

Сводка и группировка статистических данных

Понятие. Задачи. Этапы. Виды.

Статистическая сводка -

- комплекс последовательных операций по первичной обработке данных с целью выявления типичных черт и закономерностей, присущих изучаемому явлению.
- научно-организованная обработка материалов наблюдения, включающая подсчет групповых и общих итогов, получение производных показателей (средних, относительных величин), систематизацию, группировку данных и составление таблиц и графиков
- систематизация единичных фактов, позволяющая перейти к обобщающим показателям, относящимся по всей изучаемой совокупности и ее частям и осуществить анализ и прогнозирование изучаемых явлений и процессов.

Задача сводки

дать характеристику объекту исследования с помощью системы статистических показателей, выявить и измерить его существенные черты и особенности:

- определение групп и подгрупп;
- определение системы показателей;
- определение видов таблиц.

Этапы сводки стат.данных:

- **1 этап:** осуществляется систематизация, группировка материалов, собранных при наблюдении.
- **2 этап:** уточняется предусмотренная планом система показателей, с помощью которых количественно характеризуются свойства и особенности изучаемого предмета.
- **3 этап:** исчисляются сами показатели; обобщенные данные представляется в таблицах, статистических рядах, графиках, диаграммах.

Статистическая сводка:

- **Простая сводка** - это операция по подсчету общих итогов по совокупности единиц наблюдения.
- **Сложная сводка** - это комплекс операций, включающих группировку единиц наблюдения, подсчет итогов по каждой группе и по всему объекту и представление результатов группировки и сводки в виде статистических таблиц.

Статистическая сводка:

- *Подлежащее сводки* составляют группы или части, на которые разбивается совокупность.
- *Сказуемое сводки* составляют показатели, характеризующие каждую группу и совокупность в целом.

Группировка -

- это метод, при котором вся исследуемая совокупность разделяется на группы по какому-то существенному признаку.

Функции группировки:

- выделение социально-экономических типов явлений;
- изучение структуры и структурных сдвигов, происходящих в социально-экономических явлениях;
- анализ взаимосвязей между явлениями;
- выявление влияния отдельных единиц на средние итоговые показатели

Категории метода группировки:

- Группировочный признак
- Интервал группировки
- Число групп

1. Группировочный признак (основание группировки) -

- признак, по которому осуществляется группировка (по которому происходит объединение отдельных единиц совокупности в однородные группы)
- По количеству группировочных признаков, выделяют:
 - простую (монотетическую) группировку (если для построения группировки используется один группировочный признак);
 - сложную, многомерную (политетическую) группировку (если для построения группировки используется несколько группировочных признаков).
Частным случаем многомерной группировки является *комбинационная группировка*, базирующаяся на двух и более признаках, *взятых во взаимосвязи, в комбинации*

В зависимости от решаемых задач различают:

- **Типологическую группировку** - разделение исследуемой совокупности на однородные группы.
- **Структурную группировку** - в которой происходит разделение однородной совокупности на группы, характеризующие ее структуру по какому-то варьирующему признаку. Изменение структуры изучаемых явлений - **структурные сдвиги**.
- **Аналитическую (факторную) группировку** - позволяет выявить взаимосвязи между изучаемыми явлениями и их признаками (показатели: обобщающий (результативный) и факторные)

2. Интервал группировки - значение варьирующего признака, лежащее в определенных пределах

- **Нижняя граница** - значение наименьшего признака в интервале.
- **Верхняя граница** - наибольшее значение в интервале.
- **Величина (ширина) интервала** - разница между верхней и нижней границами.

Интервалы бывают:

- *равные*, когда разность между максимальным и минимальным значениями в каждом из интервалов одинакова или значение количественного признака внутри совокупности изменяется равномерно;
- неравные*, когда значения признака варьируют неравномерно и в значительных размерах (**прогрессивно возрастающие, прогрессивно - убывающие, произвольные и специализированные**);
- *открытые*, у которых указана только одна граница, либо верхняя, либо нижняя;
- *закрытые*, когда имеются и нижняя, и верхняя границы.

Величина равных интервалов определяется по формуле:

$$i = (X_{\max} - X_{\min}) / n, \text{ или } i = R / n,$$

где **R** – размах вариации = $X_{\max} - X_{\min}$

i — величина интервала

X_{max} - максимальное значение признака в совокупности

X_{min} — минимальное значение признака в совокупности

n — число групп

Правила округления интервалов:

- Если интервал имеет один знак ДО запятой, то полученное значение округляется до десятых ($0,88 = 0,9$; $8,715 = 8,7$)
- Если величина интервала имеет два знака ДО запятой, то полученное значение округляется до целых ($11,11 = 11$; $29,98 = 30$)
- Если интервал трех, четырех и более значимое число, то интервал принимают кратным 50 или 100 ($h=246 \approx 250$)

Величина неравных интервалов определяется по формулам:

- Величина интервалов, изменяющихся в арифметической прогрессии:

$$h_{i+1} = h_i + a, \text{ где}$$

a - константа, которая будет положительной при прогрессивно возрастающих интервалах и отрицательной при прогрессивно убывающих;

- Величина интервалов, изменяющихся в геометрической прогрессии:

$$h_{i+1} = h_i * q, \text{ где}$$

q - положительная константа, которая при прогрессивно возрастающих интервалах больше 1, а при прогрессивно убывающих интервалах меньше 1.

**3. Определение числа групп n , при N
единицах совокупности
(формула Стерджесса)**

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

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

или

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

Пример. Распределение предприятий по региона по величине розничного товарооборота в текущем году*

№ п/п	Розничный товарооборот, тыс.руб.	№ п/п	Розничный товарооборот, тыс.руб.
1.	151331	16.	21253
2.	56440	17.	47248
3.	99212	18.	92955
4.	34088	19.	178291
5.	43520	20.	68865
6.	38196	21.	9767
7.	208492	22.	60674
8.	104518	23.	23944
9.	82972	24.	127725
10.	45561	25.	24559
11.	137445	26.	21946
12.	28970	27.	44876
13.	51387	28.	117021
14.	156775	29.	33775
15.	65680	30.	36637

$$\lg 30 = 1,477121254719$$

Величину интервала $i=h$ определяют по следующей формуле:

$$h = \frac{R}{n} = \frac{x_{\max} - x_{\min}}{n} = \frac{208492 - 9767}{1 + 3.322 \lg 30} = \frac{198725}{6} \approx 33121$$

Группы предприятий по размеру розничного товарооборота, тыс.руб.	Число предприятий	№ предприятия	Общий товарооборот по группе, тыс.руб.	Средний размер товарооборота по группам, тыс.руб.
9767 – 42888	10	4, 6, 12, 16, 21, 23, 25, 26, 29, 30	273135	27313,5
42888 – 76009	9	2, 5, 10, 13, 15, 17, 20, 22, 27	484251	23805,7
76009 – 109130	4	3, 8, 9, 18	379657	94914,3
109130 – 142251	3	11, 24, 28	382191	127397
142251 – 175372	2	1, 14	308106	154053
175372 – 208493	2	7, 19	386783	193391,5
Итого:	30		2214123	73804,1

ПРИМЕР:

Данные о стоимости основных фондов 50 предприятий:

12.9	12.6	6.7	9.7	9.3	10.8	15	7	13	9.5
9.4	8	6.3	10	15	8.2	7.3	9.2	5.8	8.7
5.2	13.2	8.1	7.5	11.8	14.6	8.5	7.8	10.5	6
5.1	6.8	8.3	7.7	7.9	9	10.1	8	12	14
8.2	9.8	13.5	12.4	5.5	7.9	9.2	10.8	12.1	12.4

***Ряд распределения –
ряд чисел, показывающий,
как распределяются единицы
некоторой совокупности
по изучаемому признаку.***

***Атрибутивные, построенные
по качественному признаку***

***Вариационные, построенные
по количественному признаку***

Вариационные ряды распределения

- дискретные – основаны на прерывной вариации признака
- интервальные - непрерывные

Варианта – отдельное значение варьируемого признака, который он принимает в ряду распределения

Частоты - это численности отдельных вариантов или каждой группы вариационного ряда. Частоты, выражающиеся в долях единицы, или в % к итогу, называются **частостями**.

- Сумма всех частот называется **объемом совокупности**

Данные о стоимости основных фондов 50 предприятий:

Группы	F_i (частоты) сколько раз встречается значение признака	P_i (частоты) = $F:N$ (отношение частоты к числу опытов)	W_i (сумма частот, накопленная численность, объем совокупности)
5,1-6.5	6	0.12	0.12
6.6-8	11	0.22	0.34
8,1-9.5	12	0.24	0.58
9.6-11	7	0.14	0.72
11,1-12.5	5	0.1	0.82
12.6-14	6	0.12	0.94
14,1-15.5	3	0.06	1

**Пример дискретного ряда распределения
(распределение магазинов по числу товарных секций)**

Число товарных секций	Число магазинов	В % к итогу	Число магазинов	В % к итогу
	На 01.01.2014 г		На 01.01.2015 г	
1	3	6	6	10
2	10	20	16	27
3	15	30	20	33
4	12	24	12	20
5	7	14	4	7
6	3	6	2	3
Итого	50	100	60	100

Пример интервального ряда распределения

Выработка продавцов (тыс.руб)	Число продавцов (чел.)	В % к итогу	Коммулятивная (накопленная) численность продавцов
80-100	5	10	5
100-120	10	20	15
120-140	20	40	35
140-160	10	20	45
160-180	5	10	50
ИТОГО	50	100	

Статистические таблицы и графики

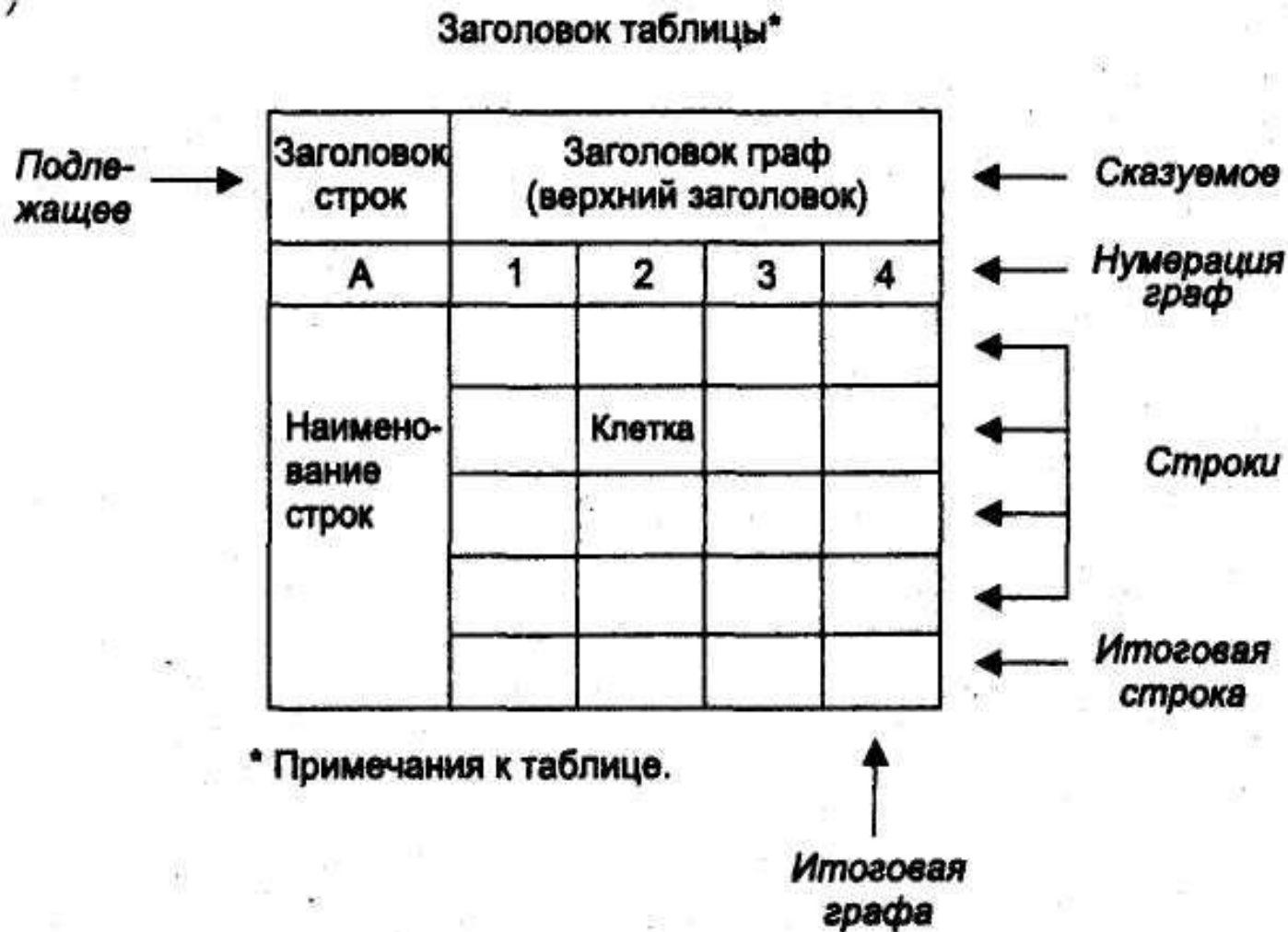


Рис. 3.1. Схема статистической таблицы

Составные части таблицы

- *Подлежащее* таблицы (обычно помещается слева и составляет содержание строк) показывает объект изучения, характеризующийся рядом показателей (*сказуемым*).
- *Сказуемое* таблицы (записывается сверху и составляет содержание граф) показывает, какими признаками (показателями) характеризуется объект (*подлежащее*).

Статистические графики

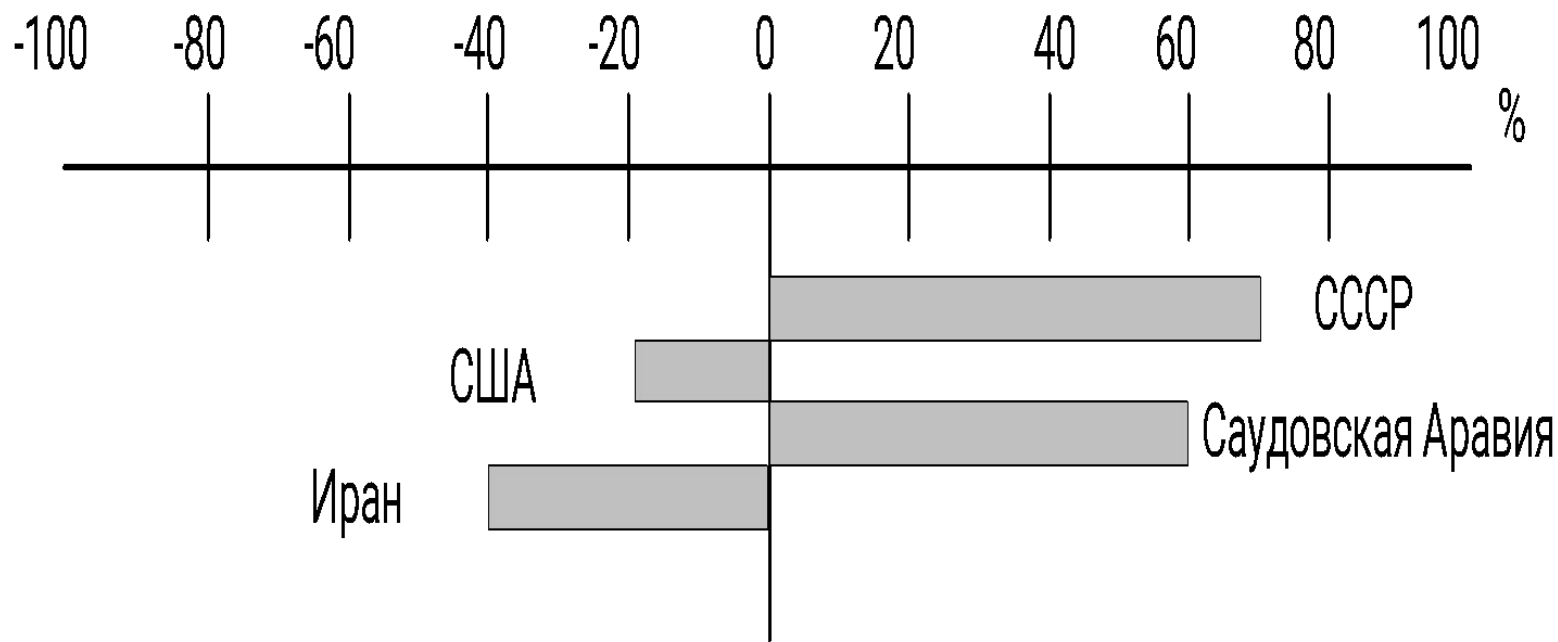
- По *содержанию* или назначению : графики сравнения в пространстве, графики различных относительных величин, графики вариационных рядов, графики размещения по территории, графики взаимосвязанных показателей.
- По *способу построения* делятся на диаграммы, картодиаграммы и картограммы.
- По *характеру графического образа* различают графики точечные, линейные, плоскостные (столбиковые, почасовые, квадратные, круговые, секторные, фигурные) и объемные.

Диаграмма



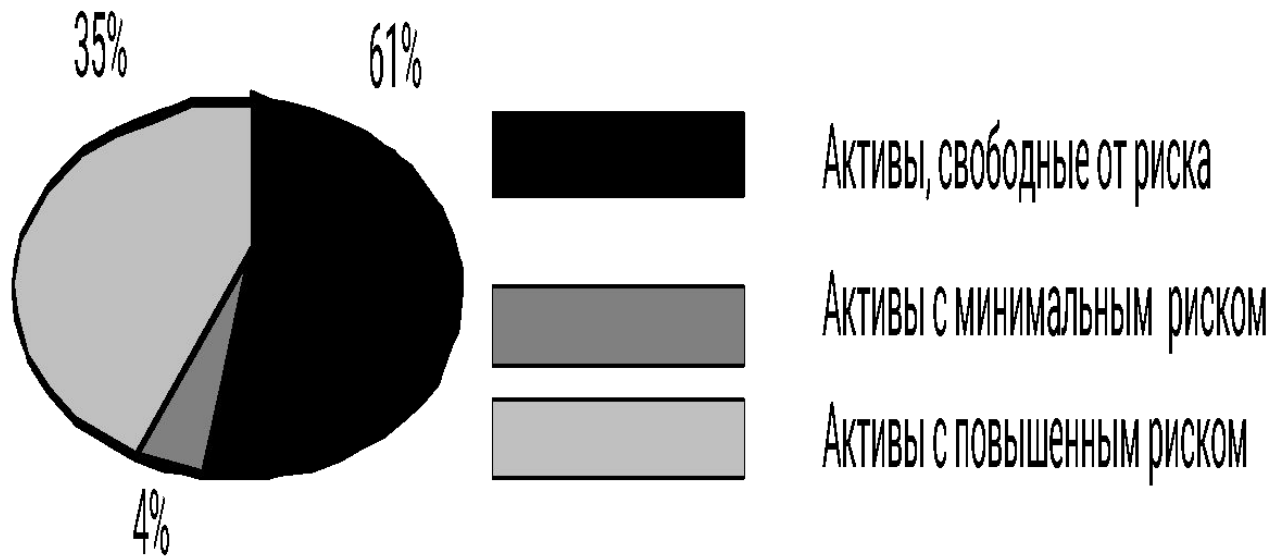
Запасы нефти в отдельных странах в 1987 г.

Ленточная диаграмма



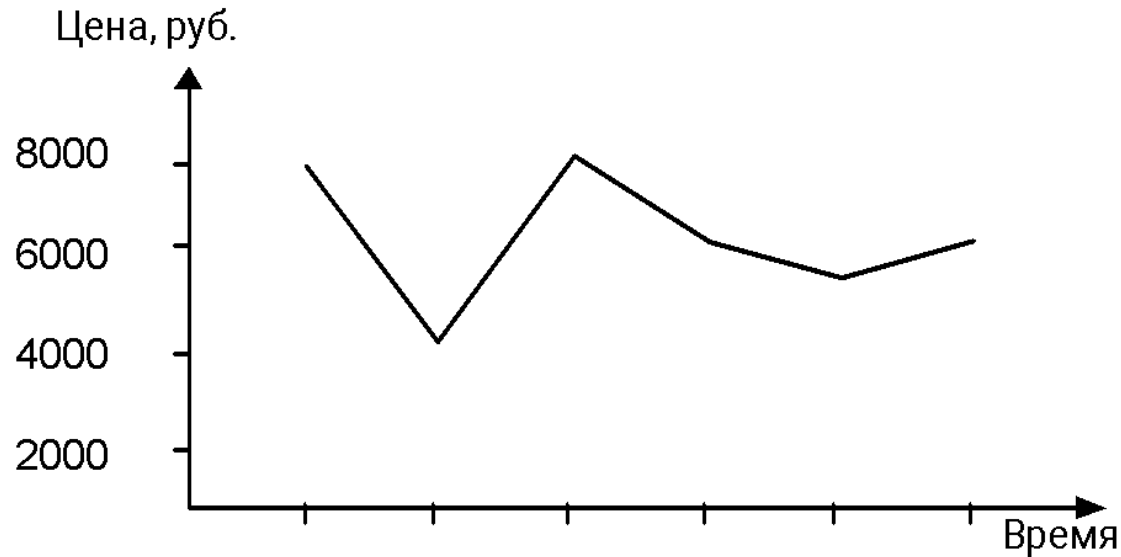
Добыча нефти в отдельных странах в 1986 г. по сравнению с 1970 г.

Секторная диаграмма



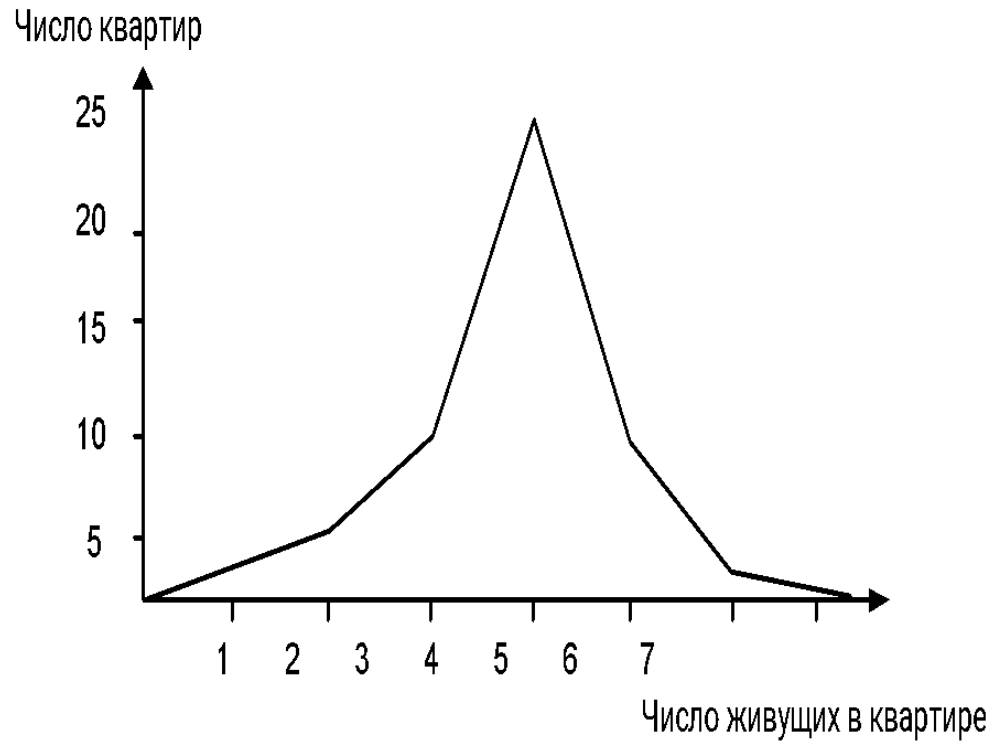
Структура активов коммерческого банка по степени риска.

Линейная координатная диаграмма



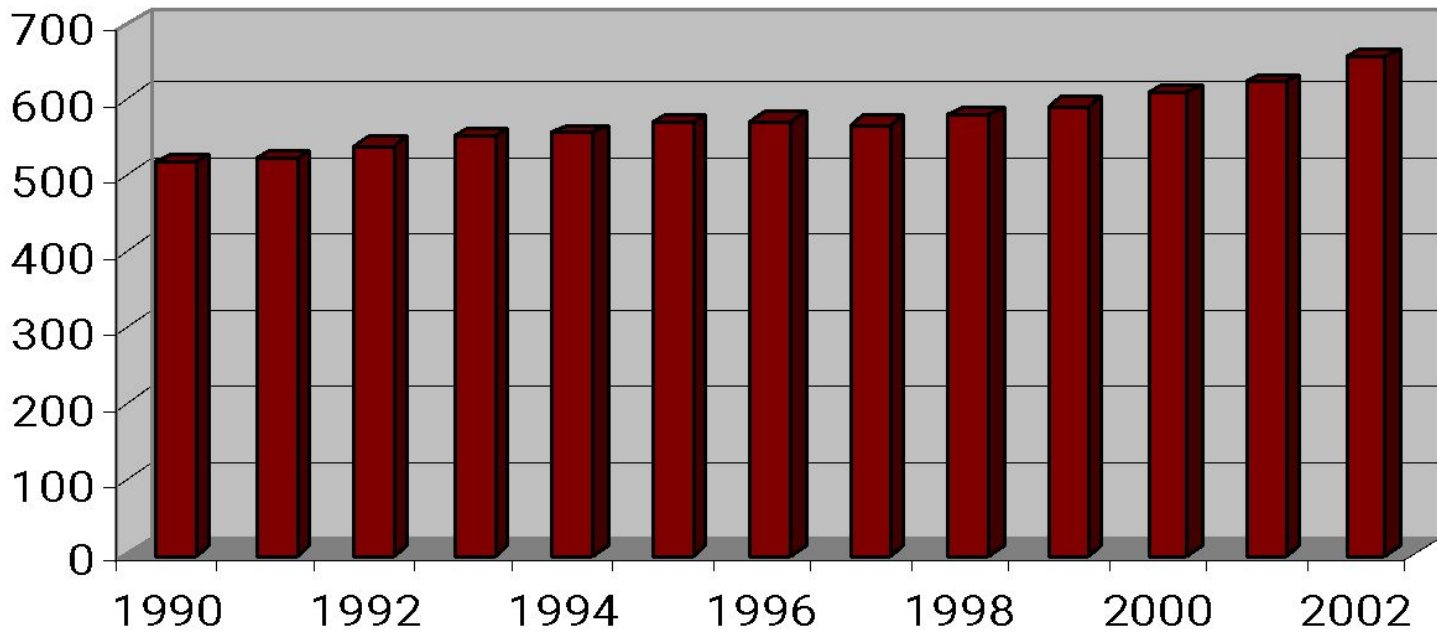
Уровень средней цены приватизационных чеков на торгах РТСБ, руб.

Полигон распределения



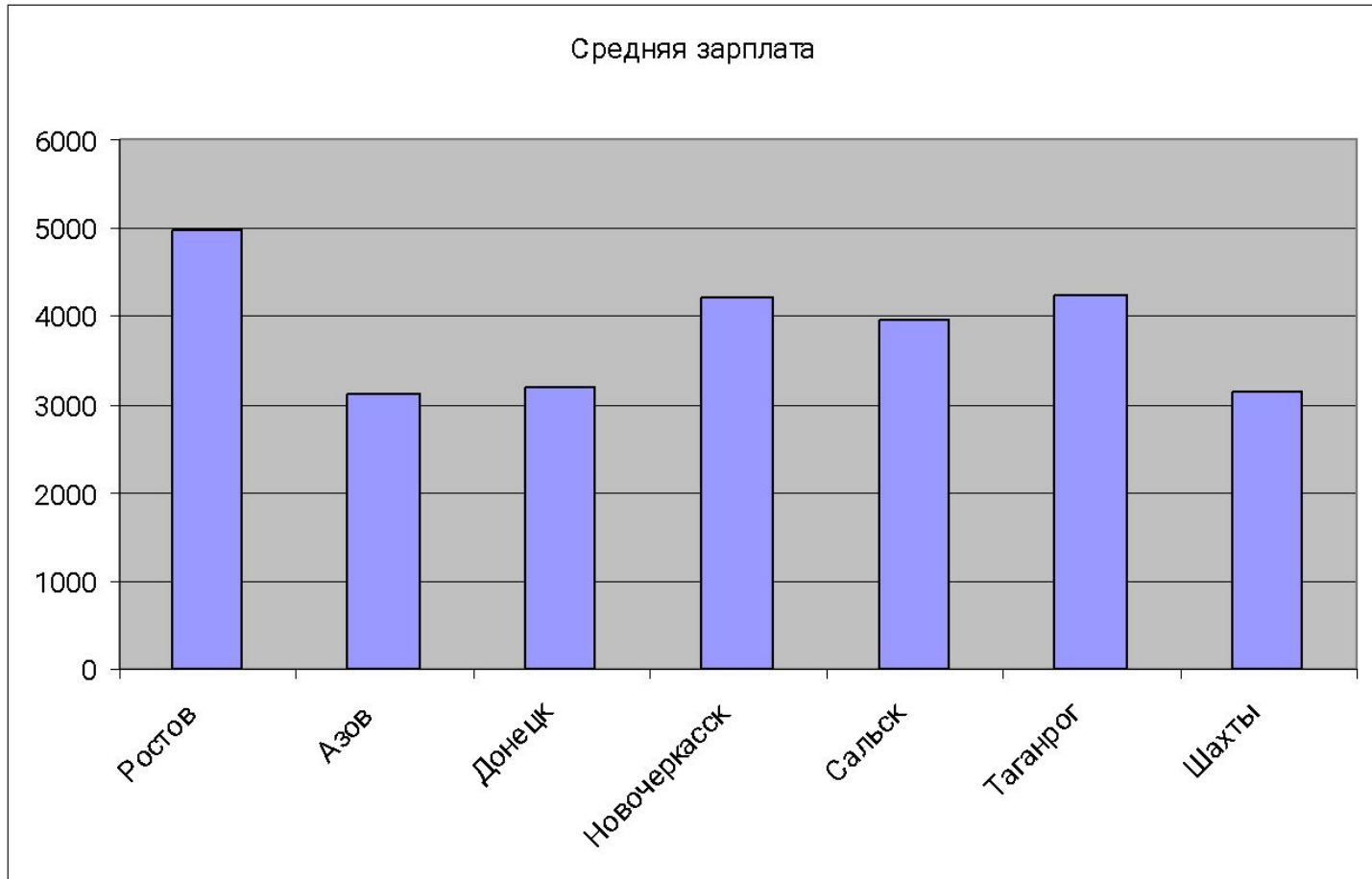
Гистограмма

Итого государственных Вузов



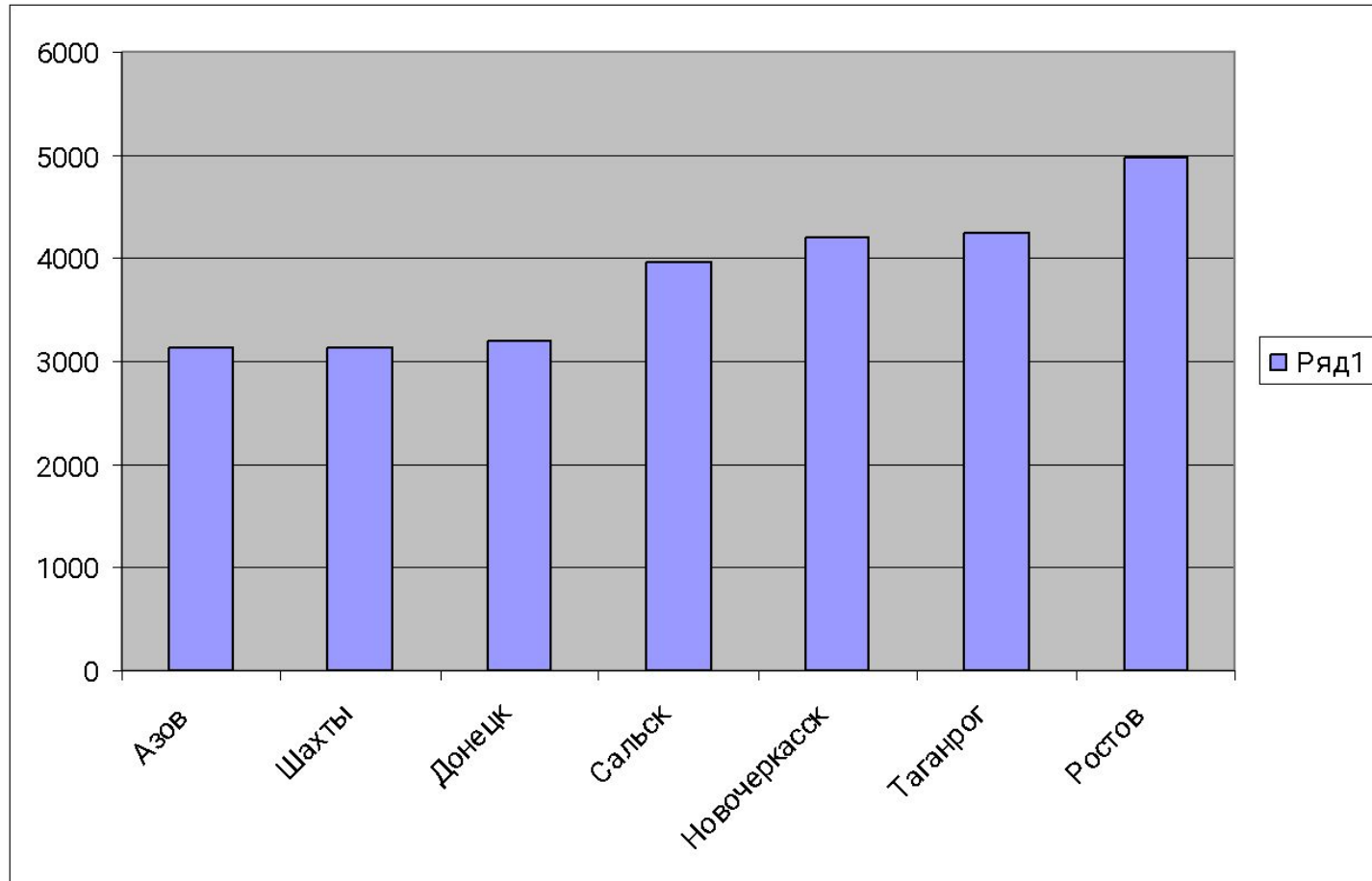
Количество государственных вузов в России

Диаграмма казусов



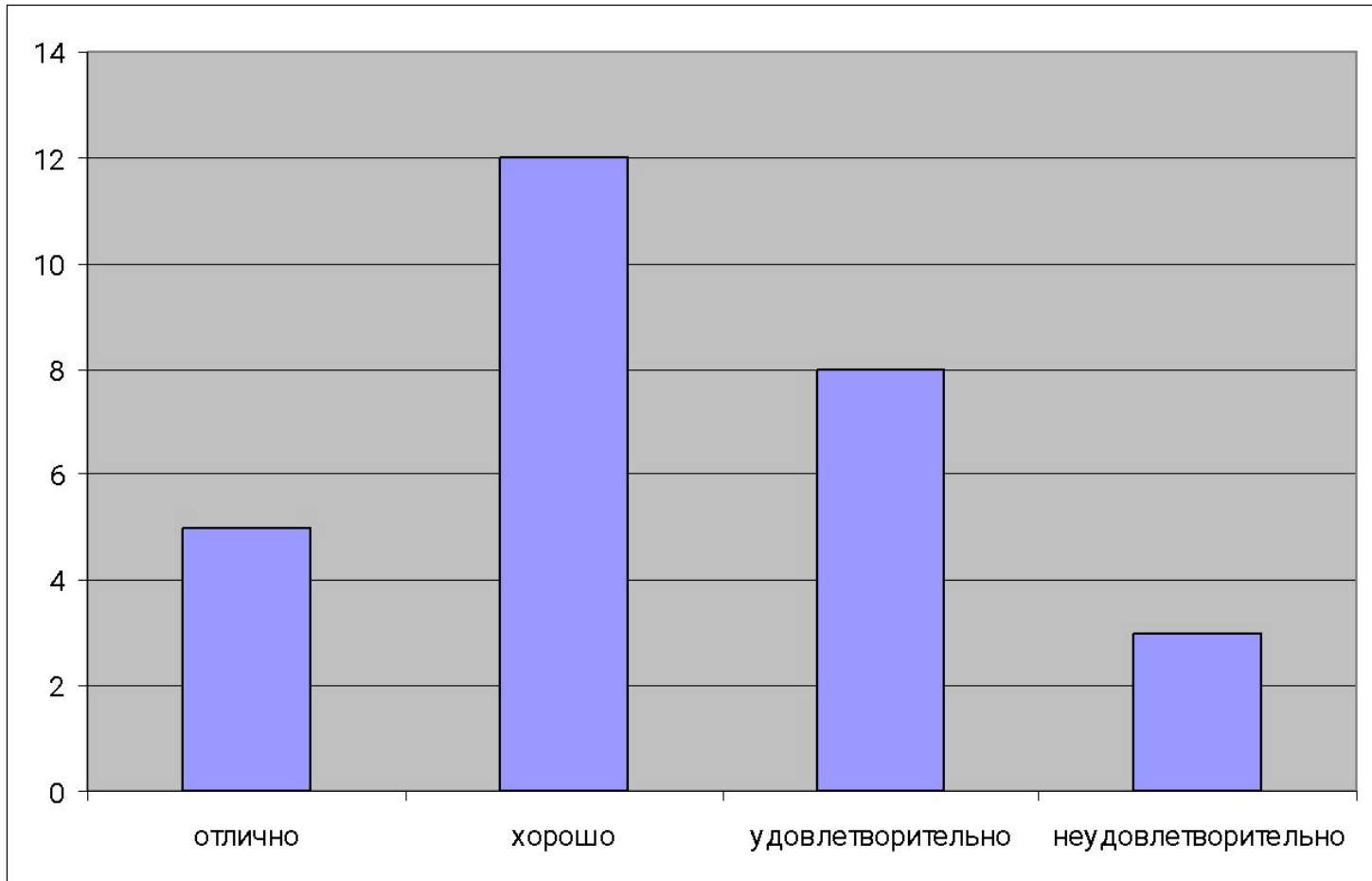
Средняя заработная плата по Ростовской области, 2003 год

Огива



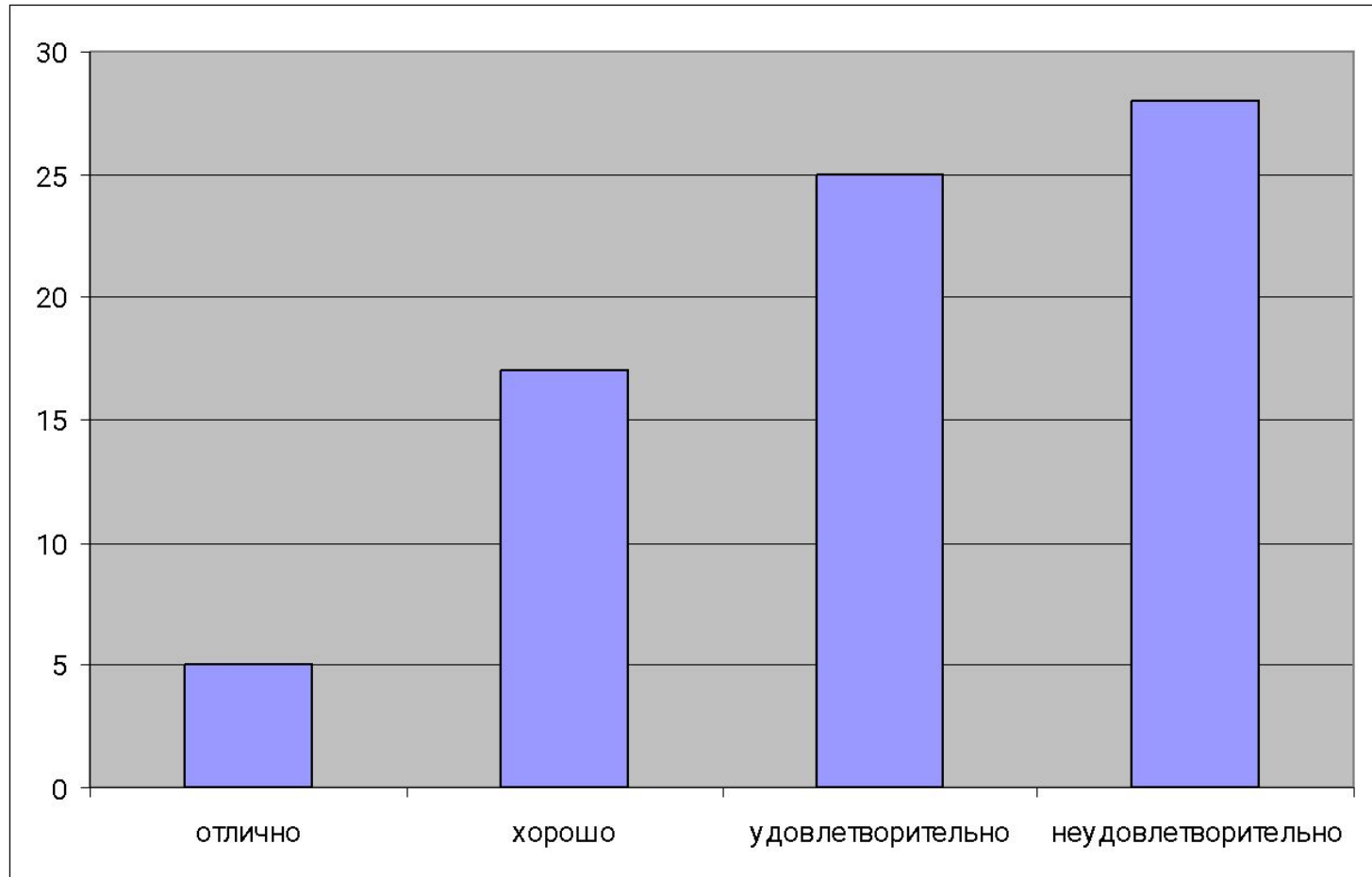
Средняя заработная плата по Ростовской области, 2003 год

Полигон



Результаты контрольной по дисциплине

Кумулята



Результаты контрольной по дисциплине
«Статистика»