

Математическая статистика



Статистика – наука, которая занимается получением, обработкой и анализом количественных данных о разнообразных массовых явлениях, происходящих в природе и обществе. Слово « статистика» происходит от латинского слова *status*, которое означает « состояние, положение вещей»

Статистика знает всё!

Известно, сколько, какой пищи съедает в год в среднем гражданин страны. Сколько в стране охотников, балерин, артистов, рабочих и т. д.

Результаты статистических исследований широко используются для практических и научных выводов. Статистические характеристики применяют для нахождения средней урожайности пшеницы с 1 га в данном районе, среднего суточного удоя молока от одной коровы на ферме и т.п.

Математическая статистика

- это раздел математики, посвященный методам сбора, анализа и обработки статистических данных для научных и практических целей.



описательная статистика

- методы описания статистических данных, представления их в форме таблиц, распределений и пр.

аналитическая статистика (теория статистических выводов)

- обработка данных, полученных в ходе эксперимента, и формулировка выводов, имеющих прикладное значение для конкретной области человеческой деятельности.

- Теория статистических выводов тесно связана с другой математической наукой — теорией вероятностей и базируется на ее математическом аппарате.

ТАБЛИЧНОЕ ПРЕДСТАВЛЕНИЕ СТАТИСТИЧЕСКИХ ДАННЫХ

Для обобщения и систематизации данных, полученных в результате статистического наблюдения, их по какому-либо признаку разбивают на группы, и результаты группировки сводят в таблицы.

Статистическая таблица

- — это своего рода статистическое предложение, которое состоит из статистического подлежащего и статистического сказуемого.
- Подлежащее показывает о чем идет речь, в таблице обычно располагают слева и представляют собой название строк.
- Сказуемое располагают сверху таблицы и представляют собой название граф, т.е. показывают какими признаками обладает объект исследования.



Правила построения таблиц:

1. Таблица должна иметь общий заголовок, в котором выражается: сам объект, признаки объекта, время и место, к которому относится статистический материал, единица измерения, если они общие для всей таблицы.
2. Число признаков в сказуемом должно быть ограничено.
3. Округление должно быть проведено с одинаковой степенью точности (максимально 4 знака после запятой).
4. Отсутствие данных могут быть обусловлены разными причинами и это по-разному отражается в таблице:



Прогноз численности населения России.

Год	Росстат		Институт демографии	ООН	Год	Росстат		Институт демографии	ООН
	отчёт	прогноз	прогноз	прогноз		отчёт	прогноз	прогноз	прогноз
	млн. чел.	млн. чел.	млн. чел.	млн. чел.		млн. чел.	млн. чел.	млн. чел.	млн. чел.
1897		67,5			2007	142,2	141,4		
1917		91			2008		140,8		
1926		92,7			2009		140,2		
1939		108,4			2010		139,6	138,7	
1959		117,2			2011		139,1		
1970		129,9			2020		135,9	130,9	
1990		147,7			2025		133,8	125,8	
2000		146,9		144,8	2050			98,6	101,5
2005		143,5		141,7	2070				
2006		142,8	142,1		2100				79,5

Пример

Фирма планирует купить в интернет-магазине 14 ежедневников для сотрудников. Цены на ежедневники и стоимость доставки всего купленного объема приведены в таблице.

Поставщик №	Цена за 1 шт., руб.	Стоимость доставки, руб.	Дополнительные условия
1	300	200	
2	250	700	При объеме покупки на сумму более 3000 рублей доставка бесплатно
3	240	600	При объеме покупки на сумму более 5000 рублей доставка бесплатно

Сколько рублей нужно заплатить за самую дешевую покупку с доставкой?

Пример 2.

Администрация школы решила проверить математическую подготовку девятиклассников. С этой целью был составлен тест, содержащий 9 заданий. Работу выполняли 40 учащихся. При проверке каждой работы учитель отмечал число верно выполненных заданий. Полученные данные представим в виде таблицы.

Число верно выполненных заданий	0	1	2	3	4	5	6	7	8	9
Частота появлений	1	1	1	2	5	6	8	7	5	4

Такую таблицу называют *таблицей частот*.

Число верно выполненных заданий	0	1	2	3	4	5	6	7	8	9
Относительная частота, %	2,5	2,5	2,5	5	12,5	15	20	17,5	12,5	10

Отношение частоты к общему числу данных, выраженное в процентах, называют относительной частотой, а саму таблицу – *таблицей относительных частот*.

ГРАФИЧЕСКОЕ ПРЕДСТАВЛЕНИЕ СТАТИСТИЧЕСКИХ ДАННЫХ

Для наглядного представления данных, полученных в результате статистического исследования, широко используются различные способы их изображения: **полигон частот, гистограмму, диаграммы.**

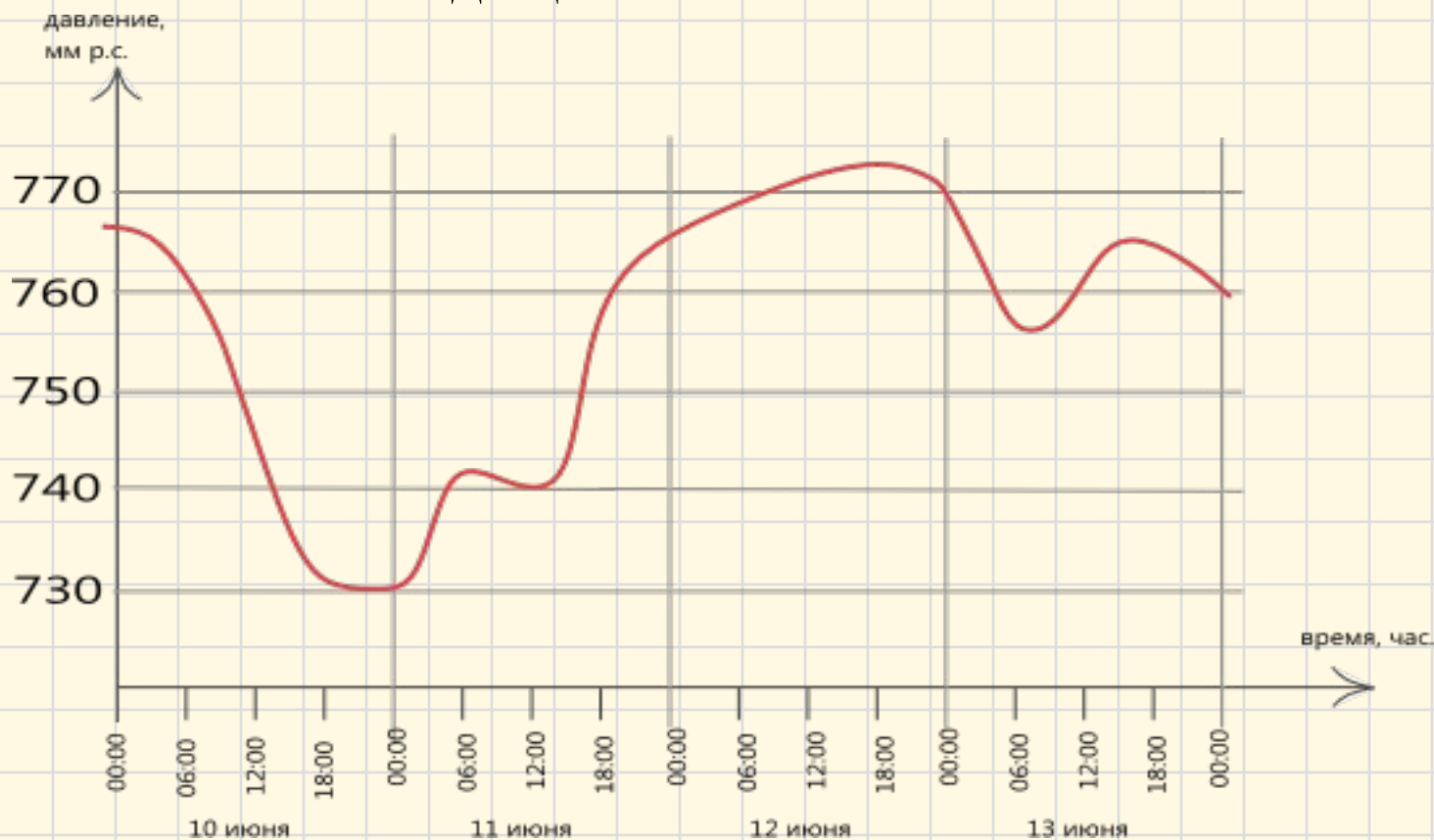
Полигон - это ломаная, полученная при соединении соответствующих точек.

- Для построения полигона отмечают в координатной плоскости точки, абсциссами которых служат моменты времени, а ординатами — соответствующие им статистические данные. Соединив последовательно эти точки отрезками, получают ломаную, которую называют полигоном.



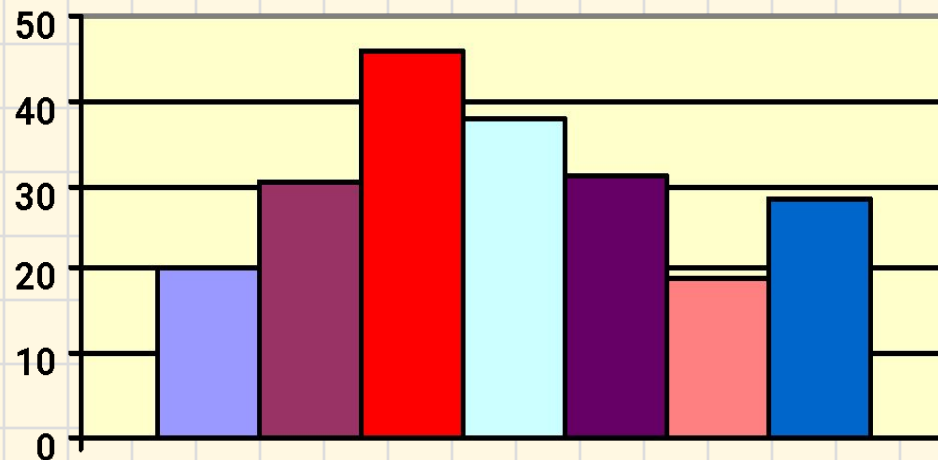
Пример

На графике изображена зависимость атмосферного давления от времени. На оси абсцисс отмечается время суток в часах, а также день и месяц. На оси ординат – атмосферное давление в миллиметрах ртутного столба. Определите по графику наименьшее давление 11 июня. Ответ запишите в виде целого числа.



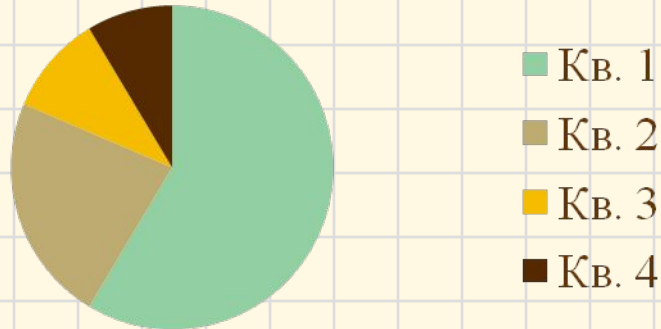
Гистограмма представляет собой ступенчатую фигуру, составленную из сомкнутых прямоугольников.

- Основание каждого прямоугольника равно длине интервала, а высота — частоте или относительной частоте.
- В гистограмме основания прямоугольников выбираются не произвольно, а строго определены длиной интервала.

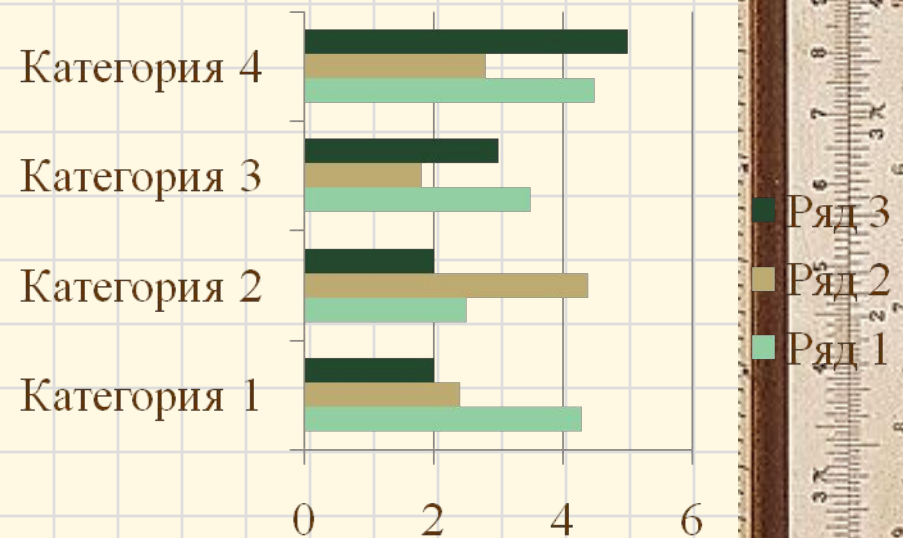
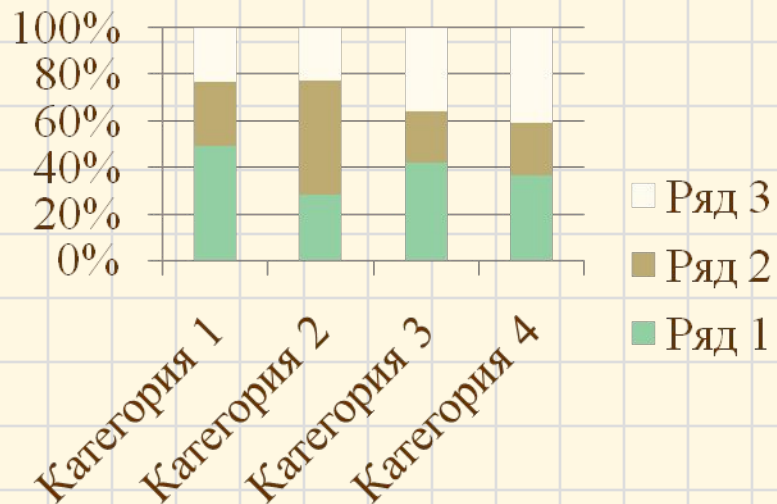
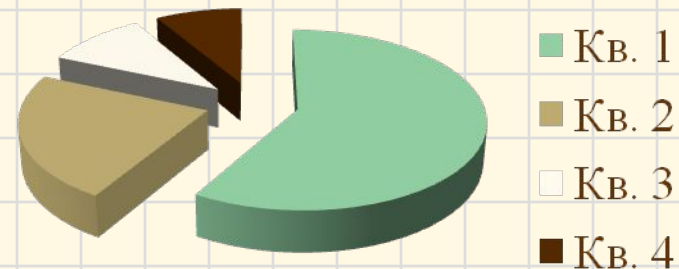


Диаграммы

Продажи



Продажи



Числовые характеристики рядов данных

- Пример.

12 девятиклассников попросили отметить в определенный день время (в минутах), затраченное на выполнение домашнего задания по математике. Получили такие данные: 23, 18, 25, 20, 25, 25, 32, 37, 34, 26, 34, 25.

Средним арифметическим ряда чисел называется частное от деления суммы этих чисел на число слагаемых.

$$\bar{x}_B = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{1}{n} \sum_{i=1}^n x_i$$

$$\bar{X} = \frac{23 + 18 + 25 + 20 + 25 + 25 + 32 + 37 + 34 + 26 + 34 + 25}{12} = \frac{324}{12} = 27$$

Размахом ряда чисел называется разность между наибольшим и наименьшим из этих чисел.

$$R = x_{\text{макс}} - x_{\text{мин}}.$$

- $R = 37 - 18 = 19$

Модой ряда чисел называется число, наиболее часто встречающееся в данном ряду.

- $M_o = 25$.
- Ряд может иметь более одной моды или не иметь моды совсем.



Медиана.

- Медианой упорядоченного ряда чисел с нечетным числом членов называется число, записанное посередине.
- Медианой упорядоченного ряда чисел с четным числом членов называется среднее арифметическое двух чисел, записанных посередине.

18, 20, 23, 25, 25, 25, 25, 26, 32, 34, 34, 37

- $Me = 25$