FREQUENCY TABLE

A **frequency table** is a table with two columns.
One column lists _____, and another for the _____

(how many items fit into each category).

Example
A local movie theater chain wants to understand which genres of movies are most popular among their customers to guide future movie screenings. They collect data on ticket sales for different genres over the past three months. The data is summarized in the following frequency table:

Genre	Frequency (Tickets Sold)
Action	102
Comedy	85
Drama	76
Horror	58
Sci-Fi	64
Romance	43

Example
A hospital is analyzing the ages of patients admitted for routine check-ups in the past month to better understand the age demographics of their visitors.

24 26 29 31 33 36
36 38 38 42 45 47
48 50 52 54 56 59
60 61 63 65 65 67
70 73 75 77 79 82

Age Range	Frequency
20-29	
30-39	
40-49	
50-59	
60-69	
70-79	
80-89	

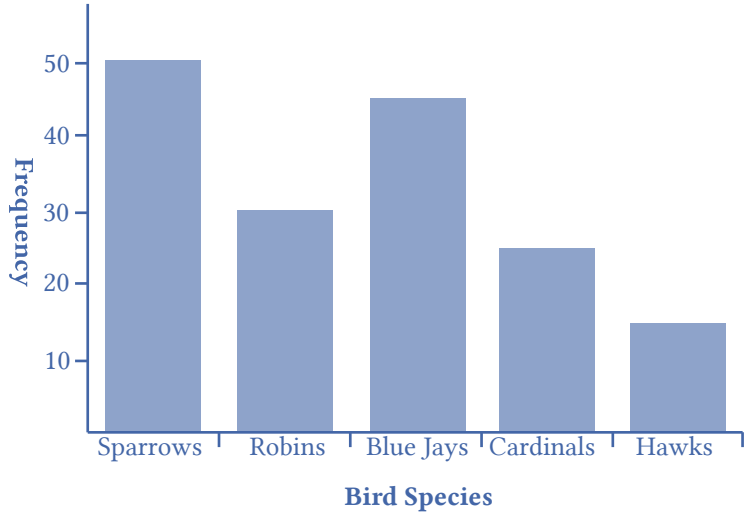
BAR GRAPH

A _____ is a graph that displays a bar for each category with the length of each bar indicating the frequency of that category.

Example

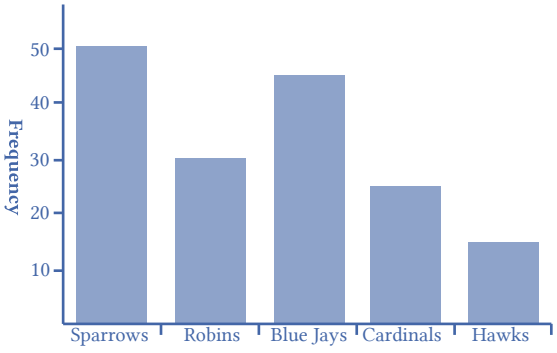
A local wildlife reserve is studying the behavior of different bird species visiting their park. Over the course of a month, park rangers record the number of sightings for five common bird species. They want to use this data to decide which areas of the park to focus conservation efforts on, based on which species are most commonly seen. The data collected is shown below:

Bird Species	Frequency
Sparrows	50
Robins	30
Blue Jays	45
Cardinals	25
Hawks	15

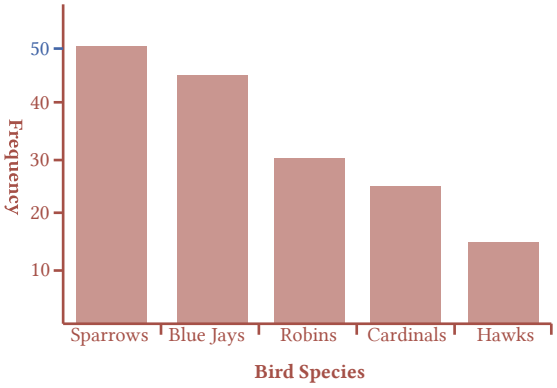


PARETO CHART

A **Pareto chart** is a bar graph ordered from



Pareto chart

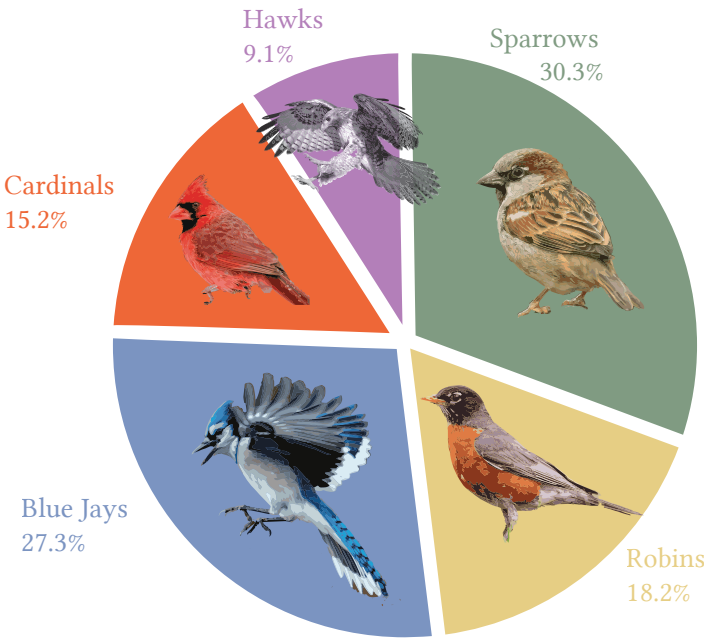


PIE CHART

A _____ is a circle with wedges cut of varying sizes marked out like slices of pie or pizza.

The relative sizes of the wedges correspond to the relative frequencies of the categories.

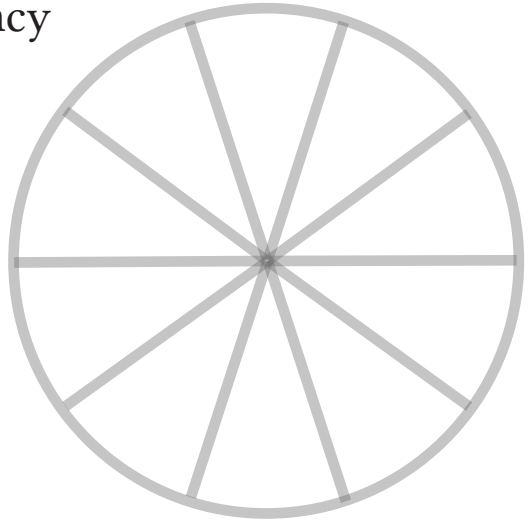
Bird Species	Frequency
Sparrows	50
Robins	30
Blue Jays	45
Cardinals	25
Hawks	15



Question

Create a pie chart for the following frequency table:

Workshop Type	Frequency
Painting	3
Sculpture	1
Music	2
Dance	4



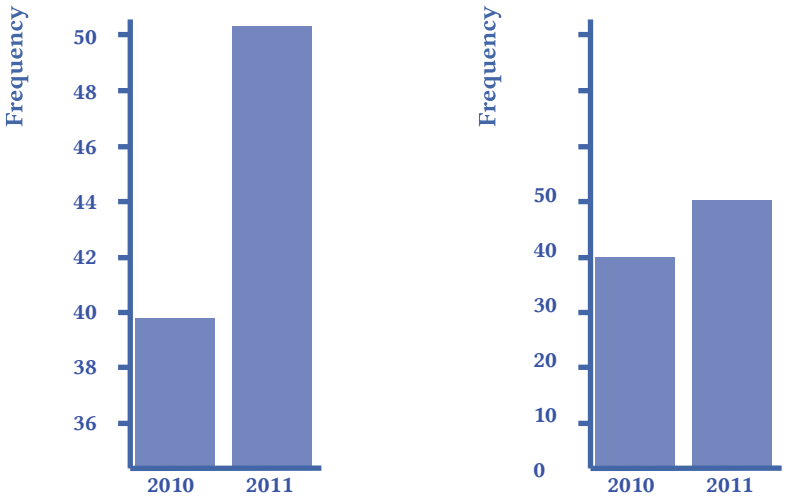
PICTOGRAM

A **pictogram** is a statistical graphic in which

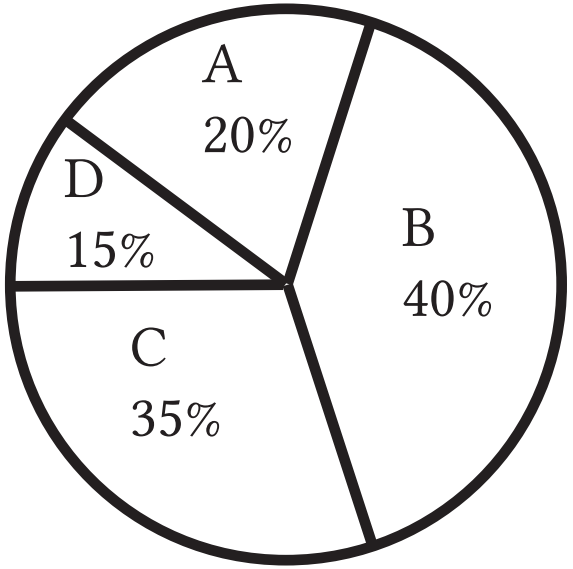
Beware:

It is easy to be misled with charts

Can you compare the following bar charts? What is misleading here?

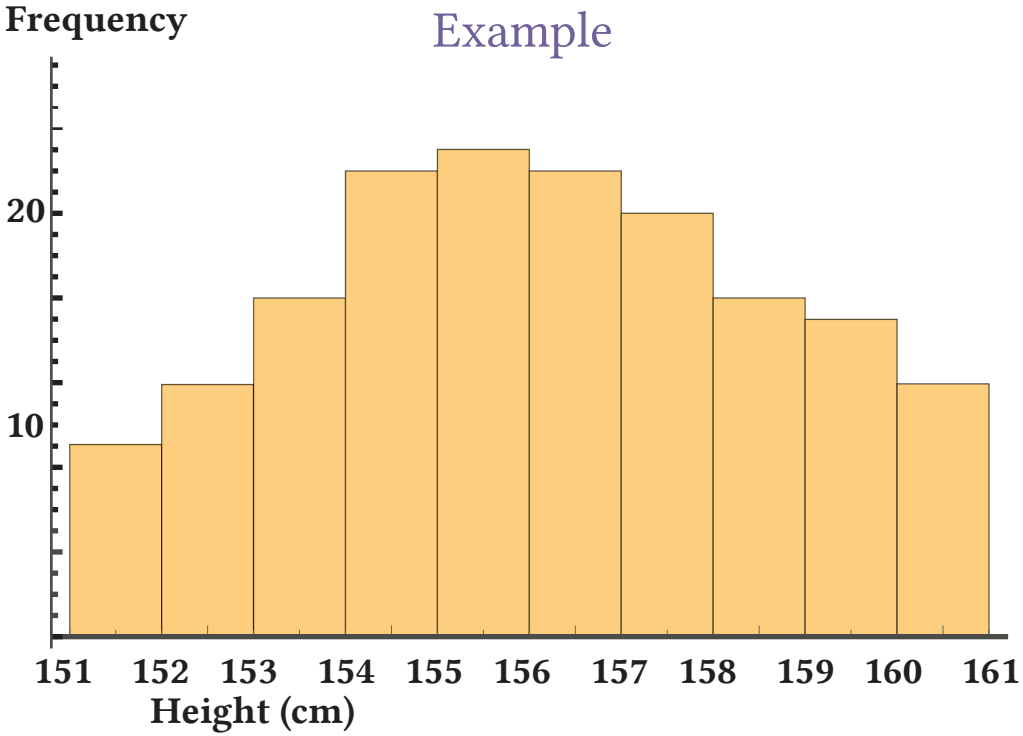


What's wrong with this pie chart?



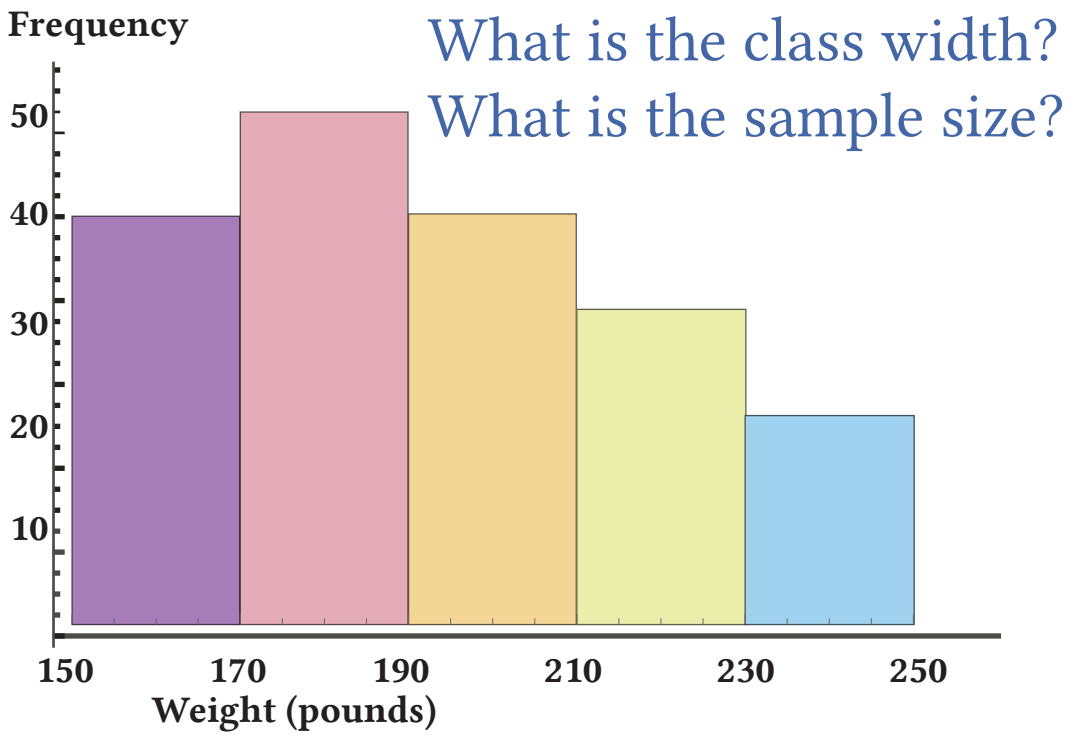
HISTOGRAM

A **histogram** is like a bar graph, but where the



CLASS INTERVALS

Class intervals are _____.
In general, we define class intervals so that each interval is _____.
For example, if the first class contains values from 120-129, the second class should include values from 130-139.
We have somewhere between 5 and 20 classes, typically, depending upon the number of data we're working with.

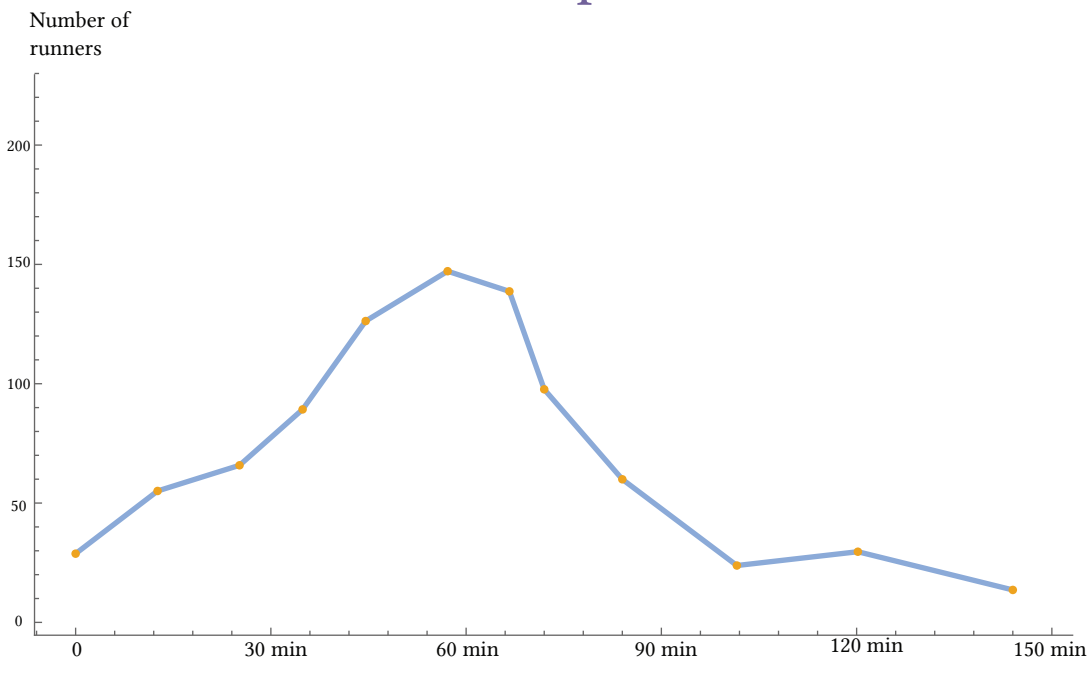


What is the class width?
What is the sample size?

FREQUENCY POLYGON

An alternative representation is a **frequency polygon**. A frequency polygon starts out like a histogram, but instead of drawing a bar, a point is placed in the midpoint of each interval at height equal to the frequency. Typically the points are connected with straight lines to emphasize the distribution of the data.

Example



Question

A tourism agency conducted a survey to analyze the ages of participants in a group travel program. The ages of 30 participants are listed below:

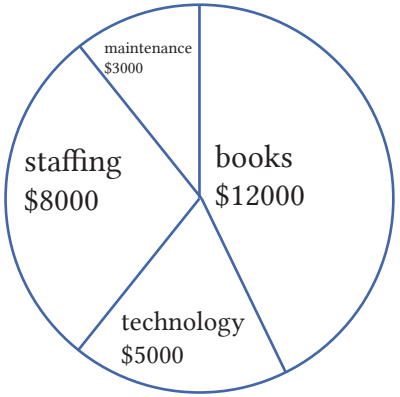
20, 21, 22, 23, 25, 27, 28, 29, 31, 32, 34, 35, 37, 39, 41, 43, 45, 46, 48, 50, 52, 53, 54, 56, 58, 60, 62, 64, 66, 68

Complete the frequency distribution for the data.

Age	Frequency
20-29	
30-39	
40-49	
50-59	
60-69	
70-79	

Question

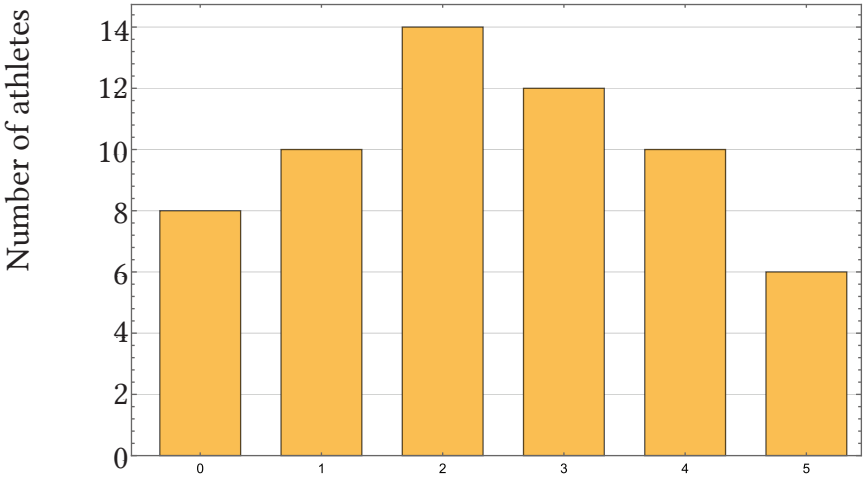
A local library tracks its annual spending in different areas: Books, Technology, Staffing, and Maintenance. For this year, the spending in each category is as follows:



Calculate the percentage of the total budget that was spent on Books.

Question

The data in the figure below represents the number of medals won by 150 athletes in a sports competition.



How many athletes won exactly 2 medals?

MEAN

The **mean** of a set of data is the

divided by the

number of values.

Example

During the track season, Marci ran four races, with times of 12.8 seconds, 13.4 seconds, 12.6 seconds, and 13.1 seconds. What is the mean of her race times?

Question

Marci baked four batches of cookies, with the following quantities: 24, 30, 27, and 33 cookies. What is the mean number of cookies per batch?

Question

A survey was conducted in a town where 80 households were asked about their monthly grocery expenses, rounded to the nearest $10. The results are summarized in the table below.

Grocery Expenses (dollars)	Frequency
150	5
160	10
170	14
180	18
190	16
200	9
210	5
220	3

What is the mean average monthly grocery expense for households in this town?

MEDIAN

The **median** of a set of data is the value

To find the median, begin by listing the data in order from

If the number of data values, N, is odd, then the median is the middle data value. This value can be found by rounding N/2 up to the next whole number.

Example

Steve has the following list of his daily step counts for the last 30 days and wants to calculate the median:

4,500 4,600 4,800 5,000 5,200 5,500 5,700 5,800 6,000 6,200 6,500 6,500 6,800 7,000 7,100
7,200 7,400 7,600 7,800 8,000 8,200 8,300 8,500 8,600 8,800 9,000 9,200 9,400 9,600 10,000

What is the median step count value?

Question

A group of employees tracked the number of emails they sent in a day, listed below:

5, 13, 7, 8, 17, 10, 11, 12, 19, 14, 15, 16, 9, 18, 6

What is the median number of emails sent?

Example

A city conducted a survey on the number of hours people spent volunteering in a month. The results are summarized in the table below.

Hours Volunteered	Frequency
5	8
10	12
15	15
20	18
25	14
30	10
35	7
40	6

What is the mean number of hours spent volunteering?

MODE

The **mode** is the element of the data set that occurs _____

It is possible for a data set to have more than one mode if several categories have the same frequency, or no modes if every category occurs only once.

Example

In a survey of favorite fruits among a group of friends, the following data was collected:

Fruit	Frequency
Apple	4
Banana	6
Cherry	2
Date	3
Elderberry	5

Which fruit is the mode?

Question

In a survey of customer satisfaction with a new service, the following data was collected on the number of issues reported:

Number of Issues	Frequency
0	12
1	15
2	8
3	5
4	3

Find:
The average number of issues reported
The median number of issues reported
The mode of the number of issues

RANGE

The **range** is the difference between the _____ and the _____ of the data set.

Example

In a study of the number of hours worked per week by employees in a department, the following data was collected:

Hours Worked per Week: 38, 42, 45, 37, 50, 41, 39, 48

What is the range of this data?

STANDARD DEVIATION

The **standard deviation** is a measure of variation based on measuring how far each data value deviates, or is different,

A few important characteristics:

Standard deviation is always

Standard deviation will be zero if all the data values are equal, and will get larger as the data spreads out.

Standard deviation has the same units as the original data.

STANDARD DEVIATION

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2}$$

TO COMPUTE THE STANDARD DEVIATION

1. Calculate the mean

2. Find the deviation of each data from the mean. In other words, subtract the mean from the data value.

3. Square each deviation.

4. Add the squared deviations.

5. Divide by n, the number of data values, if the data represents a whole population; **divide by n – 1 if the data is from a sample**.

6. Compute the square root of the result.

Example

The temperatures recorded in five different cities for a particular day were 72°F, 68°F, 75°F, 70°F, and 80°F. Find the standard deviation of the temperatures.

Question

The number of miles run each day for a week by a runner were: 4, 6, 5, 7, 8, 6, and 5. Calculate the standard deviation of the number of miles run.

Question

The number of items sold by a small shop over seven days were: 20, 25, 22, 30, 27, 24, and 26. Calculate the standard deviation of the number of items sold.

Question

The number of attendees at five different workshops were: 45, 50, 55, 60, and 52. Calculate the standard deviation of the number of attendees.

Question

The number of calls received by a customer service center each hour over an 8-hour shift were: 32, 27, 35, 30, 28, 31, 29, and 34. Calculate the standard deviation of the number of calls received.

Question

Consider the data set

9 6 2 4 3 8

Find the average (mean):

Find the median:

Question

A survey was conducted to determine the number of hours spent on hobbies per week by a group of adults. The results were:

Hours spent	0	1	2	3	4	5
Frequency	2	8	5	1	9	3

Calculate:

The average (mean) number of hours spent on hobbies per week

The median number of hours spent on hobbies per week

Two sports teams were evaluated in a series of practice drills:

Team X had a mean score of 75 and a standard deviation of 5
Team Y had a mean score of 73 and a standard deviation of 7

Which team performed better on average?

Which team had more consistent scores?

Question

A sample was done, collecting the data below. Calculate the standard deviation, to one decimal place.

x
6
14
17
4
20

Question

Calculate the average (mean) of the data shown, to two decimal places:

1, 4, 7, 2, 0, 4, 8, 6

QUARTILES

Quartiles are values that divide the data in

The first quartile (Q1) is the value so that

the third quartile ( ) is the value so that *75% of the data values are below it.* You may have guessed that the second quartile is the same as the median, since the median is the value so that 50% of the data values are below it.

This divides the data into quarters; 25% of the data is between the minimum and Q1, 25% is between Q1 and the median, 25% is between the median and Q3, and 25% is between Q3 and the maximum value.

FIVE NUMBER SUMMARY

The five number summary takes this form:

, , , ,

TO FIND THE FIRST QUARTILE, Q1

Begin by ordering the data from smallest to largest
Compute the locator: L =
If L is a decimal value:
Round up to L+
Use the data value in the L+th position
If L is a :
Find the mean of the data values in the Lth and L+1th positions.

TO FIND THE THIRD QUARTILE, Q3

Use the same procedure as for Q1, but with locator: L = 0.75n

Example

Suppose a group of 10 athletes have their running speeds (in meters per second) recorded, and their speeds sorted from slowest to fastest are:

5.8, 6.2, 6.4, 6.6, 6.8, 7.0, 7.2, 7.4, 7.6, 7.8

What are the first and third quartiles of their running speeds?

Q1: 0.25 x 10 = 2.5
Rounding up, we get 3, the third position. So Q1 = 6.4 m/s

Q3: 0.75x10 = 7.5
Rounding up, we get 8, the eighth position. So Q3 = 7.4 m/s

Example

Suppose a group of 10 athletes have their running speeds (in meters per second) recorded, and their speeds sorted from slowest to fastest are:

5.8, 6.2, 6.4, 6.6, 6.8, 7.0, 7.2, 7.4, 7.6, 7.8

What is the 5 number summary?

Median: mean of 5th and 6th position

$$\text{Median} = \frac{6.8 + 7.0}{2} = 6.9$$

Min = 5.8 , Max = 7.8
Five number summary: 5.8, 6.4, 6.9, 7.4, 7.8

Question

The monthly rent paid by 24 individuals was recorded, and the amounts sorted from lowest to highest are:

$800, $850, $850, $900, $950, $950, $975, $1000, $1025, $1050, $1075, $1100, $1125, $1150, $1200, $1250, $1300, $1300, $1350, $1400, $1450, $1500, $1550, $1600

Find the 5-number summary of this data.

Example

Using the data on weekly working hours from a group of employees, create the five-number summary.

Weekly Working Hours (hours)	Frequency
30	4
35	6
40	10
45	15
50	18
55	12
60	8
65	5

The total amount of employees is the total sum of the frequencies:
4 + 6 + 10 + 15 + 12 + 8 + 5 = 78
Min = 30, Max = 65
Q1: 78/4 = 19.5, rounding up we get 20. 40 hours is in the 20th position and so Q1 = 40.
Median: 78/2 = 39. The mean of the 39th and 40th position is 50.
Q3: 3x78/4 = 58.5, rounding up we get 59. 55 hours is in the 59th position and so Q3 = 55.
5 number summary: 30, 40, 50, 55, 65

Question

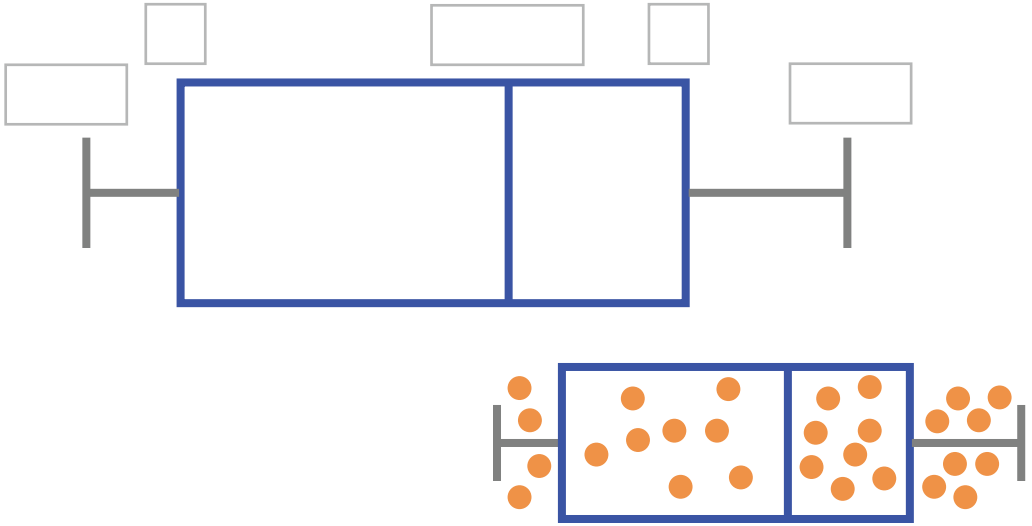
The daily caloric intake (in kilocalories) of 60 individuals was recorded, and their intake data is grouped as follows:

Calories (kcal)	Frequency
1800	5
2000	8
2200	12
2400	15
2600	10
2800	6
3000	4

Using this data, calculate the five-number summary.

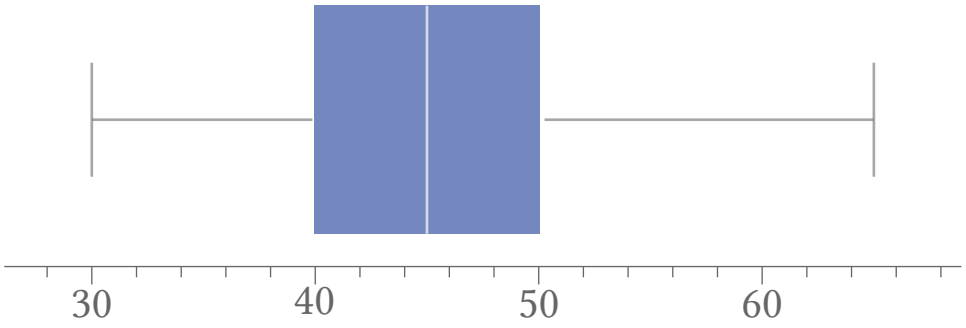
BOX PLOT

A _____ is a graphical representation of a five-number summary.



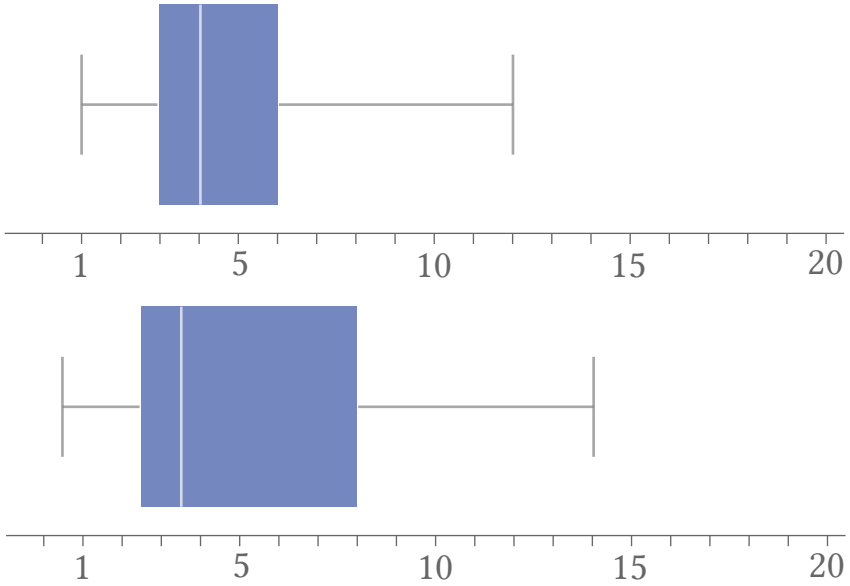
Example

five-number summary:
30, 40, 45, 50, 65



Example

The box plot of delivery times in days for two retailers is shown below:



Which retailer would you choose for a faster delivery?

Bank of Questions Chapter 11

1. The following data represents the daily temperature (in °C) recorded over 30 days in a coastal city:

   ```
   12 14 15 15 16 17
   17 18 18 19 20 20
   21 21 22 23 24 25
   25 26 26 27 28 29
   30 31 32 33 34 35
   ```

   Complete the frequency distribution for the data.

Temperature (°C)	**Frequency**
10-14	
15-19	
20-24	
25-29	
30-35	

2. The following data represents the monthly electric bills (in dollars) for 30 households:

   45, 50, 52, 55, 60, 63, 65, 68, 70, 72, 75, 77, 80, 82, 85, 88, 90, 92, 95, 98, 100, 105, 108, 110, 115, 118, 120, 125, 130, 135.

   Complete the frequency distribution for the data.

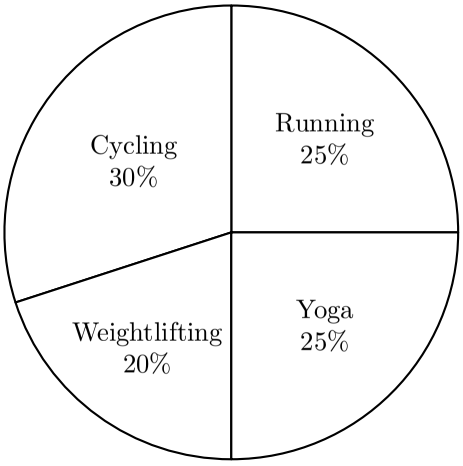
Electric Bill ($)	**Frequency**
40-59	
60-79	
80-99	
100-119	
120-139	

3. The following data represents the time (in minutes) spent commuting to work by 30 employees:

   5, 10, 12, 15, 18, 20, 20, 22, 25, 28, 30, 32, 35, 37, 40, 42, 45, 48, 50, 52, 55, 58, 60, 62, 65, 68, 70, 72, 75, 80.

   Complete the frequency distribution for the data.

Commute Time (minutes)	**Frequency**
0-19	
20-39	
40-59	
60-79	
80-99	



If James exercised for a total of 14 hours this week, how many hours did he spend on Cycling?

_____ hours

7. The following data represents the number of hours spent studying by 30 students during a week:

```
1 2 2 3 3 4 5
5 6 7 8 8 9 10
10 10 11 12 13
13 14 15 16 17
18 19 20 21 22
```

Complete the frequency distribution for the data.

Value Range	**Frequency**
1 - 5	
6 - 10	
11 - 15	
16 - 20	
21 - 25	

8. The following data represents the scores of 30 students in a mathematics exam:

55	62	62	65	67	68
70	73	75	76	78	80
82	84	85	87	88	90
91	92	93	94	95	96
98	99	100			

Complete the frequency distribution for the data.

Score Range	**Frequency**
50-59	
60-69	
70-79	
80-89	
90-99	
100-109	

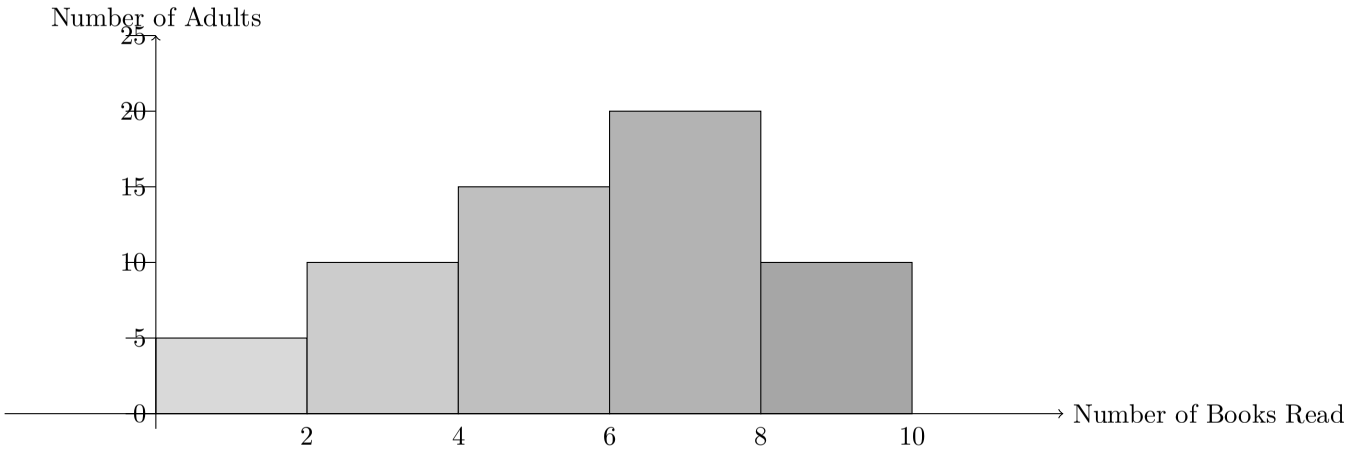
9. The following data represents the ages of 30 participants in a local community fitness program:

22	24	25	26	27	28
29	30	31	32	33	34
35	36	37	38	39	40
41	42	43	44	45	46
47	48	49	50	51	52

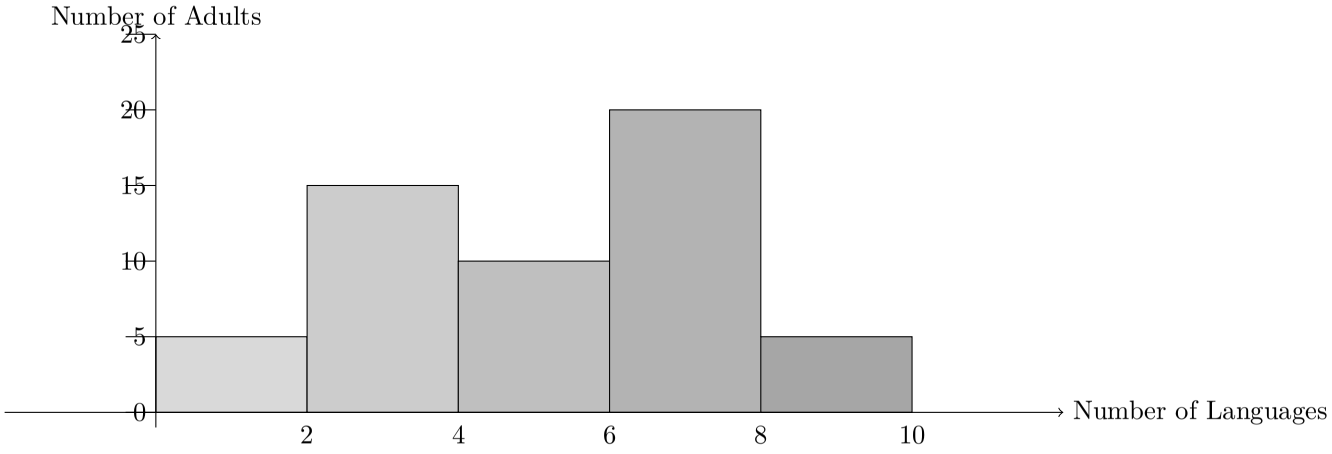
Complete the frequency distribution for the data.

Age Range	**Frequency**
20-29	
30-39	
40-49	
50-59	

12. A survey was conducted to find out how many books adults read in a month. The bar graph below shows the number of adults who indicated each number of books read.



13. A survey was conducted to find out how many languages adults can speak. The histogram below shows the number of adults who indicated each number of languages spoken.



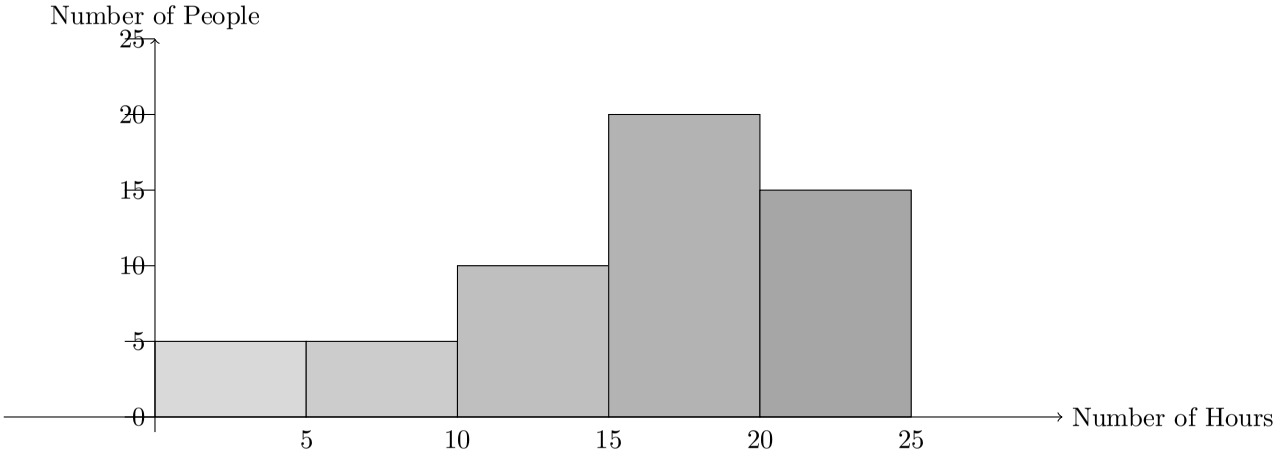
Based on the histogram above, what is the class width?

Class width = _____

What is the sample size?

Sample size = _____

14. A survey was conducted to find out how many hours people spend watching television each week. The histogram below shows the number of individuals who indicated each range of hours spent watching TV.



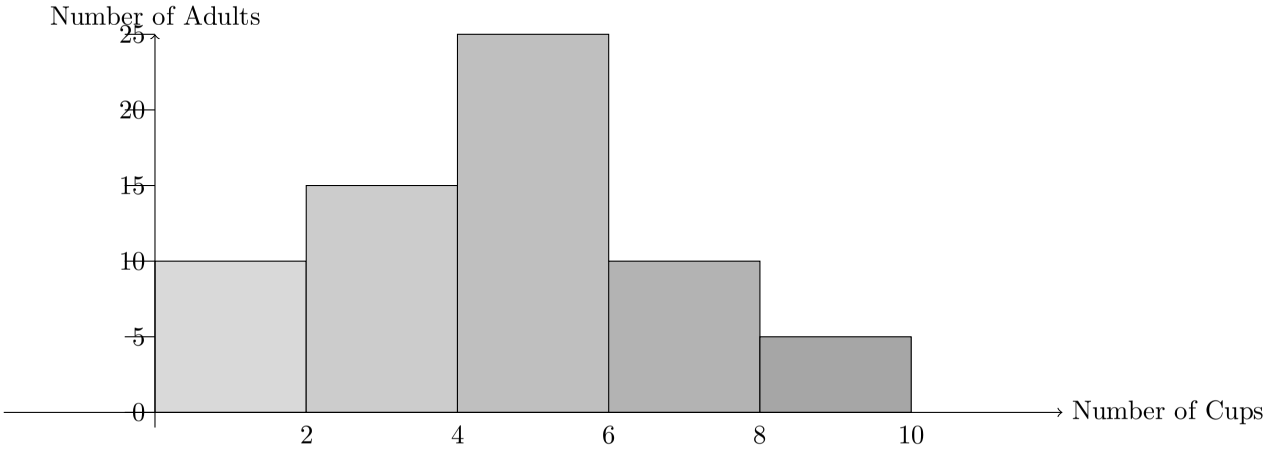
Based on the histogram above, what is the class width?

Class width = _____

What is the sample size?

Sample size = _____

15. A survey was conducted to find out how many cups of coffee adults drink in a week. The bar graph below shows the number of individuals who indicated the number of cups they consume weekly.



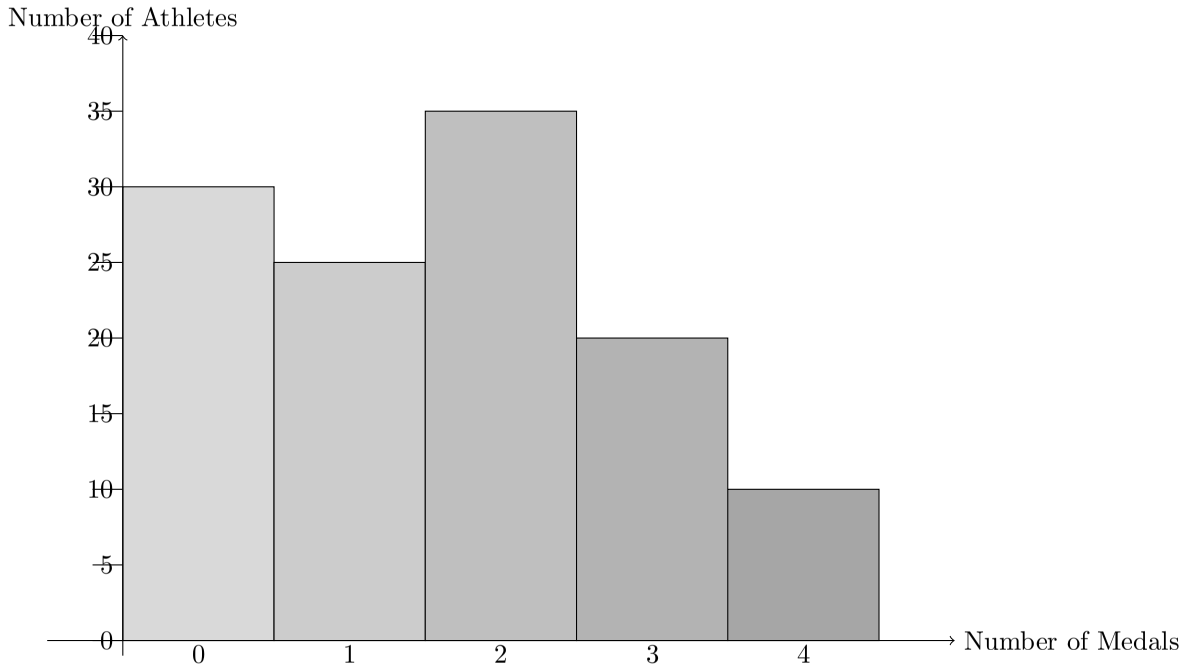
Based on the histogram above, what is the class width?

Class width = _____

What is the sample size?

Sample size = _____

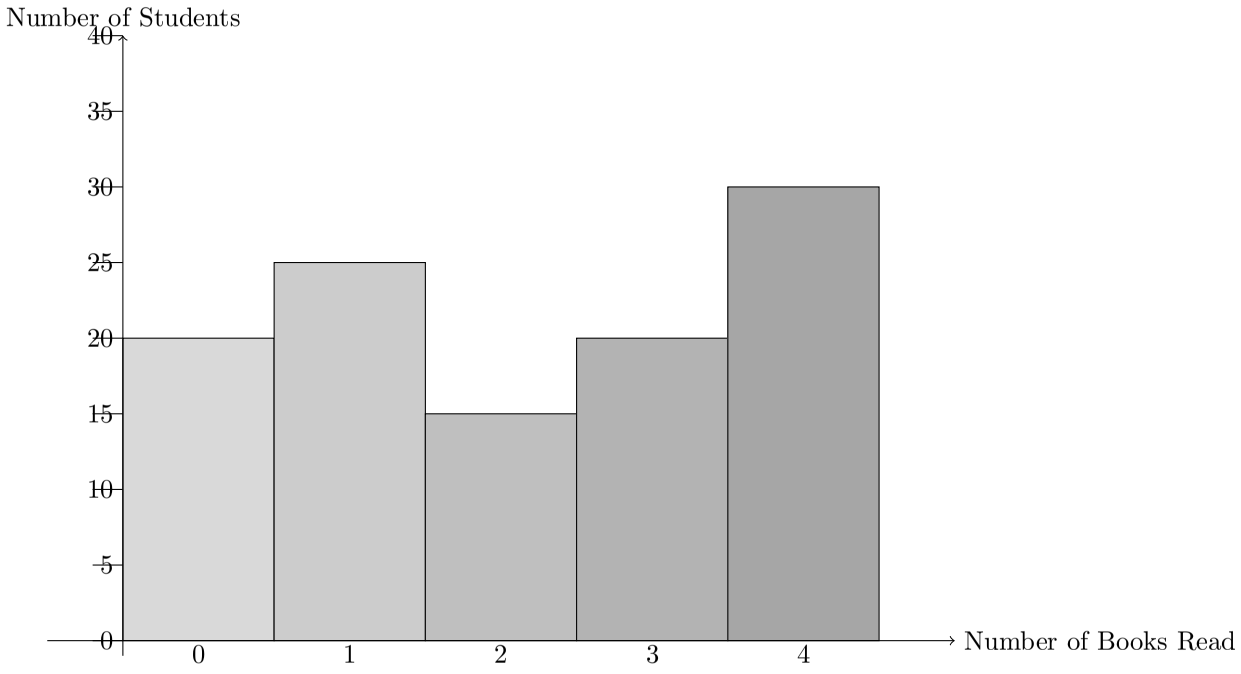
16. The data in the figure below represents the number of medals earned by 120 athletes in a sports competition.



How many athletes earned more than 2 medals?

Number of athletes = _____

17. The data in the figure below represents the number of books read by 100 students over the summer.



Question: How many students read an even number of books?

Number of students = _____

18. A survey was conducted on the monthly rent paid by 80 tenants in an apartment complex, rounded to the nearest $100. The results are summarized in a frequency table below.

Rent (dollars)	**Frequency**
600	5
700	10
800	12
900	16
1000	18
1100	11
1200	6
1300	2

Question: What is the mean (average) rent paid by tenants in this apartment complex?

Mean rent = _____

What is the median rent paid by tenants in this apartment complex?

Median rent = _____

19. A car dealership collected data on the number of cars sold by each of its 85 salespeople in the past month. The results are summarized in a frequency table below.

Cars Sold	**Frequency**
1	6
2	9
3	12
4	15
5	18
6	14
7	7
8	4

What is the mean average number of cars sold per salesperson last month?

Mean cars sold = _____

20. The following data represents the ages of participants in a marathon race:

    12, 15, 18, 21, 25, 29, 33, 37, 41, 44, 50, 55, 61, 67, 72, 75, 81, 85.

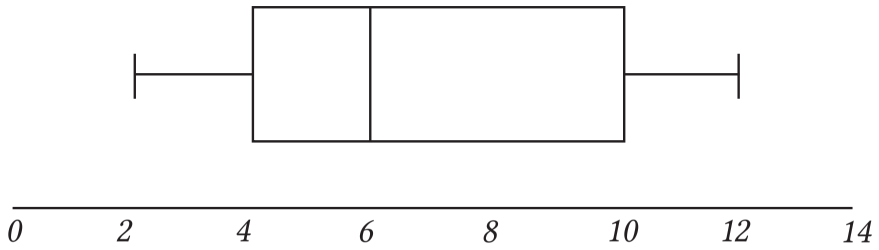
    Find the 5-number summary for the data shown.

21. The following data represents the monthly rainfall (in millimeters) recorded in a city over the past year. Find the 5-number summary for the data shown:

    22 28 35 40 45 49 52 57 61 66 70 75 81 85 92 98

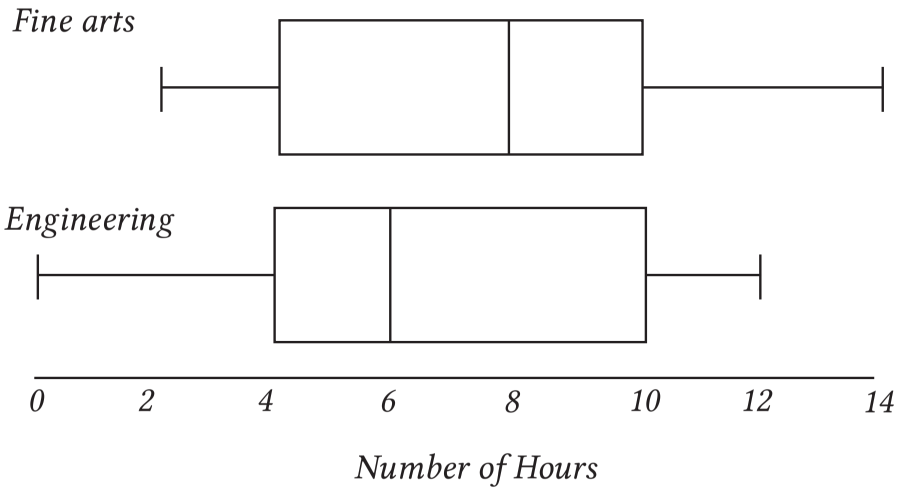
22. Based on the boxplot below, identify the 5 number summary.

23. Based on the boxplot below, identify the 5 number summary.



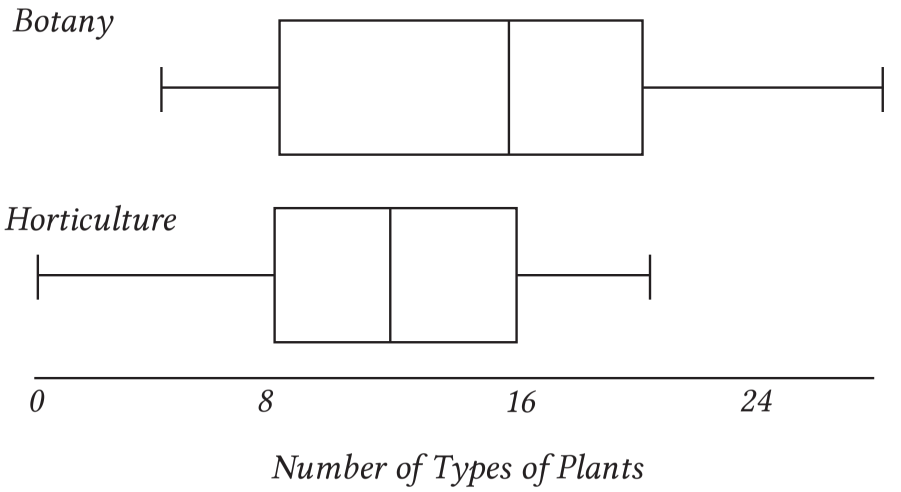
24. The boxplot below shows the number of hours spent on creative projects by students in the Fine Arts and Engineering programs during the past semester. Lily spent the median number of hours on creative projects for Fine Arts students. Ethan spent the median number for Engineering students.

    Who dedicated more time to creative projects?



25. The boxplot below shows the number of different types of plants in gardens managed by students in the Botany and Horticulture programs. Emma has the median number of types of plants in her Botany garden. Liam has the maximum number of types of plants in his Horticulture garden.

    Who has a more diverse garden?



26. Two local farmers reported the number of different vegetable types they grew this season.

    Farmer A grew a mean of 15.4 different vegetable types with a standard deviation of 3.2. Farmer B grew a mean of 14.8 different vegetable types with a standard deviation of 2.1.

    Which farmer grew a greater variety of vegetables on average?

    Which farmer had a more consistent variety of vegetables?

27. A research team collected data on the number of hours spent volunteering by a group of community members over the last year. The data collected is as follows:

    2, 5, 3, 8, 4, 6, 7, 1, 9, 10

    Calculate the standard deviation of the number of hours spent volunteering, rounded to one decimal place.

28. A group of friends recorded the number of different hobbies they each have. The data collected is as follows:

    3, 4, 2, 5, 6, 1, 3, 8, 4, 7

    Calculate the standard deviation of the number of hobbies, rounded to one decimal place.

29. A group of students recorded the number of plants they care for at home. The data collected is as follows:

    2, 5, 7, 3, 4, 6, 8, 1, 10, 4

    Calculate the average (mean) number of plants cared for, rounded to two decimal places.