

**«Статистика знает всё»**

Ильф и Петров «Двенадцать стульев»

# **Начальные сведения о статистике**



# Определение статистики



**СТАТИСТИКА** (от лат. *status* - состояние) - наука, изучающая, обрабатывающая и анализирующая количественные данные о самых разнообразных массовых явлениях окружающей нас жизни.

**Статистика** изучает численность отдельных групп населения страны и ее регионов, производство и потребление разнообразных видов продукции, перевозку грузов и пассажиров различными видами транспорта, природные ресурсы и т. п.



**На основании статистических данных можно делать научно – обоснованные выводы.**

**Для этого статистические данные определенным образом должны быть систематизированы и обработаны.**

**Математическая статистика изучает математические методы систематизации, обработки и использования статистических данных для научных и производственных целей.**

## **Выборочные исследования**

*При невозможности провести сплошное исследование выполняют – выборочное:*

-из всей изучаемой совокупности данных, называемой **генеральной совокупностью**, выбирается определённая её часть, т.е. составляется **выборочная совокупность (выборка)**, которая подвергается исследованию.

Выборка должна быть **представительной (репрезентативной)** т.е. достаточной по объёму и отражающей характерные особенности исследуемой генеральной совокупности.

# содержание



1. Характеристики среднего
2. Мода набора
3. Медиана набора
4. Размах набора
5. Наглядное представление статистической информации



# Характеристики среднего



- Характеристики среднего (или средние характеристики) описывают положение всего статистического ряда на числовой прямой.
- Наиболее известной и употребляемой такой характеристикой является *среднее арифметическое* всех членов данного ряда, т.е.

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_N}{N}$$

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

# ЗАДАЧА

## [среднее арифметическое]

№1

Ученик получил в течение первой учебной четверти следующие отметки по географии: 5, 2, 4, 5, 5, 4, 4, 5, 5, 5. Найдем его **средний балл**, т.е. **среднее арифметическое** всех членов ряда:

$$\bar{x} = \frac{5+2+4+5+5+4+4+5+5+5}{10} = 4,4$$



# Мода ряда



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

- *Ряд чисел может иметь более одной моды или не иметь моды совсем.*
- Для наборов, где каждое значение встречается только один раз или одинаковое число раз (скажем, два), говорят, что мода отсутствует. Если несколько значений в наборе (но не все) встречаются с одинаковой с одинаковой наибольшей частотой, то говорят, что мода принимает несколько значений.
- Например, в наборе чисел 1, 2, 2, 4, 4, 5, 7, 7 мода принимает одновременно три значения **2, 4, 7**.
- В наборе чисел 5, 7, 1, 7, 1, 5, 5, 1, 7 мода отсутствует.



# Задача [мода ряда]

№1

- На соревнованиях по фигурному катанию судьи поставили спортсмену следующие оценки:

**5,2; 5,4; 5,5; 5,4; 5,1; 5,1; 5,4; 5,5; 5,3**

Решение:

Оценка	Встречается (раз)
<b>5,1</b>	2
<b>5,2</b>	1
<b>5,3</b>	1
<b>5,4</b>	3
<b>5,5</b>	2

Ответ: 5,4.



# Медиана набора



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

- Чтобы найти медиану числового ряда, сначала его нужно ранжировать Чтобы найти медиану числового ряда, сначала его нужно ранжировать и получить вариационный ряд.

# ЗАДАЧА

## [медиана набора]

№1

Найдите медианы наборов чисел:

- а) **686; 478; 834; 706; 843; 698; 549;**  
б) **686; 478; 834; 706; 843; 698; 549; 112.**

Ответьте на следующие вопросы.

- а) Чем отличаются наборы чисел в задании 1?  
б) Сравните получившиеся значения медиан этих двух наборов.  
в) На сколько изменилась медиана?  
г) Можно ли считать, что появление нового, относительно небольшого числа в наборе сильно изменило найденную медиану?



# ЗАДАЧА

## [медиана набора]

№1

Решение:

а) 478; 549; 686; 698; 706; 834; 843

Ответ: 698

б) 112; 478; 549; 686; 698; 706; 834; 843

$$(686 + 698) : 2 = 692$$

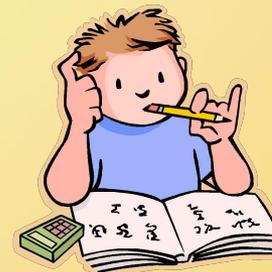
Ответ: 692

а) первый набор чисел состоит из 7 чисел, второй – из 8;

б)  $698 > 692$ ;

в)  $698 - 692 = 6$ ;

г) нет.



# Наибольшее и наименьшее значения. Размах набора



*Размах – это разность наибольшего и наименьшего значений выборки.*

Размах находят тогда, когда хотят определить, как велик разброс данных в наборе чисел .

- *Так, для температуры на Меркурии, где средняя температура, напомним, около  $+15^\circ$ , размах равен  $350^\circ - (-150^\circ) = 500^\circ$ .*

## ЗАДАЧА

[Наибольшее и наименьшее значения.  
Размах набора]

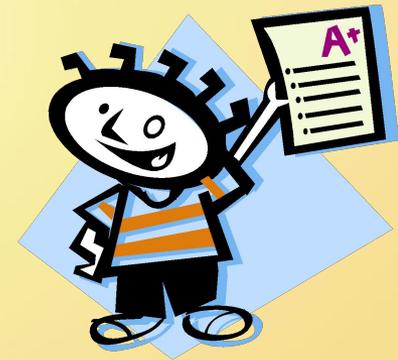
№1

Укажите наибольшее, наименьшее значения и  
размах набора чисел: **0; - 2; 14.**

$$x_{\max} = 14, x_{\min} = -2$$

$$x_{\max} - x_{\min} = 14 - (-2) = 16$$

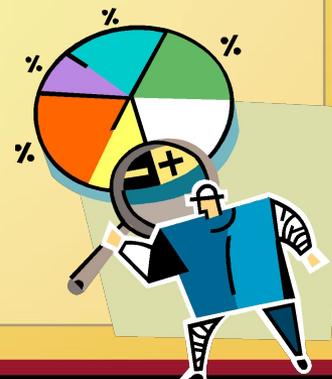
Ответ: 16.



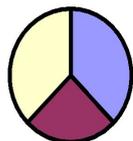
# Статистические исследования



- Для изучения различных общественных и социально-экономических явлений, а также некоторых процессов, происходящих в природе, проводятся специальные статистические исследования.
  - Для обобщения и систематизации данных, полученных в результате статистического наблюдения, их по какому-либо признаку разбивают на группы и результаты группировки сводят в таблицы.
- Для наглядного представления данных, полученных в результате статистического исследования, широко используются различные способы их изображения.

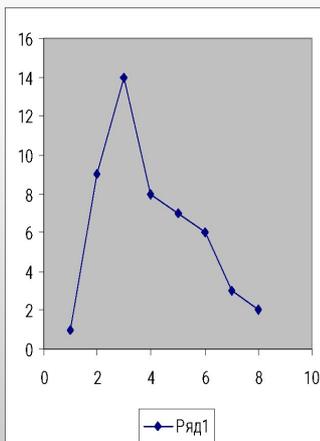


# Наглядное представление статистической информации

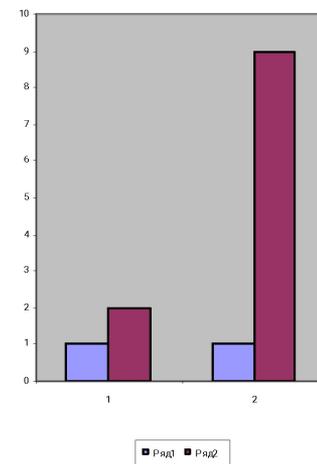


■ 1 ■ 2 ■ 3

*Круговая диаграмма*



*Полигон*



*Гистограмма*

# *Способы изображения статистического исследования.*

- ❖ **I. Столбчатые диаграммы** ( используют тогда, когда хотят проиллюстрировать динамику изменения данных во времени или распределение данных, полученных в результате статистического исследования).
- ❖ **II. Круговые диаграммы** (используют для наглядного изображения соотношения между частями исследуемой совокупности ).
- ❖ **III. Полигон** (иллюстрирует динамику изменения статистических данных во времени).
- ❖ **IV. Гистограмма** ( изображает интервальные ряды данных)

## ***Круговые диаграммы***

Если результат статистического исследования представлен в виде таблицы относительных частот, то для построения круговой диаграммы круг разбивается на секторы, центральные углы которых пропорциональны относительным частотам, определенным для каждой группы данных.

Круговая диаграмма сохраняет свою наглядность и выразительность лишь при небольшом числе частей совокупности. В противном случае её применение малоэффективно.

# *Полигон*

- Динамику изменения статистических данных во времени часто иллюстрируют с помощью полигона.
- Построение полигона:
  - отмечают в координатной плоскости точки, абсциссами которых служат моменты времени, ординатами - соответствующие им статистические данные;
  - соединив последовательно эти точки отрезками;
  - получим **ЛОМАНУЮ**, которую называют **ПОЛИГОНОМ**.

# Гистограммы

- **Гистограмма** - ступенчатая фигура, составленная из сомкнутых прямоугольников.
- **Основание** каждого **прямоугольника** равно **длине интервала**;
- **Высота** равна **частоте** или **относительной частоте**.
- В гистограмме **основание** **прямоугольников** **выбираются** не произвольно, а **строго определены** **длиной интервала**.

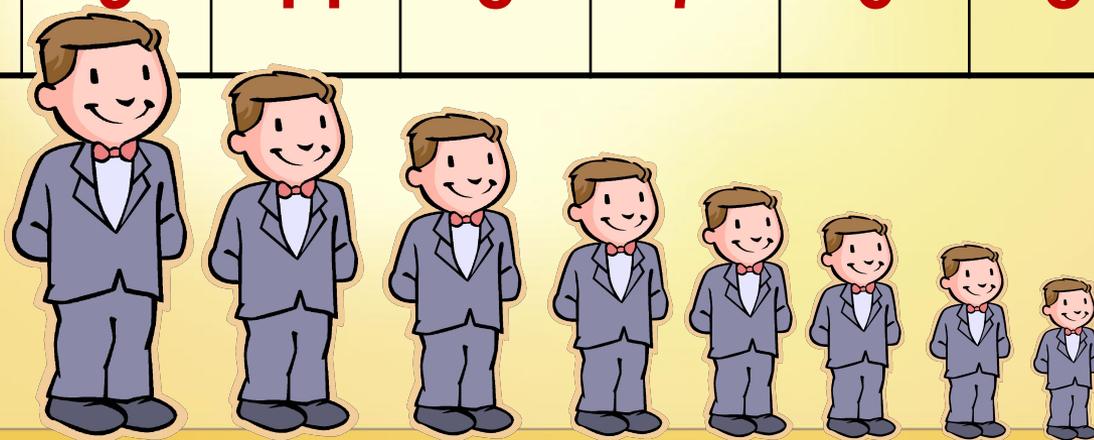
**Измерив рост 50 старшеклассников в сантиметрах, результаты записали в таблицу:**



<b>149</b>	<b>150</b>	<b>150</b>	<b>151</b>	<b>151</b>	<b>152</b>	<b>152</b>	<b>153</b>	<b>154</b>	<b>154</b>
<b>155</b>	<b>155</b>	<b>155</b>	<b>156</b>	<b>156</b>	<b>157</b>	<b>157</b>	<b>157</b>	<b>158</b>	<b>158</b>
<b>159</b>	<b>159</b>	<b>159</b>	<b>159</b>	<b>161</b>	<b>161</b>	<b>161</b>	<b>162</b>	<b>162</b>	<b>162</b>
<b>162</b>	<b>162</b>	<b>165</b>	<b>166</b>	<b>166</b>	<b>166</b>	<b>167</b>	<b>167</b>	<b>169</b>	<b>170</b>
<b>171</b>	<b>171</b>	<b>173</b>	<b>173</b>	<b>173</b>	<b>175</b>	<b>176</b>	<b>178</b>	<b>180</b>	<b>182</b>

Сгруппировав данные по классам 145-149, 150-154, ..., 180-184, представить частотное распределение учащихся по этим группам с помощью :1) таблицы;

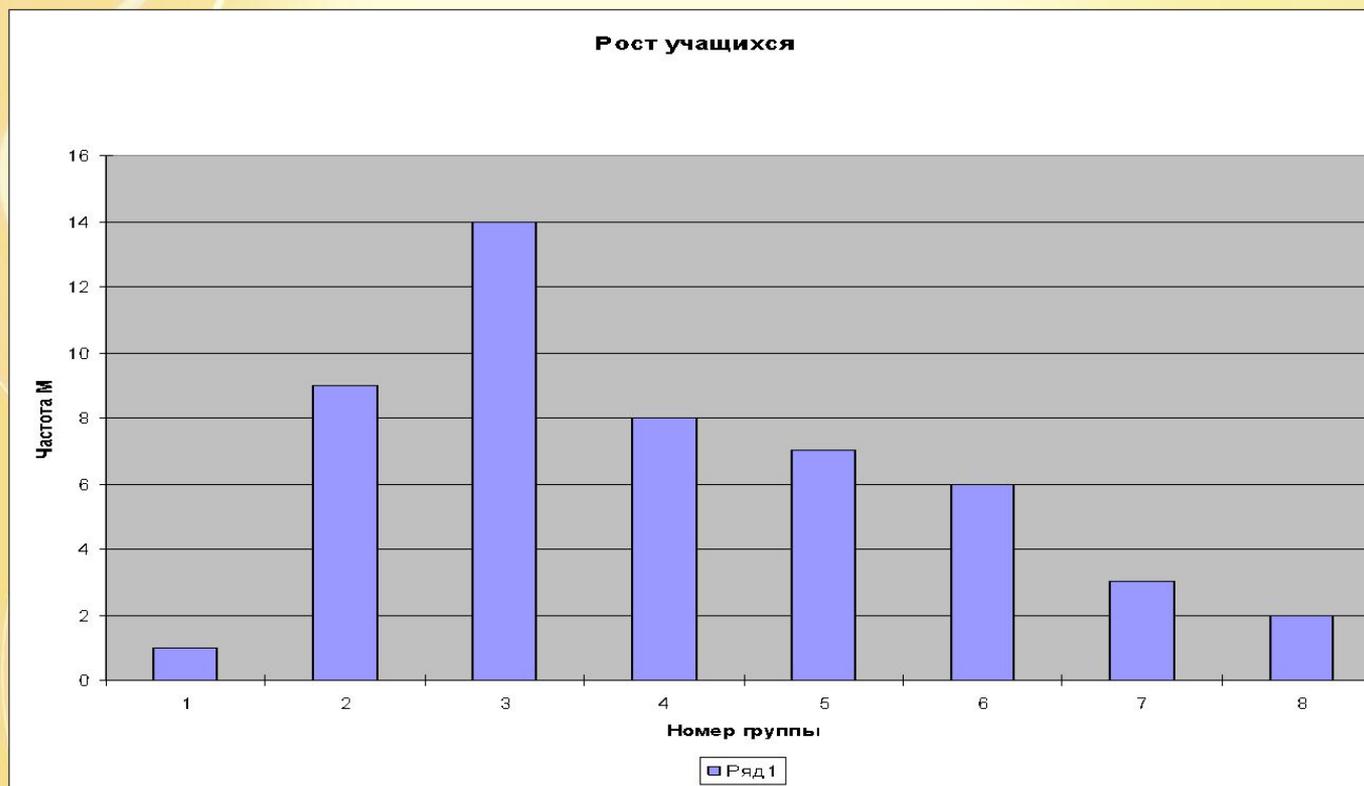
№ группы	1	2	3	4	5	6	7	8
Рост (см)	145-149	150-154	155-159	160-164	165-169	170-174	175-179	180-184
Кол-во человек	1	9	14	8	7	6	3	2



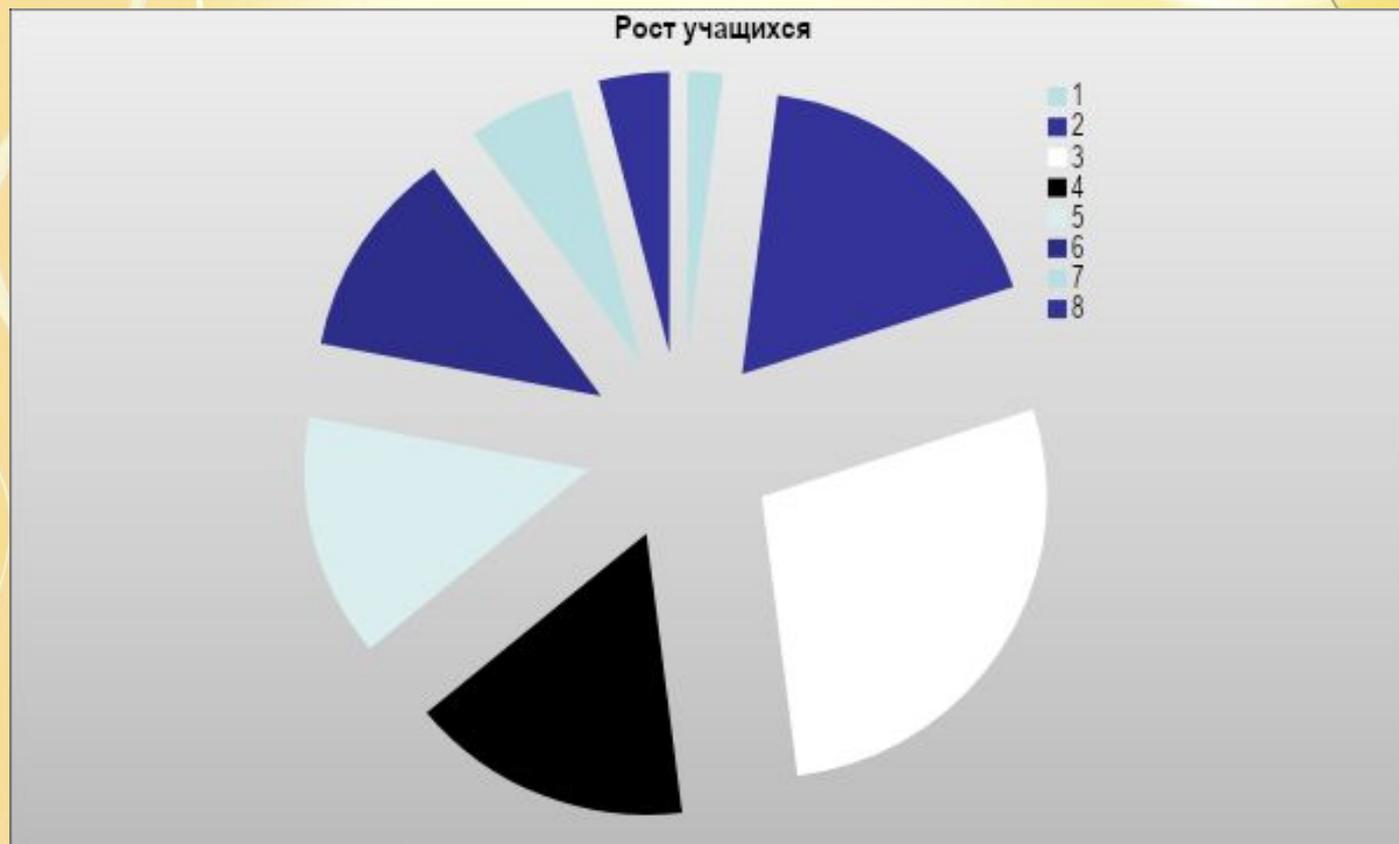
## 2) полигона частот;



### 3) столбчатой диаграммы (гистограмма)



# Круговая диаграмма

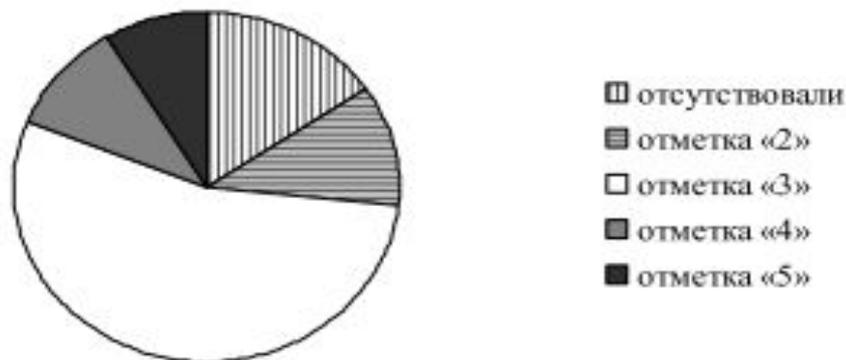


# Задача демонстрационного варианта ГИА 2017г.

18

Завуч школы подвёл итоги контрольной работы по математике в 9-х классах. Результаты представлены на круговой диаграмме.

*Результаты контрольной работы по математике.  
9 класс*



Какие из утверждений относительно результатов контрольной работы **верны**, если всего в школе 120 девятиклассников?

В ответе укажите номера верных утверждений.

- 1) Более половины учащихся получили отметку «3».
- 2) Около половины учащихся отсутствовали на контрольной работе или получили отметку «2».
- 3) Отметку «4» или «5» получила примерно шестая часть учащихся.
- 4) Отметку «3», «4» или «5» получили более 100 учащихся.

# Задача демонстрационного варианта ГИА 2017г.



- 15 На графике изображена зависимость атмосферного давления (в миллиметрах ртутного столба) от высоты над уровнем моря (в километрах). На какой высоте (в км) летит воздушный шар, если барометр, находящийся в корзине шара, показывает давление 540 миллиметров ртутного столба?



### Пример.

Для выборки определить объем, размах, найти статистический ряд и выборочное распределение:

3, 8, -1, 3, 0, 5, 3, -1, 3, 5

Объем:  $n = 10$ , размах =  $8 - (-1) = 9$

### Статистический ряд:

$x_i$	-1	0	3	5	8
$n_i$	2	1	4	2	1

### Выборочное распределение:

$x_i$	-1	0	3	5	8
$\frac{n_i}{n}$	0,2	0,1	0,4	0,2	0,1

(убеждаемся  $0,2 + 0,1 + 0,4 + 0,2 + 0,1 = 1$ )

# Выборочные характеристики

Для выборки  $x_1, x_2, \dots, x_n$  объема  $n$

**Выборочное статистическое ожидание**  
(выборочное среднее) – это среднее  
арифметическое значений выборки

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n}$$

Если выборка задана статистическим рядом, то

$$\bar{x} = \frac{n_1 x_1 + n_2 x_2 + \dots + n_n x_n}{n}$$

**Выборочная дисперсия** – это среднее арифметическое квадратов отклонений значений выборки от выборочного среднего

$$S_0 = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{n}$$

Если выборка задана статистическим рядом, то

$$S_0 = \frac{n_1(x_1 - \bar{x})^2 + n_2(x_2 - \bar{x})^2 + \dots + n_n(x_n - \bar{x})^2}{n}$$

**Несмещенная выборочная дисперсия**

$$S = \frac{n}{n-1} \cdot S_0$$

**Пример.** Для выборки найти  $\bar{x}, S_0, S$   
Выборка: 4, 5, 3, 2, 1, 2, 0, 7, 7, 3;  $n = 10$

$$\bar{x} = \frac{4 + 5 + 3 + 2 + 1 + 2 + 0 + 7 + 7 + 3}{10} = \frac{34}{10} = 3,4$$

$$S_0 = \frac{(4 - 3,4)^2 + (5 - 3,4)^2 + (3 - 3,4)^2 + (2 - 3,4)^2 + (1 - 3,4)^2 + (2 - 3,4)^2 + (0 - 3,4)^2 + (7 - 3,4)^2 + (7 - 3,4)^2 + (3 - 3,4)^2}{10} = \frac{50,4}{10} = 5,04$$

$$S = \frac{10}{9} \cdot 5,04 = \frac{50,4}{9} = 5,6$$

# Словарь



- ***Ранжирование*** – упорядочение данных, полученных в выборке;
- ***Вариационный ряд*** – упорядоченный по возрастанию статистический ряд;
- ***Выборка*** – ряд данных (чаще всего числовых), полученных в результате статистического наблюдения. Такой ряд называют ***статистическим***;

