элементы статистической таблицы



Макет составных частей статистической таблицы

Виды матриц



$$A = \begin{Bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{Bmatrix}$$

$$D = \begin{Bmatrix} d_1 & 0 & \dots & 0 \\ 0 & d_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & d_n \end{Bmatrix}$$

Общая схема таблиц частот 2 х 2

Признак	B_1	B_2	Всего
A_1	f_{11}	f_{12}	f_{10}
A_2	f_{21}	f_{22}	f_{20}
Всего	f_{01}	f_{02}	f_{00}

Общая схема таблицы сопряженности большей размерности

Признак	B_1	B_2	…	B_j	Всего
A_1	f_{11}	f_{12}		f_{1j}	f_{10}
A_2	f_{21}	f_{22}		f_{2j}	f_{20}
…	…	…	…	…	…
A_i	f_{i1}	f_{i2}		f_{ij}	f_{i0}
Всего	f_{01}	f_{02}		f_{0j}	f_{00}

Классификация статистических таблиц по характеру подлежащего

- Статистические таблицы
  - Простые
    - Монографические
    - Перечневые
      - Простые
      - Территориальные
      - Временные
  - Групповые
  - Комбинационные

Классификация статистических таблиц по разработке сказуемого

Статистическая таблица

Простая

не подразделяется на подгруппы

Сложная

подразделяется на подгруппы

Показатель, формирующий сказуемое

Характеристика выпуска государственных краткосрочных облигаций (ГКО) в РФ в 2005 г. (цифры условные)

простая

	Объем поданных заявок, шт.	Объем выпуска, млн. руб.		Доля ГКО, приобретенная сторонними инвесторами
		объявленный	реальный	
Государственные краткосрочные облигации	476 354	295 000	230 569	34,7

перечневая

Номер ГКО	Объем поданных заявок, шт.	Объем выпуска, млн руб.		Доля ГКО, приобретенная сторонними инвесторами
		объявленный	реальный	
21003 RMFS7	40 256	90 000	37 020	21,0
21004 RMFS5	164 609	55 000	49 848	37,2
21005 RMFS2	271 489	150 000	143 701	44,7
Всего	476 354	295 000	230 569	34,5

Распределение предприятий, выставивших акциина чековые аукционы РФ в 2005 г., по величине уставного капитала(цифры условные)

Групповая

№ п.п.	Группа предприятий по величине уставного капитала	Число предприятий	Количество акций, шт.
1	1215 – 2340	14	7 395
2	2340 – 3465	4	3 402
3	3465 – 4590	4	4 058
4	4590 – 5715	2	3 004
5	5715 – 6840	5	8 587
6	6840 – 7965	1	2 194
	Итого	30	28 667

Группировка предприятий, выставивших акции на чековыеаукционы РФ в 1996 г., по величине уставного капитала и числу занятых(цифры условные)

Комбинационная

№ п.п.	Группы предприятий по величине уставного капитала, млн руб.	Группы предприятий по числу занятых, человек	Число предприятий	Количество проданных акций, шт.
1	1235 – 2340	14 – 33 33 – 52 52 - 71	3 7 4	1 206 4 729 1 390
	Итого по группе	-	14	7 325
2	2340 – 3465	14 – 33 33 – 52 52 - 71	3 - 1	2 508 - 894
	Итого по группе	-	4	3 402
3	3465 – 4590	14 – 33 33 – 52 52 - 71	1 - 3	761 - 3 324
	Итого по группе	-	4	4 085
	Итого по подгруппам	14 – 33 33 – 52 52 - 71	7 7 8	4 475 4 729 5 608
	Всего		22	14 812

Распределение акций среди работников приватизированных предприятий промышленности.

При простой разработке сказуемого

Предприятия	Приобретено акций, всего	В том числе			
		Приватизированные типа А	Обыкновенные	На льготных условиях	По цене, определенной Госкомимуществом

При сложной разработке сказуемого

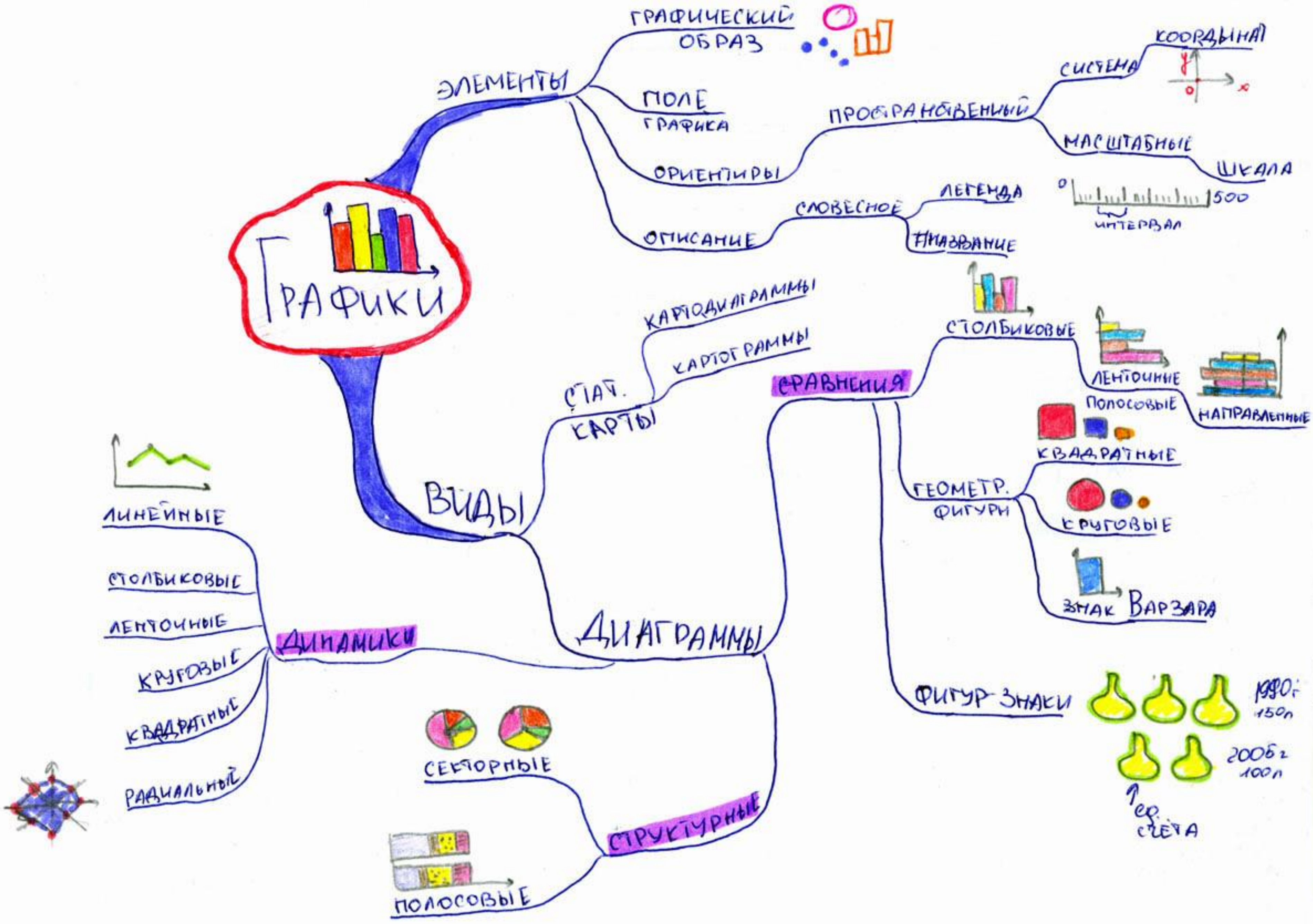
Предприятия	Приобретено акций, всего	В том числе			
		На льготных условиях		По цене, определенной Госкомимуществом	
		Привилегированные типа А	Обыкновенные	Привилегированны е типа А	Обыкновенные

Этапы анализа статистических таблиц

Анализ статистической таблицы

- Структурный
  - ед. наблюдения
  - совокупности
  - вида таблицы
  - решаемых задач
  - признаков
    - содержания
    - комбинаций
    - соотношения
- Содержательный
  - подлежащего
  - сказуемого
  - подлежащего по признакам сказуемого
  - формулировка выводов: по группам, по совокупностям и т.д.
- Проверка данных
  - логическая
  - счетная

ГРАФИЧЕСКИЕ ИЗОБРАЖЕНИЯ В СТАТИСТИКЕ



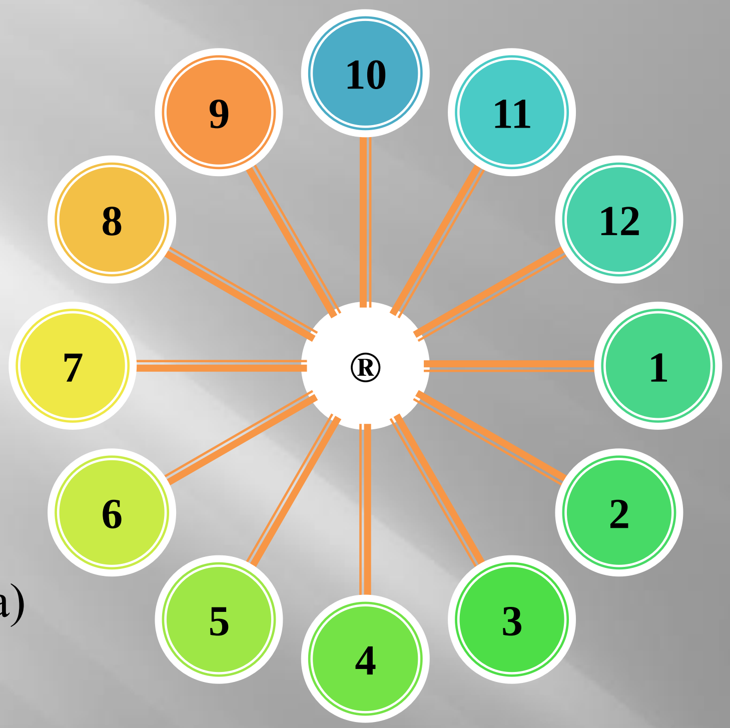
ГРАФИКИ
ЭЛЕМЕНТЫ
ГРАФИЧЕСКИЙ ОБРАЗ
ПОЛЕ ГРАФИКА
ОРИЕНТИРЫ
ПРОСТРАНСТВЕННЫЙ
СИСТЕМА
КООРДИНАТ
МАСШТАБНЫЕ
ШКАЛА
0
500
ИНТЕРВАЛ
ОПИСАНИЕ
СЛОВЕСНОЕ
ЛЕГЕНДА
НАЗВАНИЕ
ВИДЫ
СТАТ. КАРТЫ
КАРТОДИАГРАММЫ
КАРТОГРАММЫ
ДИАГРАММЫ
СРАВНЕНИЯ
СТОЛБИКОВЫЕ
ЛЕНТОЧНЫЕ
ПОЛОСОВЫЕ
НАПРАВЛЕННЫЕ
ГЕОМЕТР. ФИГУРЫ
КВАДРАТНЫЕ
КРУГОВЫЕ
ЗНАК ВАРЗАРА
ФИГУР-ЗНАКИ
1990 г 150 л
2005 г 100 л
ед. счета
ДИНАМИКИ
ЛИНЕЙНЫЕ
СТОЛБИКОВЫЕ
ЛЕНТОЧНЫЕ
КРУГОВЫЕ
КВАДРАТНЫЕ
РАДИАЛЬНЫЕ
СТРУКТУРНЫЕ
СЕКТОРНЫЕ
ПОЛОСОВЫЕ

Классификация статистических графиков по способу построения и задачам изображения

Статистические графики

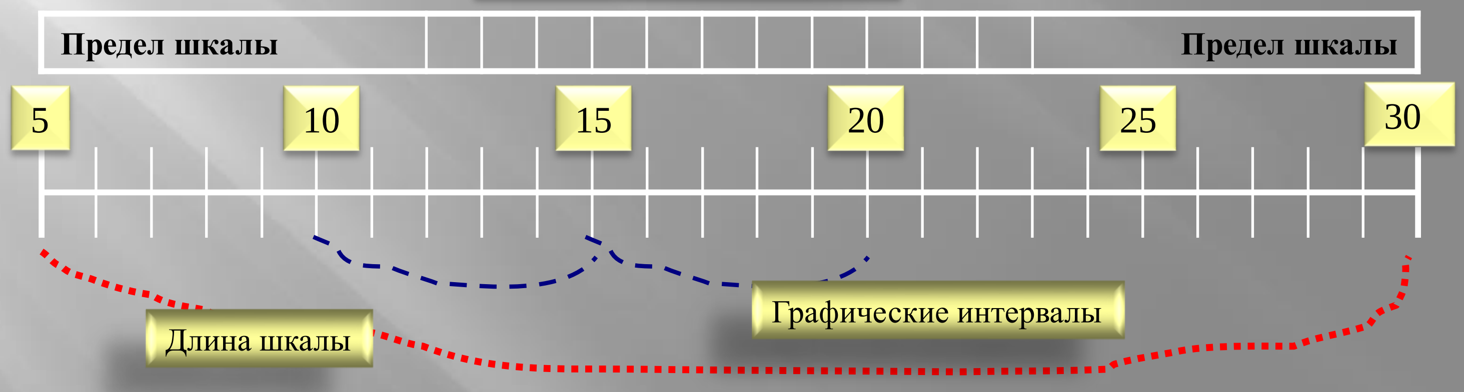
- Диаграммы
  - Сравнения
    - Столбиковые
    - Полосовые
  - Динамики
  - Структуры
    - Столбиковые
    - Сравнения
    - Сравнения
  - Связи
- Картограммы
- Картодиаграммы

Полярная система координат



а)

Числовые интервалы



б)

М а с ш т а б ы

0 1	Масштаб 50 мм
0 1 2 3 4 5	Масштаб 10 мм
0 10 20 30 40 50	Масштаб 1 мм
0 100 200 300 400 500	Масштаб 0,1 мм

Логарифмические и числовые шкалы

0 0,5 1,0

0 1 2 3

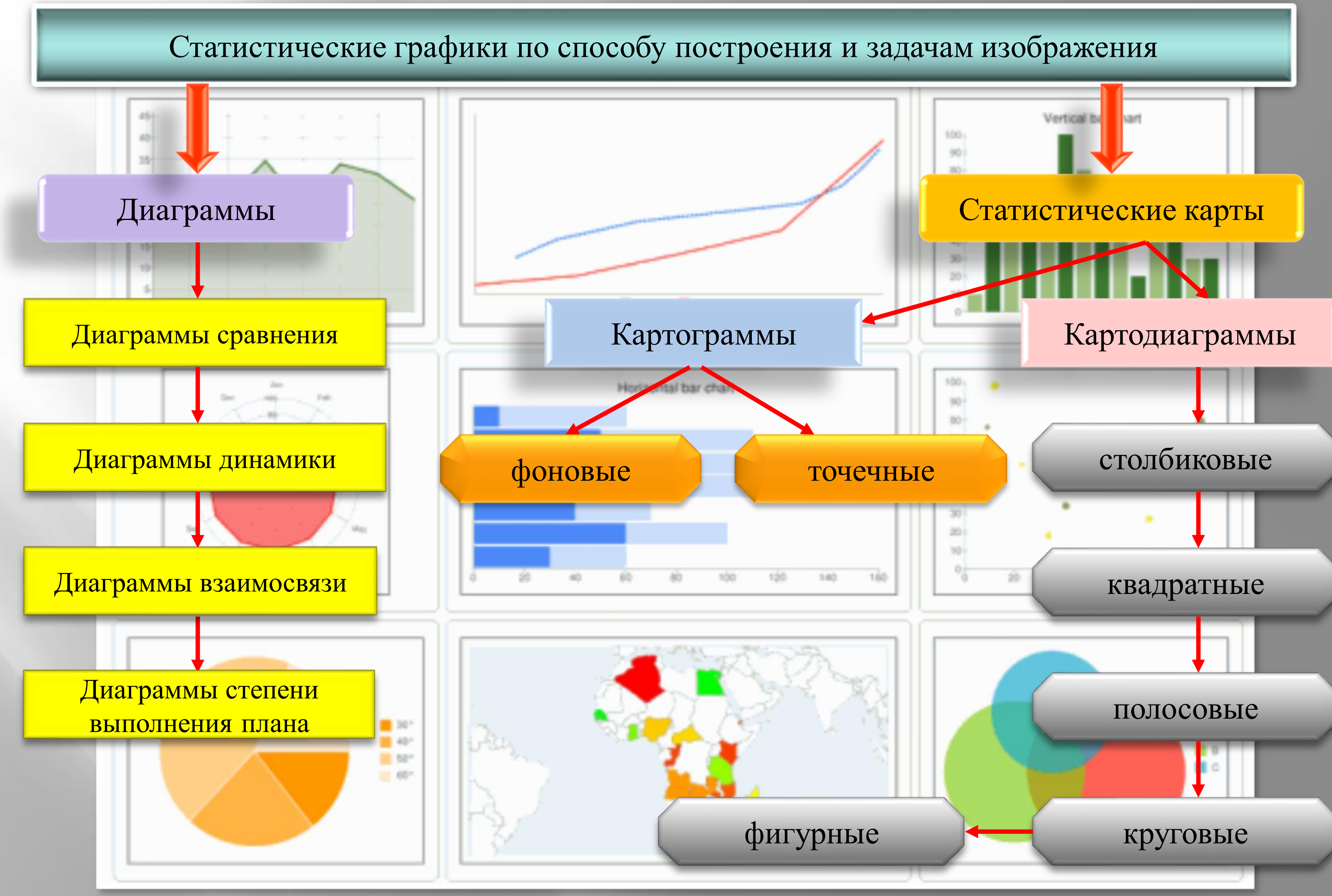
Числа

0 10 100 1000

0 1 2 3

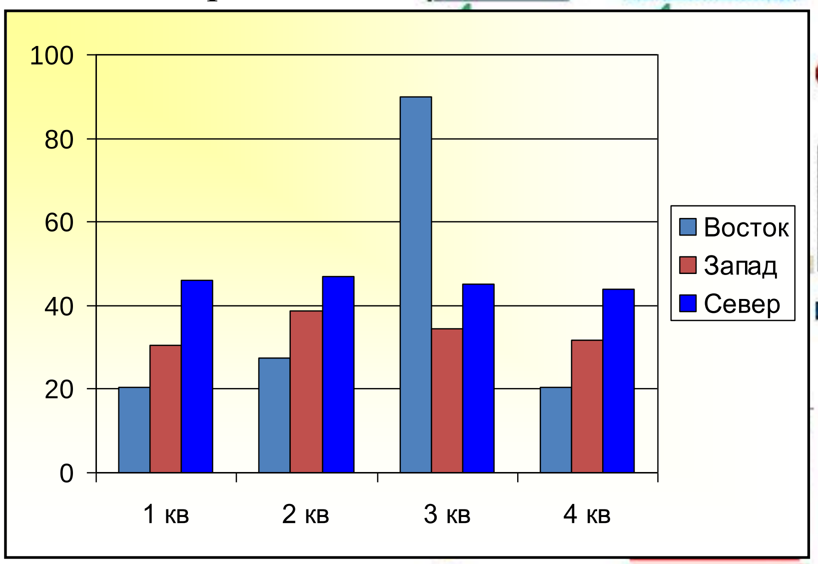
Логарифмы чисел

Классификация статистических графиков по способу построения и содержанию изображаемых данных

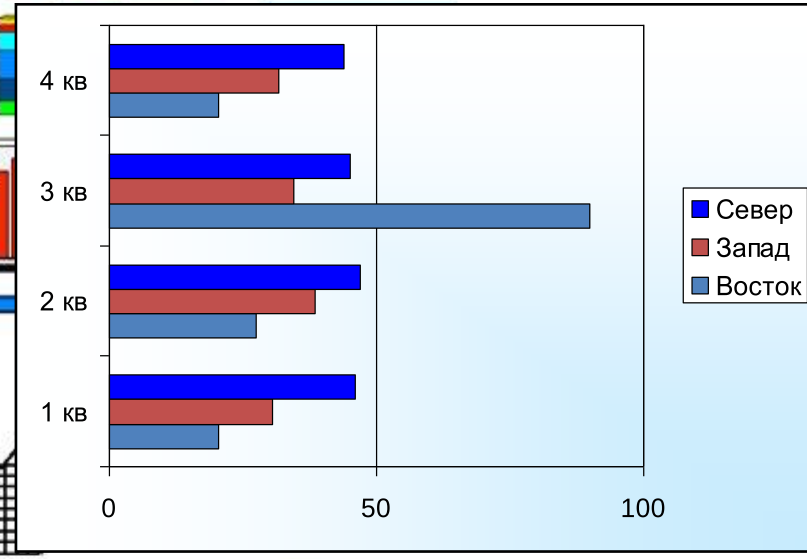


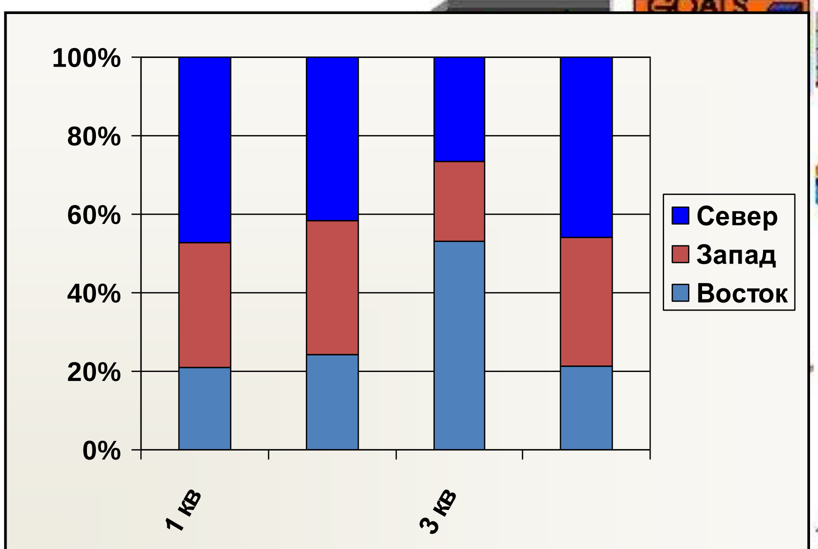
Диаграммы сравнения

Гистограмма

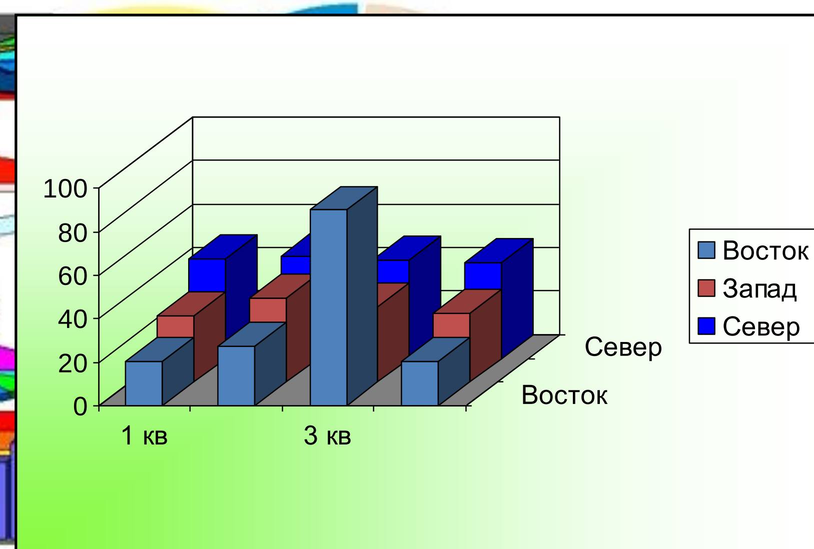


Линейчатая





Нормированная гистограмма



Трехмерная гистограмма

Картограмма

Стоимость минимального набора продуктов питания

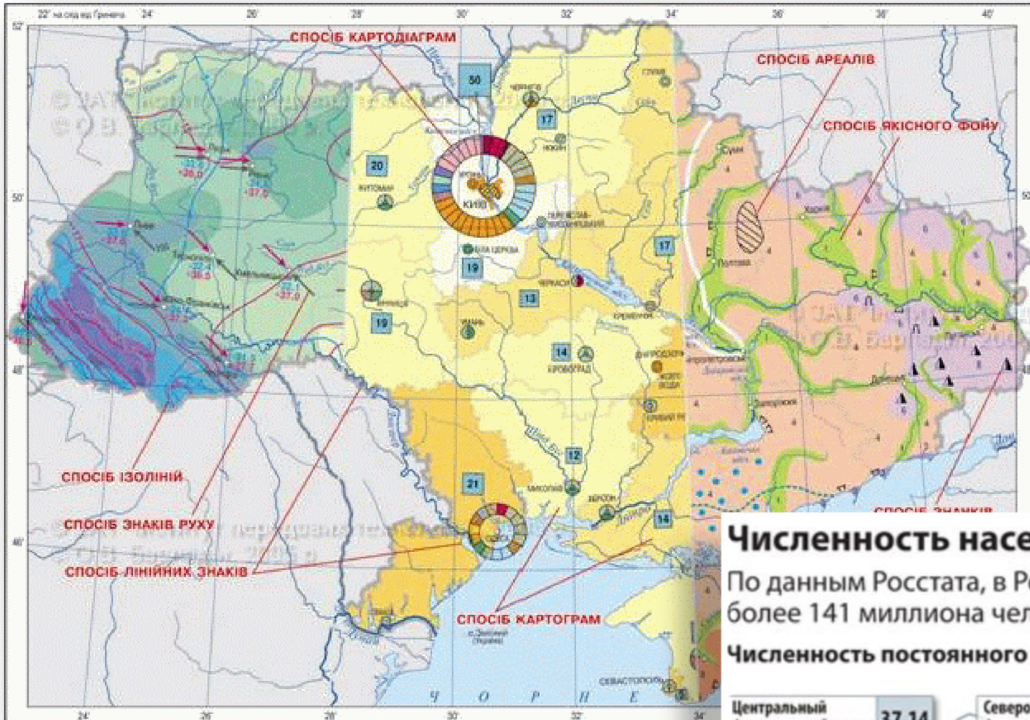
в декабре 2007 года

Цифрами на карте обозначены:

1 - Белгородская область
2 - Владимирская область
3 - Воронежская область
4 - Ивановская область
5 - Калужская область
6 - Костромская область
7 - Курская область
8 - Липецкая область
9 - Московская область
10 - Орловская область
11 - Рязанская область
12 - Тамбовская область
13 - Тульская область
14 - Ярославская область
15 - Вологодская область
16 - Новгородская область
17 - Республика Адыгея
18 - Республика Ингушетия
19 - Кабардино-Балкарская Республика
20 - Карачаево-Черкесская Республика
21 - Республика Северная Осетия-Алания
22 - Чеченская Республика
23 - Краснодарский край
24 - Ставропольский край
25 - Республика Башкортостан
26 - Республика Марий Эл
27 - Республика Мордовия
28 - Республика Татарстан
29 - Удмуртская Республика
30 - Чувашская Республика
31 - Кировская область
32 - Нижегородская область
33 - Пензенская область
34 - Пермский край
35 - Самарская область
36 - Саратовская область
37 - Ульяновская область

* Данные рассчитаны по субъекту, включая АО

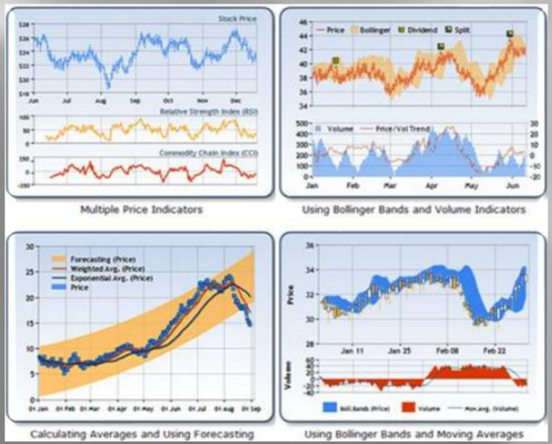
Картодиаграмма

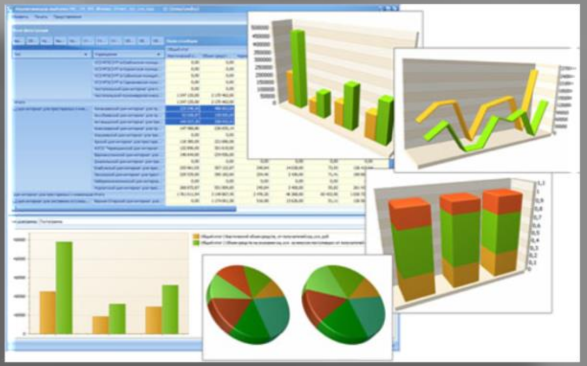


Классификация статистических графиков по форме графического образа

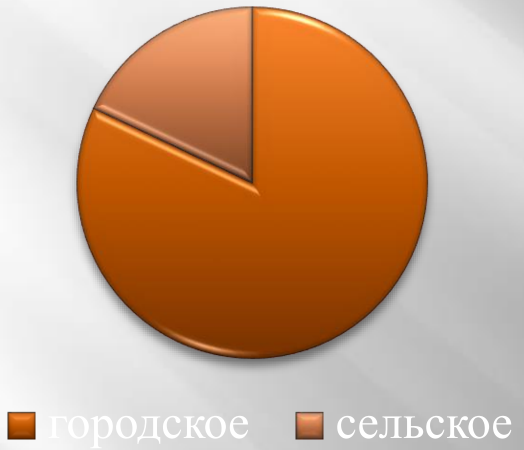
Статистические графики по форме графического образа

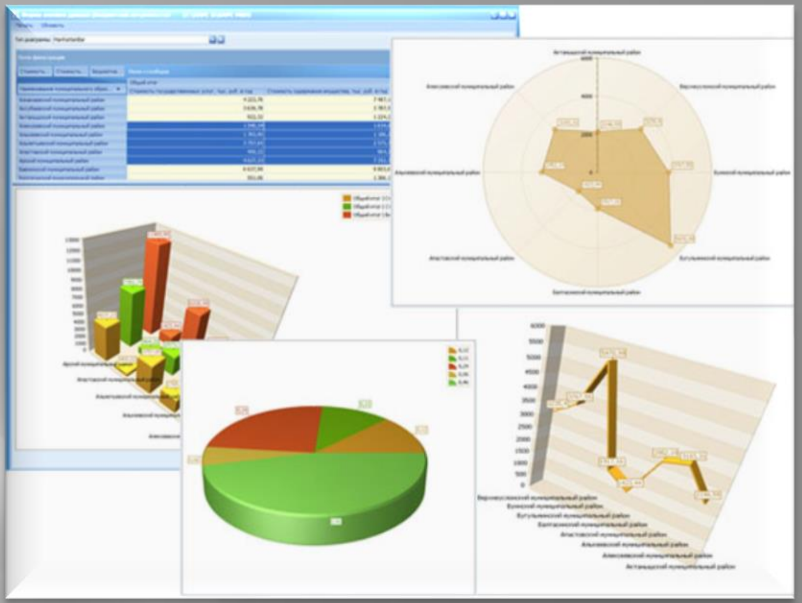
- Линейные
  - Статистические кривые
- Плоскостные
  - Столбиковые
  - Полосовые
  - Квадратные
  - Круговые
  - Секторные
  - Фигурные
  - Точечные
  - Фоновые
- Объемные
  - Поверхностного распределения

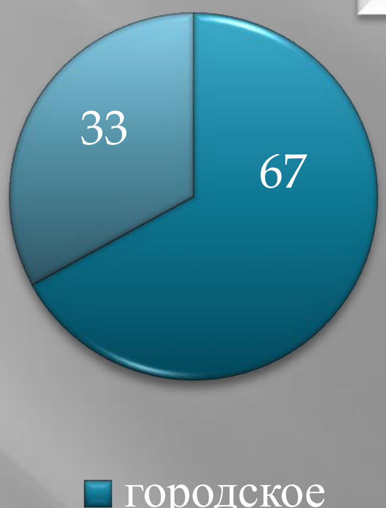




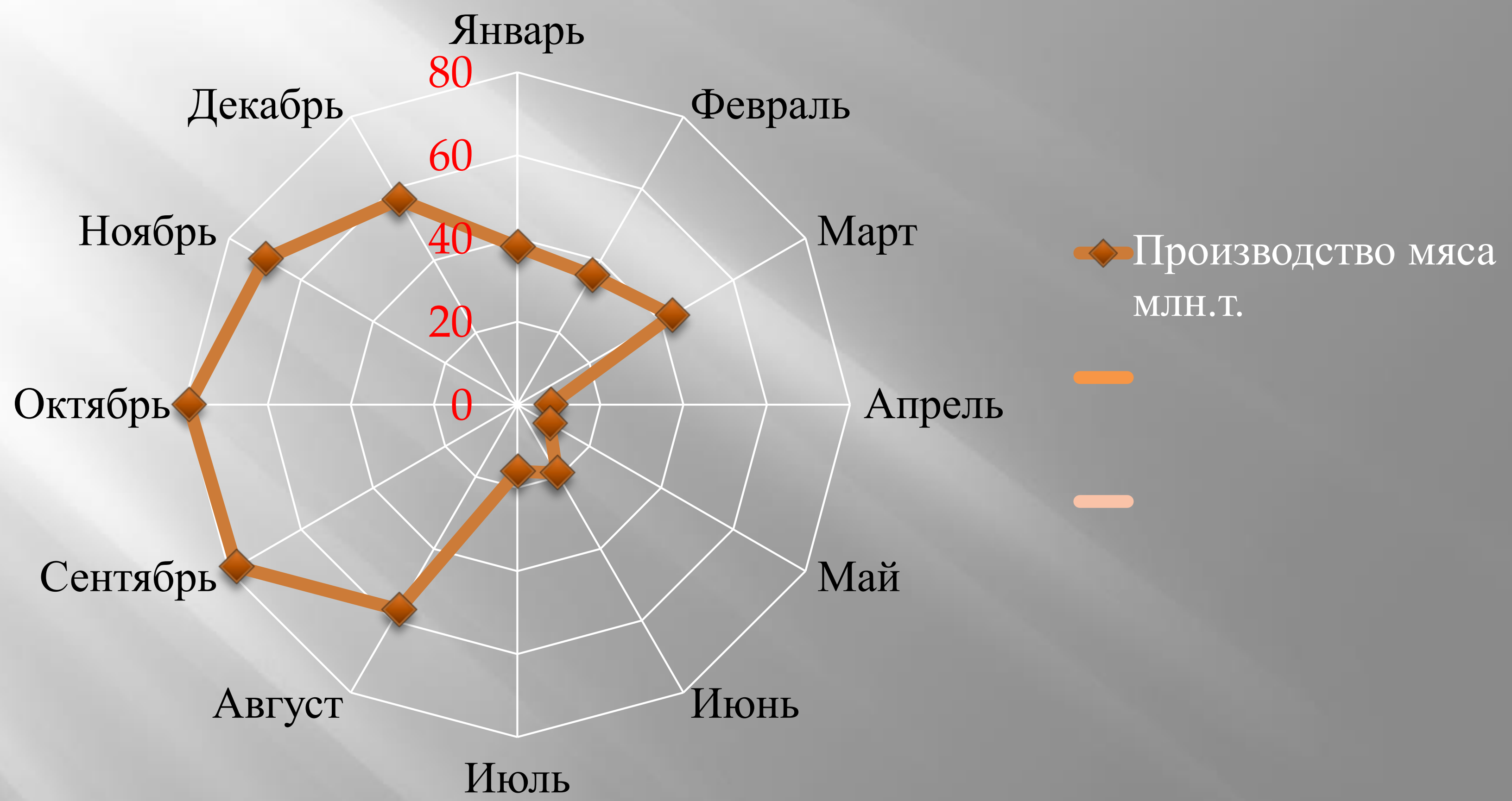
Секторная диаграмма.
Распределение населения на городское и сельское за 1970, 1985 и 2005 гг. соответственно



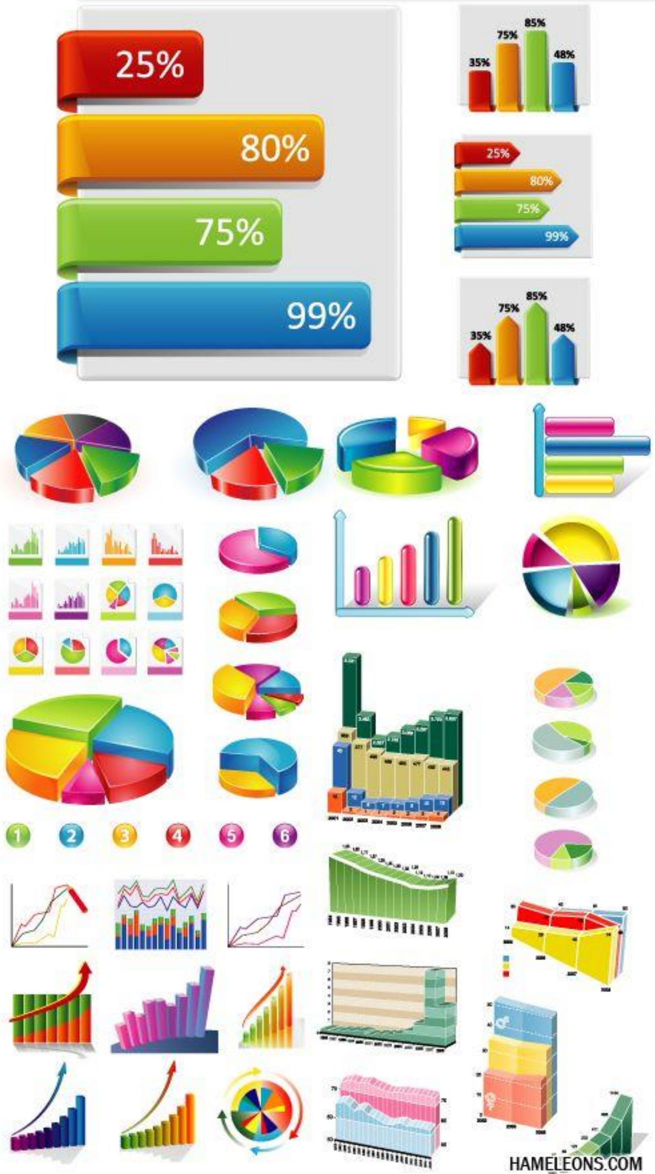




Радиальная диаграмма.



Диаграммы сравнения



Квадратная диаграмма доходов

16714

14560

4662

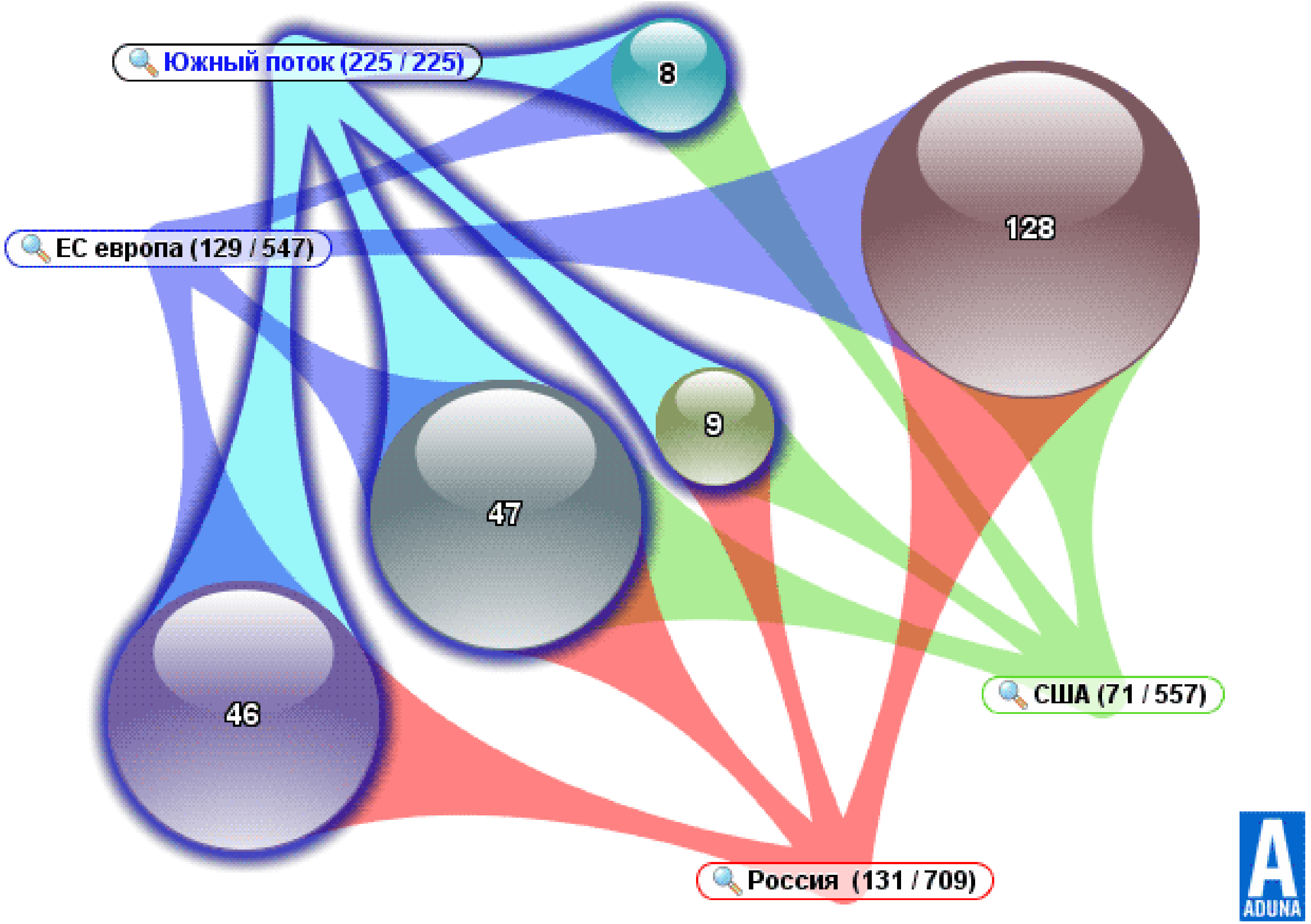
Круговая диаграмма доходов

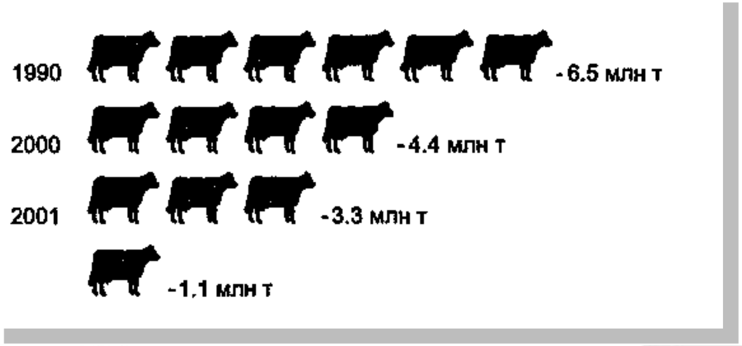
4662

14560

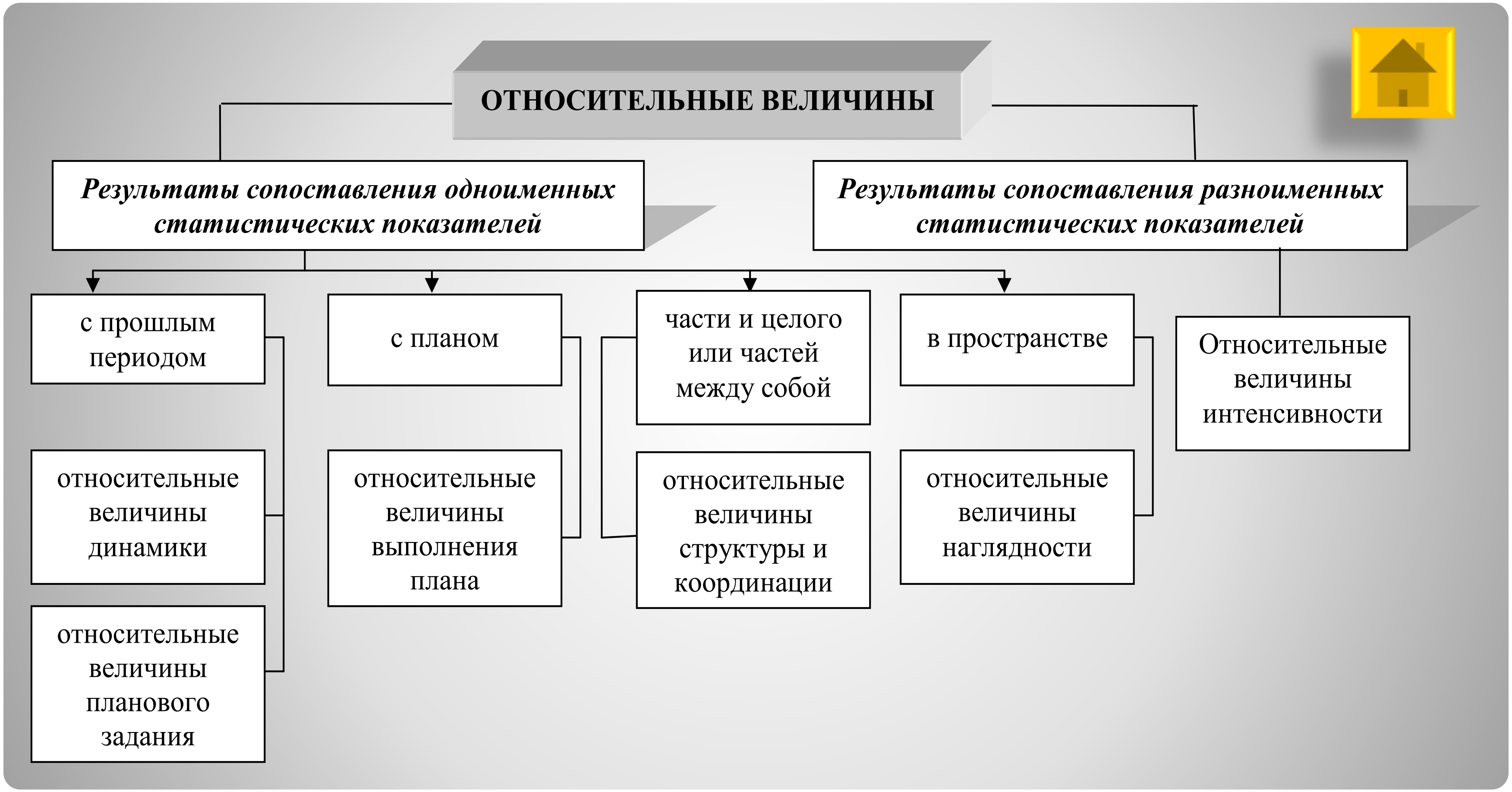
16714

Диаграмма сравнения

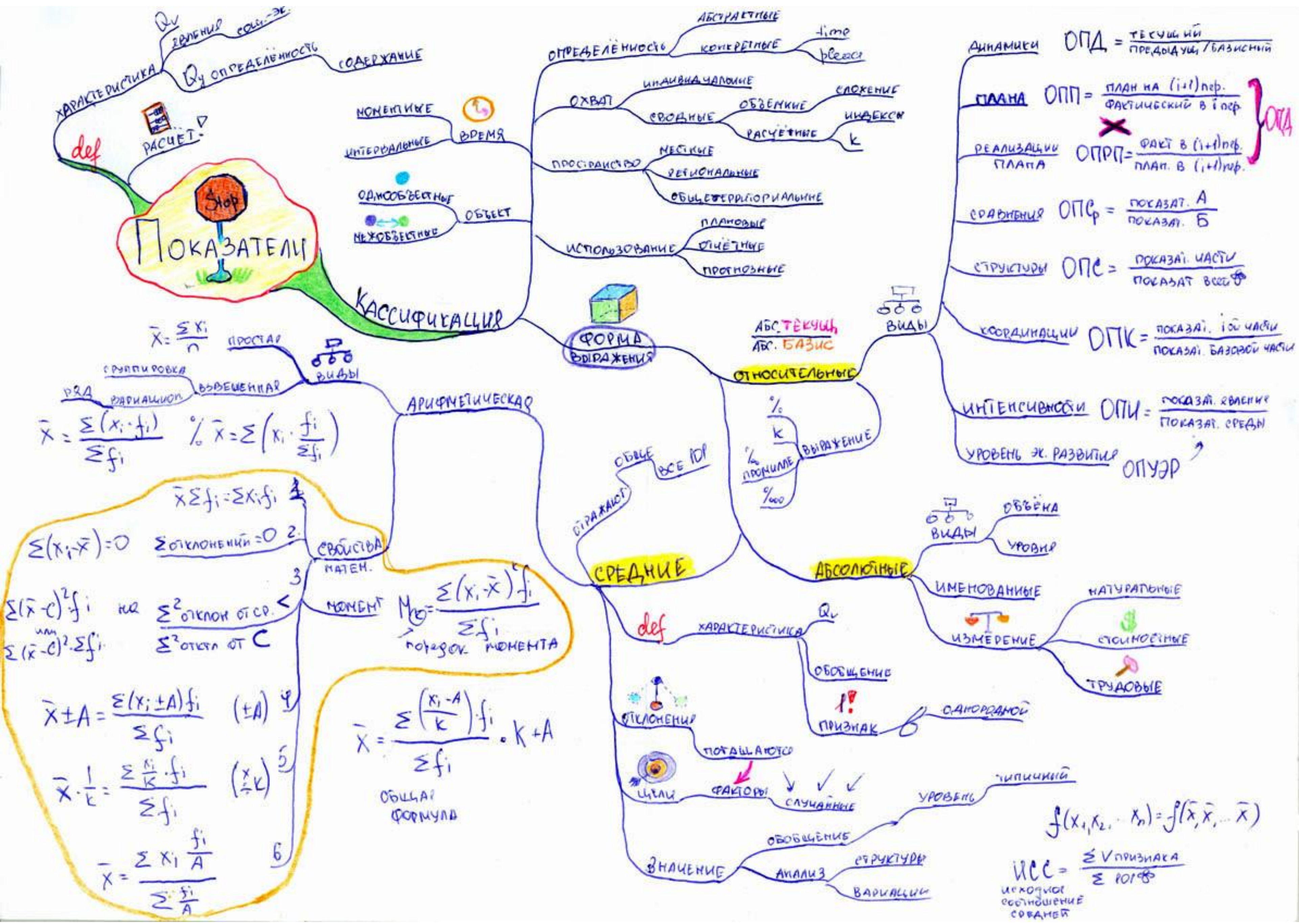




Фигурная диаграмма



Абсолютные и относительные величины
Средние величины, показатели вариации



ПОКАЗАТЕЛИ
Stop
ХАРАКТЕРИСТИКА
ЯВЛЕНИЯ
СОЦ.-ЭК.
Qv
Qy ОПРЕДЕЛЁННОСТЬ
СОДЕРЖАНИЕ
def
РАСЧЁТ
КЛАССИФИКАЦИЯ
МОМЕНТНЫЕ
ИНТЕРВАЛЬНЫЕ
ВРЕМЯ
ОДНООБЪЕКТНЫЕ
МЕЖОБЪЕКТНЫЕ
ОБЪЕКТ
ОПРЕДЕЛЁННОСТЬ
АБСТРАКТНЫЕ
КОНКРЕТНЫЕ
time
place
ОХВАТ
ИНДИВИДУАЛЬНЫЕ
СВОДНЫЕ
ОБЪЁМНЫЕ
СЛОЖЕНИЕ
РАСЧЁТНЫЕ
ИНДЕКСЫ
k
ПРОСТРАНСТВО
МЕСТНЫЕ
РЕГИОНАЛЬНЫЕ
ОБЩЕТЕРРИТОРИАЛЬНЫЕ
ИСПОЛЬЗОВАНИЕ
ПЛАНОВЫЕ
ОТЧЁТНЫЕ
ПРОГНОЗНЫЕ
ФОРМА ВЫРАЖЕНИЯ
АБС. ТЕКУЩ.
АБС. БАЗИС
ОТНОСИТЕЛЬНЫЕ
ВИДЫ
ДИНАМИКИ
ОПД = ТЕКУЩИЙ / ПРЕДЫДУЩ./БАЗИСНЫЙ
ПЛАНА
ОПП = ПЛАН НА (i+1) пер. / ФАКТИЧЕСКИЙ В i пер.
РЕАЛИЗАЦИИ ПЛАНА
ОПРП = ФАКТ В (i+1) пер. / ПЛАН. В (i+1) пер.
ОПД
СРАВНЕНИЯ
ОПСр = ПОКАЗАТ. А / ПОКАЗАТ. Б
СТРУКТУРЫ
ОПС = ПОКАЗАТ. ЧАСТИ / ПОКАЗАТ. ВСЕГО
КООРДИНАЦИИ
ОПК = ПОКАЗАТ. i-ой ЧАСТИ / ПОКАЗАТ. БАЗОВОЙ ЧАСТИ
ИНТЕНСИВНОСТИ
ОПИ = ПОКАЗАТ. ЯВЛЕНИЯ / ПОКАЗАТ. СРЕДЫ
УРОВЕНЬ ЭК. РАЗВИТИЯ
ОПУЭР
ВЫРАЖЕНИЕ
%
k
‰ ПРОМИЛЛЕ
‱
АБСОЛЮТНЫЕ
ВИДЫ
ОБЪЁМА
УРОВНЯ
ИМЕНОВАННЫЕ
ИЗМЕРЕНИЕ
НАТУРАЛЬНЫЕ
СТОИМОСТНЫЕ
ТРУДОВЫЕ
СРЕДНИЕ
ОТРАЖАЮТ
ОБЩЕЕ
ВСЕ ГОР
def
ХАРАКТЕРИСТИКА
Qv
ОБОБЩЕНИЕ
ПРИЗНАК
ОДНОРОДНОЙ
ОТКЛОНЕНИЯ
ПОГАШАЮТСЯ
ЦЕЛИ
ФАКТОРЫ
СЛУЧАЙНЫЕ
ЗНАЧЕНИЕ
ОБОБЩЕНИЕ
УРОВЕНЬ
ТИПИЧНЫЙ
АНАЛИЗ
СТРУКТУРЫ
ВАРИАЦИИ
$f(x_1, x_2, \dots x_n) = f(\bar{x}, \bar{x}, \dots \bar{x})$
ИСС = Σ V ПРИЗНАКА / Σ ЕДИН.
ИСХОДНОЕ СООТНОШЕНИЕ СРЕДНЕЙ
АРИФМЕТИЧЕСКАЯ
ВИДЫ
ПРОСТАЯ
$\bar{x} = \frac{\sum x_i}{n}$
ВЗВЕШЕННАЯ
ГРУППИРОВКА
РЯД ВАРИАЦИОН.
$\bar{x} = \frac{\sum (x_i \cdot f_i)}{\sum f_i}$
$\% \; \bar{x} = \sum \left(x_i \cdot \frac{f_i}{\sum f_i} \right)$
СВОЙСТВА МАТЕМ.
1. $\bar{x} \sum f_i = \sum x_i f_i$
2. $\sum (x_i - \bar{x}) = 0$ Σ ОТКЛОНЕНИЙ = 0
3. $\sum (\bar{x} - c)^2 f_i$ или $\sum (\bar{x} - c)^2 \cdot \sum f_i$ на Σ² ОТКЛОН. ОТ СР. < Σ² ОТКЛОН. ОТ C
4. $\bar{x} \pm A = \frac{\sum (x_i \pm A) f_i}{\sum f_i}$ $(\pm A)$
5. $\bar{x} \cdot \frac{1}{k} = \frac{\sum \frac{x_i}{k} \cdot f_i}{\sum f_i}$ $(\frac{x}{\div k})$
6. $\bar{x} = \frac{\sum x_i \frac{f_i}{A}}{\sum \frac{f_i}{A}}$
МОМЕНТ
$M_k = \frac{\sum (x_i - \bar{x})^k f_i}{\sum f_i}$
ПОРЯДОК МОМЕНТА
$\bar{x} = \frac{\sum \left(\frac{x_i - A}{k} \right) \cdot f_i}{\sum f_i} \cdot k + A$
ОБЩАЯ ФОРМУЛА

Виды относительных величин

Относительный показатель динамики

1 $$\text{ОПД} = \frac{\text{Текущий уровень}}{\text{Предшествующий или базисный уровень}}$$

Относительные показатели плана и **реализации плана**.

2 $$\text{ОПП} = \frac{\text{Уровень, планируемый на i - тый период}}{\text{Уровень, достигнутый в i - м периоде}}$$

3 $$\text{ОПРП} = \frac{\text{Уровень, планируемый на i - тый период}}{\text{Уровень, планируемый на i + 1 - й период}}$$

Относительный показатель структуры

4 Î Ï Ñ = Ï î êàçàòåëü, õàðàêòåðèçóþ ù èé ÷àñòü ñî âî êóï í î ñòè / Ï î êàçàòåëü ï î âñåé ñî âî êóï í î ñòè â öåëî ì

Относительный показатель координации

5 Î Ï Ê = Ï î êàçàòåëü, õàðàêòåðèçóþ ù èé i - þ ÷àñòü ñî âî êóï í î ñòè / Ï î êàçàòåëü, õàðàêòåðèçóþ ù èé ÷àñòü ñî âî êóï í î ñòè âû áðàí í óþ â êà÷åñòâå áàçû äëÿ ñðàâí åí èÿ

Относительный показатель интенсивности

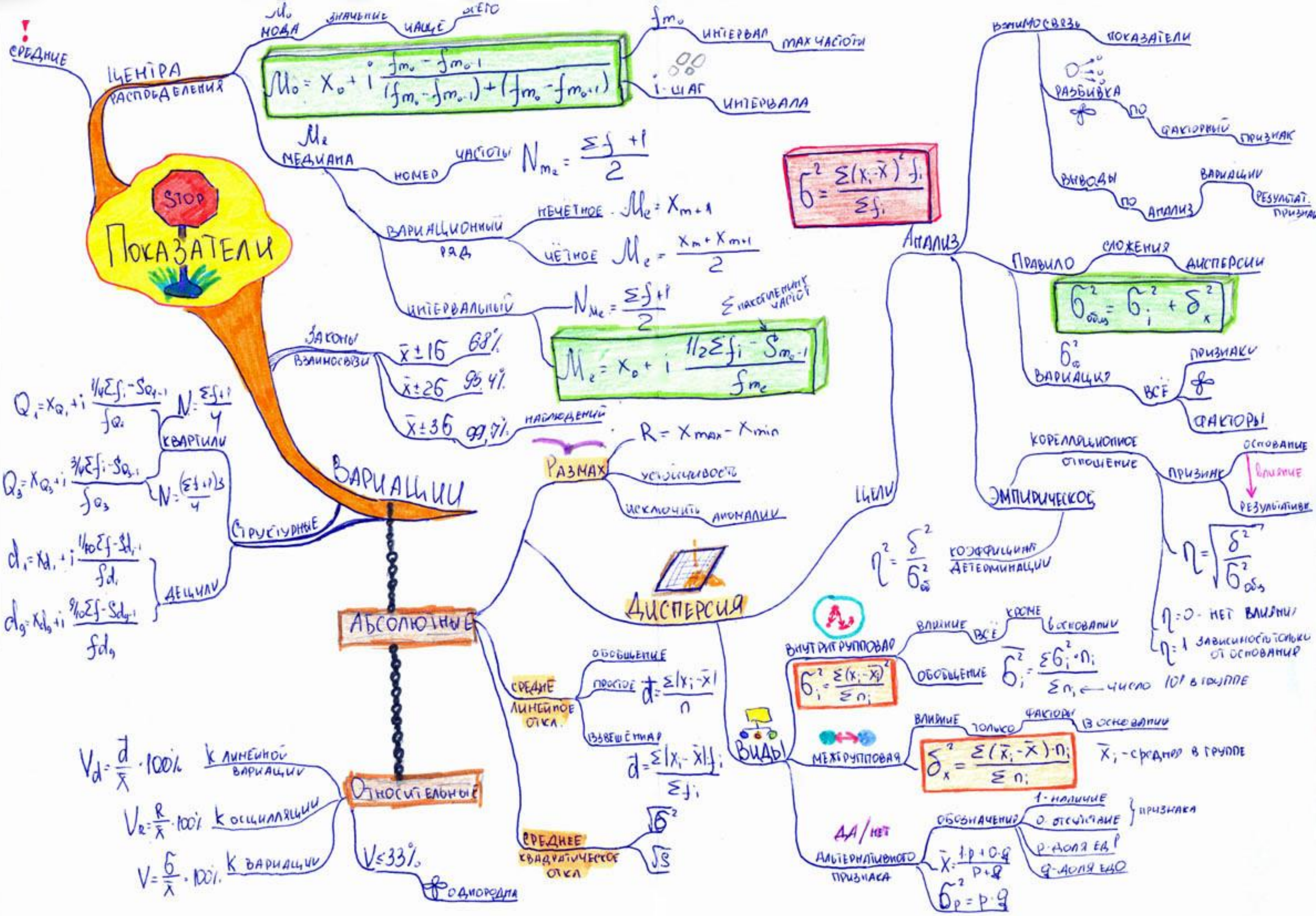
6 Î Ï È = Ï î êàçàòåëü, õàðàêòåðèçóþ ù èé ÿâëåí èå À / Ï î êàçàòåëü, õàðàêòåðèçóþ ù èé ñðåäó ðàñï ðî ñòðàí åí èÿ ÿâëåí èÿ À

Относительный показатель сравнения

7 $$\text{ОПСр} = \frac{\text{Показатель, ххарактерзующий объект А}}{\text{Показатель, ххарактерзующий объект Б}}$$

Средние величины

- Математические
  - Арифметическая
  - Геометрическая
  - Гармоническая
- Структурные
  - Мода
  - Медиана



Показатели
STOP
средние
центра распределения
Mo мода
значение чаще всего
$Mo = X_0 + i\frac{fm_0 - fm_{0-1}}{(fm_0 - fm_{0-1}) + (fm_0 - fm_{0+1})}$
fm_0 интервал max частоты
i - шаг интервала
Me медиана
номер частоты
$N_{Me} = \frac{\Sigma f + 1}{2}$
вариационный ряд
нечетное $Me = X_{m+1}$
чётное $Me = \frac{X_m + X_{m+1}}{2}$
интервальный
$N_{Me} = \frac{\Sigma f + 1}{2}$
Σ накопленных частот
$Me = X_0 + i\frac{1/2\Sigma f_i - S_{m_0-1}}{fm_e}$
законы взаимосвязи
$\bar{x} \pm 1\sigma$ 68%
$\bar{x} \pm 2\sigma$ 95,4%
$\bar{x} \pm 3\sigma$ 99,7% наблюдений
$Q_1 = X_{Q_1} + i\frac{1/4\Sigma f_i - S_{Q_1-1}}{f_{Q_1}}$
$N = \frac{\Sigma f + 1}{4}$
квартили
$Q_3 = X_{Q_3} + i\frac{3/4\Sigma f_i - S_{Q_3-1}}{f_{Q_3}}$
$N = \frac{(\Sigma f + 1)3}{4}$
структурные
$d_1 = X_{d_1} + i\frac{1/10\Sigma f - S_{d_1-1}}{f_{d_1}}$
децили
$d_9 = X_{d_9} + i\frac{9/10\Sigma f - S_{d_9-1}}{f_{d_9}}$
вариации
абсолютные
размах
$R = X_{max} - X_{min}$
устойчивость
исключить аномалии
средние линейное откл.
обобщение
простое $\bar{d} = \frac{\Sigma|x_i - \bar{x}|}{n}$
взвешенная $\bar{d} = \frac{\Sigma|x_i - \bar{x}| f_i}{\Sigma f_i}$
среднее квадратическое откл.
$\sqrt{\sigma^2}$
$\sqrt{S}$
относительные
$V_d = \frac{\bar{d}}{\bar{x}} \cdot 100\%$ к линейной вариации
$V_R = \frac{R}{\bar{x}} \cdot 100\%$ к осцилляции
$V = \frac{\sigma}{\bar{x}} \cdot 100\%$ к вариации
$V \leq 33\%$ однородна
дисперсия
$\sigma^2 = \frac{\Sigma(x_i - \bar{x})^2 f_i}{\Sigma f_i}$
анализ
цели
взаимосвязь показатели
разбивка по факторный признак
выводы по анализ вариации результат. признак
правило сложения дисперсии
$\sigma^2_{общ} = \bar{\sigma}^2_i + \delta^2_x$
σ^2_φ вариация все признаки факторы
корелляционное отношение
признак основание влияние результатив.
эмпирическое
$\eta^2 = \frac{\delta^2}{\sigma^2_{об}}$ коэффициент детерминации
$\eta = \sqrt{\frac{\delta^2}{\sigma^2_{общ}}}$
$\eta = 0$ - нет влияния
$\eta = 1$ зависимость только от основания
виды
внутригрупповая
$\sigma^2_i = \frac{\Sigma(x_i - \bar{x}_i)^2}{\Sigma n_i}$
влияние все кроме в основании
обобщение $\bar{\sigma}^2_i = \frac{\Sigma\sigma^2_i \cdot n_i}{\Sigma n_i}$ ← число ед. в группе
межгрупповая
влияние только факторы в основании
$\delta^2_x = \frac{\Sigma(\bar{x}_i - \bar{x}) \cdot n_i}{\Sigma n_i}$
$\bar{x}_i$ - средняя в группе
да / нет
альтернативного признака
обозначения
1 - наличие
0 - отсутствие } признака
p - доля ед. P
q - доля ед. Q
$\bar{x} = \frac{1 \cdot p + 0 \cdot q}{p + q}$
$\sigma^2_p = p \cdot q$

Простая степенная средняя

$$\bar{x} = \sqrt[k]{\frac{\sum_{i=1}^{n} x_i^k}{n}}$$

Взвешенная степенная средняя

$$\bar{x} = \sqrt[k]{\frac{\sum_{i=1}^{n}(x_i^k f_i)}{\sum_{i=1}^{n} f_i}}$$

$$\overline{X}_{\text{гарм.}} \le \overline{X}_{\text{геом.}} \le \overline{X}_{\text{арифм.}} \le \overline{X}_{\text{квадр.}} \le \overline{X}_{\text{куб.}}..$$

Вид степенной средней	Показатель степени (k)	Формула расчета	
		Простая	Взвешенная
Гармоническая	**-1**	$\overline{X} = \dfrac{n}{\sum \frac{1}{x}}$	$\overline{X} = \dfrac{\sum w_i}{\sum \frac{w_i}{x}}$ $w = x_i f_i$
Геометрическая	**0**	$\overline{X} = \sqrt[k]{Ï\ x} = \sqrt[k]{x_1 x_2 K\ x_n}$	$\overline{X} = \sqrt[\sum f]{\Pi x^f} = \sqrt[\sum f]{x_1^{f_1} x_2^{f_2} K\ x_n^{f_n}}$
Арифметическая	**1**	$\overline{X} = \dfrac{\sum x_i}{n}$	$\overline{X} = \dfrac{\sum x_i f_i}{\sum f_i}$
Квадратическая	**2**	$\overline{X} = \sqrt{\dfrac{\sum x_i^{\ 2}}{n}}$	$\overline{X} = \sqrt{\dfrac{\sum x_i^{\ 2} f_i}{\sum f_i}}$
Кубическая	**3**	$\overline{X} = \sqrt[3]{\dfrac{\sum x_i^{\ 3}}{n}}$	$\overline{X = \sqrt[3]{\dfrac{\sum x_i^{\ 3} f}{\sum f_i}}}$

Предельные теоремы теории вероятностей

Закон Больших Чисел устанавливает связь между абстрактными моделями теории вероятностей и основными ее понятиями и средними значениями, полученными при статистической обработке выборки ограниченного объема из генеральной совокупности. P, F(x), M(x), D(x).

ЗБЧ доказывает, что средние выборочные значения при $n\to\infty$ стремятся к соответствующим значениям генеральной совокупности: $h_n(A)\to P$, $X_{ср}\to M(X)$, $\sigma_{ср}^2\to D(X)$, $F^*(X)\to F(X)$.

$P(Y\geq\varepsilon)\leq M(x)/\varepsilon$, $P(Y<\varepsilon)\geq 1-M(x)/\varepsilon$.

Доказательство. Рассмотрим Y и $Y_\varepsilon = \begin{cases} 0, Y<\varepsilon \\ \varepsilon, Y\geq\varepsilon \end{cases}$: $Y_\varepsilon\leq Y$, $M(Y_\varepsilon)\leq M(Y)$

$M(Y_\varepsilon)=0\cdot P(Y<\varepsilon)+\varepsilon\cdot P(Y\geq\varepsilon)=\varepsilon\cdot P(Y\geq\varepsilon)$

$M(Y)\geq M(Y_\varepsilon)=\varepsilon\cdot P(Y\geq\varepsilon)$.

Лемма позволяет сделать оценку вероятности наступления события по математическому ожиданию этой СВ.

Неравенство Чебышева. Для любой СВ с ограниченными первыми двумя моментами (есть МО и D) и для любого $\varepsilon>0$:

$$P(|X-m|\geq\varepsilon)\leq\frac{D(x)}{\varepsilon^2};\qquad P(|X-m|<\varepsilon)\geq 1-\frac{D(x)}{\varepsilon^2}$$

Доказательство. По лемме Маркова: рассмотрим не отрицательную СВ Y

$Y=(X-m)^2$ $M(Y)=M(X-m)^2=D(x)$

$P(|X-m|\geq\varepsilon)=P((X-m)^2\geq\varepsilon^2)=P(Y\geq\varepsilon^2)\leq M(Y)/\varepsilon^2=D(x)/\varepsilon^2$.

Требуется только знание дисперсии СВ при любом законе распределения.

Средняя арифметическая простая

$$\bar{x} = \frac{x_1 + x_2 + \ldots + x_n}{n} = \frac{\sum x_i}{n}$$

x_1 x_2 , , … , x_n - индивидуальные значения варьирующего признака

n – число единиц совокупности

Средняя арифметическая взвешенная

$$\bar{x} = \frac{x_1 f_1 + x_2 f_2 + \ldots + x_n f_n}{f_1 + f_2 + \ldots + f_n} = \frac{\sum x_i f_i}{\sum f_i}$$

Средняя арифметическая суммы варьирующих величин равна сумме средних арифметических величин.

$$\bar{x} = \frac{\sum x_i}{n} = \frac{\sum (y_i + z_i)}{n} = \frac{\sum y_i}{n} + \frac{\sum z_i}{n} = \bar{y} + \bar{z}$$

Алгебраическая сумма отклонений индивидуальных значений варьирующего признака от средней равна нулю, так как сумма отклонений в одну сторону погашается суммой отклонений в другую сторону, то есть

$$\sum(x_i - \bar{x}) = 0$$, потому что

$$\sum(x_i - \bar{x}) = \sum x_i - \sum \bar{x} = \sum x_i - n\bar{x} = \sum x_i - n\frac{\sum x_i}{n}$$

Если все варианты ряда уменьшить или увеличить на одно и то же число *а*, то средняя уменьшится или увеличится на это же число *а*.

$$\bar{x} = \frac{\sum x_i f_i}{\sum f_i} = \frac{\sum (x_i \mu a)}{\sum f_i} \pm a$$

Если все варианты ряда уменьшить или увеличить в *А* раз, то средняя также соответственно уменьшится или увеличится в *А* раз.

$$\bar{x} = \frac{\sum x_i f_i}{\sum f_i} = \frac{\sum \frac{x_i}{A} f_i}{\sum f_i} \times A = \frac{\sum A x_i f_i}{\sum f_i} \div A$$

Если все частоты ряда разделить или умножить на одно и то же число *d*, то средняя не изменится

$$\bar{x} = \frac{\sum x_i f_i}{\sum f_i} = \frac{\sum x_i \dfrac{f_i}{d}}{\sum \dfrac{f_i}{d}} = \frac{\sum x_i f_i d_i}{\sum f_i d_i}$$

«Способ моментов»

$$\overline{x}_{\text{àð}} = m_1 i + A = \frac{\sum\left(\dfrac{x-A}{i}\right)f}{\sum f} + A$$

$$m_1 = \frac{\overline{x_{\text{àð}}} - A}{i} \qquad m_1 = \frac{\sum x_1 f}{\sum f} \qquad x_1 = \frac{x-A}{i}$$

Средняя гармоническая взвешенная

$$\boxed{x_i \times f_i} = w_i$$

$$\bar{x}_{\tilde{a}\grave{a}\eth} = \frac{\sum x_i f_i}{\sum f_i} = \frac{\sum w_i}{\sum \frac{w_i}{x_i}} = \frac{w_1 + w_2 + \ldots + w_n}{\frac{w_1}{x_1} + \frac{w_2}{x_2} + \ldots + \frac{w_n}{x_n}}$$

$$\boxed{f_i} = \frac{w_i}{x_i}$$

Средняя гармоническая простая

$$\bar{x}_{\text{ãàð}} = \frac{1+1+\ldots+1}{\dfrac{1}{x_1}+\dfrac{1}{x_2}+\ldots+\dfrac{1}{x_n}} = \frac{n}{\sum \dfrac{1}{x_n}}$$

где - $\dfrac{1}{x}$ отдельные варианты обратного признака, встречающиеся по одному разу;

n – число вариантов.

Взвешенная гармоническая средняя из групповых средних

$$\bar{x}_{гар} = \frac{n_1 + n_2}{\dfrac{n_1}{\bar{x}_{1гар}} + \dfrac{n_2}{\bar{x}_{2гар}}}$$

Средняя геометрическая

$$\overline{K} = \sqrt[n-1]{\frac{y_n}{y_i}}$$

$$\bar{\hat{E}} = \sqrt[n]{\frac{y_2}{y_1} \times \frac{y_3}{y_2} \times \frac{y_4}{y_3} \times \ldots \times \frac{y_n}{y_{n-1}}}$$

где n – число лет, а не коэффициентов.

Средняя геометрическая простая и взвешенная

$$\bar{x}_г = \sqrt[n]{x_1 x_2 x_3 \ldots x_n} = \sqrt[n]{\Pi x_i}$$

$$\bar{x}_г = \sqrt[\sum f]{\Pi (x_i)^{f_i}}$$

Средняя квадратическая простая

$$\bar{x}_{кв} = \sqrt{\frac{x_1^2 + x_2^2 + \ldots + x_n^2}{n}} = \sqrt{\frac{\sum x_i^{\ 2}}{n}}$$

Средняя квадратическая взвешенная

$$\bar{x}_{кв} = \sqrt{\frac{\sum x_i^2 f_i}{\sum f_i}}$$

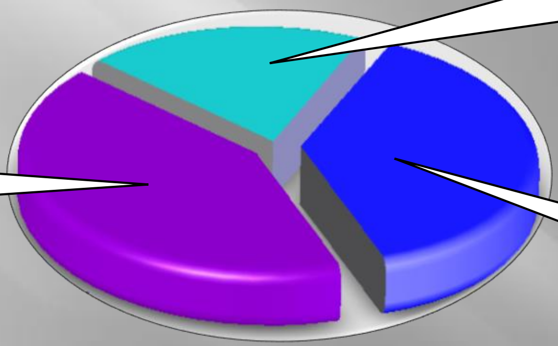
Средняя кубическая

простая: ***взвешенная:***

$$\bar{x}_{куб} = \sqrt[3]{\frac{\sum x^3}{n}}$$

$$\bar{x}_{куб} = \sqrt[3]{\frac{\sum x_i^3 f_i}{\sum f_i}}$$

Деление совокупности на несколько групп



$$\overline{\tilde{O}1} = \frac{\sum \tilde{O}ifi}{\sum fi}$$

$$\overline{\tilde{O}2}$$

$$\overline{\tilde{O}3}$$

$$\bar{X}_{\hat{\imath}\,\acute{a}\grave{u}\,.} = \frac{\sum \overline{\tilde{O}i} * Fi}{\sum Fi}$$

М о д а

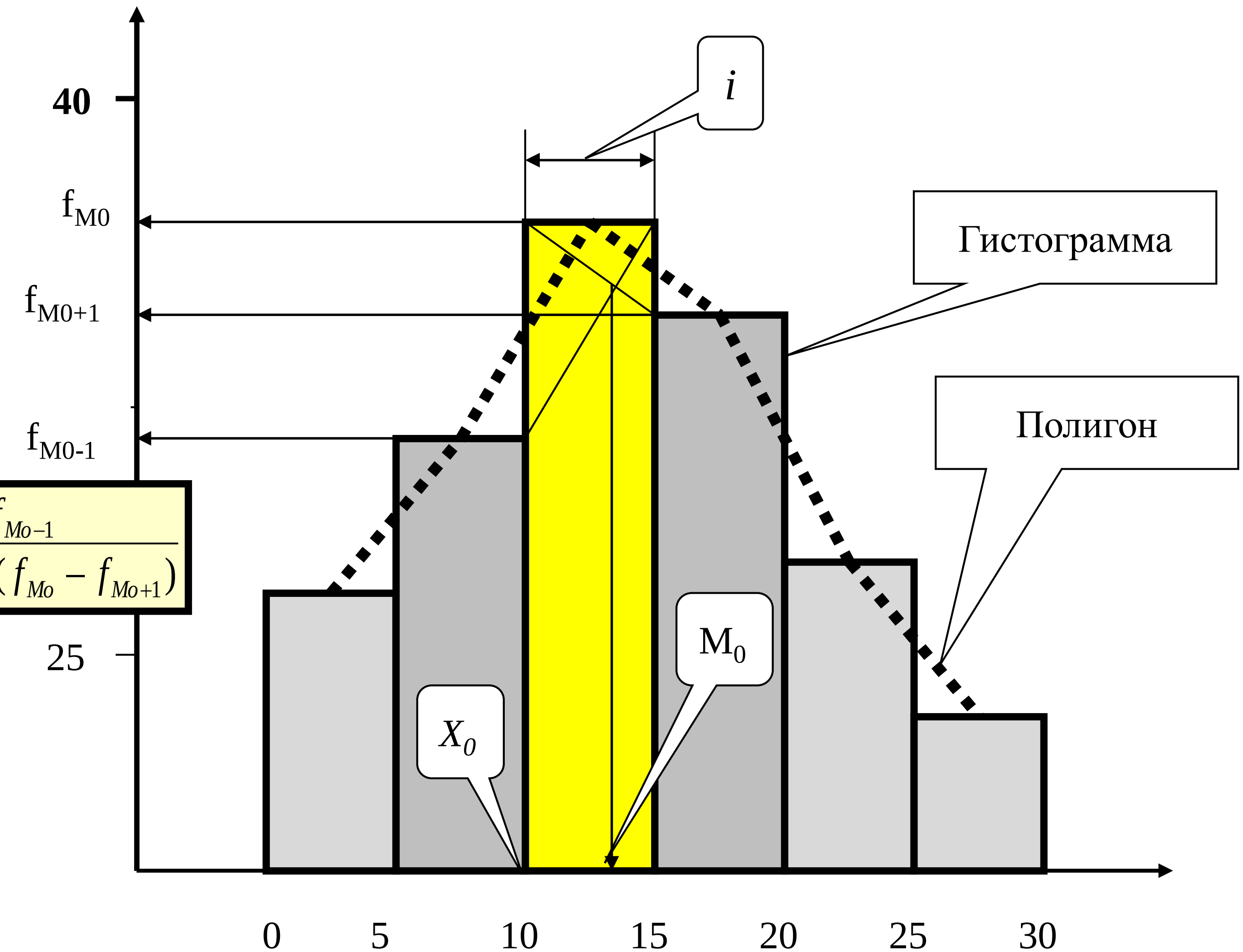
$$M_o = x_{Mo} + i_{Mo} \frac{f_{Mo} - f_{Mo-1}}{(f_{Mo} - f_{Mo-1}) + (f_{Mo} - f_{Mo+1})}$$

Где X_{Mo} - нижняя граница модального интервала

i_{Mo} - величина модального интервала

f_{Mo}, f_{Mo-1}, f_{Mo+1} - частоты в модальном, предыдущем и следующим за модальным интервалах соответственно

Графическое отображение моды

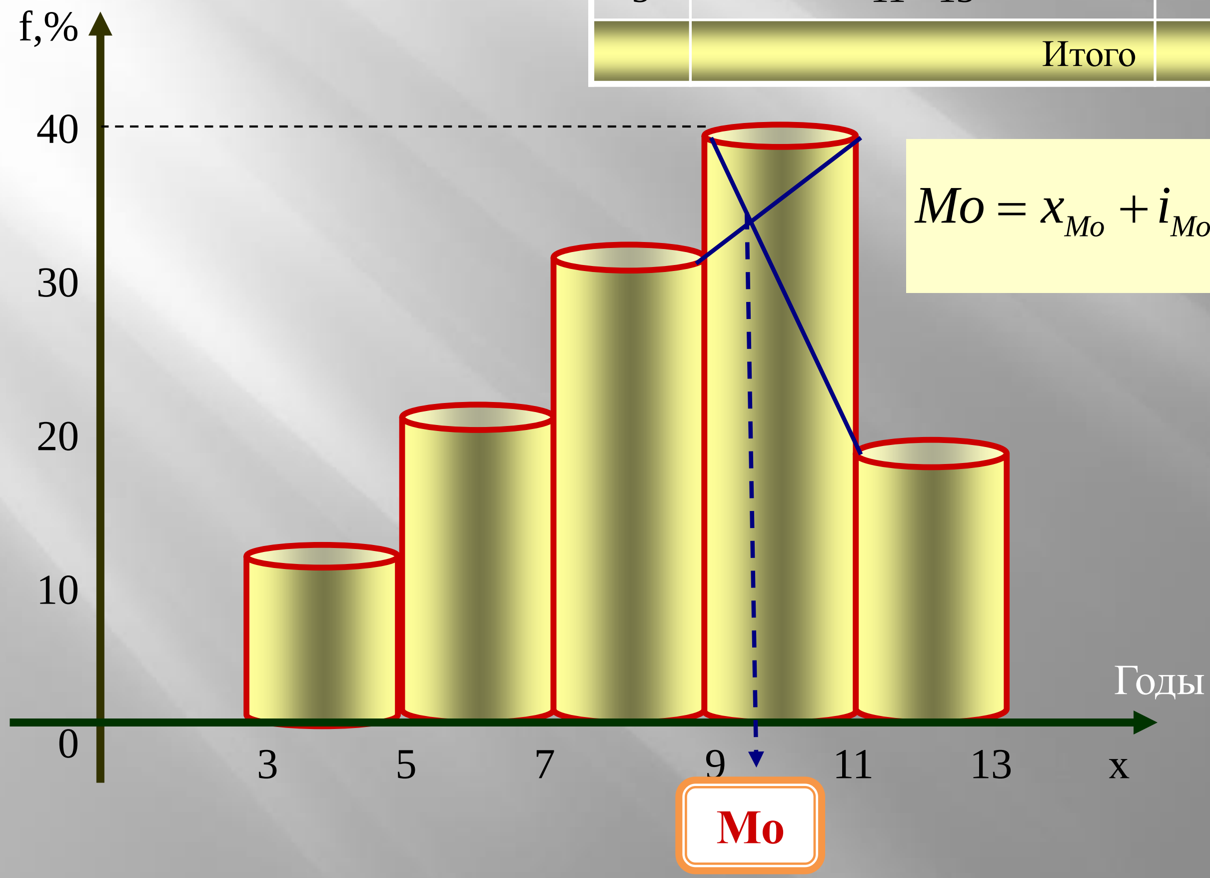


$$M_o = x_{Mo} + i_{Mo} \frac{f_{Mo} - f_{Mo-1}}{(f_{Mo} - f_{Mo-1}) + (f_{Mo} - f_{Mo+1})}$$

Мода

№ п/п	Группы семей по размеру жилой площади, на одного человека, м^2 x	Число семей с данными размерами жилой площади f	Накопленное число семей S
1	3 – 5	10	10
2	5 – 7	20	30
3	7 – 9	30	60
4	9 – 11	40	100
5	11 - 13	15	115
	Итого	115	-

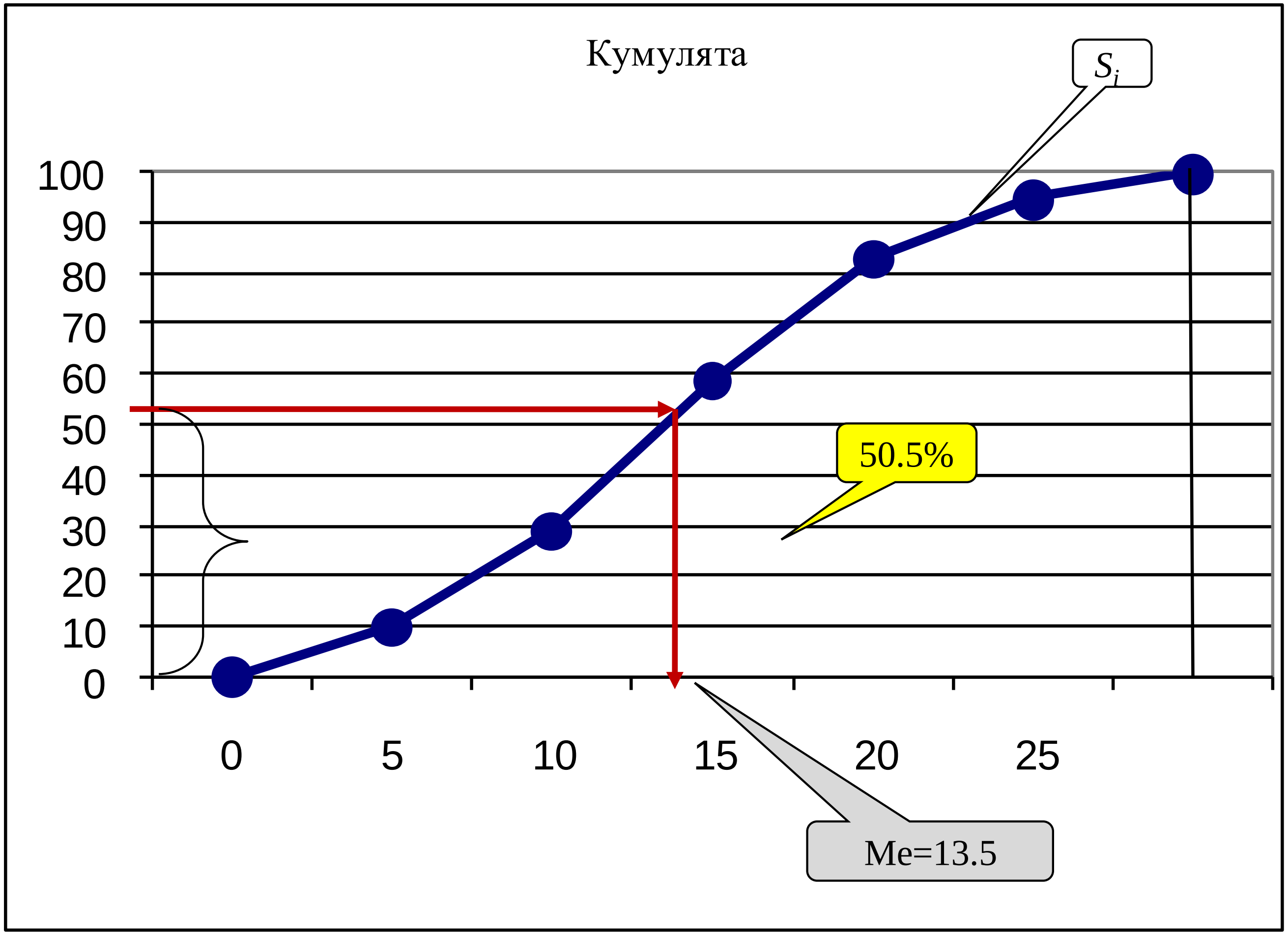
$$Mo = x_{Mo} + i_{Mo} \times \frac{f_{Mo} - f_{Mo-1}}{\left(f_{Mo} - f_{Mo-1}\right) + \left(f_{Mo} - f_{Mo+1}\right)}$$



Медиана интервального ряда распределения

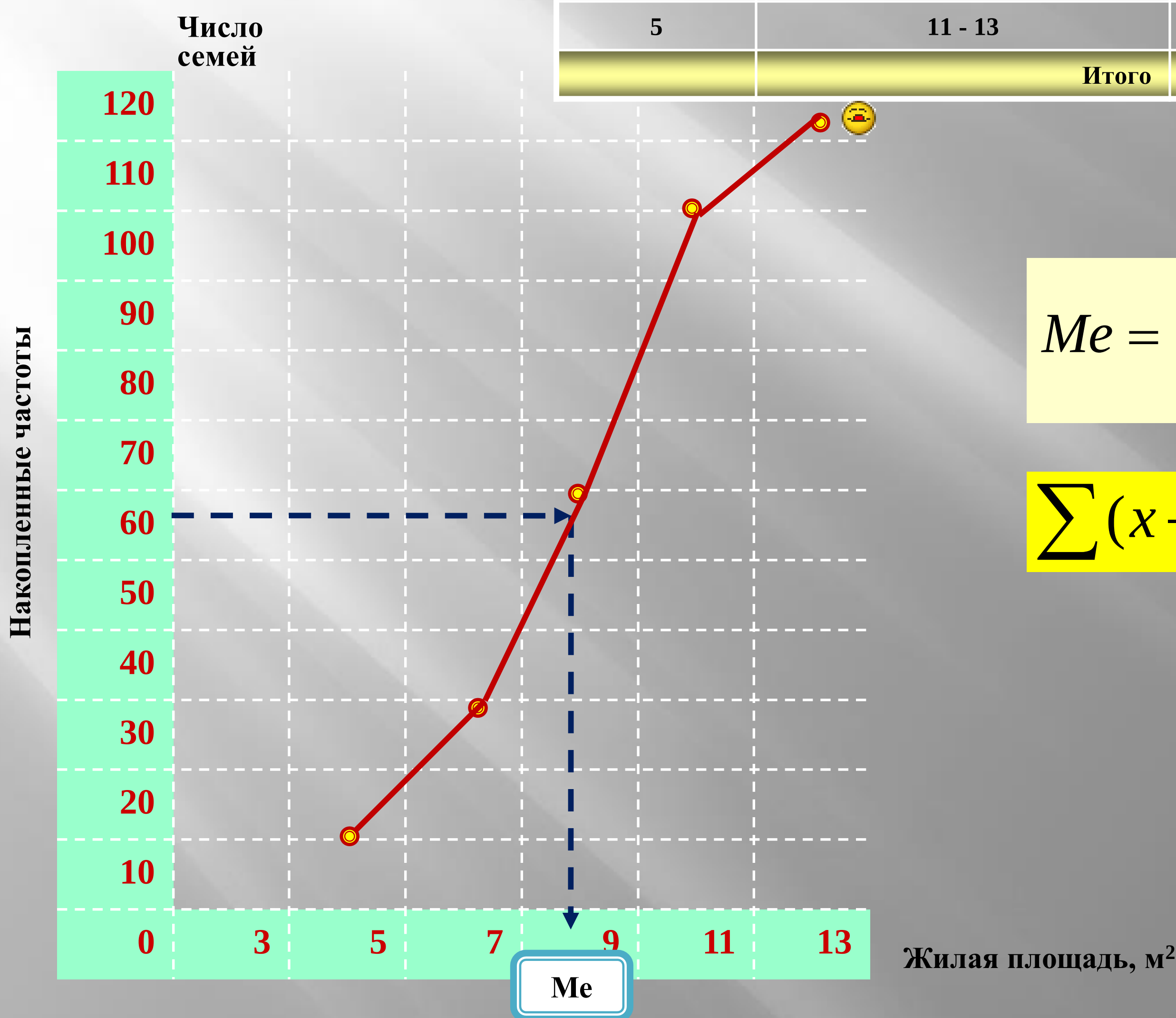
$$Me = x_{Me} + i_{Me} \frac{\frac{\sum f}{2} - S_{Me-1}}{f_{Me}}$$

Графическое отображение медианы (кумулята)



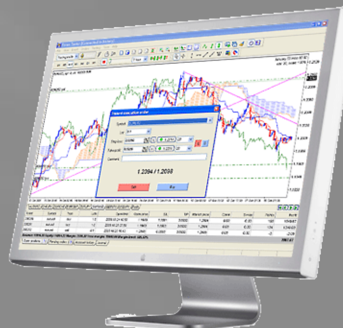
Медиана

№ п/п	Группы семей по размеру жилой площади, на одного человека, $м^2$ x	Число семей с данными размерами жилой площади f	Накопленное число семей S
1	3 – 5	10	10
2	5 – 7	20	30
3	7 – 9	30	60
4	9 – 11	40	100
5	11 - 13	15	115
	Итого	115	-



$$Me = x_{Me} + i_{Me}\frac{0{,}5f - S_{Me-1}}{f_{Me}}$$

$$\sum(x - Me) = \min$$




Показатели вариации
Линейные
Квадратичные
Абсолютные
Средние
Относительные
Размах
Среднее линейное отклонение
Среднее относительное отклонение
Относительные
Средние
Абсолютные
Коэффициент вариации
Дисперсия
Среднее квадратическое отклонение
Сумма квадратов отклонений

Среднее линейное отклонение для несгруппированных данных

$$\bar{d} = \frac{\sum \left| x_i - \bar{x} \right|}{n}$$

где *n* – число членов ряда

$$d = \frac{\sum / x_i - \overline{x} /}{n} = \frac{/ x_1 - \overline{x} /+ / x_2 - \overline{x} /+ \ldots + / x_n - \overline{x} /}{n}$$

- ***Порядок расчета среднего линейного отклонения следующий:***

1) по значениям признака исчисляется средняя арифметическая:

$$\overline{x} = \frac{\sum x_i}{n}$$

2) определяются отклонения каждой варианты x_i от средней $x_i - \overline{x}$

3) рассчитывается сумма абсолютных величин отклонений: $\sum / x_i - \overline{x} /$

4) сумма абсолютных величин отклонений делится на число значений:

$$\frac{\sum / x_i - \overline{x} /}{n}$$

Среднее линейное отклонение для сгруппированных данных

$$\bar{d} = \frac{\sum |x_i - \bar{x}| f_i}{\sum f_i}$$

где $\sum f$ - сумма частот вариационного ряда

$$d = \frac{\sum / x_i - \overline{x} / f_i}{\sum f_i} = \frac{/ x_1 - \overline{x} / f_1 + / x_2 - \overline{x} / f_2 + \ldots + / x_n - \overline{x} / f_i}{f_1 + f_2 + \ldots + f_i}$$

- ***Порядок расчета среднего линейного отклонения взвешенного следующий:***

1) вычисляется средняя арифметическая взвешенная: $\frac{\sum x_i f_i}{\sum f_i}$
2) определяются абсолютные отклонения вариант от средней $/ x_i - \overline{x} /$;
3) полученные отклонения умножаются на частоты ;
4) находится сумма взвешенных отклонений без учета знака: $\sum / x_i - \overline{x} / f_i$
5) сумма взвешенных отклонений делится на сумму частот:

$$\frac{\sum / x_i - \overline{x} / f_i}{\sum f_i}$$

Простая дисперсия для несгруппированных данных

$$\sigma^2 = \frac{\sum (x_i - \bar{x})^2}{n}$$

Порядок расчета дисперсии простой:

1. определяют среднюю арифметическую $\bar{x} = \frac{\sum x}{n}$

2. возводят в квадрат среднюю арифметическую $\bar{x}^2 = \left(\frac{\sum x}{n}\right)^2$

3. возводят в квадрат каждую варианту ряда x_i^2

4. находим сумму квадратов вариант $\sum x_i^2$

5. делят сумму квадратов вариант на их число, т.е. определяют средний квадрат $\overline{x^2} = \frac{\sum x_i^2}{n}$

6. определяют разность между средним квадратом признака и квадратом средней $\overline{x^2} - \bar{x}^2$

Взвешенная дисперсия для вариационного ряда

$$\sigma^2 = \frac{\sum (x_i - \bar{x})^2 f_i}{\sum f_i}$$

Порядок расчета дисперсии взвешенную:

1. определяют среднюю арифметическую взвешенную $\dfrac{\sum x_i f_i}{\sum f_i}$
2. определяются отклонения вариант от средней $\left(x_i - \bar{x}\right)$
3. возводят в квадрат отклонение каждой варианты от средней $\left(x_i - \bar{x}\right)^2$
4. умножают квадраты отклонений на веса (частоты) $\left(x_i - \bar{x}\right)^2 f_i$
5. суммируют полученные произведения $\sum \left(x_i - \bar{x}\right)^2 f_i$
6. полученную сумму делят на сумму весов $\dfrac{\sum \left(x_i - \bar{x}\right)^2 f_i}{\sum f_i}$

Расчет дисперсии и среднего квадратического отклонения по индивидуальным данным и в рядах распределения.

- дисперсия невзвешенная (простая)

$$\sigma^2 = \frac{\sum (x_i - \overline{x})^2}{\sum n}$$

- дисперсия взвешенная

$$\sigma^2 = \frac{\sum (x_i - \overline{x})^2 f_i}{\sum f_i}$$

- среднее квадратическое отклонение невзвешенное

$$\sigma = \sqrt{\frac{\sum (x_i - \overline{x})^2}{n}}$$

- среднее квадратическое отклонение взвешенное

$$\sigma = \sqrt{\frac{\sum (x_i - \overline{x})^2 f_i}{\sum f_i}}$$

Дисперсия, исчисленная способом моментов

$$\sigma^2 = i^2(m_2 - m_1) = i^2\left(\frac{\sum x_1^2 f_i}{\sum f_i} - \left(\frac{\sum x_1 f_i}{\sum f_i}\right)^2\right)$$

где i - величина интервала;

$$x_1 = \frac{x - A}{i}$$

- новые (преобразованные) значения вариант (*A* – условный ноль, в качестве которого удобно использовать середину интервала, обладающего наибольшей частотой);

$$m_2 = \frac{\sum x_1^2 f}{\sum f}$$ - момент второго порядка;

$$m_1 = \left(\frac{\sum x_1 f}{\sum f}\right)^2$$ - квадрат момента первого порядка;

Среднее квадратическое отклонение

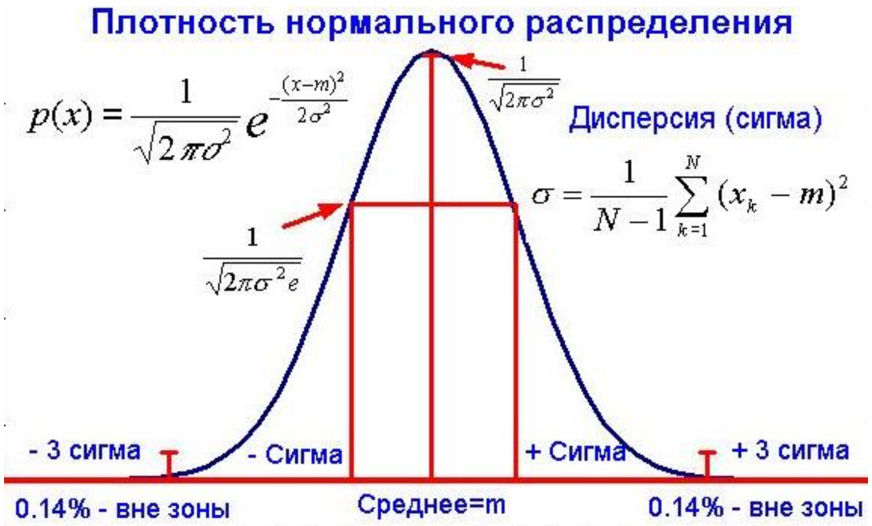
Для несгруппированных данных:

$$\sigma = \sqrt{\frac{\sum \left(x_i - \bar{x} \right)^2}{n}}$$

Для вариационного ряда:

$$\sigma = \sqrt{\frac{\sum \left(x_i - \bar{x} \right)^2 f}{\sum f}}$$

Правила трех сигм

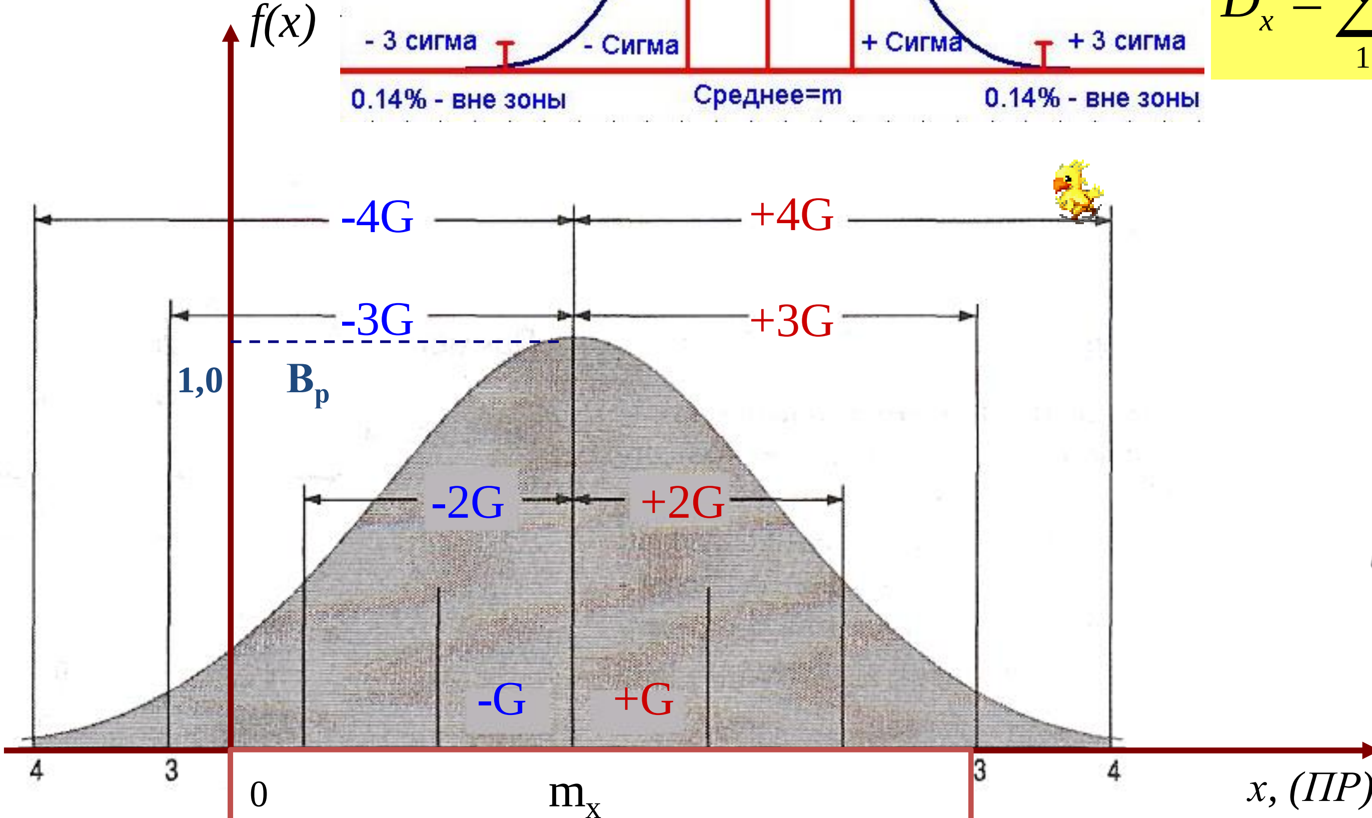


$$f(x) = \frac{1}{\sigma_x\sqrt{2\pi}} \exp\left[-\frac{(x-m_x)^2}{2\sigma_x^2}\right]$$

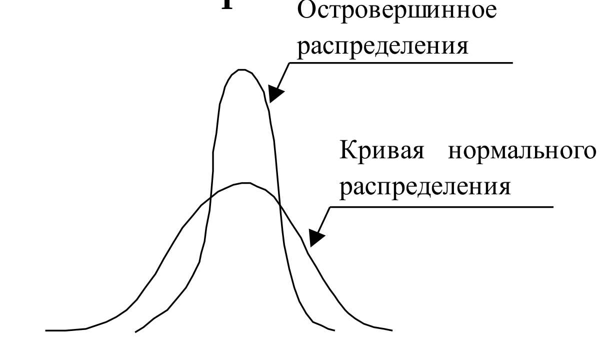
$$m_x = \sum_{i=1}^{k} x_i' \ p_i$$

$$\sigma_x = \sqrt{D_x}$$

$$D_x = \sum_{1}^{k}(x_i' - m_x)^2 \ p_i$$

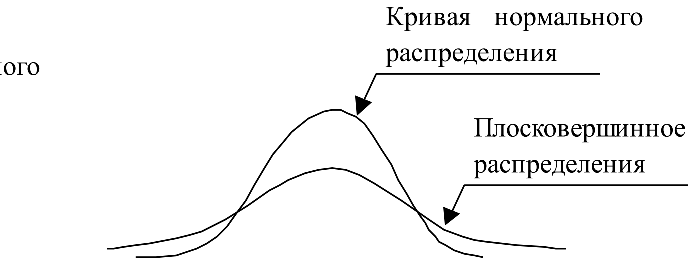


I

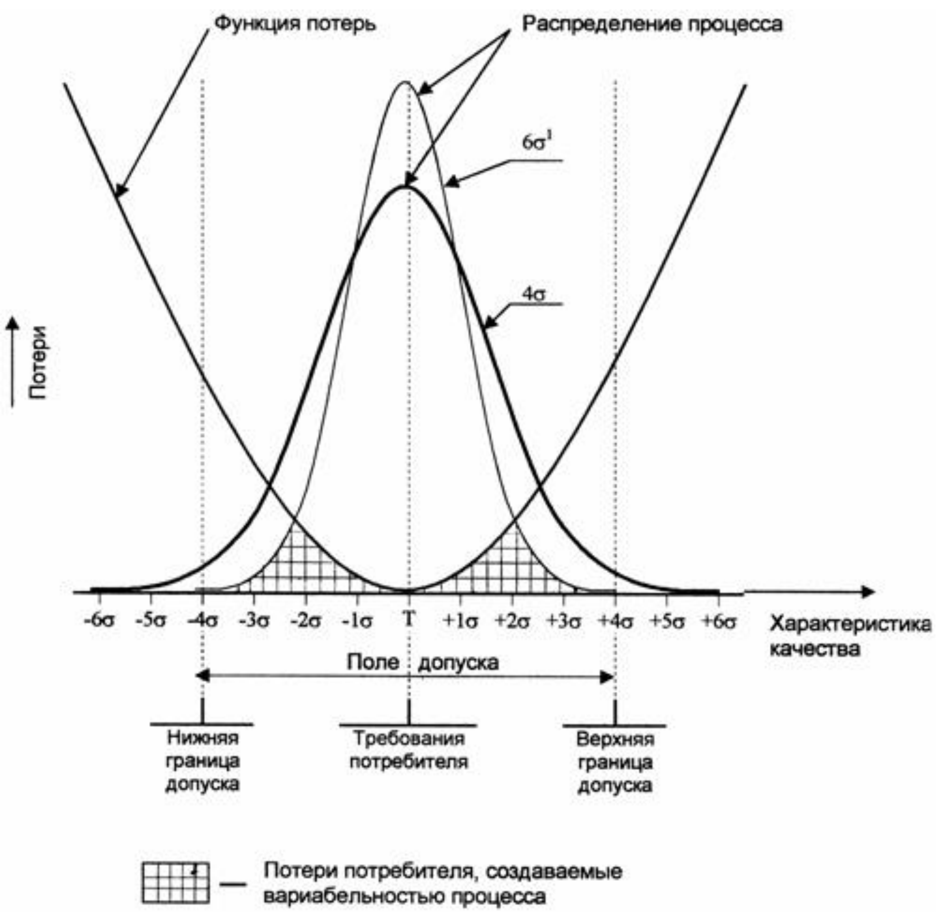


Островершинное распределения
Кривая нормального распределения

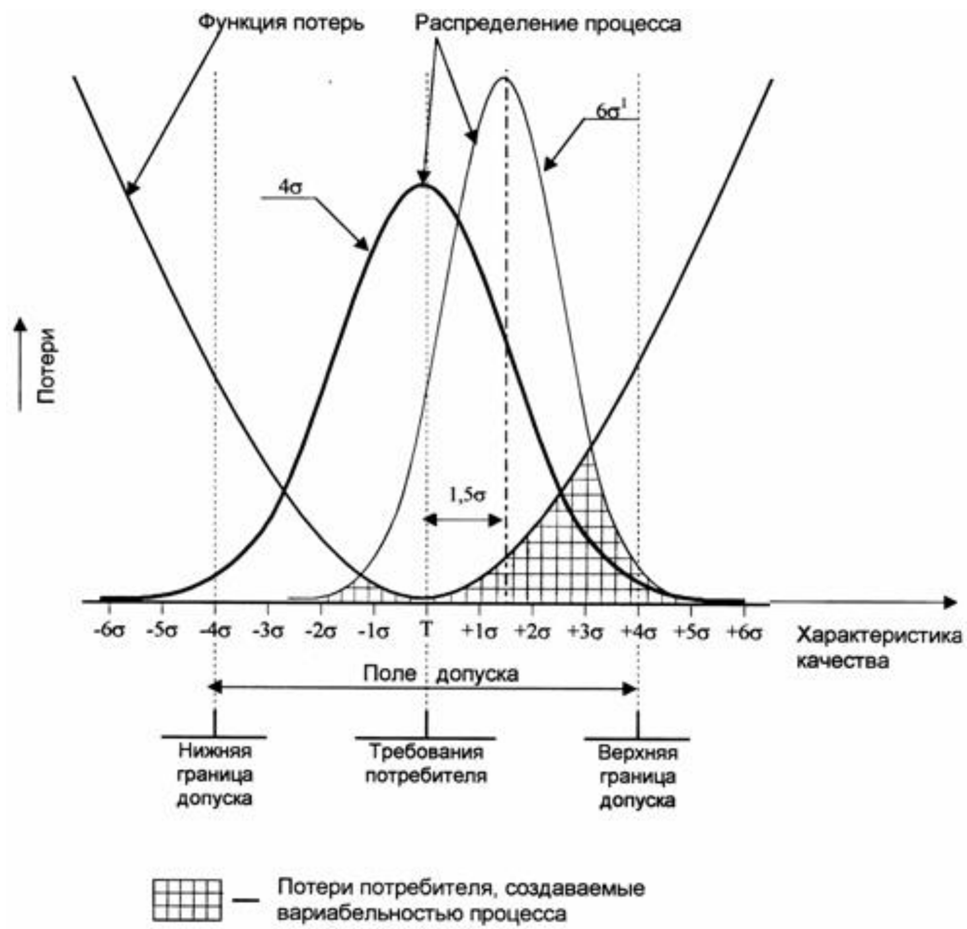
II



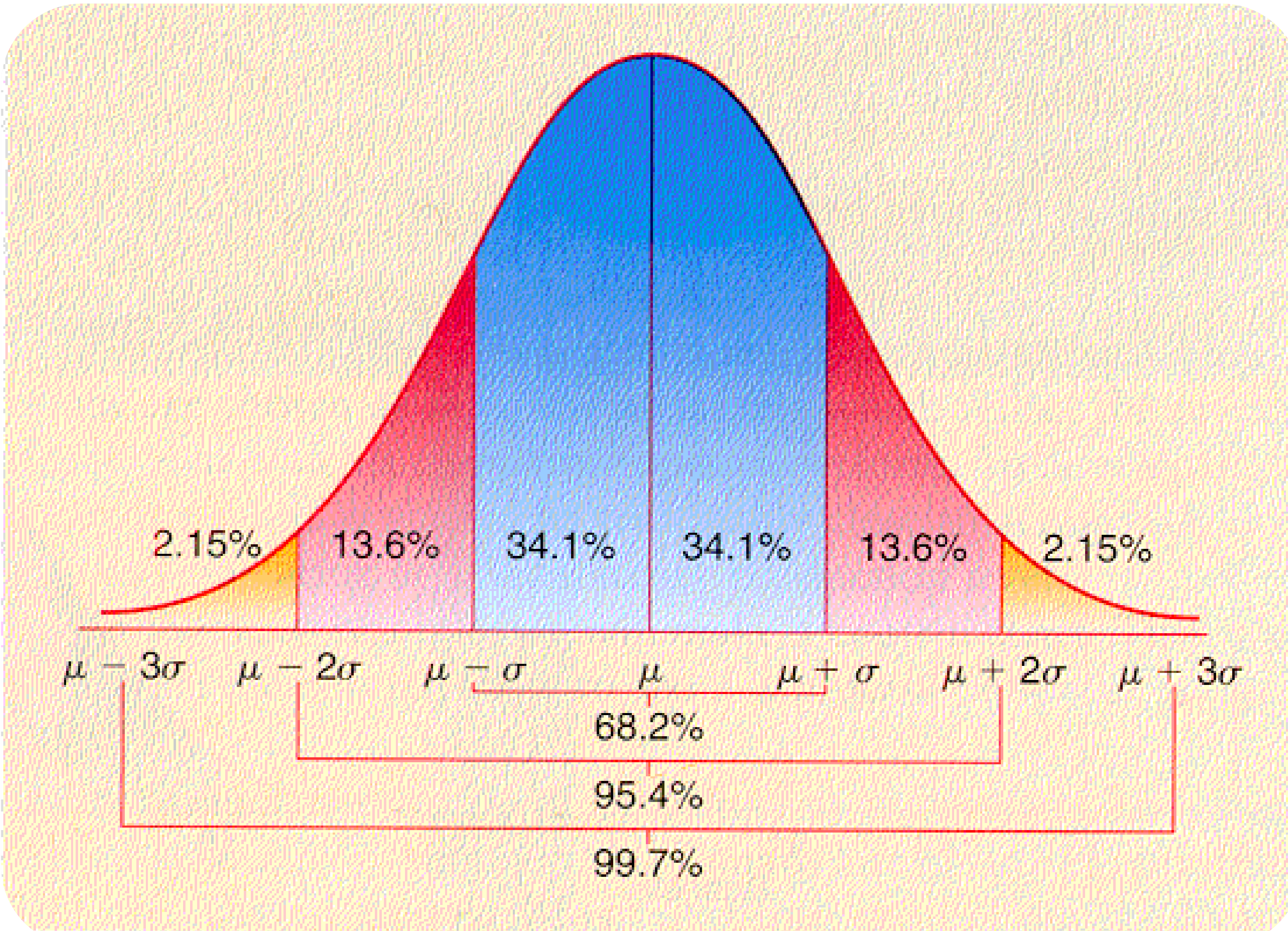
Кривая нормального распределения
Плосковершинное распределения



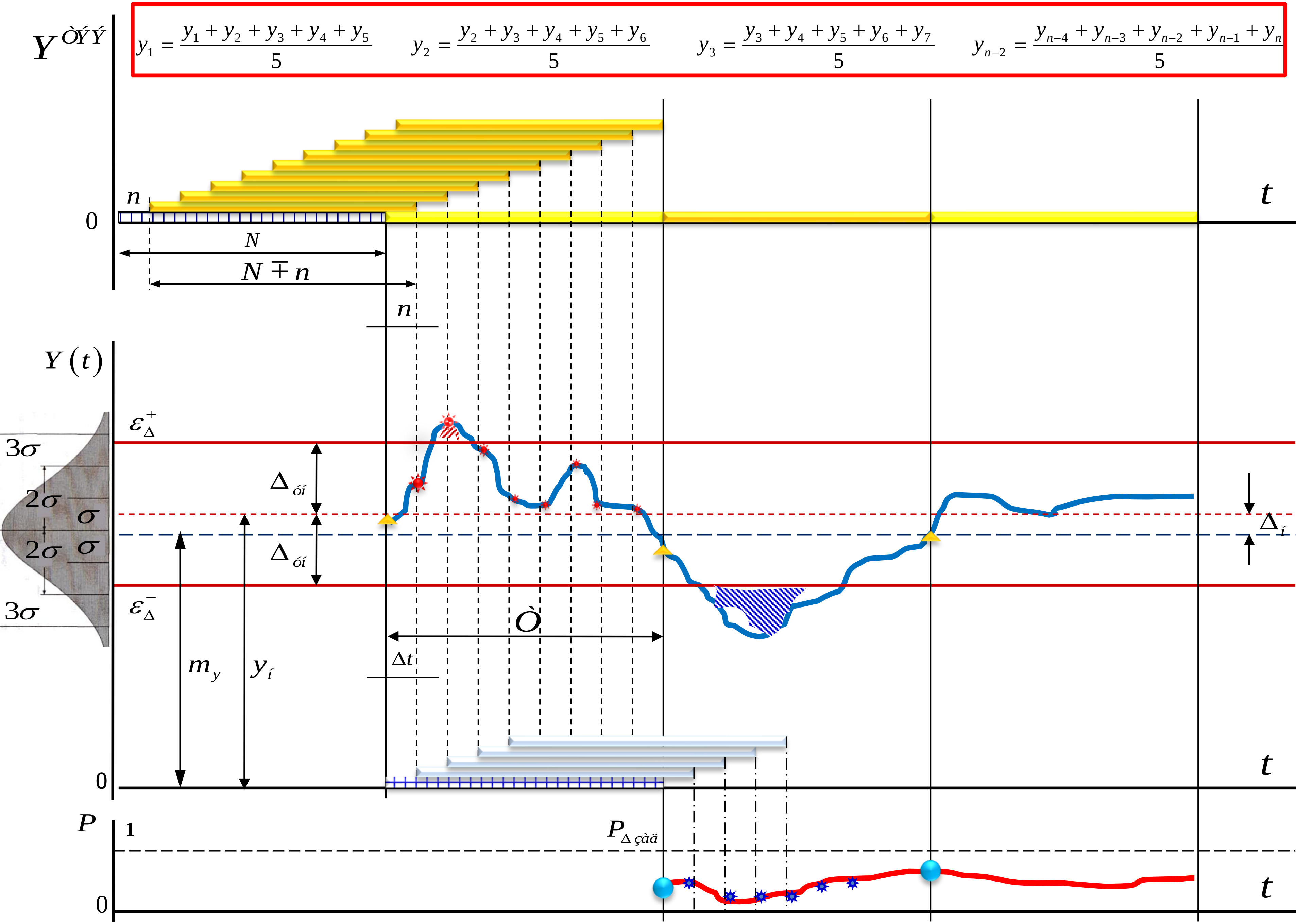
Функция потерь
Распределение процесса
$6\sigma^1$
4σ
Потери
-6σ -5σ -4σ -3σ -2σ -1σ T +1σ +2σ +3σ +4σ +5σ +6σ
Характеристика качества
Поле допуска
Нижняя граница допуска
Требования потребителя
Верхняя граница допуска
— Потери потребителя, создаваемые вариабельностью процесса



Функция потерь
Распределение процесса
$6\sigma^1$
4σ
Потери
1,5σ
-6σ -5σ -4σ -3σ -2σ -1σ T +1σ +2σ +3σ +4σ +5σ +6σ
Характеристика качества
Поле допуска
Нижняя граница допуска
Требования потребителя
Верхняя граница допуска
— Потери потребителя, создаваемые вариабельностью процесса



2.15%
13.6%
34.1%
34.1%
13.6%
2.15%
$\mu - 3\sigma$
$\mu - 2\sigma$
$\mu - \sigma$
μ
$\mu + \sigma$
$\mu + 2\sigma$
$\mu + 3\sigma$
68.2%
95.4%
99.7%



$Y^{ÒÝÝ}$
$y_1 = \frac{y_1 + y_2 + y_3 + y_4 + y_5}{5}$
$y_2 = \frac{y_2 + y_3 + y_4 + y_5 + y_6}{5}$
$y_3 = \frac{y_3 + y_4 + y_5 + y_6 + y_7}{5}$
$y_{n-2} = \frac{y_{n-4} + y_{n-3} + y_{n-2} + y_{n-1} + y_n}{5}$
n
0
t
N
$N \mp n$
n
$Y(t)$
ε_Δ^+
3σ
2σ
σ
$\Delta_{óí}$
$\Delta_í$
ε_Δ^-
$Ò$
Δt
m_y
$y_í$
0
t
P
1
$P_{\Delta\, çàä}$
0
t

Среднее значение альтернативного признака

$$\bar{x} = \frac{\sum xf}{\sum f} = \frac{1 \times p + 0 \times q}{p + q} = p$$

Дисперсия альтернативного признака

$$\sigma^2_p = \frac{\sum \left(x - \bar{x}\right)^2 f}{\sum f} = \frac{\left(1 - p\right)^2 p + \left(0 - p\right)^2 q}{p + q}$$

Среднее квадратическое отклонение альтернативного признака

$$\sigma_p = \sqrt{pq} = \sqrt{p\left(1 - p\right)}$$

Межгрупповая дисперсия

$$\sigma^2 = \frac{\sum\left(\bar{x}_i - \bar{x}\right)^2 f}{\sum f}$$

Внутригрупповая дисперсия

$$\sigma_i^2 = \frac{\sum \left(x - \bar{x}_i\right)^2}{n}$$

$$\sigma_i^2 = \frac{\sum \left(x - \bar{x}_i\right)^2 f}{\sum f}$$

Общая средняя из внутригрупповых дисперсий

$$\overline{\sigma_i^2} = \frac{\sum \sigma_i^2 f}{\sum f}$$

Правило сложения дисперсий

$$\sigma^2 = \overline{\sigma_i^2} + \delta^2$$

Эмпирический коэффициент детерминации

$$\eta^2 = \frac{\delta^2}{\sigma^2}$$

Эмпирическое корреляционное отношение

$$\eta = \sqrt{\frac{\delta^2}{\sigma^2}}$$

Соотношения Чэддока
(качественная оценка тесноты связи)

$\eta_э$	Сила связи
0,1- 0,3	Слабая
0,3- 0,5	Умеренная
0,5- 0,7	Заметная
0,7- 0,9	Тесная
0,9- 0,99	Весьма тесная

Коэффициент вариации

$$V_\sigma = \frac{\sigma}{\bar{x}} \times 100$$

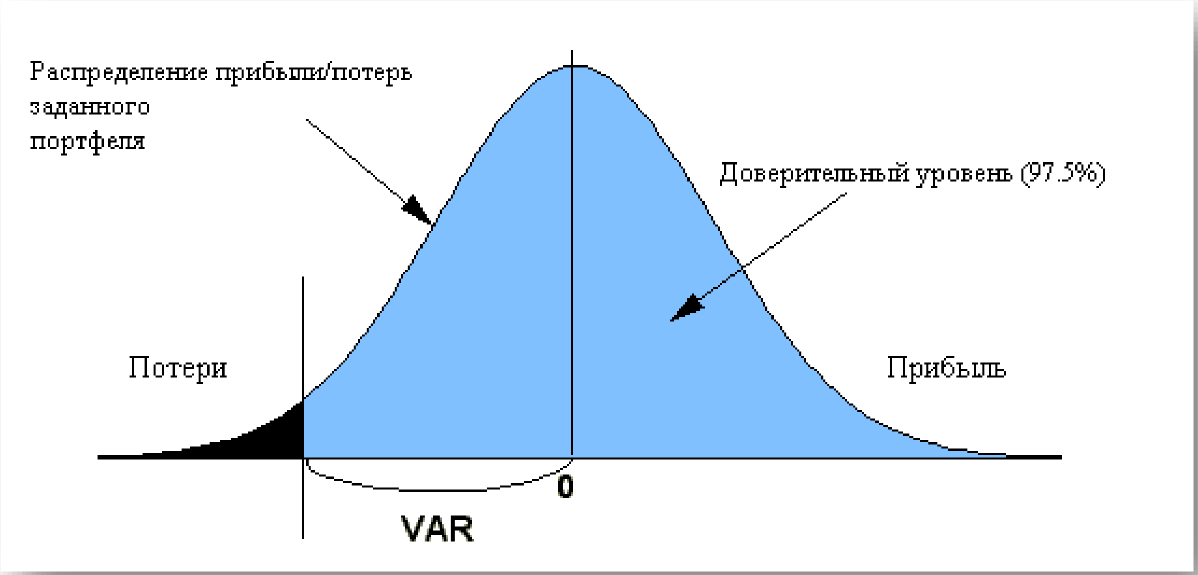
Линейный коэффициент вариации

$$V_d = \frac{\bar{d} \times 100}{\bar{x}}$$

Коэффициент осцилляции

$$V_R = \frac{R \times 100}{\bar{x}}$$

Применение VAR для управления рыночным риском



Кривая на рисунке задает распределение вероятностей прибылей и потерь для заданных портфеля и периода поддержания позиций. Заштрихованная светлым область соответствует выбранному доверительному уровню (97,5%) в том смысле, что ее площадь составляет 97,5% от общей площади под кривой. VAR представляет собой величину возможных потерь, отвечающих заданному доверительному уровню.


ИНДЕКСЫ
Объемных показателей
Качественных показателей
Общие
Групповые
Индивидуальные
Агрегатные
Средние из индивидуальных
Цепные
Базисные

ИНДЕКСЫ

Классификация экономических индексов

Индивидуальный индекс цен

$$i_p = \frac{p_1}{p_0}$$

Индивидуальный индекс физического объема

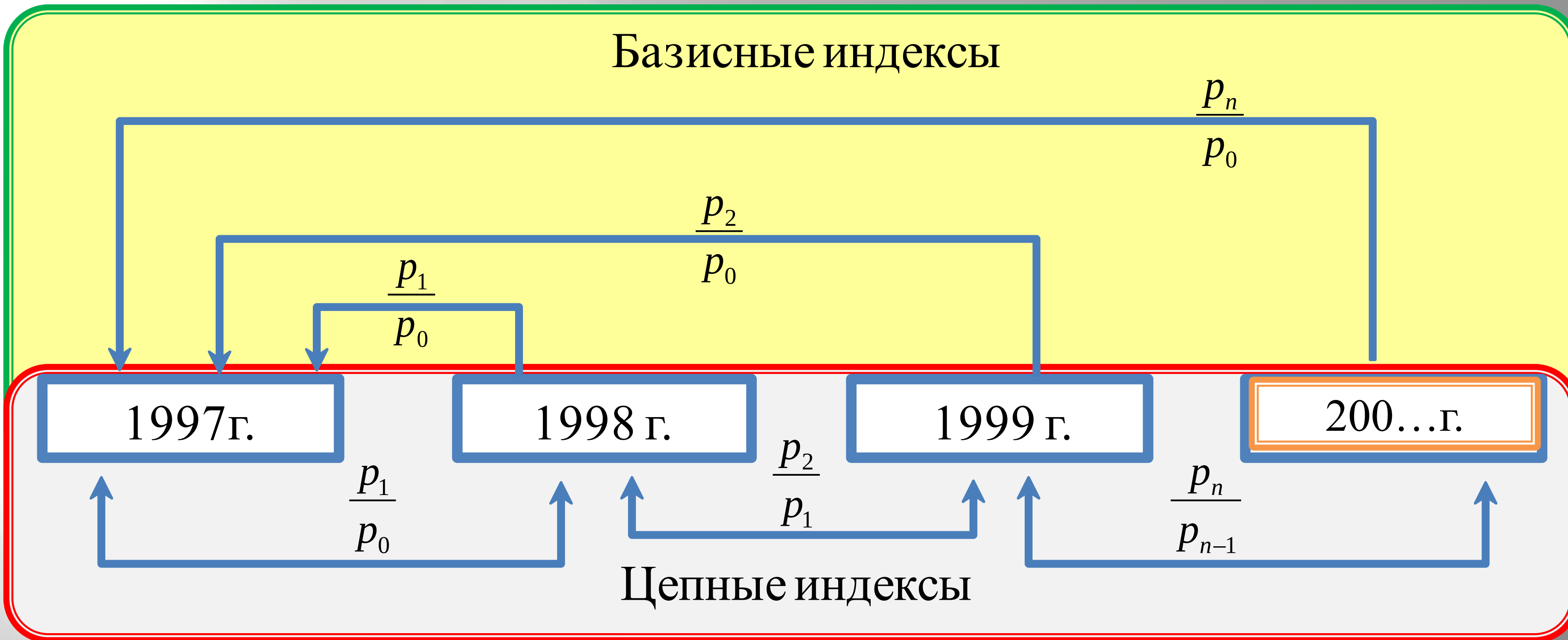
$$i_q = \frac{q_1}{q_0}$$

Индивидуальный индекс выручки

$$i_{pq} = \frac{p_1 q_1}{p_0 q_0}$$

Индивидуальный индекс выручки, рассчитанный через взаимосвязь индексов

$$i_{pq} = i_p i_q$$



$$\frac{p_2}{p_1} = \frac{p_2}{p_0} : \frac{p_1}{p_0}$$

или

$$\frac{q_2}{q_1} = \frac{q_2}{q_0} : \frac{q_1}{q_0}$$

Система индивидуальных индексов

Название индивидуального индекса	Система индексов	
	базисных	*цепных*
Индекс стоимости	$\frac{p_1q_1}{p_0q_0}$; $\frac{p_2q_2}{p_0q_0}$;...; $\frac{p_nq_n}{p_0q_0}$	$\frac{p_1q_1}{p_0q_0}$; $\frac{p_2q_2}{p_1q_1}$;...; $\frac{p_nq_n}{p_{n-1}q_{n-1}}$
Индекс физического объема	$\frac{q_1}{q_0}$; $\frac{q_2}{q_0}$;...; $\frac{q_n}{q_0}$	$\frac{q_1}{q_0}$; $\frac{q_2}{q_1}$;....; $\frac{q_n}{q_{n-1}}$
Индекс цен	$\frac{p_1}{p_0}$; $\frac{p_2}{p_0}$; ... ; $\frac{p_n}{p_0}$	$\frac{p_1}{p_0}$; $\frac{p_2}{p_1}$; ... ; $\frac{p_n}{p_{n-1}}$

Взаимосвязь между индексами.

1) Произведение общих цепных индексов дает базисный индекс последнего периода. Пусть мы имеем 3 периода 1997, 1998, 1999.

$$I_q^{98/97} \times I_q^{99/98} = I_q^{99/97}$$

Эта взаимосвязь имеет место лишь в цепных индексах физического объема (индексах с постоянными весами). В индексах цен, так же и в других индексах с переменными весами, такой взаимосвязи нет.

2) Отношение последующего базисного индекса к предшествующему равно цепному индексу последующего периода:

$$I_q^{99/97} \div I_q^{98/97} = I_q^{99/98}$$

3) Поскольку величина объема продукции равна произведению количества продукции на цену, то индекс физического объема (I_q), умноженный на индекс цен (I_p) дает индекс стоимости продукции в фактических ценах (I_{qp}):

$$I_q \times I_p = I_{qp}$$

4) Индекс изменения средней величины ($I_{пер}$) равен произведению индекса в неизменной структуре ($I_{пост}$) на индекс, отображающий влияние изменения структуры явления на динамику средней величины ($I_{стр}$):

$$I_{ñòð} \times I_{ïîñò.ñ} = I_{ïåð.ñ}$$

или

$$I_{\bar{p}} = I_{\bar{p}}^{p} \times I_{\bar{p}}^{q}$$

Формула агрегатного индекса (Шарль Дюто, 1738 г)

$$I_p^{Д} = \frac{\sum P_1}{\sum P_0}$$

Формула средних индексов (итальянский экономист Карли, 1764 г)

$$I_p^K = \frac{\sum i_p}{n}$$

Агрегатный индекс стоимости продукции

$$I_{pq} = \frac{\sum p_1 q_1}{\sum p_0 q_0}$$

где $\sum p_1 q_1$ - фактическая стоимость продукции отчетного периода;

$\sum p_0 q_0$ - фактическая стоимость продукции базисного периода.

Индекс Ласпейреса

$$I_q = \frac{\sum p_0 q_1}{\sum p_0 q_0}$$

Индекс Ласпейреса

$$I_p = \frac{\sum p_1 q_0}{\sum p_0 q_0}$$

Индекс Пааше

$$I_q = \frac{\sum p_1 q_1}{\sum p_1 q_0}$$

Индекс Пааше

$$I_p = \frac{\sum p_1 q_1}{\sum p_0 q_1}$$

«Идеальный» индекс
И. Фишера

$$I_p^{\Phi} = \sqrt{I_p^{Л} I_p^{П}}$$

Основные формулы исчисления сводных, или общих индексов.

Наименование индекса	Индекс физического объема продукции	Индекс цен	Индекс стоимости продукции (товарооборота)
Формула расчета индекса	$I_q = \frac{\Sigma q_1 p_0}{\Sigma q_0 p_0}$	$I_{ð} = \frac{\Sigma p_1 q_1}{\Sigma p_0 q_1}$	$I_{pq} = \frac{\Sigma p_1 q_1}{\Sigma p_0 q_0}$
Наименование индекса	Индекс физического объема продукции	Индекс себестоимости продукции	Индекс издержек производства
Формула расчета индекса	$I_q = \frac{\Sigma q_1 z_0}{\Sigma q_0 z_0}$	$I_z = \frac{\Sigma z_1 q_1}{\Sigma z_0 q_1}$	$I_{zq} = \frac{\Sigma z_1 q_1}{\Sigma z_0 q_0}$
Наименование индекса	Индекс физического объема продукции	Индекс производительности труда по трудовым затратам	Индекс затрат времени на производство продукции
Формула расчета индекса	$I_q = \frac{\Sigma q_1 t_0}{\Sigma q_0 t_0}$	$I_t = \frac{\Sigma t_1 q_1}{\Sigma t_0 q_1}$	$I_{tq} = \frac{\Sigma t_1 q_1}{\Sigma t_0 q_0}$

Средний гармонический индекс цен

$$I_p^{\Pi} = \frac{\sum p_1 q_1}{\sum p_0 q_1} = \frac{\sum p_1 q_1}{\sum \frac{p_1 q_1}{i_p}}$$

Средний арифметический индекс цен

$$I_p^{Л} = \frac{\sum p_1 q_0}{\sum p_0 q_0} = \frac{\sum i_p p_0 q_0}{\sum p_0 q_0}$$

Средний арифметический индекс физического объема продукции

$$I_q^{Л} = \frac{\sum i_q p_0 q_0}{\sum p_0 q_0}$$

Средний гармонический взвешенный индекс физического объема продукции:

$$I_q^n = \frac{\sum p_1 q_1}{\sum \frac{p_1 q_1}{i_q}}$$

Индекс переменного состава

$$I_{перем} = \frac{\bar{x}_1}{\bar{x}_0} = \frac{\sum x_1 f_1}{\sum f_1} \div \frac{\sum x_0 f_0}{\sum f_0}$$

Индекс постоянного состава

$$I_{пост} = \frac{\sum x_1 f_1}{\sum f_1} \div \frac{\sum x_0 f_1}{\sum f_1}$$

Индекс структурных сдвигов

$$I_{стр} = \frac{\sum x_0 f_1}{\sum f_1} \div \frac{\sum x_0 f_0}{\sum f_0}$$

Интегральный коэффициент структурных различий

$$K_c = \sqrt{\frac{\sum \left(\frac{d_1 - d_0}{d_1 + d_0} \right)^2}{n}}$$

где d_1, d_0 - относительные показатели структуры изучаемых совокупностей в отчетном и базисном периодах соответственно;

n – число структурных составляющих (групп)

Взаимосвязь индексов постоянного, переменного состава и структурных сдвигов

$$I_{перем} = I_{пост} I_{стр}$$

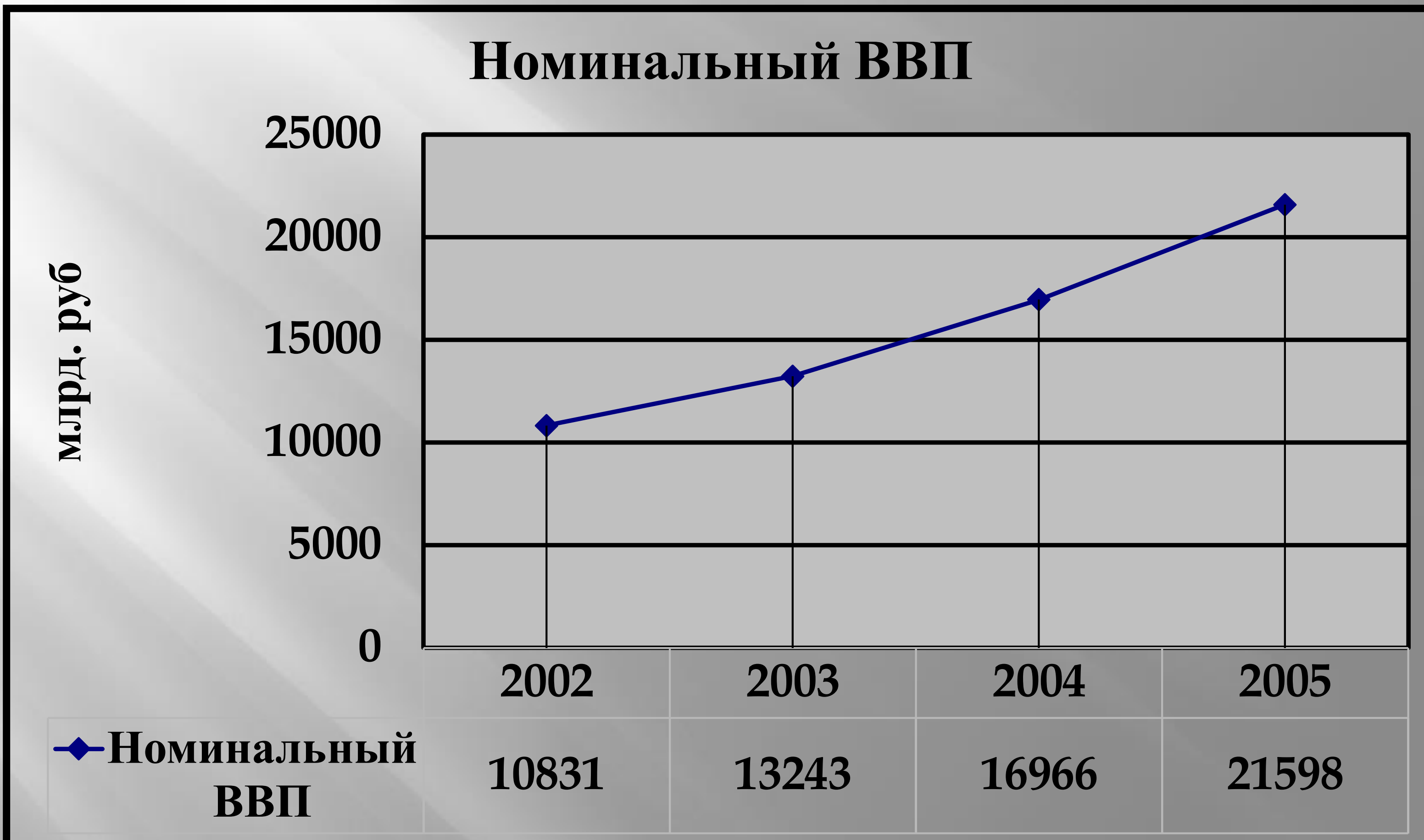
Общий абсолютный прирост (уменьшение)среднего уровня признака в целом по совокупности

$$\Delta\bar{x} = \bar{x}_1 - \overline{x_0} = \frac{\sum x_0 f_1}{\sum f_1} - \frac{\sum x_0 f_0}{\sum f_0}$$

или

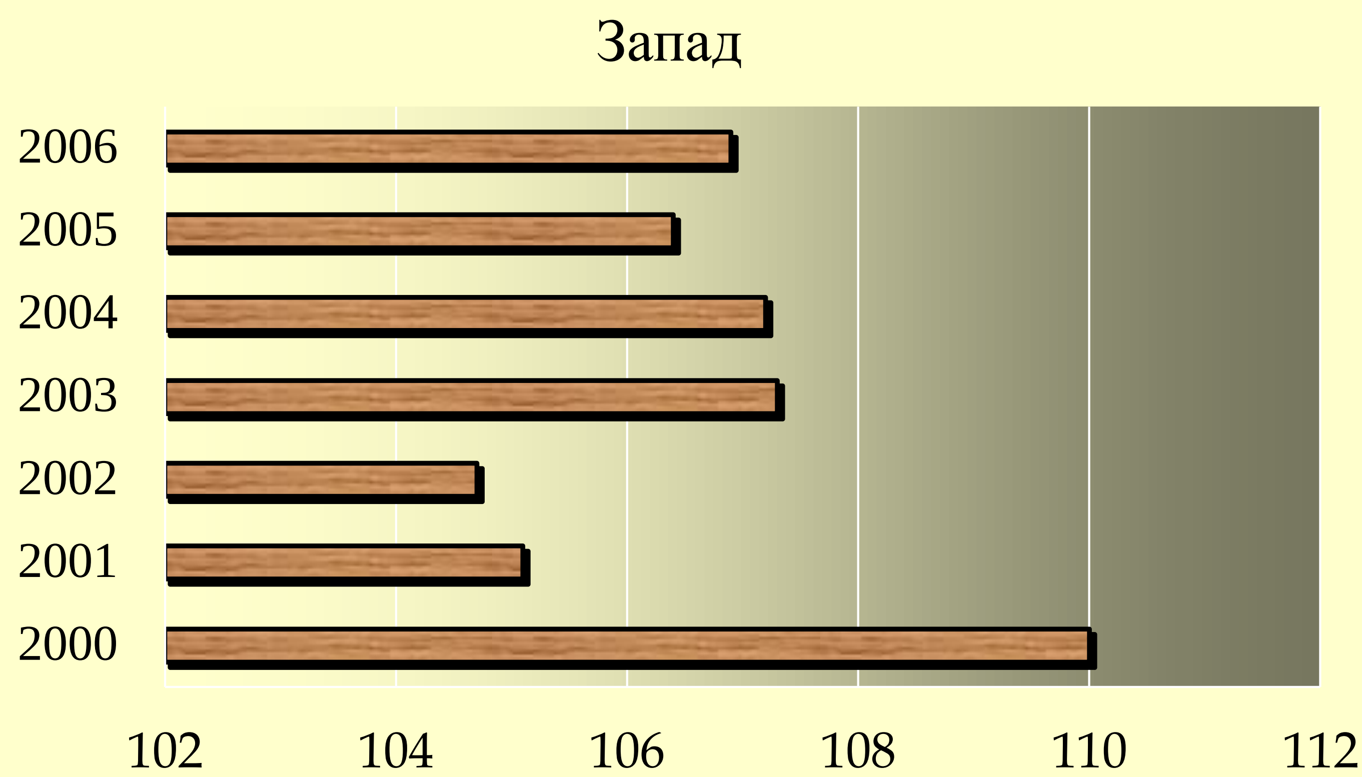
$$\Delta\bar{x} = \sum x_1 d_1 - \sum x_0 d_0$$

Динамика роста номинального ВВП
с 2002 по 2005 гг.

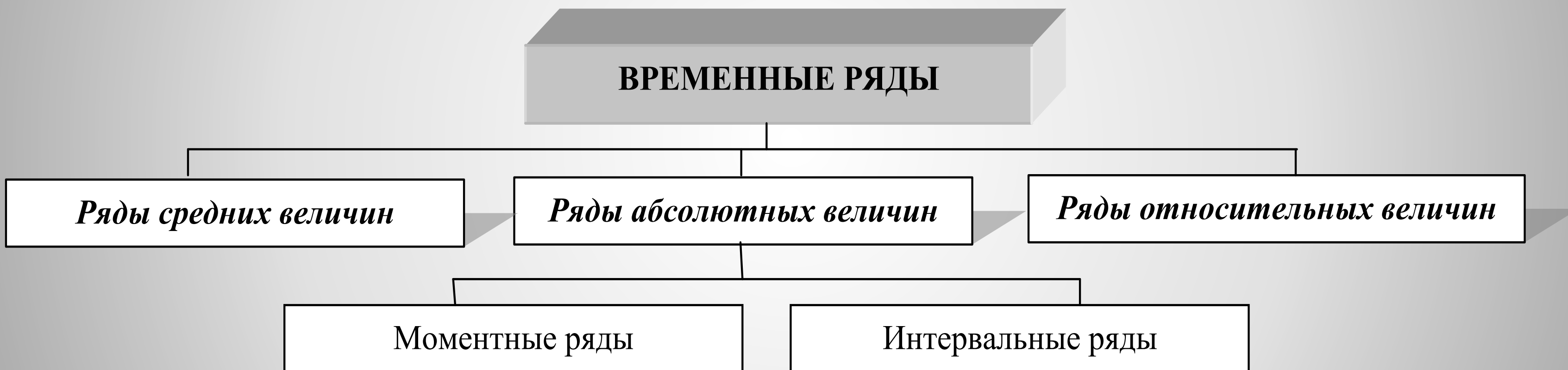


Динамика ВВП России

(в сопоставимых ценах, % к предшествующему году)



	2000	2001	2002	2003	2004	2005	2006
Запад	110	105,1	104,7	107,3	107,2	106,4	106,9



РЯДЫ ДИНАМИКИ

Классификация видов рядов динамики

Классификационные признаки — **Виды рядов динамики**

- Способ выражения уровней рядов
  - Ряды абсолютных величин
  - Ряды средних величин
  - Ряды относительных величин
- Способ представления хронологии
  - Моментный ряд
  - Интервальный ряд
- Расстояние между периодами /датами/
  - Равноотстоящие уровни во времени
  - Неравноотстоящие уровни во времени
- Наличие основной тенденции в ряду
  - Стационарные ряды
  - Нестационарные ряды
- Число показателей
  - Изолированный ряд
  - Многомерный ряд

Динамика означает изменение процессов во времени, поэтому ряд статистических показателей, характеризующий изменение общественных явлений во времени называется динамическим рядом.

Показатели, из которых состоит динамический ряд называются уровнями динамического ряда и обозначаются - y, а период времени, за который они представлены - t.

В теории статистики различают следующие виды динамических рядов:

Моментные ряды динамики. Моментным называется ряд, уровни которого характеризуют размеры социально-экономических явлений по состоянию на определенную дату или определенный момент времени.

Дата	1.01	1.04	1.07	1.10	1.01
Год	2004 г.	2004 г.	2004 г.	2004 г.	2005 г.
Число работников, чел.	192	190	195	198	200

Периодические (интервальные) ряды динамики. Периодический ряд - это такой ряд, уровни которого характеризуют размеры общественно-экономических явлений за определенный период (интервал) времени.

Год	2000	2001	2002	2003	2004
Объем розничного товарооборота, тыс. руб.	885,7	932,6	980,1	1028,7	1088,4

Показатели анализа динамики

- В интервальном ряду динамики расчет производится по методу средней арифметической простой:

$$\bar{y} = \frac{\sum_{t=1}^{n} y_t}{n}$$

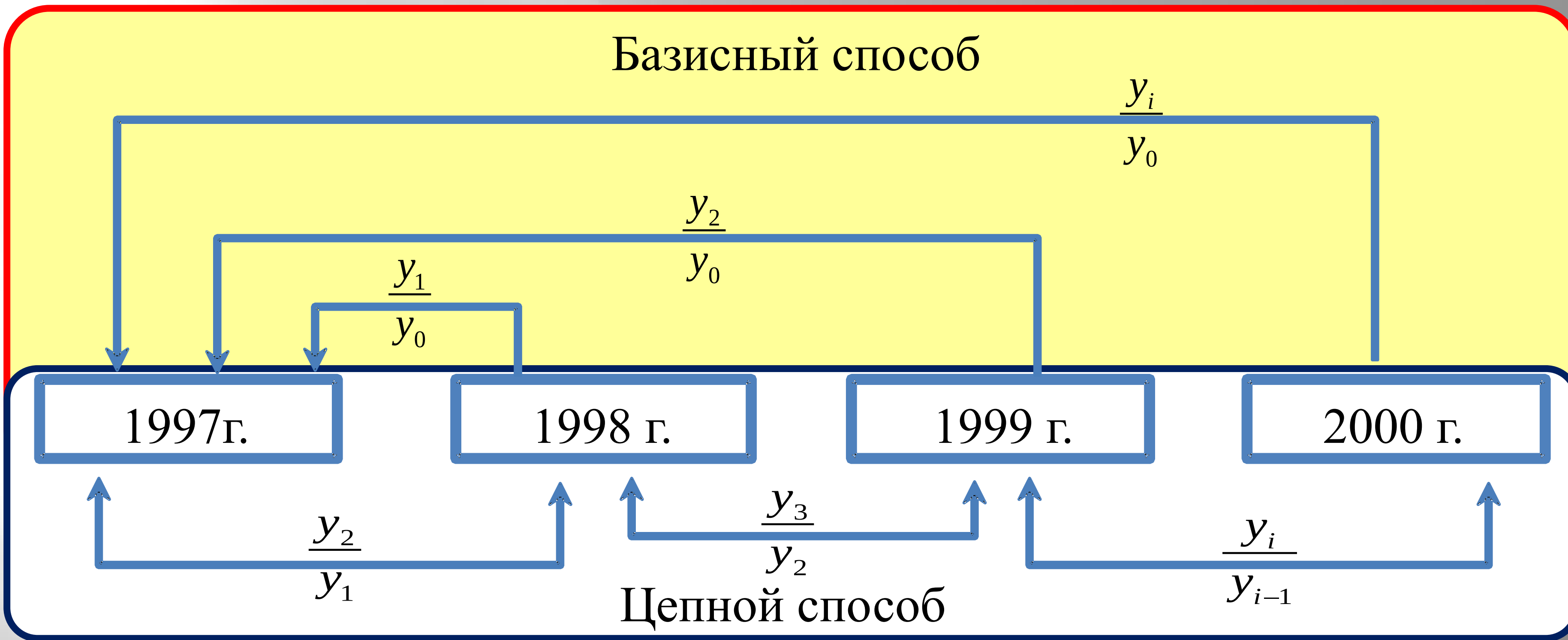
- Для моментного ряда расчет среднего уровня ряда производится по формуле:

$$\bar{y} = \frac{\frac{1}{2} y_1 + y_2 + \ldots + y_{n-1} + \frac{1}{2} y_n}{n-1}$$

Группировка показателей, характеризующих скорость и интенсивность изменения уровней ряда динамики

Показатели абсолютной скорости и интенсивности рядов динамики

- Абсолютные
  - Абсолютный прирост (цепной и базисный)
  - Абсолютное ускорение
  - Абсолютное значение 1% прироста
- Относительные
  - Темп роста (цепной и базисный)
  - Темп прироста (цепной и базисный)
  - Относительное ускорение
- Обобщающие
  - Средний абсолютный прирост
  - Средний темп роста
  - Средний темп прирост



$$K_p^{\text{ö}} = \frac{y_i}{y_{i-1}}$$

или

$$K_p^{\text{á}} = \frac{y_i}{y_0}$$

Абсолютный прирост

(абсолютное изменение)

базисный: **цепной:**

$$\Delta_y^{б} = y_i - y_0$$

$$\Delta_y^{ц} = y_{i} - y_{i-1}$$

где y_i - уровень сравниваемого периода;

y_{i-1} - уровень предшествующего периода;

y_0 - уровень базисного периода.

Коэффициент роста

цепной: **базисный:**

$$K_p^{ц} = \frac{y_i}{y_{i-1}}$$

$$K_p^{б} = \frac{y_i}{y_0}$$

где y_i - уровень сравниваемого периода;

y_{i-1} - уровень предшествующего периода;

y_0 - уровень базисного периода.

Темп роста

цепной: **базисный:**

$$T_p^{ц} = \frac{y_i}{y_{i-1}} \times 100\%$$

$$T_p^{б} = \frac{y_i}{y_0} \times 100\%$$

где y_i - уровень сравниваемого периода;

y_{i-1}- уровень предшествующего периода;

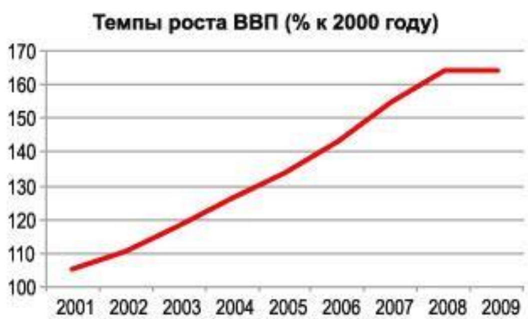
y_0 - уровень базисного периода.

$$\grave{O}_{p}^{\ddot{o}} = \frac{y_i}{y_{i-1}} \times 100\%$$

США

Великобритания

Россия



$$I = I_0 \frac{\sum P_t Q_0}{\sum P_0 Q_0}$$

Темп прироста

базисный:

$$T_{ПР}^{б} = \frac{y_i - y_0}{y_0} \times 100\% = \frac{\Delta_y^{б}}{y_0} \times 100\%$$

цепной:

$$T_{ПР}^{ц} = \frac{y_i - y_{i-1}}{y_{i-1}} \times 100\% = \frac{\Delta_y^{ц}}{y_{i-1}} \times 100\%$$

$$T_{ПР} = T_p - 100\%$$

Темп наращивания
(пункт прироста)

$$T_{н} = \frac{y_i - y_{i-1}}{y_0} \times 100\% = \frac{\Delta_y^{ц}}{y_0} \times 100\%$$

Абсолютное значение одного процента прироста

$$A\% = \frac{\boxed{y_i - y_{i-1}}}{\boxed{\dfrac{y_i - y_{i-1}}{y_{i-1}}}} \times 100\%$$

Средний абсолютный прирост

цепной:

$$\overline{\Delta} = \frac{\sum \Delta_{y}^{ц}}{n-1}$$

базисный:

$$\overline{\Delta} = \frac{y_n - y_1}{n-1}$$

где n - количество уровней ряда;

y_n - самое последнее значение уровня ряда;

y_1 - самое первое значение;

Средний темп роста

базисный:

$$\overline{T}_p = \sqrt[n-1]{\frac{y_n}{y_1}} \times 100\%$$

цепной:

$$\overline{T}_p = \sqrt[n-1]{\Pi T_p^{ц}} \times 100\%$$

где n - количество уровней ряда;

y_n - самое последнее значение уровня ряда;

y_1 - самое первое значение;

Средний уровень интервального ряда

$$\bar{y} = \frac{\sum y_i}{n}$$

Средний уровень интервального ряда (с неравными интервалами)

$$\overline{y} = \frac{\sum y_i t_i}{\sum t_i}$$

где t_i - количество дней между смежными датами;

Средний уровень моментного ряда

$$\bar{y}_{хрон} = \frac{\frac{1}{2}y_1 + y_2 + \ldots + y_{n-1} + \frac{1}{2}y_n}{n-1}$$

Средний уровень моментного ряда динамики с неравно отстающими интервалами

$$\bar{y}_{хрон} = \frac{(y_1 + y_2)t_1 + (y_2 + y_3)t_2 + \ldots + (y_{n-1} + y_n)t_{n-1}}{\sum t_1}$$

Метод укрупнения интервалов основан на укрупнении периодов времени, к которым относятся уровни ряда динамики (одновременно уменьшается количество интервалов). Средняя, исчисленная по укрупненным интервалам, позволяет выявить направление и характер (ускорение или замедление роста) основной тенденции развития, в то время как слишком малые интервалы между наблюдениями приводят к появлению ненужных деталей в динамике процесса, засоряющих общую тенденцию.

Месяц	Объем выпуска, млн.руб.	Месяц	Объем выпуска, млн.руб.
Январь	5,1	Июль	5,6
Февраль	5,4	Август	5,9
Март	5,2	Сентябрь	6,1
Апрель	5,3	Октябрь	6,0
Май	5,6	Ноябрь	5,9
Июнь	5,8	Декабрь	6,2

Квартал	Объем производства, млн.руб.	
	в квартал	в среднем в месяц
1	15,7	5,23
2	16,7	5,57
3	17,6	5,87
4	18,1	6,03

После укрупнения интервалов основная тенденция роста производства стала очевидной: 5,23<5,57<5,87<6,03 млн.руб.

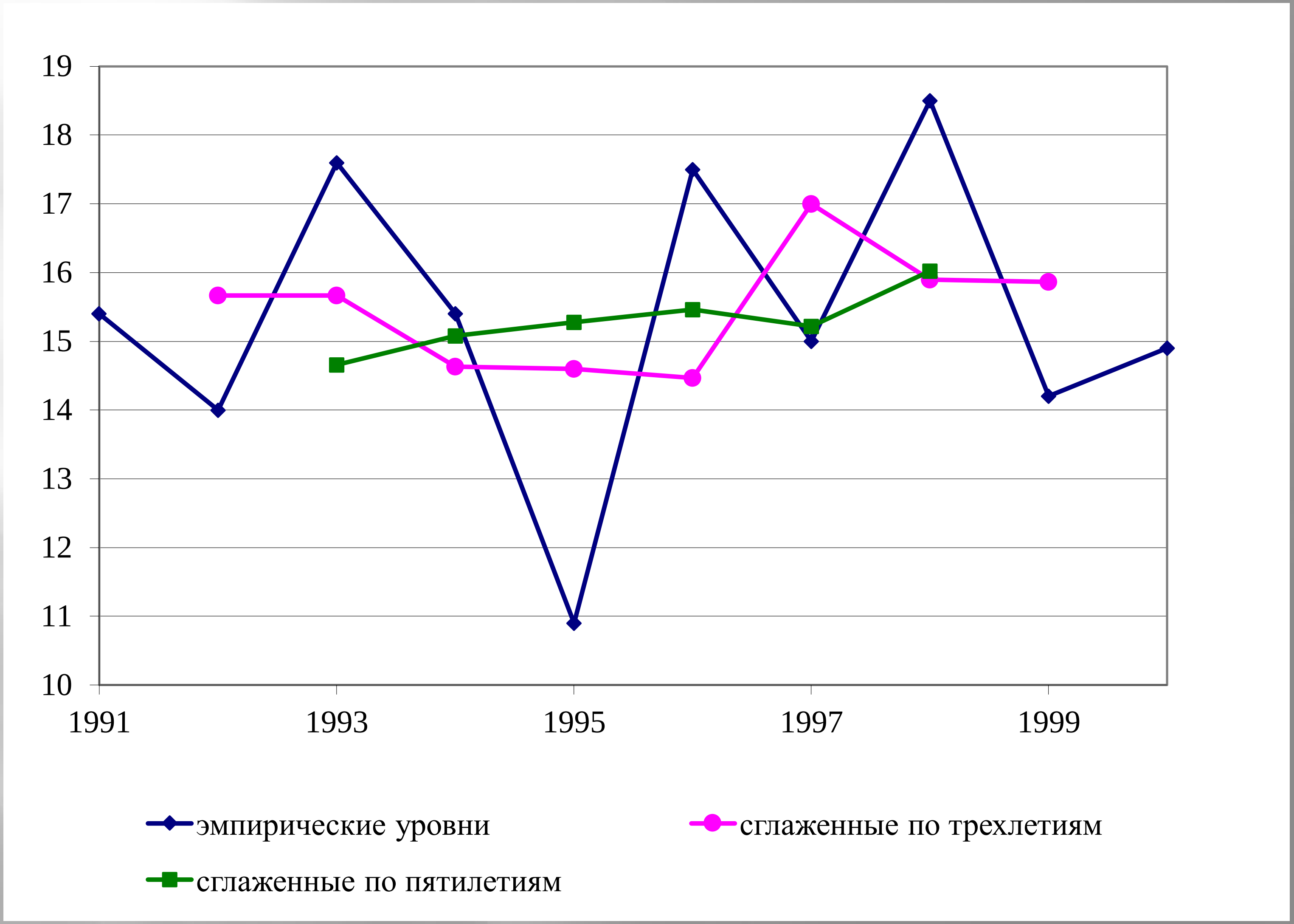
Метод скользящей средней заключается в том, что исчисляется средней уровень из определенного числа (обычно нечетного) первых по счету уровней ряда, затем – из такого же числа уровней, но начиная со второго по счету, далее – начиная с третьего и т.д.

$$\bar{y}_1 = \frac{y_1 + y_2 + y_3}{3} = \frac{15,4 + 14,0 + 17,6}{3} = \frac{47}{3} = 15,7$$

$$\bar{y}_2 = \frac{y_2 + y_3 + y_4}{3} = \frac{14,0 + 17,6 + 15,4}{3} = \frac{47}{3} = 15,7$$

Год	Урожайность, ц/га	Скользящая средняя	
		трехлетняя	пятилетняя
1991	15,4	–	–
1992	14,0	15,4+14,0+ +17,6)/3 = 15,7	–
1993	17,6	14,0+17,6+ +15,4)/3 = 15,7	14,7
1994	15,4	14,6	15,1
1995	10,9	14,6	15,3
1996	17,5	14,5	15,5
1997	15,0	17,0	15,2
1998	18,5	15,9	16,0
1999	14,2	15,9	–
2000	14,9	–	–
Итого	**153,4**		

Эмпирические и сглаженные уровни ряда динамики



Аналитическое выравнивание ряда динамики используется для того, чтобы дать количественную модель, выражающую основную тенденцию изменения уровней ряда динамики во времени. Общая тенденция развития рассчитывается как функция времени: $\hat{y}_t = f(t)$,

$$\begin{cases} na_0 + a_1 \sum t = \sum y; \\ a_0 \sum t + a_1 \sum t^2 = \sum t \cdot y, \end{cases}$$

y – фактические (эмпирические) уровни ряда;
t – время (порядковый номер периода или момента времени).

$\Sigma\, t = 0$, так что система нормальных уравнений принимает вид:

$$\begin{cases} na_0 = \sum y; \\ a_1 \sum t^2 = \sum t \cdot y. \end{cases}$$

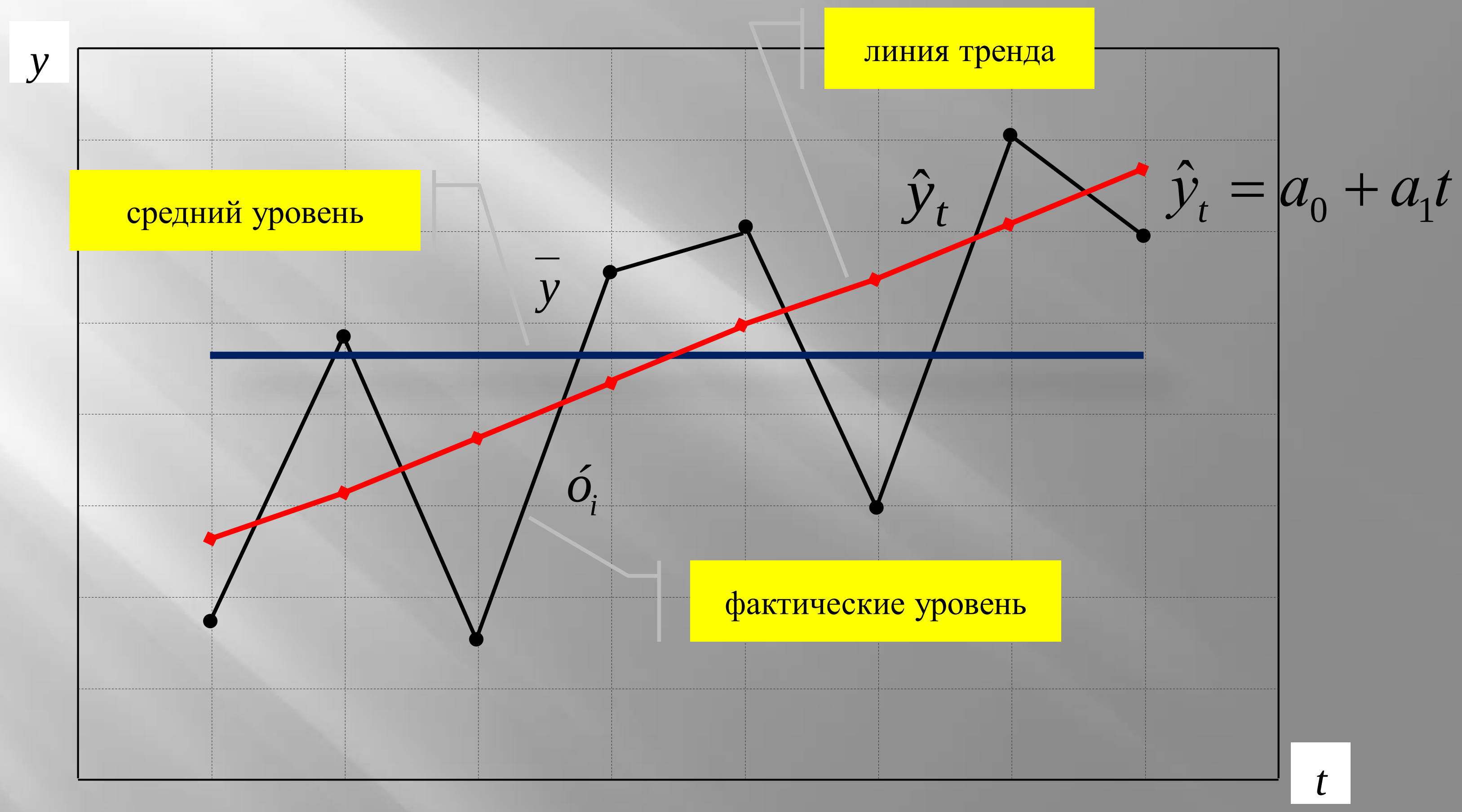
Отсюда можно выразить коэффициенты регрессии:

$$a_0 = \frac{\sum y}{n} \qquad a_1 = \frac{\sum t \cdot y}{\sum t^2}$$

Если расчеты выполнены правильно, то $\Sigma\, y = \Sigma\, \hat{y}t$.

$$\hat{y}_t = a_0 + a_1 t$$

Колебания уровней ряда носят различный характер. Наряду с трендом выделяют *циклические* (долгопериодические), *сезонные* (обнаруживаемые в рядах, где данные приведены за кварталы или месяцы) и *случайные* колебания.



Индекс сезонности

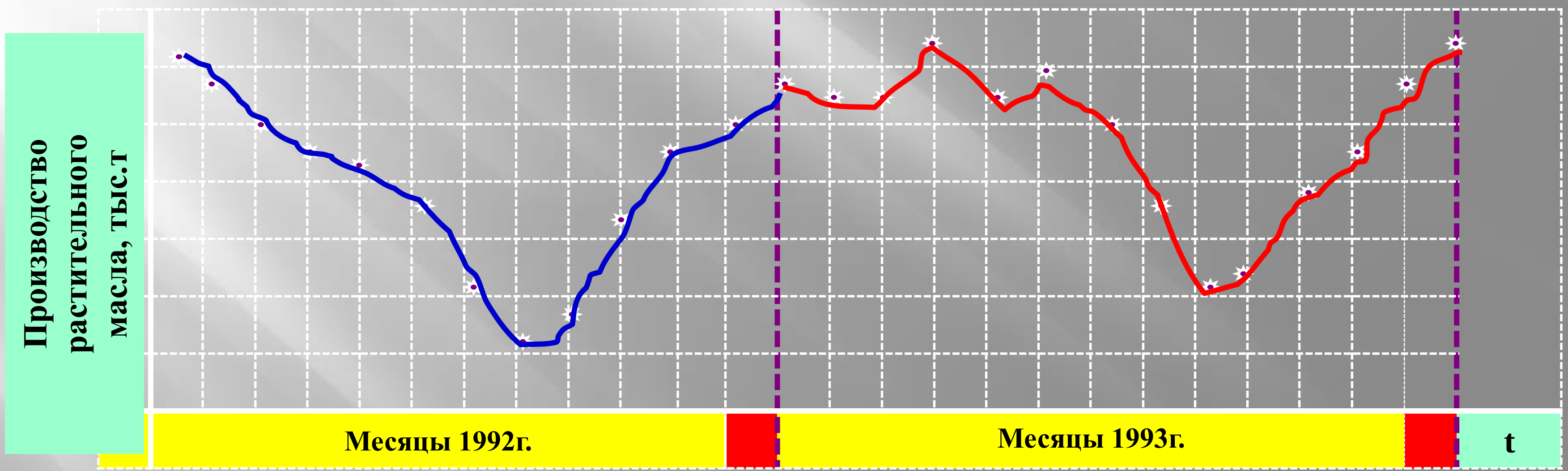
$$I_s = \frac{\overline{y_i}}{\overline{y}} \times 100\%$$

где $\overline{y}_i$ - средняя для каждого месяца минимум за три года;

$\overline{y}$ - среднемесячный уровень для всего ряда.

Динамика производства растительного масла в России за 1992 – 1993гг. по месяцам

Год	Месяц											
1992	109,5	102,7	86,6	82,3	76,6	70,0	57,6	24,5	36,3	70,7	95,2	104,5
1993	97,6	95,5	114,2	101,3	105,6	94,6	75,2	38,6	38,9	78,7	96,5	111,0



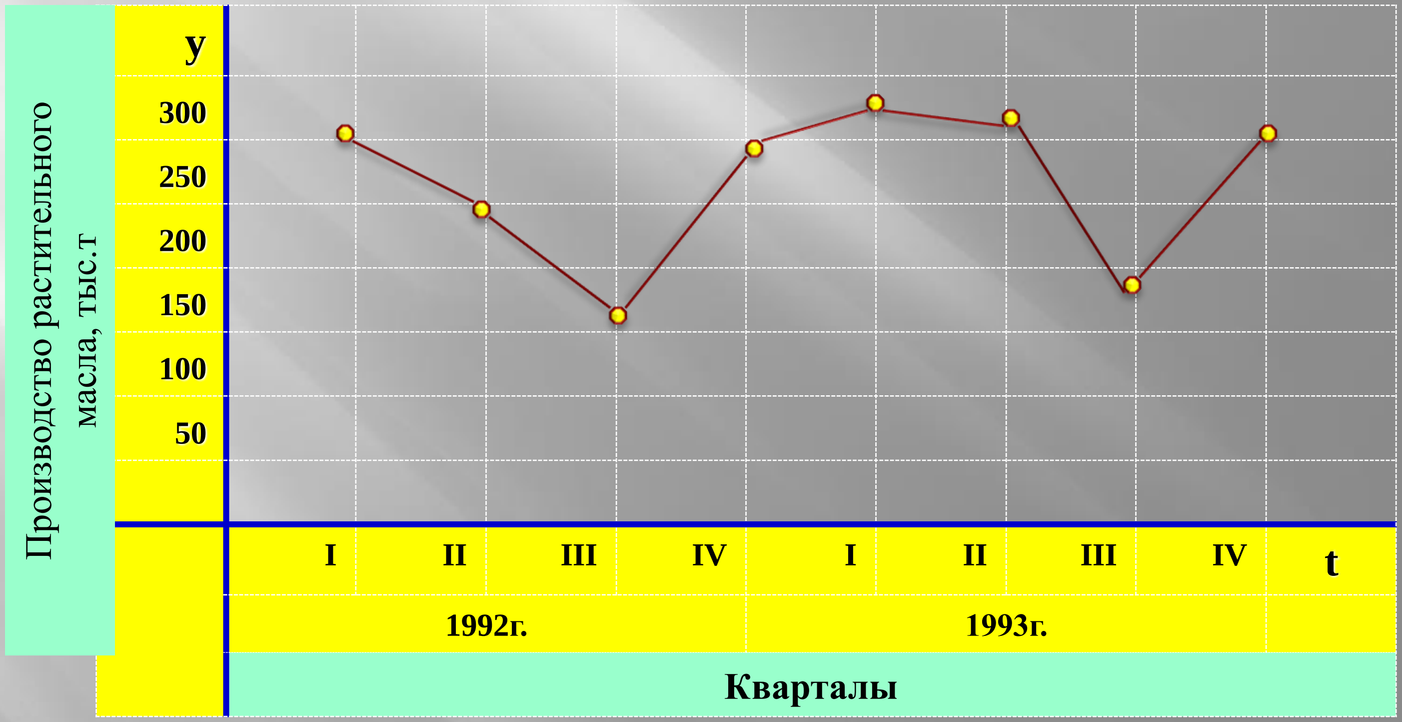
$$I_s = \frac{\overline{y}_i}{\overline{y}} \times 100\%$$

– средний уровень для каждого месяца;

– среднемесячный уровень для всего ряда.

Динамика производства растительного масла в России за 1992 – 1993гг. по кварталам

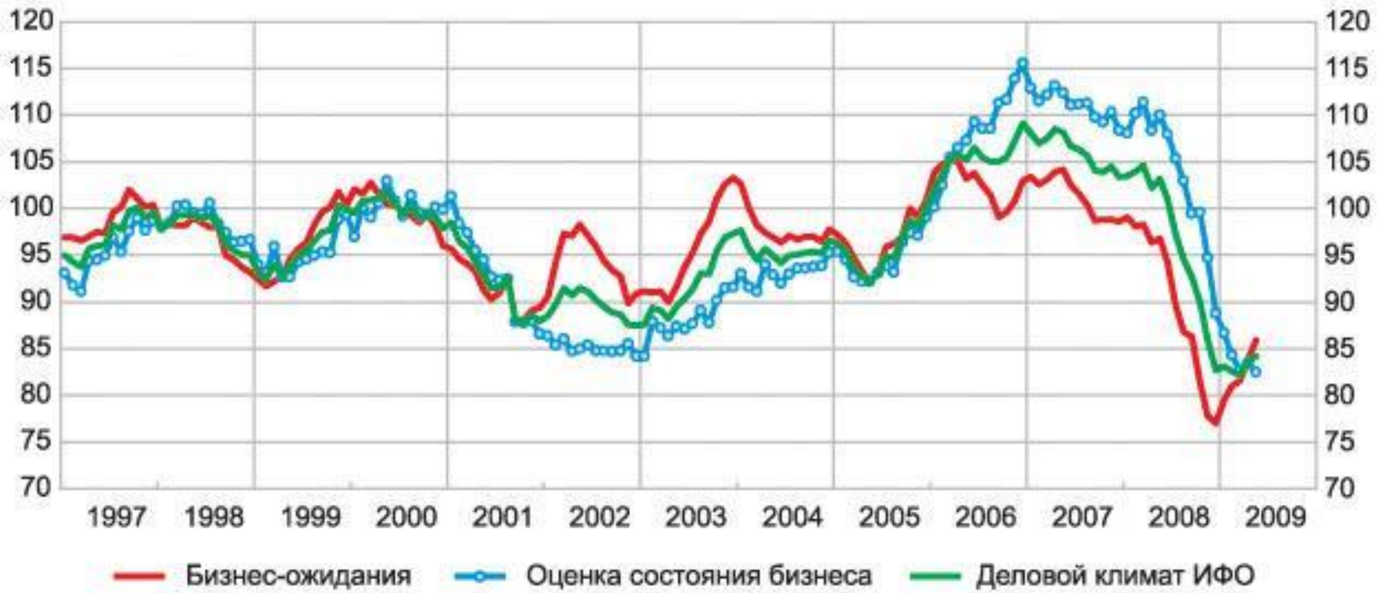
Год	1992				1993			
Квартал	I	II	III	IV	I	II	III	IV
Произведено тыс.т	298.8	228.9	118.4	270.4	307.3	301.5	152.7	286.2



Сальдо платежного баланса с учетом сезонных колебаний (EUR)

	1	2	3	4	5	6	7	8	9	10	11	12	1	2	3	4	5	6	7	8	9	10	11	12	1
Актуальное	-16	-7,3	-12,7	-8,1	-6,5	-5,9	-1,2	-5,3	6,6	-1,3	-5,4	-4,6	0,1	1,9	-8,1	-3,9	1,7	-5,1	-5,8	-4,6	-3,8	-7,5	-13,1	-9,8	
Прогнозы	-8	-14,3	-6,7	-10,6	-7,5	-4,7	-3,6	-1,7	-4,3	1,9	0,6	-2,3	-3,1	-0,6	2,9	-5,3	-2,7	1,1	-3	-3,7	-3,7	-1,9	-2,2	-6,2	-10,2
Предыдущие	-6	-13,9	-7,6	-12,3	-7,8	-7	-6,1	-0,1	-4,3	3,7	0,6	-5	-4,6	-0,5	1,9	-8,1	-4,5	1,5	-5,6	-7,4	-3,8	-4,1	-6,9	-9,7	-9,8

Месяцы



Применение метода наименьших квадратов для определения параметров линейного тренда $y^{\varepsilon} = at + b$ дает систему двух линейных уравнений:

$$bn + a\sum t = \sum y$$

$$b\sum t + a\sum t^2 = \sum yt$$

решение которой:

$$b = \frac{\sum t^2 \sum y - \sum t \sum ty}{n\sum t^2 - \left(\sum t\right)^2}, \quad a = \frac{n\sum ty - \sum t \sum y}{n\sum t^2 - \left(\sum t\right)^2}$$

t выбирается таким образом, чтобы $\sum t = 0$. В рядах с нечетным числом членов это выполняется при условии, что для центрального члена ряда $t = 0$ и вправо $t \to +1, +2, +3\ldots$, а влево: —1, —2, —3...

В этом случае:

$$b = \bar{y} = \frac{\sum y}{n}, \quad a = \frac{\sum ty}{\sum t^2}$$

Тренд характеризует основную закономерность движения во времени, свободную в основном (но не полностью) от случайных воздействий.

$$y_t = f(t) + \varepsilon_t$$

- ***Линейный тренд*** хорошо отражает тенденцию изменений при действии множества разнообразных факторов, изменяющихся различным образом по разным закономерностям.

$$y^{\varepsilon} = at + b$$

- ***Параболическая форма тренда*** выражает ускоренное или замедленное изменение уровней ряда с постоянным ускорением.

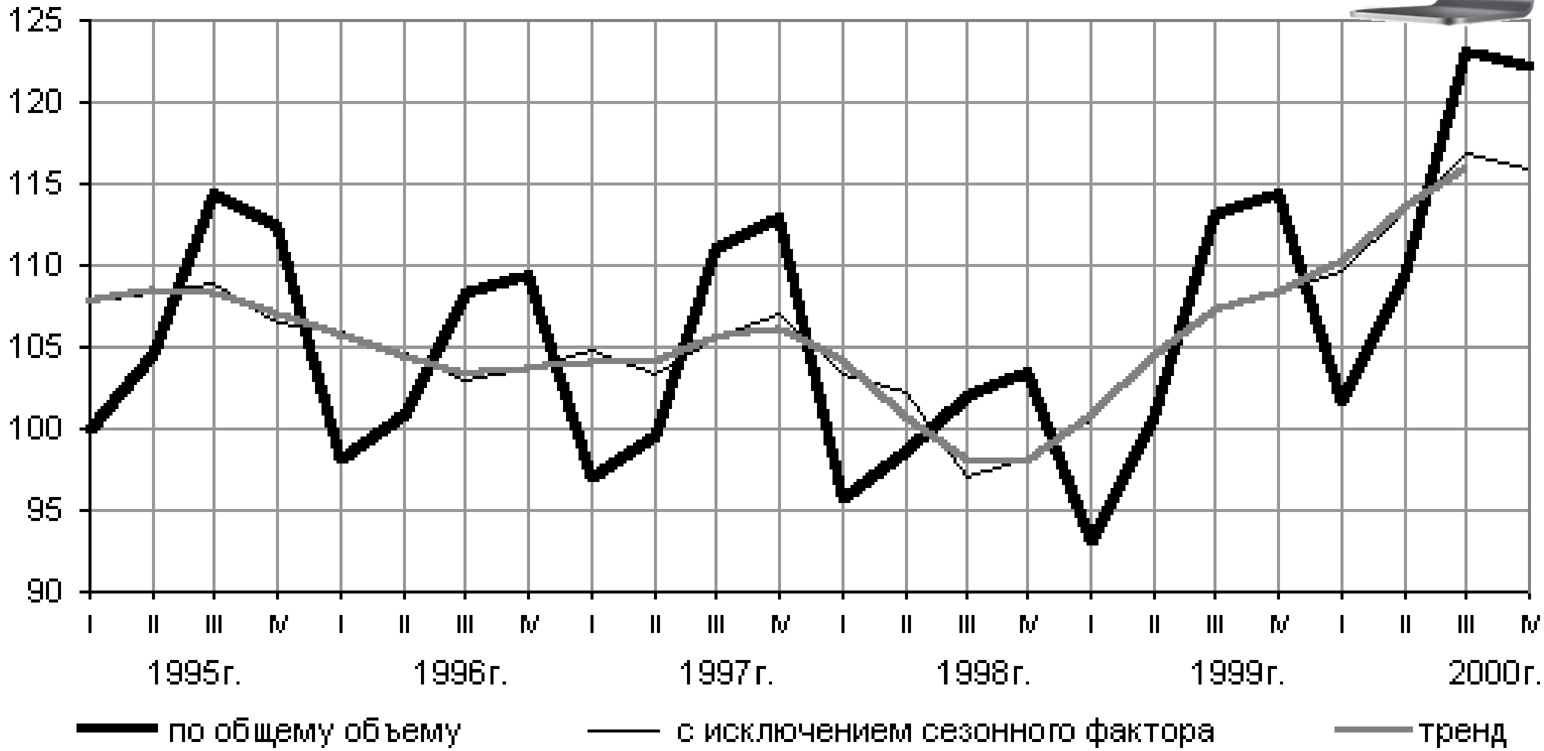
$$y^t = a + bt + ct^2$$

- ***Экспоненциальная форма тренда***:

$$y^t = ak^t$$

- ***Логарифмический тренд*** пригоден для отображения тенденции замедляющегося роста уровней при отсутствии предельного возможного значения.

$$y^t = a + d\,log(t)$$


125
120
115
110
105
100
95
90
I II III IV I II III IV I II III IV I II III IV I II III IV I II III IV
1995г.
1996г.
1997г.
1998г.
1999г.
2000г.
по общему объему
с исключением сезонного фактора
тренд

Гармонический анализ

$$\sum_{1}^{n}\left(y_i-\overline{y}_t\right)^2=\min$$

$$\overline{y_t}=a_0+\sum_{i=0}^{m}\left(a_k\cos kt+b_k\sin kt\right)$$

где k – гармоника ряда Фурье, которая может быть взята с разной степенью точности (чаще всего от 1 до 4)

Среднее квадратическое отклонение индексов сезонности

$$\sigma_{сез} = \sqrt{\frac{\sum \left(I_{сез} - 100 \right)^2}{12}}$$

Экстраполяция

$$\bar{y}_{i+1} = f(y_i, T, a_i)$$

где $\bar{y}_{i+1}$ - прогнозируемый уровень;

y_i - текущий уровень прогнозного ряда;

T - срок экстраполяции;

a_i - параметр уравнения тренда.

Метод среднего абсолютного прироста

$$\overline{y}_{i+1} = y_i + \overline{\Delta t}$$

где t – срок прогноза;

i – номер последнего уровня

Метод среднего темпа роста

$$\overline{y}_{i+1} = y_i \overline{k}$$

где y_i - последний уровень ряда динамики;

k - средний коэффициент роста.

ИЗМЕРЕНИЯ СВЯЗИ

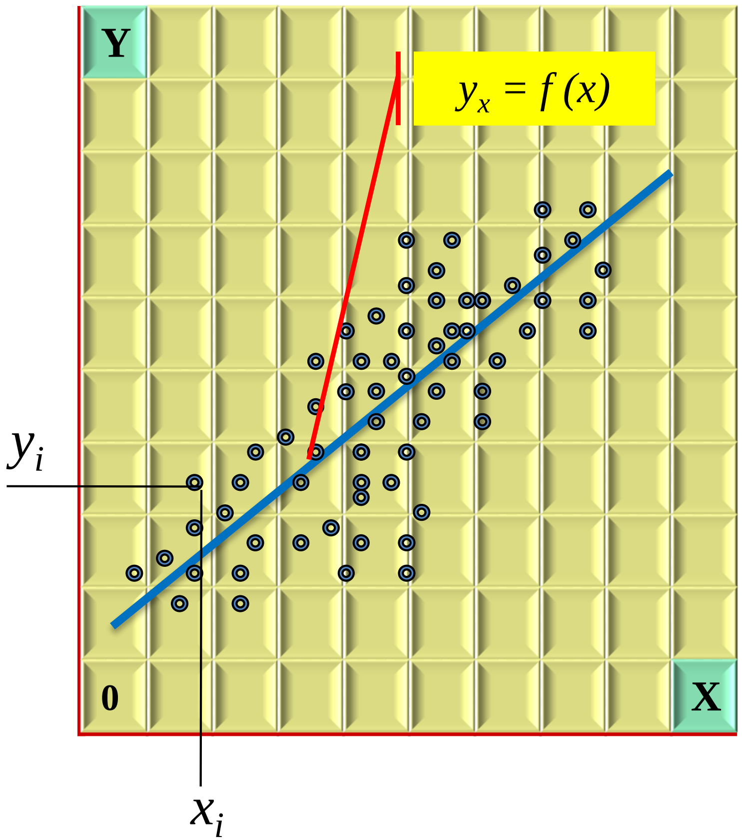

Методы изучения связей
Описательные (механические) методы
Метод аналитической группировки
Балансовый метод
Графический метод
Метод приведения параллельных рядов
Аналитические методы
Непараметрические (ранговые) методы
1)Коэффициент совпадения
2)Коэффициент ассоциации
3)Коэффициент контингенции
4)Коэффициент Спирмена
5)Коэффициент Кендалла
Параметрические методы
Корреляционный метод

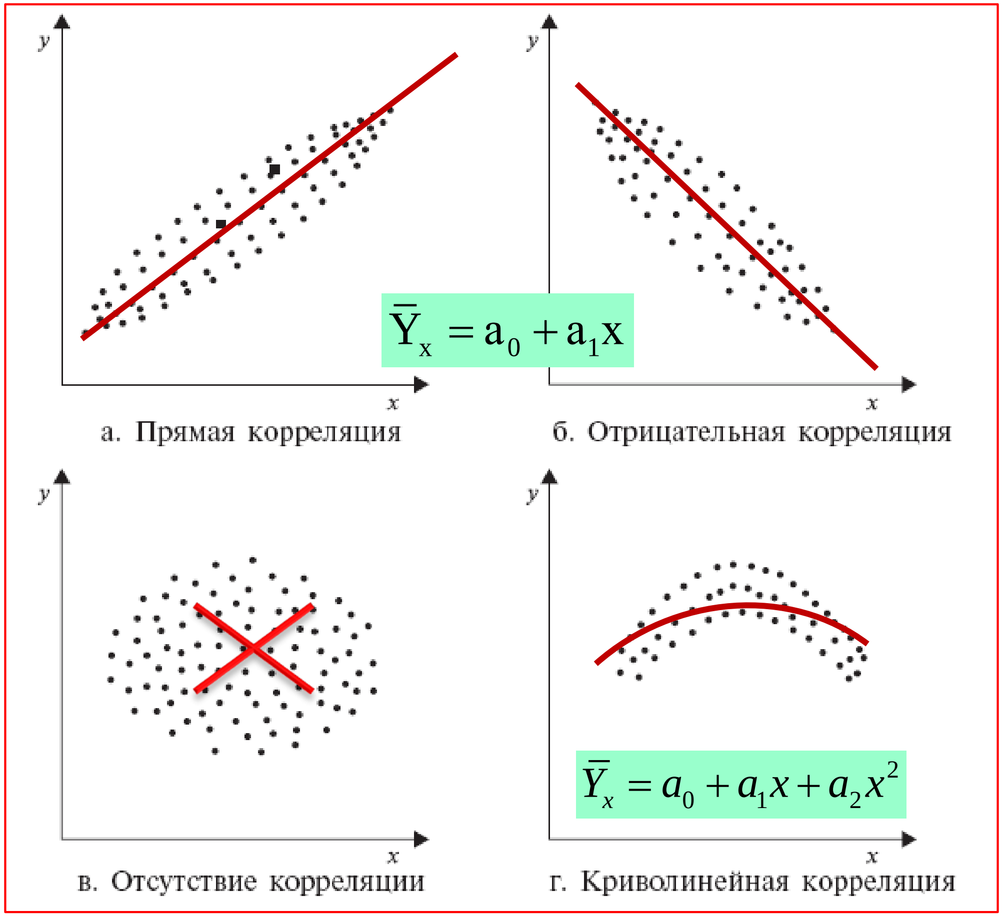
Основные показатели тесноты связи

Показатели тесноты связи

- Количественные шкалы
  - Линейный коэффициент корреляции
  - Корреляционное отношение
- Порядковые шкалы
  - Коэффициент Спирина
  - Коэффициент Кенделла
- Номинальные шкалы
  - Коэффициент ассоциации и контингенци
  - Коэффициенты Пирсона и Чупрова

График корреляционного поля





Коэффициент корреляции

$$r_{xy} = b \cdot \frac{\sigma_x}{\sigma_y} = \frac{\operatorname{cov}(x,\ y)}{\sigma_x \cdot \sigma_y} \qquad (2.6)$$

$$-1 \le r_{xy} \le 1$$

Коэффициент детерминации

$$r_{xy}^2 = 1 - \frac{\sigma_{\text{ост}}^2}{\sigma_y^2} \qquad (2.7)$$

$$\sigma_{\text{ост}}^2 = \frac{1}{n}\sum\left(y - y_x\right)^2 \qquad \sigma_y^2 = \frac{1}{n}\sum\left(y - \bar{y}\right)^2 = \overline{y^2} - \bar{y}^2$$

Степень взаимного влияния факторов в зависимости от коэффициента корреляции

Величина коэффициента корреляции	Сила связи
0,1 - 0,3	Слабая
0,3 - 0,5	Умеренная
0,5 - 0,7	Заметная
0,7 - 0,9	Тесная
0,9 - 0,99	Весьма высокая

Уравнение функциональной связи

$$\hat{y}_i = f(x_i)$$

где y_1 - результативный признак $(i = 1, \ldots, n)$

$f(x_i)$ - известная функция связи результативного и факторного признаков;

x_i - факторный признак

Модель стохастической связи

$$\hat{y}_i = f\left(x_i\right) + \varepsilon_i$$

где $\hat{y}_i$ - расчетное значение результативного признака;

$f\left(x_i\right)$ - часть результативного признака, сформировавшаяся под воздействием учтенных факторных признаков (одного или множества), находящихся в стохастической связи с признаком;

ε_i - часть результативного признака, возникшая вследствие действия неконтролируемых или неучтенных факторов, а также измерения признаков, неизбежно сопровождающегося некоторыми случайными ошибками

Схема проведения корреляционно-регрессионного анализа

Матрица исходных данных

Построение матрицы парных коэффициентов корреляции

Проверка связей между признаками на наличие мультиколлинеарности

Отбор факторных признаков

Оценка статистической значимости уравнения регрессии и коэф.регрессии

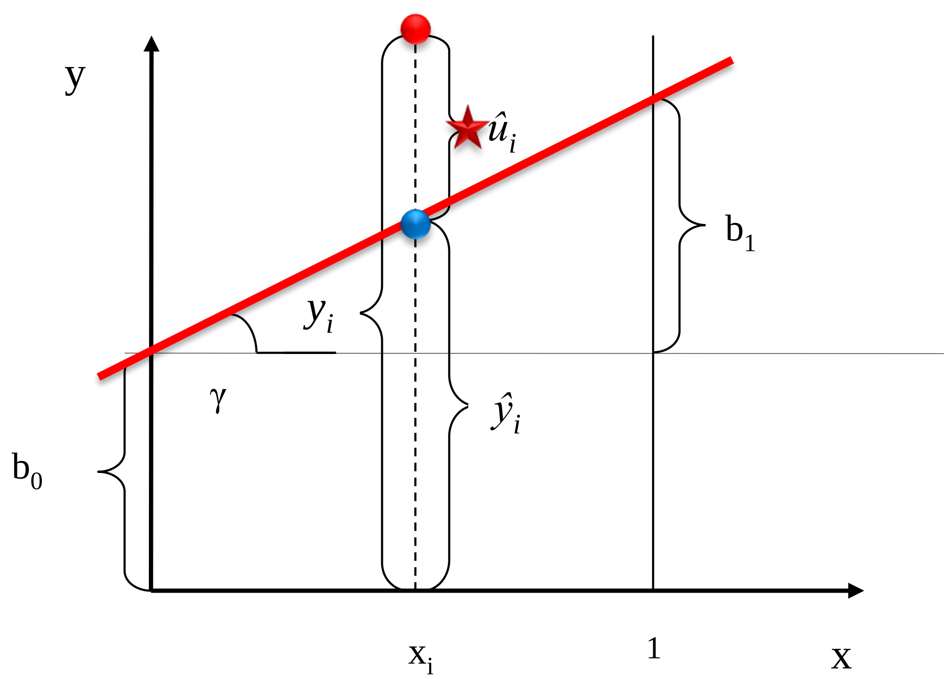
Расчет и анализ доп.показателей для расширения экономической интерпретации уравнения регрессии

Экономическая интерпретация, формулировка выводов и предложений

Простая линейная регрессия задается следующей формулой:

$$\hat{y} = b_0 + b_1 x$$

b_0 и b_1 – неизвестные параметры регрессии; имеются n наблюдений над переменной x: $x_1, x_2, \ldots, x_n$; b_0 выполняет в уравнении регрессии функцию выравнивания; b_1 характеризует наклон прямой к оси OX.



Ошибки обнаруживаются через отклонения $\hat{u}_i$ эмпирических данных от значений регрессии $\hat{y}_i$. Они являются значениями возмущающей переменной u:

$$\hat{u}_i = y_i - \hat{y}_i$$

i = 1, …, n.

Линейный коэффициент корреляции

$$r = \frac{\overline{xy} - \overline{x}\,\overline{y}}{\sigma_x \sigma_y} = \frac{\sum\left(x - \overline{x}\right)\left(y - \overline{y}\right)}{n \sigma_x \sigma_y}$$

n – число наблюдений

$$S(a,\, b) = \sum (y - a - b \cdot x)^2$$

$$\begin{cases} \dfrac{\partial S}{\partial a} = -2\sum (y - a - b \cdot x) = 0; \\ \dfrac{\partial S}{\partial b} = -2\sum x(y - a - b \cdot x) = 0. \end{cases}$$

$$\begin{cases} a \cdot n \quad + b \cdot \sum x \ = \sum y; \\ a \cdot \sum x + b \cdot \sum x^2 = \sum x \cdot y. \end{cases}$$

$$a = \bar{y} - b \cdot \bar{x} \qquad b = \frac{\operatorname{cov}(x,\ y)}{\sigma_x^2}$$

где

$$\operatorname{cov}(x, y) = \overline{x \cdot y} - \bar{x}\ \bar{y}$$ – ковариация признаков x и y

$$\sigma_x^2 = \overline{x^2} - \bar{x}^2$$ – дисперсия признака x

$$\bar{x} = \frac{1}{n}\sum x \qquad \bar{y} = \frac{1}{n}\sum y$$

$$\overline{y \cdot x} = \frac{1}{n}\sum y \cdot x \qquad \overline{x^2} = \frac{1}{n}\sum x^2$$

Система нормальных уравнений

$$\begin{cases} \sum y = na_0 + a_1 \sum x \\ \sum xy = a_0 \sum x + a_1 \sum x^2 \end{cases}$$

где:

$$a_0 = \frac{\sum y \sum x^2 - \sum xy \sum x}{n\sum x^2 - \left(\sum x\right)^2}$$

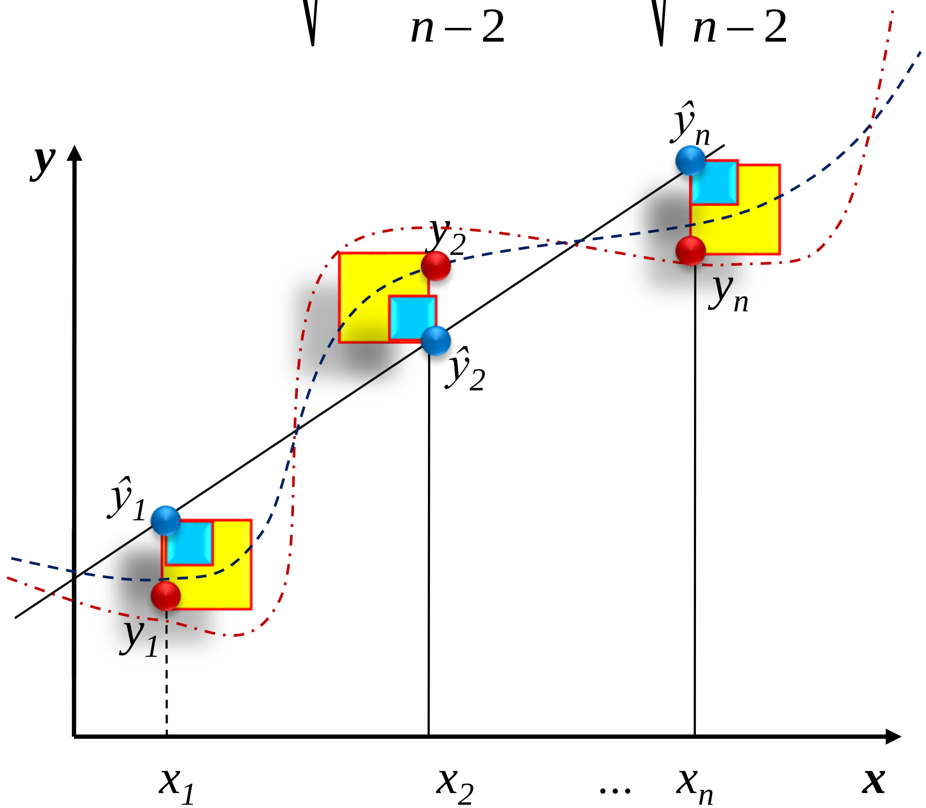
$$a_1 = \frac{n\sum xy - \sum x \sum y}{n\sum x^2 - \left(\sum x\right)^2}$$

Вычисляем выборочную дисперсию, характеризующую меру разброса опытных данных ***(x_i; y_i)*** вокруг значений регрессии, то есть дисперсию остатков

$$S_u^2 = \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n-2} = \frac{\sum_{i=1}^{n}\hat{u}_i^2}{n-2}$$

$$\sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n-2}} = \sqrt{\frac{\sum_{i=1}^{n}\hat{u}_i^2}{n-2}}$$

1 $$\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \to \min$$



Геометрическая интерпретация формулы **(1)** следующая: сумма площадей заштрихованных квадратов должна быть наименьшей

Схема построения уравнений парной регрессии

Матрица исходных данных

Графический анализ
парных зависимостей между У и Х

Уравнение регрессии
различных аналитических форм

линейное	парабола второго порядка	гипербола	. . .

Выбор формы связи между
результативным и факторным признаками

Интерпретация параметров
уравнения регрессии

Для построения многофакторной модели необходимо произвести отбор факторов по трем стадиям. Анализ факторов без особых ограничений. Сравнительная оценка и отсев части факторов путем анализа парных коэффициентов и индексов корреляции и оценки их значимости. Для этого рассчитываются парные коэффициенты корреляции, измеряющие тесноту связи каждого из факторов — признаков с результативным фактором и между собой.

	Y	x1	x2	…	xi	…	xm
Y	1	ry1	ry2	…	ryj	…	rym
x1	r1y	1	r12	…	r1j	…	r1m
x2	r2y	r21	1	…	r2j	…	r2m
…	…	…	…	1	…	…	…
xi	riy	ri1	ri2	…	1	…	rim
…	…	…	…	…	…	1	…
Xm	rmy	rm2	rm2	…	rmj	…	1

$$r_{yi} > r_{ij}, \quad r_{yj} > r_{ij}, \quad r_{ij} < 0{,}8$$

$$t_{расч.} > t_{f,a}$$

$$\left(\sigma_{\varepsilon}^{2} = \frac{\sum (y_i - y_{xi})^2}{n} = min \right)$$

$$\widetilde{Y}_i = a_0 x_0 + a_1 x_1 + a_2 x_2 + \ldots + a_m x_m$$

$$\hat{y} = b_0 + b_1 x_1 + b_2 x_2 + \ldots + b_m x_m.$$

$$\hat{y} = b_0 + b_1 x_1 + b_2 x_2.$$

$$S(b_0; b_1; b_2) = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = \sum_{i=1}^{n} \hat{u}_i^2 \to \min$$

$$S(b_0; b_1; b_2) = \sum_{i=1}^{n} (y_i - b_0 - b_1 x_{i1} - b_2 x_{i2})^2 \to \min$$

$$\begin{cases} n b_0 + b_1 \sum_{i=1}^{n} x_{i1} + b_2 \sum_{i=1}^{n} x_{i2} = \sum_{i=1}^{n} y_i \\ b_0 \sum_{i=1}^{n} x_{i1} + b_1 \sum_{i=1}^{n} x_{i1}^2 + b_2 \sum_{i=1}^{n} x_{i1} x_{i2} = \sum_{i=1}^{n} x_{i1} y_i \\ b_0 \sum_{i=1}^{n} x_{i2} + b_1 \sum_{i=1}^{n} x_{i1} x_{i2} + b_2 \sum_{i=1}^{n} x_{i2}^2 = \sum_{i=1}^{n} x_{i2} y_i \end{cases}$$

Парные коэффициенты корреляции

$$r_{yx_1} = \frac{\overline{x_1 y} - \overline{x}_1 \overline{y}}{\sigma_{x_1} \sigma_y}$$

$$r_{yx_2} = \frac{\overline{x_2 y} - \overline{x}_2 \overline{y}}{\sigma_{x_2} \sigma_y}$$

$$r_{yx_2} = \frac{\overline{x_2 x_2} - \overline{x}_1 \overline{x}_2}{\sigma_{x_{12}} \sigma_2}$$

Множественный коэффициент корреляции

$$R_{yx_1x_2} = \sqrt{\frac{r_{yx_1}^2 + r_{yx_2}^2 - 2r_{yx_1}r_{yx_2}r_{x_1x_2}}{1 - r_{x_1x_2}^2}}$$

где r_{yx_1}, $r_{x_1x_2}$, r_{yx_2} - парные линейные коэффициенты корреляции; подстрочные индексы указывают, между какими признаками они исчисляются.

Частный коэффициент эластичности

$$Э = a_i \frac{\bar{x}}{\bar{y}}$$

где a_i - коэффициент регрессии при факторе x;

$\bar{x}$, $\bar{y}$ - средние значения факторного и результативного признаков.

Бета-коэффициент

$$\beta = a_1 \frac{\sigma_{xi}}{\sigma_y}$$

где σ_{xi} - среднее квадратическое отклонение 1-го фактора;

σ_y - среднее квадратическое отклонение показателя.

Дельта-коэффициент

$$\Delta_i = \frac{\beta_i r_i}{R^2}$$

где R^2 - коэффициент множественной детерминации

Теоретическое корреляционное отношение

$$\eta = \sqrt{\frac{\sigma_{y_x}^2}{\sigma_y^2}}$$

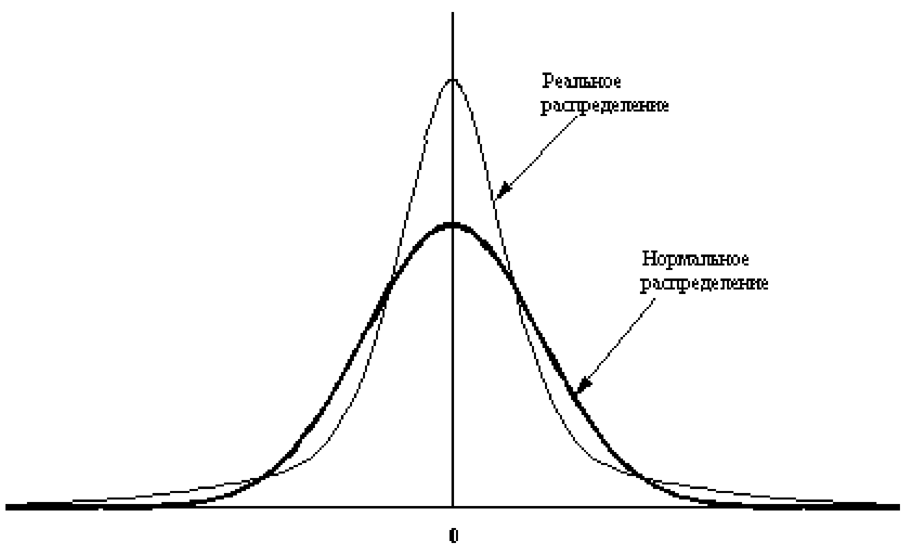
$$\eta = \sqrt{\frac{\sum \left(y_x - \bar{y}\right)^2}{\sum \left(y - \bar{y}\right)^2}}$$

где $\sigma_{y_x}^2$ - факторная дисперсия;

σ_y^2 - общая дисперсия.

Таблица интегральной функции эмпирического распределения

Интервалы группировок	Верхняя граница интервала		Интегральная функция нормального распределения	Плотность распределения	Частоты	
	Исходная	Нормированная			теоретические	фактические
i	x^b_i	U_i	$\Phi(U_i)$	φ_i	P_i	f_i



Фрагмент таблицы интегральной функции нормального распределения

$\Phi(t)$	t
0	0
0,38	0,5
0,68	1,0
0,86	1,5
0,95	1,96
0,9545	2,0
0,98	2,5
0,99	3,0

Четырёхклеточная корреляционная таблица

a	b	a + b
c	d	c + d
a + c	b + d	a + b + c + d

Коэффициент ассоциации

$$k_a = \frac{ab - bc}{ab + bc}$$

Коэффициент контингенции

$$k_k = \frac{ab - bc}{\sqrt{(a+b)(c+d)}(a+c)(b+d)}$$

Коэффициенты ассоциации и контингенции

Используются для измерения связи между двумя качественными признаками, состоящими только из двух групп.

	.....	.....	**Итого**
.....	*a*	*b*	a + b
.....	*d*	*c*	c + d
Итого	a + c	b + d	a + b+ c+ d

Оценка / Посещение	Неудовлетв.	Положит.	**Итого**
Посещали	86	14	100
Не посещали	22	28	50
Итого	108	42	150

Коэффициент ассоциации

$$\hat{E}_{\grave{a}} = \frac{ad - bc}{ad + bc} = \frac{86 \cdot 28 - 14 \cdot 22}{86 \cdot 28 + 14 \cdot 22} = 0,77$$

Коэффициент контингенции

$$K = \frac{ad - bc}{\sqrt{(a+b)(b+d)(d+c)(c+a)}} =$$

$$= \frac{86 \cdot 28 - 14 \cdot 22}{\sqrt{(86+14)(14+28)(28+22)(22+86)}} = 0,44$$

Коэффициент контингенции всегда меньше коэффициента ассоциации. Связь считается подтвержденной, если или .

Коэффициент корреляции рангов
(коэффициент Спирмена)

$$px/y = 1 - \frac{6\sum d_i^2}{n\left(n^2 - 1\right)}$$

где d_i^2 - квадраты разности рангов;

n - число наблюдений (число пар рангов).

Коэффициент Спирмана
(ранговый коэффициент)

Рассчитывается по следующей формуле:

$$\rho = 1 - \frac{6 \cdot \sum d_i^2}{n\left(n^2 - 1\right)}$$

№ п/п	***Себестоимость единицы прод.***	***Средняя з/п***	***Ранги***		***$d_i = R_z - R_f$***	***d_i^2***
			R_z	***R_f***		
1.	68,8	168,5	3	6	-3	9
2.	70,2	158,7	5	1	4	16
3.	71,4	171,7	7	8	-1	1
4.	78,5	183,9	10	10	0	0
5.	66,9	160,4	2	2	0	0
6.	69,7	165,2	4	5	-1	1
7.	72,3	175,0	8	9	-1	1
8.	77,5	170,4	9	7	2	4
9.	65,2	162,7	1	3	-2	4
10.	70,7	163,0	6	4	2	4
Итого						**40**

Коэффициент Спирмана может принимать значения от –1 до +1, причем чем ближе значение коэффициента к |1|, тем связь более тесная. Знак коэффициента говорит о направлении связи.

$$\rho = 1 - \frac{6 \cdot 40}{10 \cdot (10^2 - 1)} = 0{,}758$$

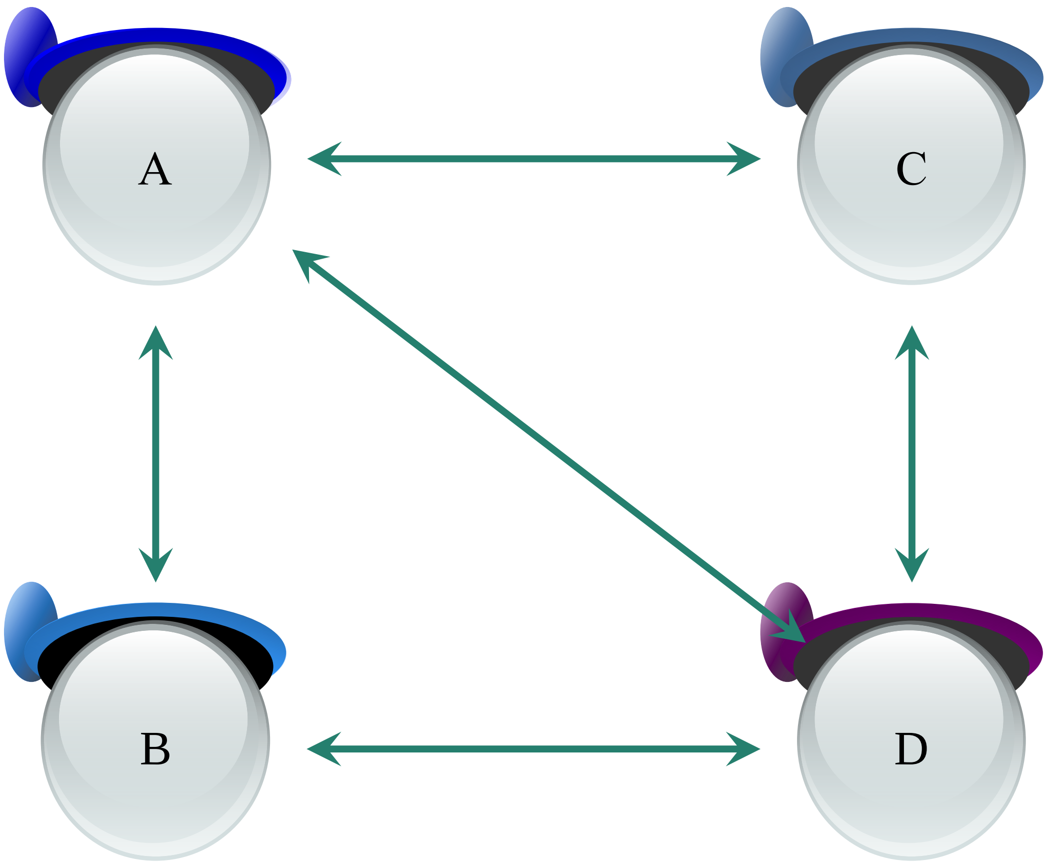
Ранговый коэффициент корреляции Кенделла

$$\tau = \frac{2S}{n(n-1)}$$

где n – число наблюдений;

S – сумма разностей между числом последовательностей и числом инверсий по второму признаку (число инверсии – естественная мера нарушения порядка объектов в одной последовательности относительно другой)

Корреляционный граф



Кластерный анализ

- Объекты считаются однородными, если

$\rho\ (x1,\ x2) < p_{предельного}$.

- Для объектов, характеризуемых числовыми признаками расстояние определяют:

$$\rho = \sqrt{\sum (xil - xjl)^2}$$

- Расстояние между объектами, описываемыми атрибутивными признаками:

$$\rho = \sum \left| xil - xjl \right|$$

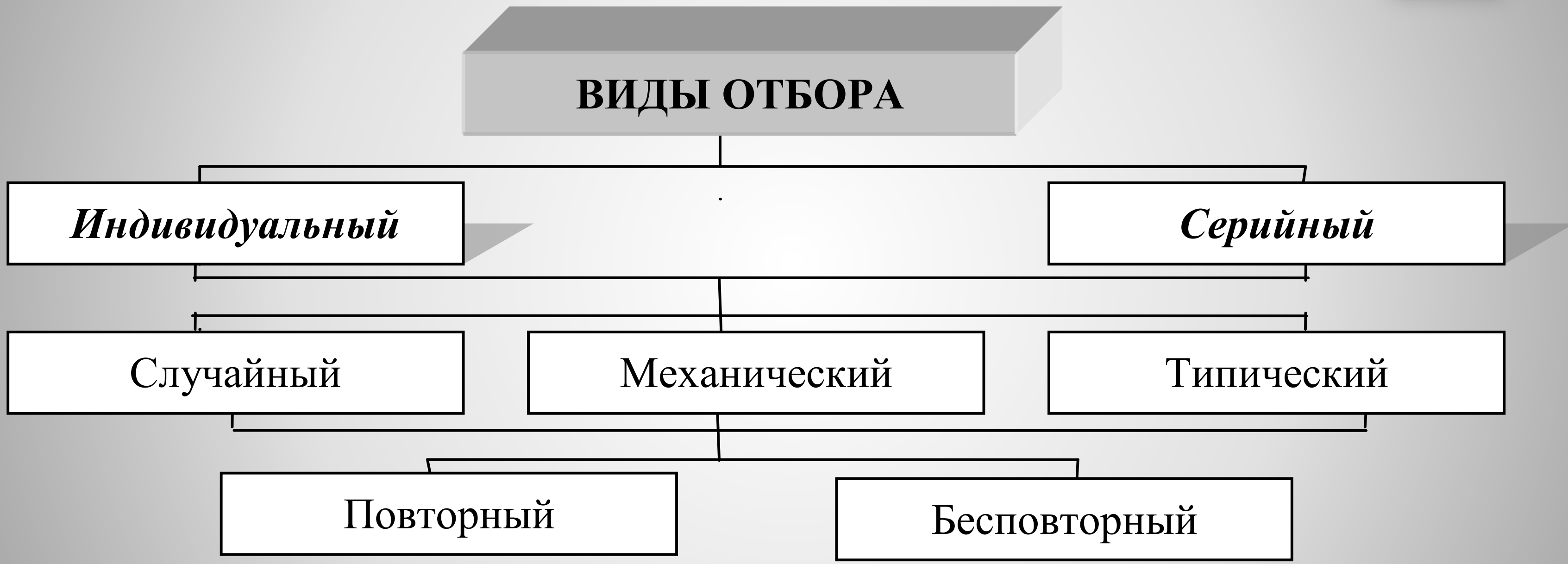
Факторный анализ

$$\bar{x}_j = \frac{1}{N}\sum_{i=1}^{N} x_{ji}; \quad \sigma_j = \sqrt{\frac{\sum_{i=1}^{N}\left(x_{ij} - \bar{x}_j\right)^2}{N}}$$

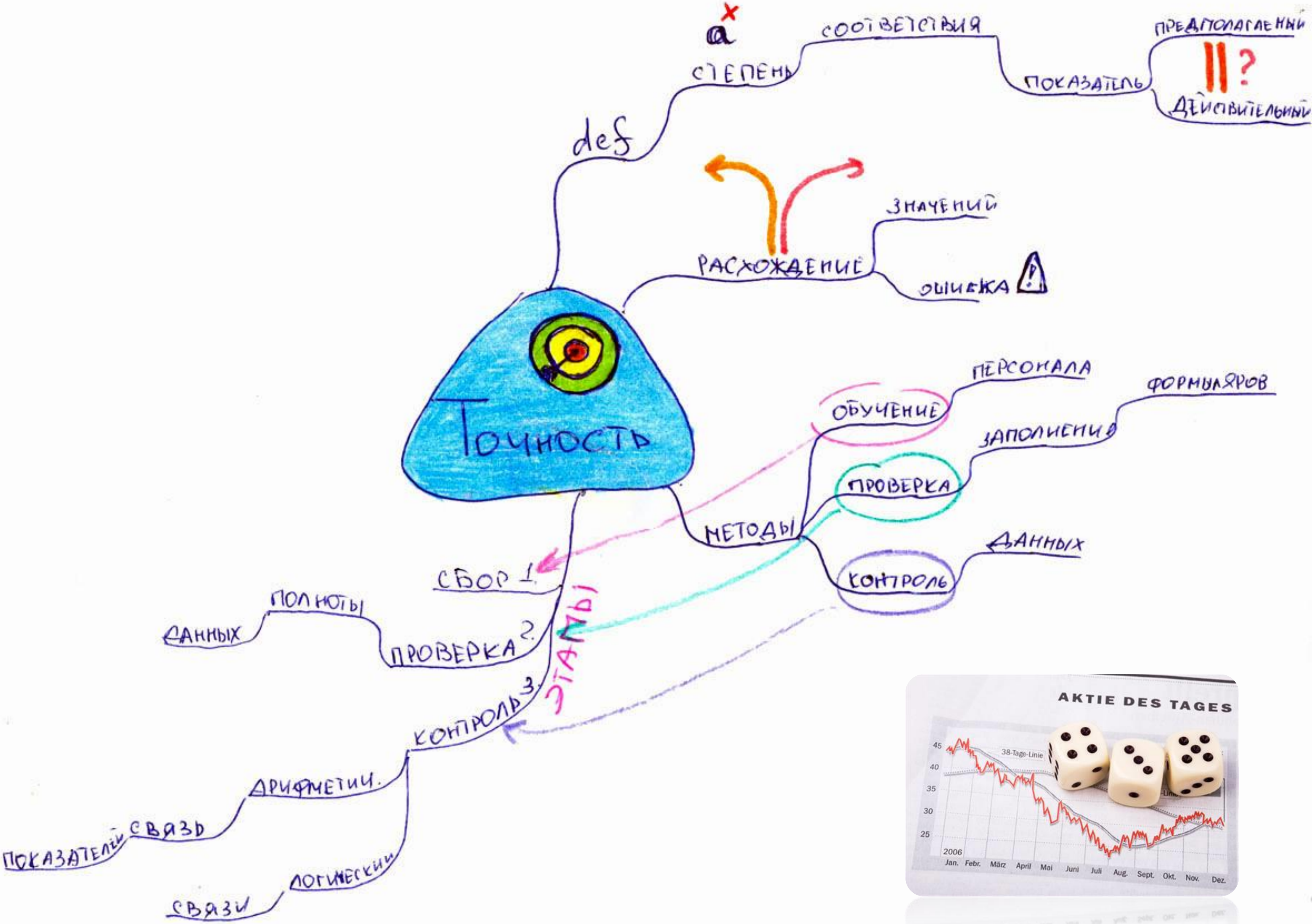
$$x_{ji} = x_{ji} - \bar{x}_i; \qquad z_{ji} = \frac{x_{ji}}{\sigma_j}$$

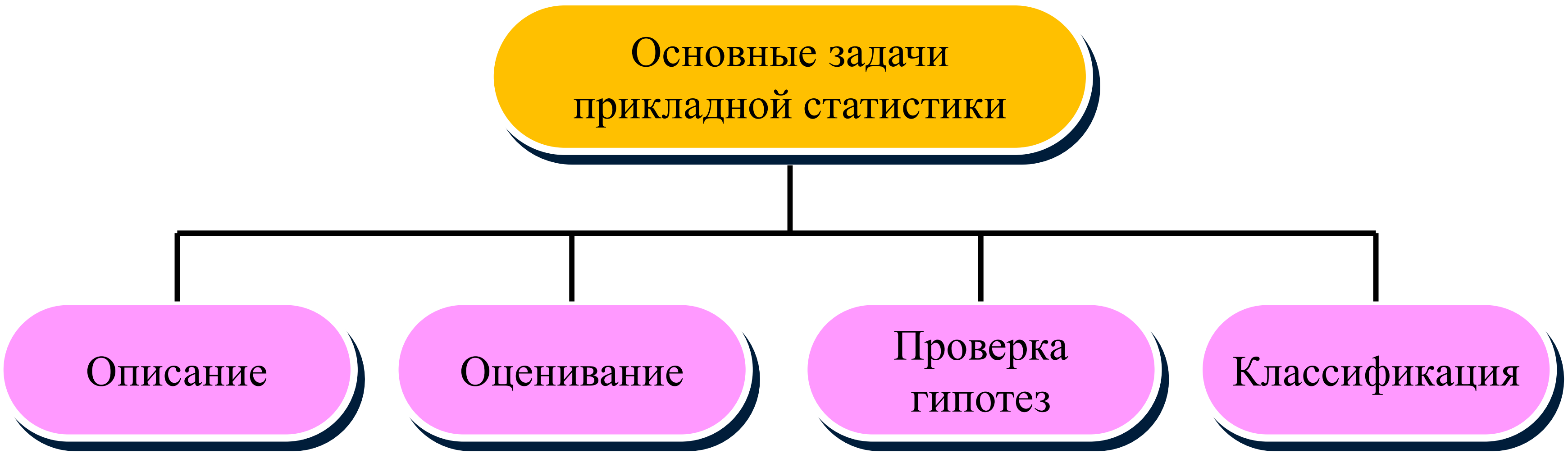
$$z_i = \sum_{\sigma=1}^{m} a_{i\sigma} F_\sigma + d_i U_i$$

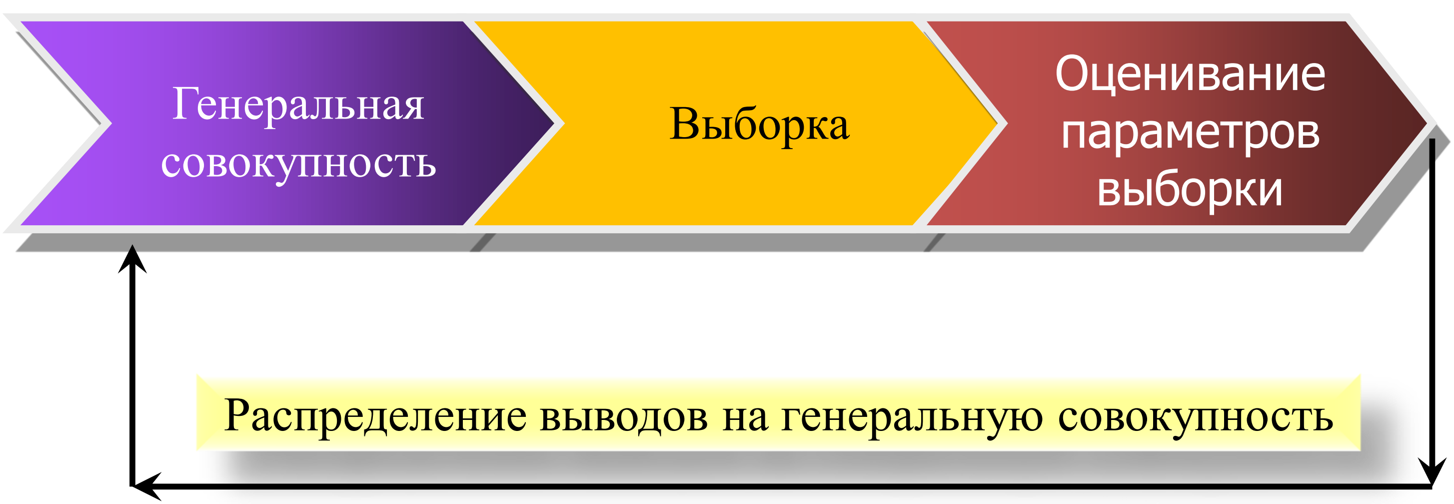
- факторы представляют собой случайные величины с нормальным законом распределения, заданные в стандартной форме;
- характерные факторы независимы как между собой, так и по отношению к общим факторам.



ОСНОВЫ ВЫБОРОЧНОГО НАБЛЮДЕНИЯ


ТОЧНОСТЬ
def
СТЕПЕНЬ
СООТВЕТСТВИЯ
ПОКАЗАТЕЛЬ
ПРЕДПОЛАГАЕМЫЙ
ДЕЙСТВИТЕЛЬНЫЙ
РАСХОЖДЕНИЕ
ЗНАЧЕНИЙ
ОШИБКА
МЕТОДЫ
ОБУЧЕНИЕ
ПЕРСОНАЛА
ПРОВЕРКА
ЗАПОЛНЕНИЯ
ФОРМУЛЯРОВ
КОНТРОЛЬ
ДАННЫХ
ЭТАПЫ
СБОР 1.
ПРОВЕРКА 2.
ПОЛНОТЫ
ДАННЫХ
КОНТРОЛЬ 3.
АРИФМЕТИЧ.
СВЯЗЬ
ПОКАЗАТЕЛЕЙ
ЛОГИЧЕСКИЕ
СВЯЗИ
AKTIE DES TAGES
38-Tage-Linie
45
40
35
30
25
2006
Jan. Febr. März April Mai Juni Juli Aug. Sept. Okt. Nov. Dez.


Основные задачи прикладной статистики
Описание
Оценивание
Проверка гипотез
Классификация


Генеральная совокупность
Выборка
Оценивание параметров выборки
Распределение выводов на генеральную совокупность

Этапы проведения выборочного наблюдения

Определение цели обследования

Установление границ генеральной совокупности

Составление программы наблюдения и программы разработки данных

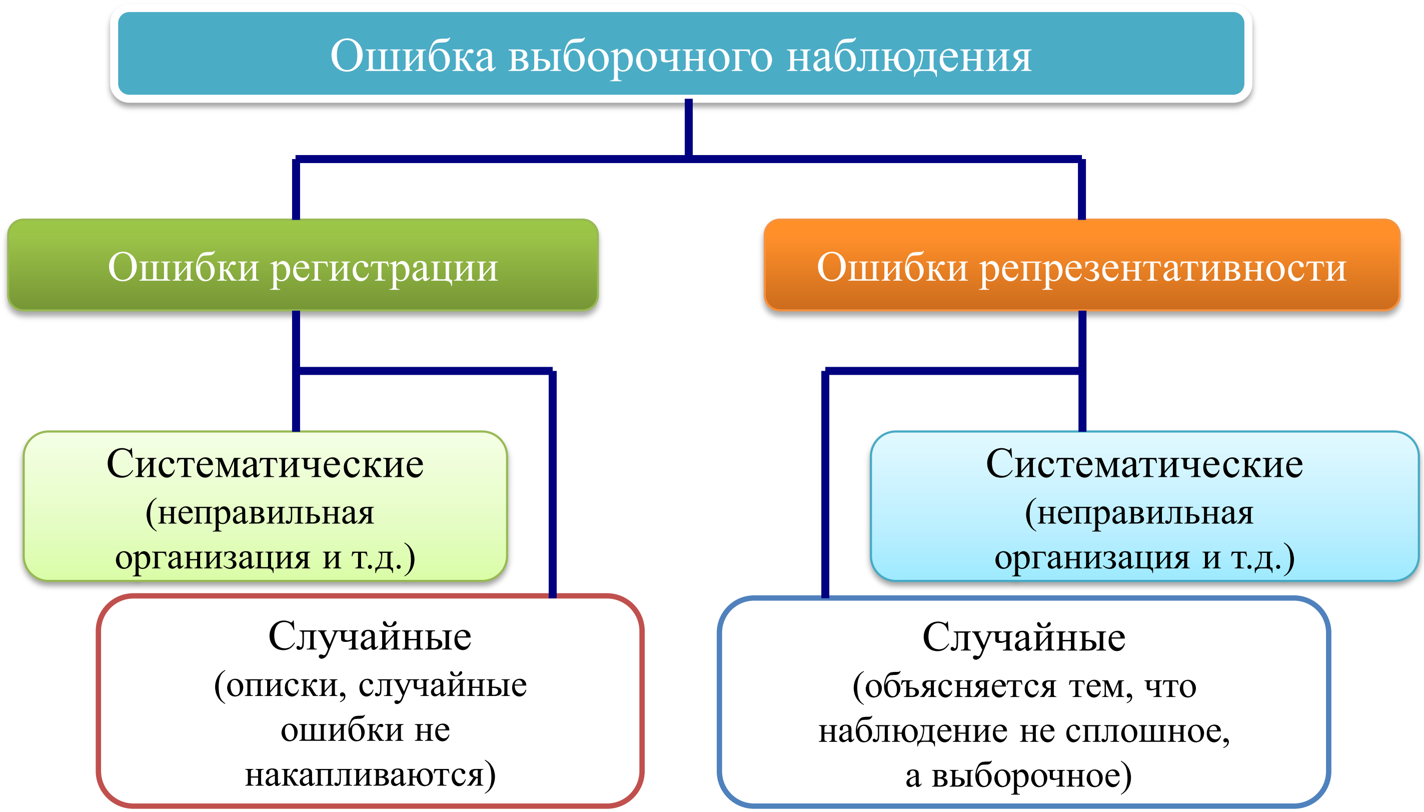
Определение вида выборки, процента отбора и метода отбора

Отбор и регистрация наблюдаемых признаков у отобранных единиц

Расчет выборочных характеристик и их ошибок

Распространение полученных результатов на генеральную совокупность

Классификация ошибок выборочного наблюдения



Ошибка выборочной доли

Выборочная доля представляет собой отношение числа единиц, обладающих данным признаком или данным его значением (m), к общему числу единиц выборочной совокупности (n)

$$w = \frac{m}{n}.$$

Ошибка выборочной доли представляет собой расхождение (разность) между долей в выборочной совокупности (w) и долей в генеральной совокупности (p), возникающее вследствие несплошного характера наблюдения. Величина ошибки выборочной доли определяется как предел отклонения w от p , гарантируемый с заданной вероятностью

$$|w - p| < \tau\mu_w ,$$

Значения гарантийного коэффициента

τ	P_τ	τ	P_τ	τ	P_τ
1,00	0,6827	1,70	0,9109	2,40	0,9836
1,10	0,7287	1,80	0,9281	2,50	0,9876
1,20	0,7699	1,90	0,9426	2,60	0,9907
1,30	0,8064	2,00	0,9545	2,70	0,9931
1,40	0,8385	2,10	0,9643	2,80	0,9949
1,50	0,8664	2,20	0,9722	2,90	0,9963
1,60	0,8904	2,30	0,9786	3,00	0,9973

Значения средней ошибки выборки определяются по формуле

$$\boxed{\mu = \sqrt{\frac{\sigma_0^2}{n}},}$$

Между дисперсиями в генеральной и выборочной совокупностях существует следующее соотношение:

$$\boxed{\sigma_0^2 = \frac{n}{n-1}\sigma^2,}$$

Если n достаточно велико, то $\dfrac{n}{n-1}$ близко к единице и дисперсию в генеральной совокупности можно заменить на дисперсию в выборке.

Средняя ошибка выборочной доли определяется по формуле $\mu_w = \sqrt{\dfrac{\sigma_w^2}{n}}$,

Для показателя доли альтернативного признака (выборочной доли) дисперсия определяется по формуле

$$\sigma_w^2 = w(1-w).$$

При бесповторном отборе численность генеральной совокупности сокращается, поэтому дисперсия умножается на коэффициент $1-\dfrac{n}{N}$.

Формулы расчета средних ошибок выборочной доли для различных способов отбора единиц из генеральной совокупности приведены в таблице.

Формулы расчета средних ошибок выборочной доли и выборочной средней

Метод отбора выборки	Средняя ошибка	
	выборочной доли	выборочной средней
Механический и собственно–случайный повторный	$\mu_w = \sqrt{\dfrac{w(1-w)}{n}}$	$\mu_x = \sqrt{\dfrac{\sigma_x^2}{n}}$
Механический и собственно–случайный бесповторный	$\mu_w = \sqrt{\dfrac{\sigma_w^2}{n}(1-\dfrac{n}{N})}$	$\mu_x = \sqrt{\dfrac{\sigma_x^2}{n}(1-\dfrac{n}{N})}$
Серийный при бесповторном отборе серий	$\mu_w = \sqrt{\dfrac{\delta_w^2}{r}(\dfrac{R-r}{R-1})}$	$\mu_x = \sqrt{\dfrac{\delta_x^2}{r}(\dfrac{R-r}{R-1})}$
Типический при повторном случайном отборе внутри групп	$\mu_w = \sqrt{\dfrac{\overline{\sigma_w^2}}{n}}$	$\mu_x = \sqrt{\dfrac{\overline{\sigma_x^2}}{n}}$
Типический при бесповторном случайном отборе внутри групп	$\mu_w = \sqrt{\dfrac{\overline{\sigma_w^2}}{n}(1-\dfrac{n}{N})}$	$\mu_x = \sqrt{\dfrac{\overline{\sigma_x^2}}{n}(1-\dfrac{n}{N})}$

Дисперсии в формулах расчета средних ошибок выборочной доли в таблице определяется следующим образом:

– межсерийная дисперсия выборочной доли $$\delta_w^2 = \frac{\sum_{j=1}^{r}(w_j - \overline{w})^2}{r},$$

– средняя из групповых дисперсий $$\overline{\sigma_w^2} = \overline{w(1-w)} = \frac{\sum_{j=1}^{k} w_j(1-w_j)n_j}{\sum_{j=1}^{k} n_j},$$

Предельное значение ошибки выборочной доли определяется по следующей формуле: $\Delta_w = \mu_w \tau.$

Ошибка выборочной средней

Ошибка выборочной средней представляет собой расхождение (разность) между выборочной средней $\tilde{x}$ и генеральной средней $\bar{x}$ возникающее вследствие несплошного выборочного характера наблюдения.

Величина ошибки выборочной средней определяется как предел отклонения $\tilde{x}$ от $\bar{x}$ гарантируемый с заданной вероятностью:

$$|\tilde{x} - \bar{x}| < \tau \mu_x ,$$

При повторном отборе средняя ошибка определяется следующим образом: $\mu_x = \sqrt{\dfrac{\sigma_x^2}{n}},$

$$\sigma_x^2 = \frac{\sum_{i=1}^{n} (x_i - \tilde{x})^2}{n},$$

или

$$\sigma_x^2 = \frac{\sum_{i=1}^{n} (x_i - \tilde{x})^2 f_i}{\sum_{i=1}^{n} f_i},$$

Межсерийная дисперсия выборочных средних δ_x^2 и средняя из выборочных дисперсий типических групп $\overline{\sigma_x^2}$ вычисляются следующим образом:

$$\delta_x^2 = \frac{\sum_{j=1}^{r} (\tilde{x}_j - \tilde{x})^2}{r}; \qquad \sigma_x^2 = \frac{\sum_{j=1}^{k} \sigma_j^2 n_j}{\sum_{j=1}^{k} n_j},$$

Предельная ошибка выражается следующим образом: $\Delta_x = \mu_x \tau$

Средняя величина количественного признака в генеральной совокупности определяется с учетом предельной ошибки выборочной средней

$$\boxed{\overline{x} = \tilde{x} \pm \Delta_x}$$

Объем выборки

Определение необходимого **объема выборки** n основывается на формулах предельных ошибок выборочной доли и выборочной средней. Например, для повторного отбора предельные ошибки равны

$$\Delta_w = \tau\sqrt{\frac{w(1-w)}{n}}; \qquad \Delta_x = \tau\sqrt{\frac{\sigma_x^2}{n}},$$

отсюда объемы выборок для расчета выборочной доли n_w и выборочной средней n_x следующие:

$$n_w = \frac{\tau^2 w(1-w)}{\Delta_w^2}; \qquad n_x = \frac{\tau^2 \sigma_x^2}{\Delta_x^2}.$$

Аналогичным образом определяются объемы выборок при различных способах отбора выборочной совокупности. Для серийного отбора определяется число отобранных серий. Формулы расчета приведены в таблице.

Формулы расчета объема выборки

Метод отбора выборки	Объем выборки или число серий для определения	
	выборочной доли	выборочной средней
Механический и собственно–случайный повторный	$n_w = \dfrac{\tau^2 w(1-w)}{\Delta_w^2}$	$n_x = \dfrac{\tau^2 \sigma_x^2}{\Delta_x^2}$
Механический и собственно–случайный бесповторный	$n_w = \dfrac{\tau^2 \sigma_w^2 N}{\Delta_w^2 N + \tau^2 \sigma_w^2}$	$n_x = \dfrac{\tau^2 \sigma_x^2 N}{\Delta_x^2 N + \tau^2 \sigma_x^2}$
Серийный при бесповторном отборе серий	$r_w = \dfrac{\tau^2 \delta_w^2 R}{\Delta_w^2 (R-1) + \tau^2 \delta_w^2}$	$r_x = \dfrac{\tau^2 \delta_x^2 R}{\Delta_x^2 (R-1) + \tau^2 \delta_x^2}$
Типический при повторном случайном отборе внутри групп	$n_w = \dfrac{\tau^2 \overline{\sigma_w^2}}{\Delta_w^2}$	$n_x = \dfrac{\tau^2 \overline{\sigma_x^2}}{\Delta_x^2}$
Типический при бесповторном случайном отборе внутри групп	$n_w = \dfrac{\tau^2 \overline{\sigma_w^2} N}{\Delta_w^2 N + \tau^2 \overline{\sigma_w^2}}$	$n_x = \dfrac{\tau^2 \overline{\sigma_x^2} N}{\Delta_x^2 N + \tau^2 \overline{\sigma_x^2}}$

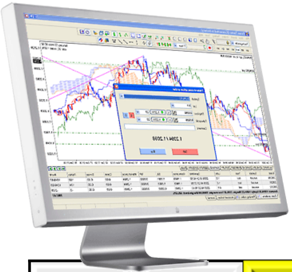
Средняя ошибка выборки при многоступенчатом отборе

$$\mu = \sqrt{\mu_1^2 + \frac{\mu_2^2}{n_1} + \frac{\mu_3^2}{n_1 n_2} + \ldots},$$

где μ_1, μ_2, μ_3 - средние ошибки выборки на отдельных ступенях отбора;

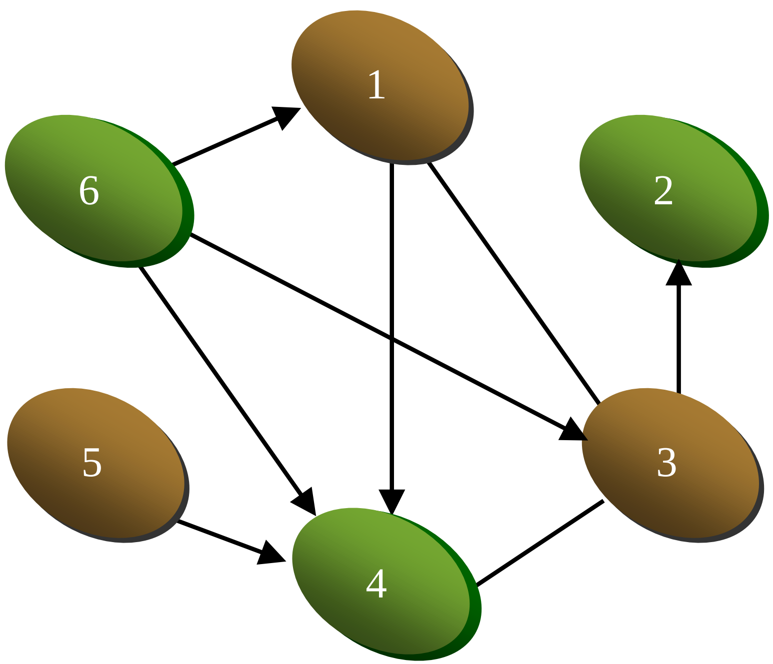
n_1, n_2 - численность выборок на соответствующих ступенях

МЕТОД КАЧЕСТВЕННОЙ ОЦЕНКИ

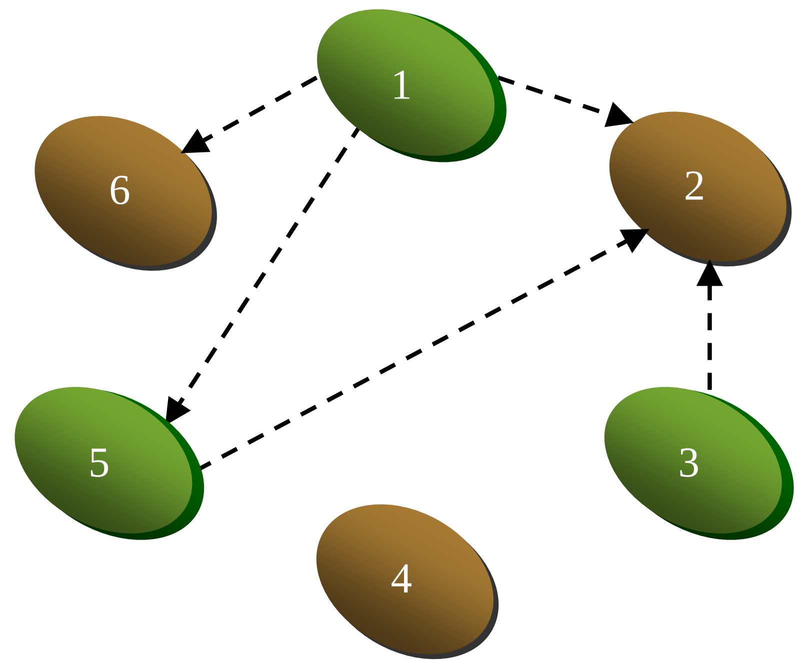


Многомерное шкалирование

№ п/п	Кто выбирает	Кого выбирают						Число отданных выборов		
		1	*2*	*3*	*4*	*5*	*6*	+	-	*Всего*
1	Алексеев	*	-	+	+	-	-	2	3	5
2	Бондарев	0	*	0	+	0	+	2	0	2
3	Михайлов	+	-	*	+	0	0	2	1	3
4	Нилов	0	0	+	*	0	+	2	0	2
5	Поляков	0	-	0	+	*	0	1	1	2
6	Чижов	+	+	+	+	0	*	4	0	4
Число полученных выборов	+	2	1	3	5	0	2	13		
	-	0	3	0	0	1	1		5	
	Всего	2	4	3	5	1	3			18



Положительные выборы по критерию



Отрицательные выборы по критерию

Построение регрессионных зависимостей