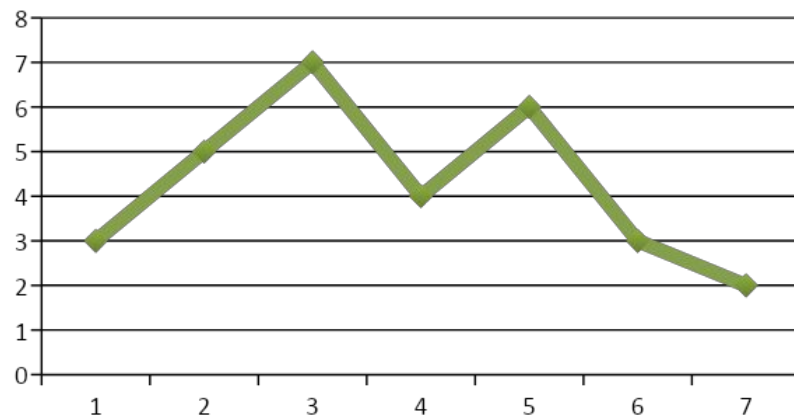


Московское суворовское военное училище

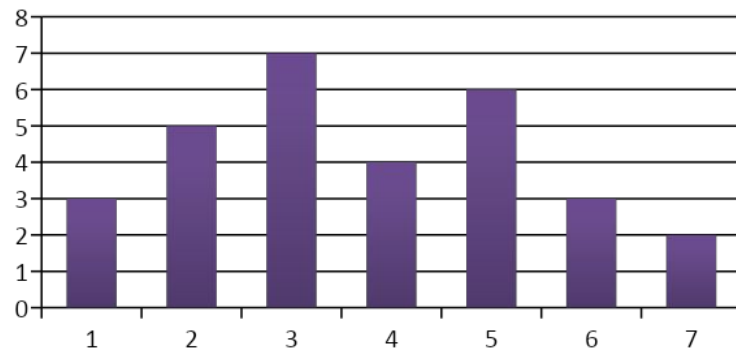
Тема
урока:



Статистическая обработка данных



Преподаватель Каримова
С.Р.



Статистика



Многие из нас участвуют в переписи населения, выборах, опросах и т. д. При этом появляется определенная информация. Задача – **ОТРАЖЕНИЕ** Этой **ИНФОРМАЦИИ** и ее **ОБРАБОТКА**. Для этого необходимо ввести некоторые статистические **ХАРАКТЕРИСТИКИ**. Рассмотрим следующий пример.



За работу на уроке пять суворовцев получили отметки:

первый – «4»;

второй – «3»;

третий – «4»;

четвертый – «5»

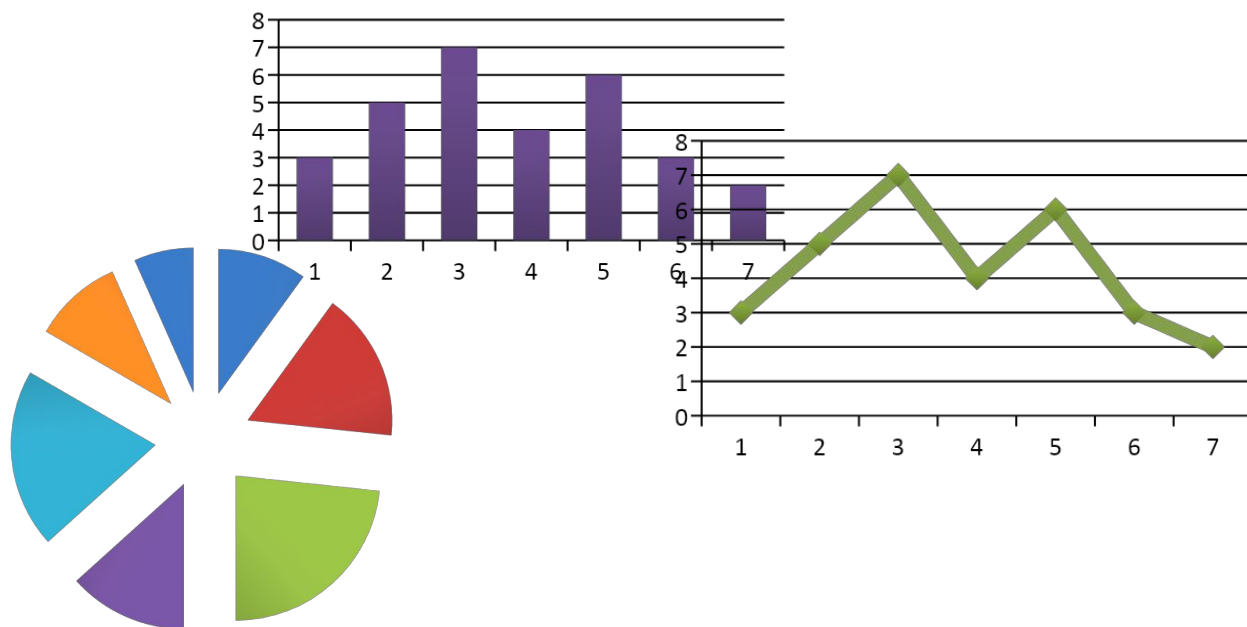
пятый – «4»

Для наглядного отражения результатов целесообразно

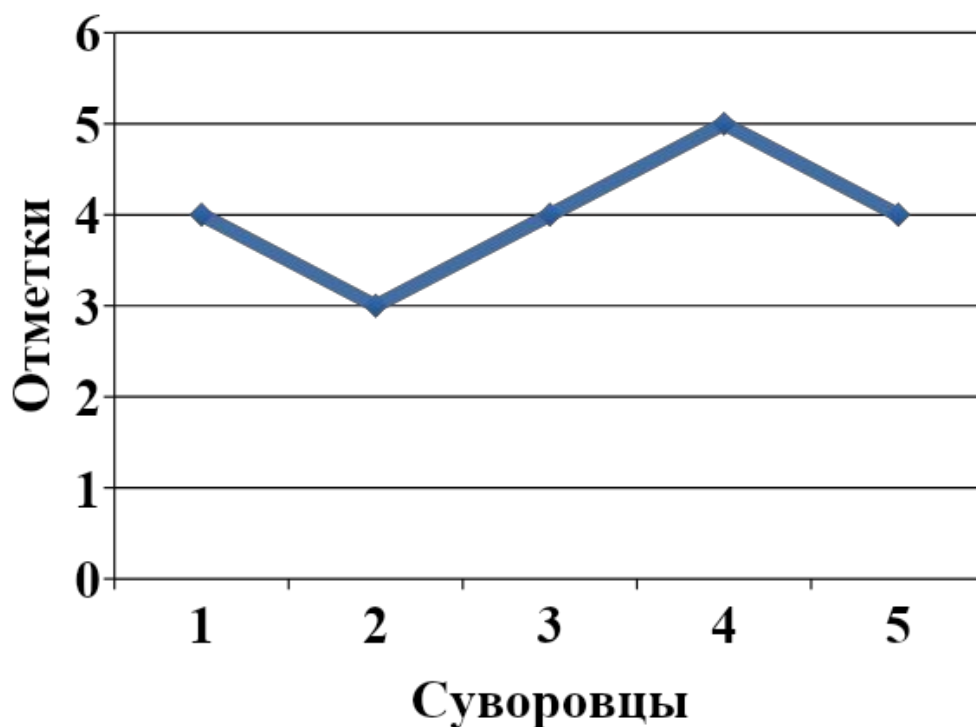
1) упорядочить и сгруппировать информацию:

Суворовец	1	2	3	4	5
Отметка	4	3	4	5	4

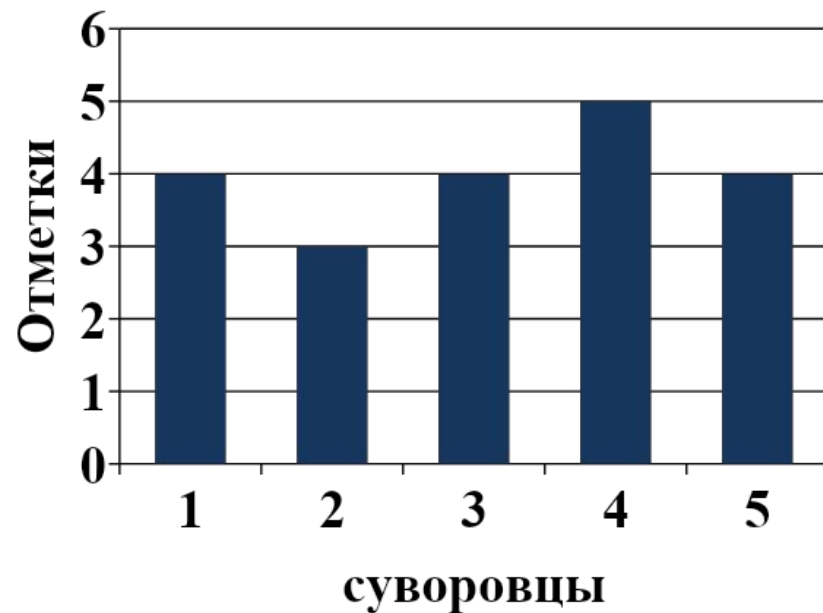
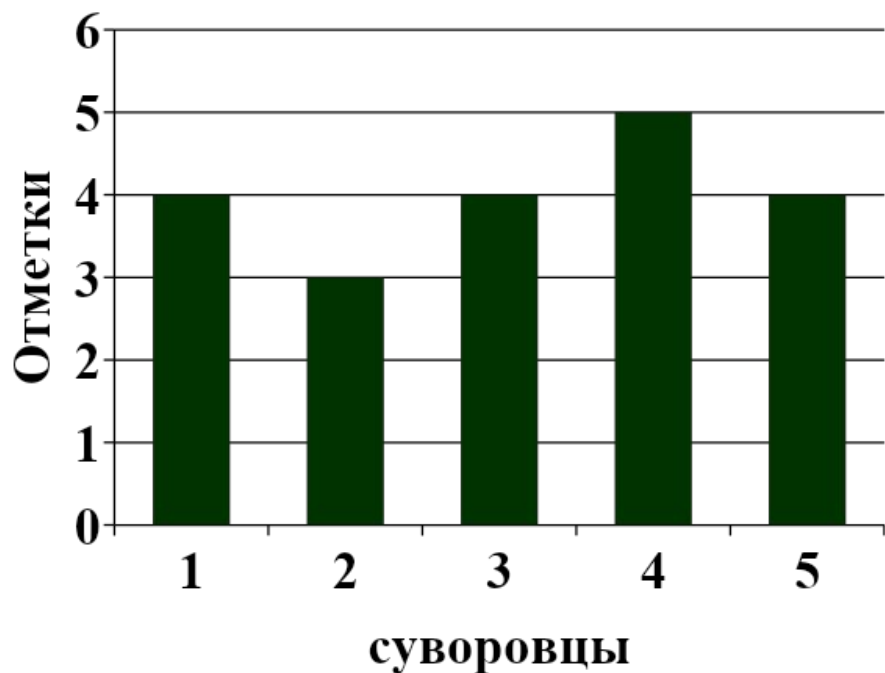
Графическая информация нагляднее табличной, поэтому для отражения результатов применяют три вида графического отражения информации – диаграммы.



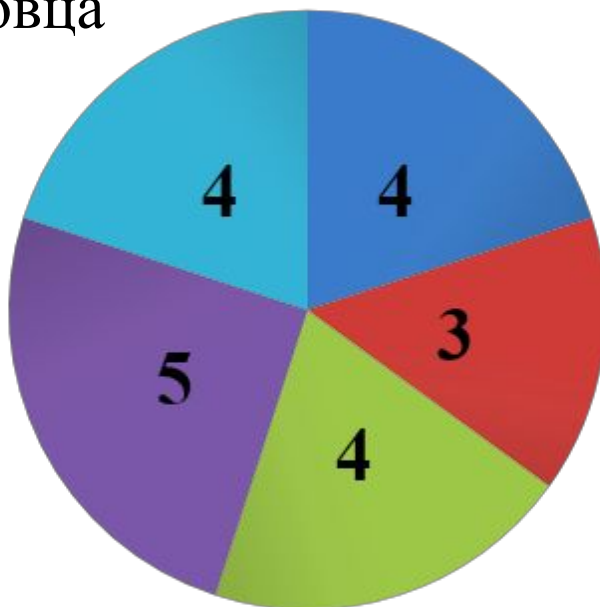
Первый вид диаграммы – **Линейная диаграмма** (или **многоугольник распределения**) Строится как обычный график. По оси абсцисс откладываются номера суворовцев, по оси ординат – отметки, т. е. точки (1;4), (2;3), (3;4), (4;5), (5;4). Для наглядности отмеченные точки соединены отрезками.



Второй вид диаграммы – **столбчатая диаграмма** (или **гистограмма распределения**) Строится следующим образом. В окрестности каждой отмеченной точки по оси абсцисс строят прямоугольник, высота которого равна соответствующей ординате. При этом обычно ширину прямоугольников делают одинаковой. Достаточно часто прямоугольники изображаются таким образом, что два соседних имеют общую сторону.



Третья диаграмма – **круговая** (или **камамбер** (по названию французского сыра)) – представляет собой круг, разделенный на **5** секторов с различными центральными углами. Сумма отметок: $4+3+4+5+4=20$, значит отметке «1» соответствует $360^\circ : 20 = 18^\circ$. Далее легко пересчитать углы секторов. Например, для первого суворовца строим сектор с углом $18^\circ * 4 = 72^\circ$. Каждый сектор маркируется номером соответствующего суворовца



Суворовец	1	2	3	4	5
Отметка	4	3	4	5	4



На практике применяют все три вида диаграмм. Итак, на конкретном примере были рассмотрены **основные этапы простейшей статистической обработки** данных:

1. Систематизация, упорядочивание и группировка.
2. Составление таблицы распределения данных.
3. Построение диаграммы распределения данных (любого вида).
4. Паспорт данных измерения (основные характеристики информации).

Обсудим некоторые **характеристики** рассматриваемого примера.

Объем измерения - количество источников. В данном случае 20.

Размах измерения - разница между наибольшим и наименьшим значениями результатов измерения. В данном случае $5-3 = 2$, т. к. наибольшая отметка – 5, наименьшая – 3.

Мода измерения - наиболее часто встречающийся результат. В данном случае 3, т. к. отметка «4» встречается наибольшее количество раз.

Среднее (или **среднее арифметическое**) - частное от деления суммы всех результатов измерения на объем измерения. Обычно его вычисляют после составления таблицы распределения.

Суворовец	1	2	3	4	5
Отметка	4	3	4	5	4

Обычно результатами измерений являются некоторые числа. **Каждое число, встретившееся в конкретном измерении**, называют **вариантой** измерения. В конкретном измерении его варианты могут быть никак не связаны (например, билетки с результатами голосования). Однако обычно результаты обрабатываются. Если записать **все варианты измерения** в некотором порядке (например, по времени поступления голосов в жюри), то получится **ряд данных измерения**.

Если записать полученные варианты в порядке их возрастания (точнее, неубывания), получим **сгруппированный ряд данных**: 1,1,..., 1; 2,2,...,2;

$$\underbrace{3,3,\dots,3}_{14}; \quad \underbrace{4,4,\dots,4}_{15}; \quad \underbrace{5,5,\dots,5}_7; \quad \underbrace{6,6,\dots,6}_4; \quad \underbrace{7,7,\dots,7}_3; \quad \overset{7}{\underbrace{8,8,\dots,8}_7}; \quad \overset{3}{\underbrace{9,9,\dots,9}_{20}};$$

Среднюю варианту в группированном ряде данных в случае нечетного количества чисел или среднее арифметическое двух стоящих посередине *вариант* в случае четного количества чисел называют **медианой измерения**.

В нашем примере суворовцу 1 соответствует отметка 4. Поэтому говорят, что **абсолютная частота** (или **кратность**) варианты 1 равна четырем. Поэтому (в другой терминологии) ранее приведенная таблица имеет вид:

Варианта	1	2	3	4	5	Сумма
Кратность	4	3	4	5	4	20

Таким образом, получаем **таблицу распределения данных** измерения. Графа «**Сумма**» добавляется для контроля: *число в этой графе обязательно равняется **объему измерения***.

Введем еще понятие **частоты данной варианты** - частное от деления кратности варианты на объем измерения.

$$\text{Частота варианты} = \frac{\text{кратность варианты}}{\text{объем измерения}}$$

$$\text{Частота варианты(в процентах)} = \frac{\text{кратность варианты}}{\text{объем измерения}} \cdot 100\%$$

III. Контрольные вопросы

1. Основные задачи статистики.
2. Виды диаграмм распределения и их построение.
3. Объем измерения.
4. Понятие размаха измерения.
5. Мода измерения.
6. Среднее арифметическое.

IV. Задание на с/п:

§ 18 (понятия, определения)

№ 18.1, 18.2