5 CONNECTING THE DOTS...

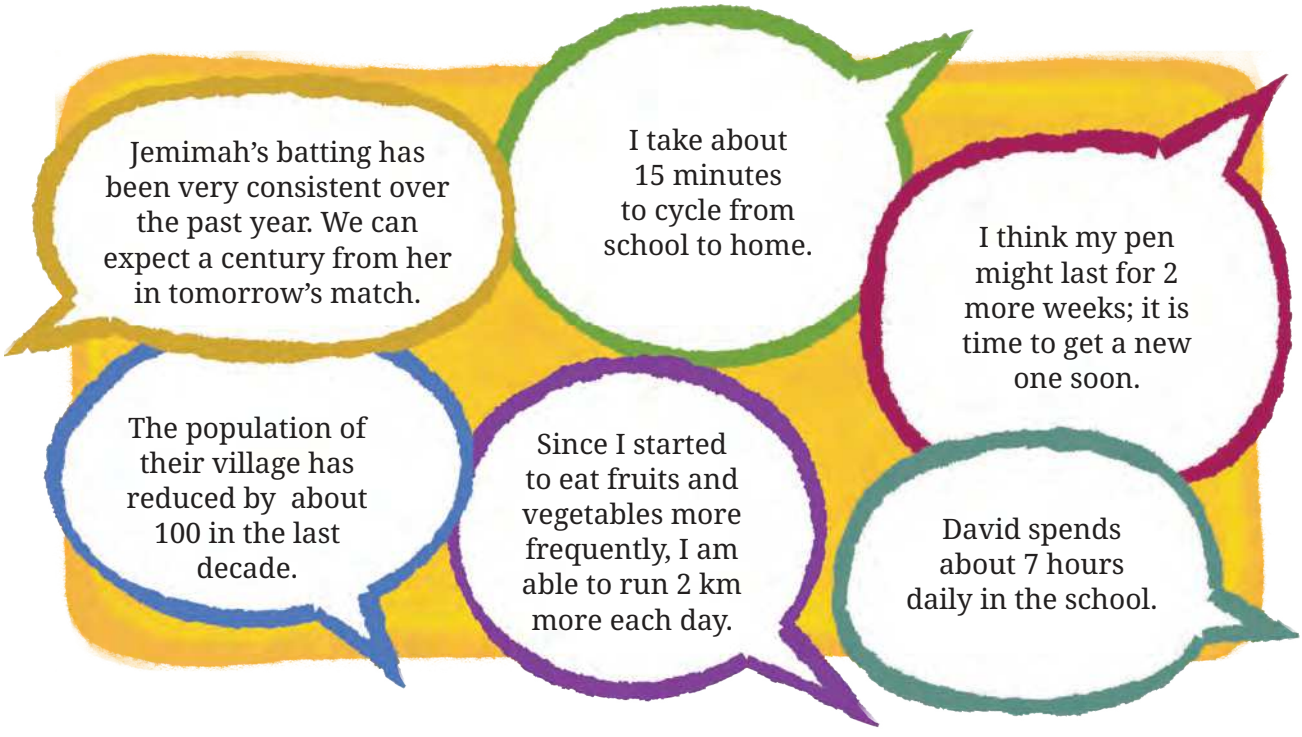
0789CH05

5.1 Of Questions and Statements

Your teacher tells you that they are meeting two of their childhood friends this evening. One is 5 feet tall and the other is 6 feet tall. What is your guess as to each friend's gender based on this information?

You might have guessed that the 5-foot-tall person is a woman and the 6-foot-tall person is a man. There is a chance that you are wrong. But experience tells us that 5-foot-tall men and 6-foot-tall women are rare. We have seen that, more often, men are taller than women.

The above is a simple example of statistical thinking. We regularly come across statements like —



We call these statistical statements. Simply put, a statistical statement is a claim or summary about some phenomenon, expressed in terms of numerical values, proportions, probabilities, or predictions.

A statistical question is a question that can be answered by collecting data. For example, "How tall are Grade 7 students in our school?" is a statistical question. We expect that not all Grade 7 students have the same height, but we can collect data, analyse it, and make conclusions about the heights that do occur. The question "Typically, are onions costlier in Yahapur or Wahapur?" is also a statistical question. Prices can vary over time. Therefore, answering this question requires us to look at data, analyse it, and come to conclusions making suitable statistical statements.

Math Talk

? Which of the following are statistical questions?

(a) What is the price of a tennis ball in India?

(b) How old are the dogs that live on this street?

(c) What fraction of the students in your class like walking up a hill?

(d) Do you like reading?

(e) Approximately how many bricks are in this wall?

(f) Who was the best bowler in the match yesterday?

(g) What was the rainfall pattern in Barmer last year?

The term statistics refers to the study of collecting, organising, analysing, interpreting, and presenting data. In this chapter, we shall encounter some statistical questions and learn how analysing data and graphs can help answer them.

5.2 Representative Values

? The runs scored by Shubman and Yashasvi in a cricket series are given in the table below. Who do you think performed better?

	Match 1	Match 2	Match 3	Match 4
Shubman	0	17	21	90
Yashasvi	67	55	18	35



Shreyas says, "Both their performances are similar since Yashasvi scored more in the first and second matches, whereas Shubman scored more in the third and fourth".

Vaishnavi says, "I think Shubman performed better because he scored the highest number of run in a match — 90!".

Shreyas says, "No! Yashasvi batted better since the total number of runs he made is 175, while Shubman made only 128".



Vaishnavi says, "Oh! Also, Yashasvi's batting is more consistent — the difference between his maximum score and minimum score is lower".

The table below shows the runs scored by these two players in another series. Who do you think performed better in this series?

	Match 1	Match 2	Match 3	Match 4	Match 5
Shubman	23	07	10	52	18
Yashasvi	26	53	02	-	15

Vaishnavi says, "Here, Shubman performed better since his total is 110 runs, while Yashasvi's total is 96 runs".

? What do you think of Vaishnavi's statement?

Shreyas says, "But Yashasvi made 96 runs in 4 matches and Shubman made 110 runs in 5 matches".

So, how do we say who performed better? It is often not simple to compare two groups of numbers and clearly say that one is better than the other.

? Can a single number act as a representative of a group of numbers? For example, can we represent Shubman's or Yashasvi's batting in this series with one number? Discuss.

Math Talk

We saw one way already — the total of the values in the group! But, if the group sizes are different, then the total may not be an appropriate measure to compare.

In some matches, a player could have scored more and in other matches less. A representative number for the group can be found by balancing out these *highs* and *lows*. For example, we can add up the runs scored in all the matches and divide the total by the number of matches played. We call this value the '**average**' or '**arithmetic mean**' of the given data.

Here, the average number of runs scored by a player in a match = [Total runs scored by the player in all the matches] ÷ [Number of matches played].

Average number of runs scored by Shubman in a match = 110 ÷ 5 = 21 runs.

Average number of runs scored by Yashasvi in a match = 96 ÷ 4 = 24 runs.

In this series, Yashasvi's average number of runs is higher than Shubman's.

The **Average** or **Arithmetic Mean** (A.M.), or simply **Mean**, is calculated as follows:

$$\text{Mean} = \frac{\text{Sum of all the values in the data}}{\text{Number of values in the data}}.$$

Average as Fair-Share

The average can also be understood as fair-share or equal-share.

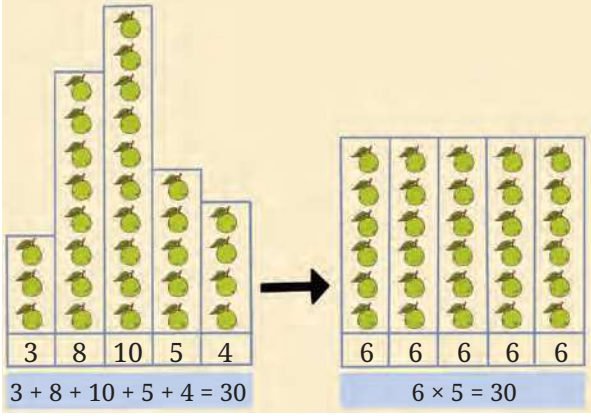
? Shreyas and 4 of his friends have collected the following numbers of guavas: 3, 8, 10, 5, and 4. Parag and 5 of his friends have collected the following numbers of guavas: 5, 4, 6, 3, 4, and 8. Each group will share their guavas equally amongst themselves. In which group will each member get a bigger share of guavas?

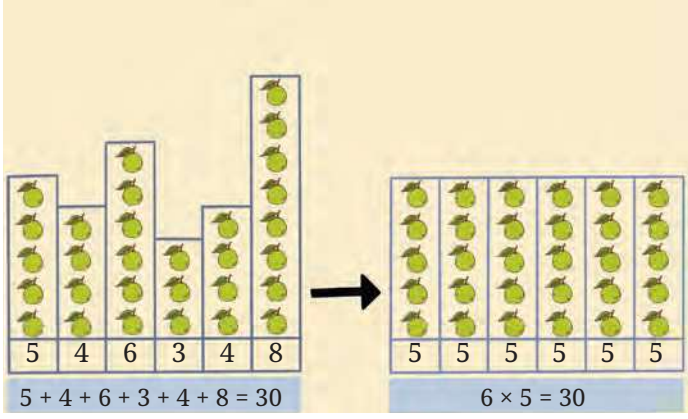
To find this out, we first find out how many guavas each group has collected. Then we divide this total by the number of people in the group to get each member's share.

Shreyas's group has collected 3 + 8 + 10 + 5 + 4 = 30 guavas. Each member of Shreyas's group gets 30 ÷ 5 = 6 guavas.

Parag's group has collected 5 + 4 + 6 + 3 + 4 + 8 = 30 guavas. Each member of Parag's group gets 30 ÷ 6 = 5 guavas.

So, the members of Shreyas's group get 1 more guava each than the members of Parag's group.





? Vaishnavi tracks the number of *Hibiscus* flowers blooming in her garden each day. The data for the last few days' is 2, 7, 9, 4, 3. What is the average number of *Hibiscus* flowers blooming per day in Vaishnavi's garden?



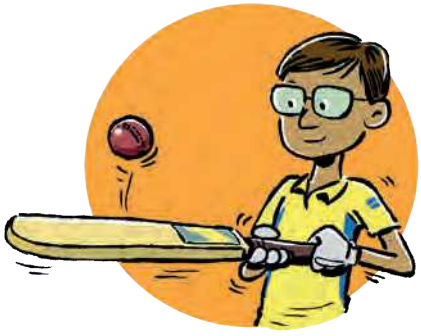
The average = (the total number of *Hibiscus* flowers bloomed) ÷ (number of days)
= (2 + 7 + 9 + 4 + 3) ÷ 5
= 5.

On an average, 5 *Hibiscus* flowers bloom daily.

In this case the average tells us the number of flowers blooming each day, if an equal number of flowers bloomed daily.

One of the terms used for the Arithmetic Mean in ancient Indian mathematics is *samamiti* (mean measure): '*sama*' means equal. Some terms used for the Arithmetic Mean in Indian texts include — *samarajju* (mean measure of a line segment) by Brahmagupta (628 CE), *samīkaraṇa* (levelling, equalising) by Mahāvīrācārya (850 CE) *sāmya* (equality, impartiality, equability towards) by Śrīpati (1039 CE) and *samamiti* (mean measure) by Bhāskarācārya (1150 CE) and Gaṇeṣa (1545 CE). The terminology shows that ancient Indian scholars perceived the Arithmetic Mean as the 'common' or 'equalising' value that is a representative measure of a collection of values.

Figure it Out



1. Shreyas is playing with a bat and a ball — but not cricket. He counts the number of times he can bounce the ball on the bat before it falls to the ground. The data for 8 attempts is 6, 2, 9, 5, 4, 6, 3, 5. Calculate the average number of bounces of the ball that Shreyas is able to make with his bat.
2. Try the activity above on your own. Collect data for 7 or more attempts and find the average.

Try This

3. Identify a flowering plant in your neighbourhood. Track the number of flowers that bloom every day over a week during its flowering season. What is the average number of flowers that bloomed per day?
4. Two friends are training to run a 100 m race. Their running times over the past week are given in seconds — Nikhil: 17, 18, 17, 16, 19, 17, 18; Sunil: 20, 18, 18, 17, 16, 16, 17. Who on average ran quicker?
5. The enrolment in a school during six consecutive years was as follows: 1555, 1670, 1750, 2013, 2040, 2126. Find the mean enrolment in the school during this period.

Know Your Onions!

The table shows the monthly price of onions, in rupees per kilogram (kg), at two towns. Where are onions costlier, according to you?

Math Talk

Month	Yahapur
January	25
February	24
March	26
April	28
May	30
June	35
July	39
August	43
September	49
October	56
November	59
December	44

Month	Wahapur
January	19
February	17
March	23
April	30
May	38
June	35
July	42
August	39
September	53
October	60
November	52
December	42

Khushboo: 'I think Wahapur is costlier because it has the highest price of ₹60.'

Nafisa: 'I added the prices of all months in each location - Yahapur's total is 458, whereas Wahapur's total is 450.'

Vishal: 'Wahapur is costlier since it has 3 numbers in the 50s'.

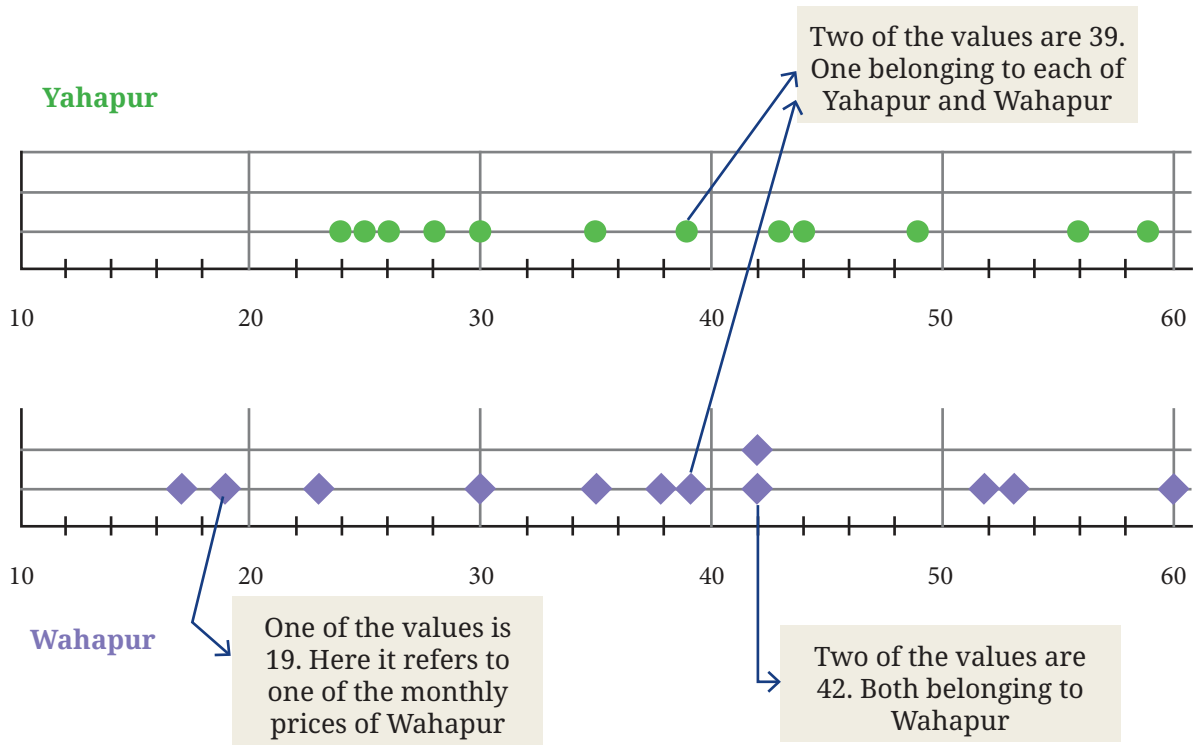
Sampat: 'I compared the prices in each month in both locations. Prices in Yahapur are higher for 6 months, prices in Wahapur are higher for 5 months, and the prices are the same for 1 month. So, I feel Yahapur is costlier.'

Jithin: 'I noticed that the difference between the highest and lowest prices in Yahapur is 59 – 24 = 35, and in Wahapur it is 60 – 17 = 43.'

Data can be described and compared by referring to its **minimum** value, **maximum** value, the average value, the sum **total** of all its values, and the difference between the maximum and minimum values.

? Can you think of any other ways to compare the data?

To study data, we can visualise it in multiple ways. One way is shown below — it is called a **dot plot**. Dot plots show data points as dots on a line, helping us visualise variability and patterns in data. In the following figure, each dot represents the monthly price of onions.



The prices in Yahapur are in green and those in Wahapur are in purple. The horizontal line shows the prices from 10 to 60 (instead of starting from 0 as there are no values from 0 – 10 or above 60). The dots on the vertical line give the number of occurrences of a data value. Notice the equal spacing between the units along the horizontal as well as the vertical lines.

? Does this visualisation capture all the data presented in the tables earlier?

? Looking at it, can we tell the price of onions in Yahapur in the month of January?

This method of presentation orders or sorts the data, but it loses the original (month-wise) sequence of the values. However, it allows us to group the data however we wish, just as Vishal did. For instance, there are 2 data values between 11 – 20 for Wahapur, while Yahapur has none. This representation makes it easier to observe the variation in the data — where and how the data is clustered or spread out. We can easily see that the prices in Wahapur are more spread out than those in Yahapur. It is also easy to spot the highest and lowest values.

We can also use the average as one of the ways to compare the prices at these two places.

Find the average price of onions at Yahapur and Wahapur.

A statement such as, "The price of onions is ₹35 per kilo", may not trigger any further questions. But looking at variations in data, like the prices of onions over a year in Yahapur and Wahapur, can spark one's curiosity. For example, one might be curious to know more about the two locations.

You might wonder —



Math Talk

What else do you wonder about?

You can discuss questions that you are curious about with your peers, teachers, or family members to find answers.



Observing and trying to make sense of data can reveal interesting things. It can also trigger our curiosity in different directions.

Averages Around Us

The Arithmetic Mean is frequently used in statistics, mathematics, experimental sciences, economics, sociology, sports, biology and diverse

disciplines as a representative of data. It is popular partly because the definition of the arithmetic mean is simple and easy to understand. Some statements involving averages in different scenarios are shown below:



The average rainfall per day in Jharkhand in the month of July is 37.2mm.



My scooty's average mileage this year is about 45 kilometers per liter.



Wheat yield averages 4.7 tonnes per hectare in Punjab vs. 2.9 tonnes per hectare in Bihar.



Smartphone users check their phone 58 times a day on average.



An average Indian citizen generates 0.45 kg of waste per day.



3126 is the average number of Indian long films released annually between 2017 – 2024.

Outliers and Medians

Does the average always give a reasonable summary of the values in a collection? If not, what is an alternative? Let us find out.

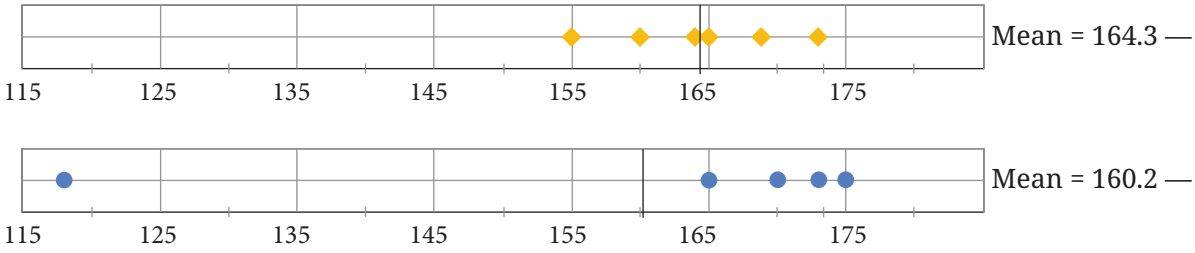
Height of a Family

The heights of the family members of Yaangba and Poovizhi are as follows:
Yaangba's family: 169 cm, 173 cm, 155 cm, 165 cm, 160 cm, 164 cm.
Poovizhi's family: 170 cm, 173 cm, 165 cm, 118 cm, 175 cm.

? Find the average height of each family. Can we say that Yaangba's family is taller than Poovizhi's family?

Math Talk

The average height of Poovizhi's family (160.2 cm) is less than that of Yaangba's family (164.3 cm). Although most members in Poovizhi's family are taller, their family's average height is less because one child is much younger and not as tall as the rest of the family. Their average height, 160.2 cm, is less than the heights of 4 out of 5 members. Here, the average doesn't seem to represent the data very well.



? Can you think of any other number that can represent the data better?

One way is to sort the data and pick the number in the middle. This number is called the **Median**.

To find the median height of Poovizhi's family, we first sort the heights —

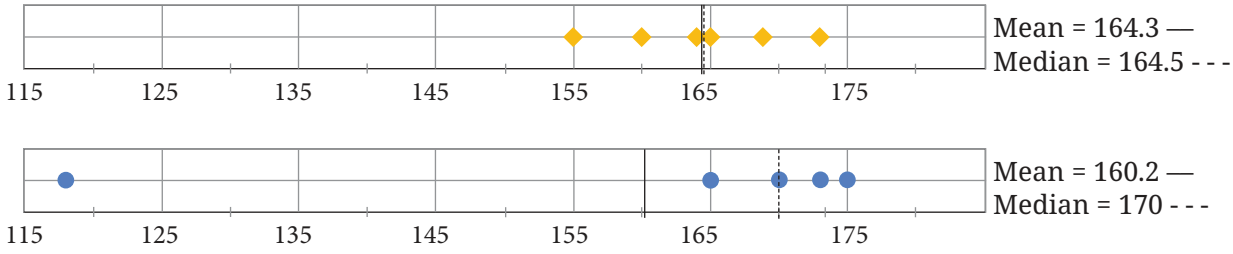
118, 165, 170, 173, 175.

The middle number in this sorted data is 170. Therefore, the median height is 170 cm.

Let us find the median height of Yaangba's family. Sorting the heights, we get

155, 160, 164, 165, 169, 173.

Since the median is the number in the middle, it will have an equal number of values less than it and greater than it. This data does not have a single middle number because it has an even number of values (6). In such cases, we take the average of the two middle numbers in the sorted data. Therefore, the median height of Yaangba's family is (164 + 165) ÷ 2 = 164.5 cm.



? In this case, does the median represent the heights of the families better than the average?

In Poovizhi's family, the height of the youngest child is quite different from the heights of the rest of the family. We call such a value an **outlier**. Outliers are values which significantly deviate from the rest of the values in the data.



Notice how the mean and the median are close to each other in Yaangba's data, in the absence of any outlier.

In Poovizhi's data, because of the outlier, the mean is much lower than the median.

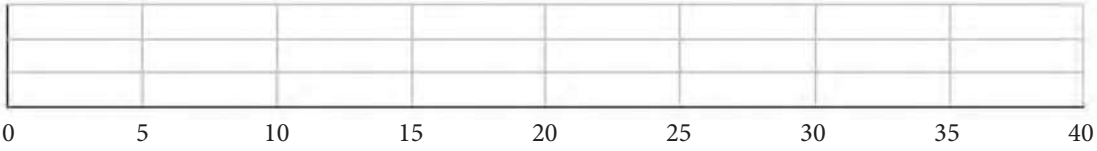
? Find the mean and median in Poovizhi's data without the outlier value 118. What change do you notice?

Are you a bookworm?

? After the summer vacation, a class teacher asked his class how many short stories they had read. Each student answered the number of stories read on a piece of paper, as shown below. Find the mean and median number of short stories read. Before calculating them, can you guess whether the mean will be less than or greater than the median?



Mark the data, the mean, and the median on the dot plot below.



The median value 6 means that half of the class members have read 6 or more stories.

? Which of the values would you consider an outlier?

? Find the mean and median in the absence of the outlier. What change do you notice?

The average may not always be an appropriate representative of data that has outliers. A very high or a very low outlier can significantly impact the sum, thus affecting the average. For example, the 118 cm height in Poovizhi's family is an outlier at the lower end of the data. And the count of 40 short stories read is an outlier at the higher end of the data. In these cases, we saw that the median was not affected much by the outliers.

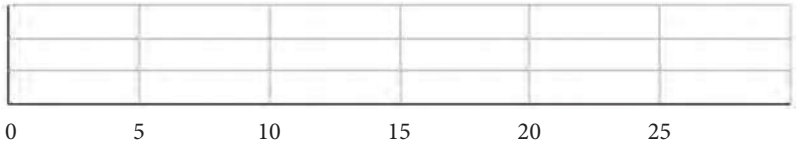
Are We on the Same Page?



? Do you read newspapers? Have you noticed how many pages a newspaper has on different days of the week — is it the same or different?

The list below shows the number of pages for a particular newspaper from Monday to Sunday: 16, 18, 20, 22, 26, 16, 10.

Mark the data, the mean, and the median on the dot plot below.



? In the three examples we considered — the heights, short-stories, and newspaper pages — observe the variability in data when:

Math Talk

(a) the mean and median are close to each other

(b) the mean and median are comparatively far apart, with mean < median

(c) the mean and median are comparatively far apart, with mean > median

When the data is more balanced or uniformly spread out the mean, and the median appear to be close to each other. When the outlier is on the lower end, the mean appears to shift in that direction, i.e., mean < median. When the outlier is on the higher end, the mean appears to shift in that direction, i.e., mean > median.

? Discuss the effect on the mean and median when outliers are present on both sides. You may take some example data to examine and explain this.

Math Talk

Mean and Median are called measures of central tendency, i.e., the tendency of the values to pile up around a particular value. In other words, they represent the 'centre' of the data.

Of Ends and the Essence

As we have just seen, the mean and the median can give different perspectives on the data. As part of analysing data, it can also be valuable to look at the variability in the given data, i.e., its extremes (minimum and maximum values).

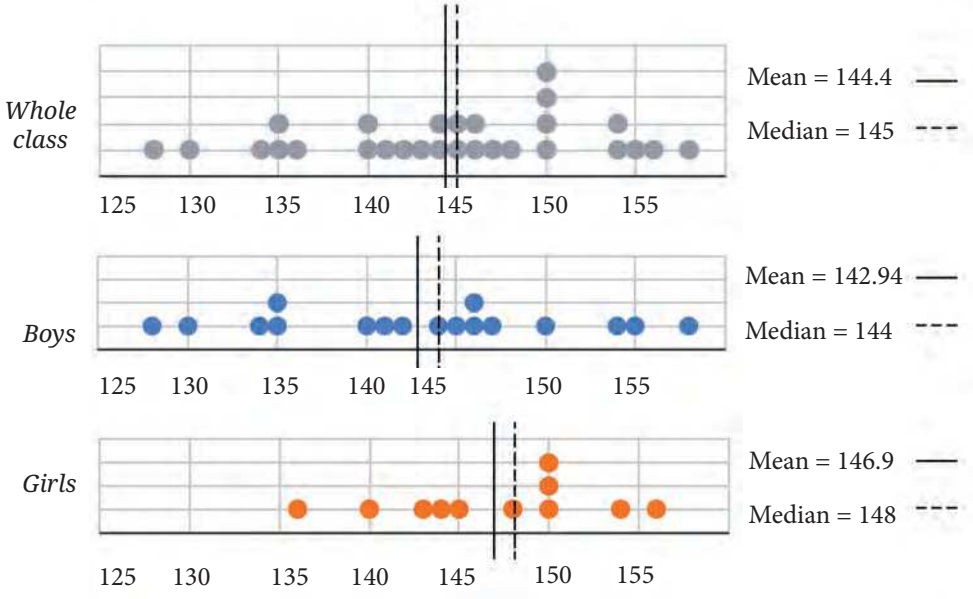
How Tall is Your Class?

Suppose you are asked the question, "How tall is your class?" What would you say?

The table below shows the heights of students in a Grade 5 class in centimeters.

Boys	147, 135, 130, 154, 128, 135, 134, 158, 155, 146, 146, 142, 140, 141, 144, 145, 150
Girls	143, 136, 150, 144, 154, 140, 145, 148, 156, 150, 150

We can visualise the data using a dot plot, identify the ends and patterns, and look at the variability. We can also find the measures of central tendency. The dot plot for the whole class, followed by the dot plots for boys and girls, respectively, are shown. The mean and the median are also shown for each collection.



What can we infer from the dot plots and the central tendency measures?

The following points can help answer the question of how tall the class is.

- The boys' heights are more spread out and are between 128 and 158. The girls' heights lie between 136 and 156. Both the tallest and shortest in the class are boys.
- Yet, the boys' average height is less than the whole class average, and also less than the girls' average height. We can say girls are taller than boys in this class. Of course, this doesn't mean every girl is taller than every boy!
- For boys' heights, mean < median (142.94 < 144) indicating a small influence of values on the lower side. For girls' heights too, mean < median (146.9 < 148) indicating a small influence of values on the lower side.

How many students are taller than the class' average height?

How many boys are taller than the class' average height?

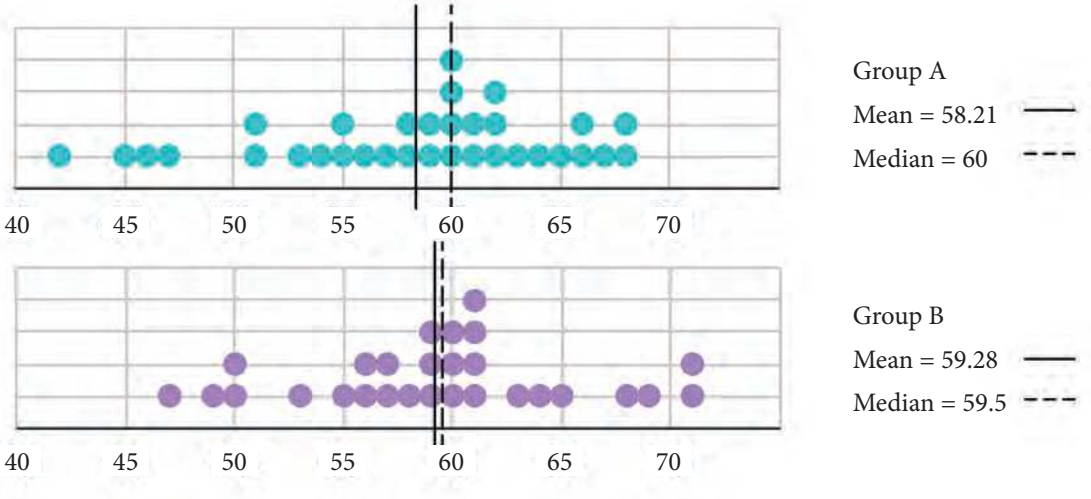
How long is a minute?



Two groups of children were asked to estimate the length of 1 minute. They start by closing their eyes and then open when they think 1 minute has passed. Of course, they are not supposed to count while their eyes are closed. The dot plots below show after how many seconds the children opened their eyes.

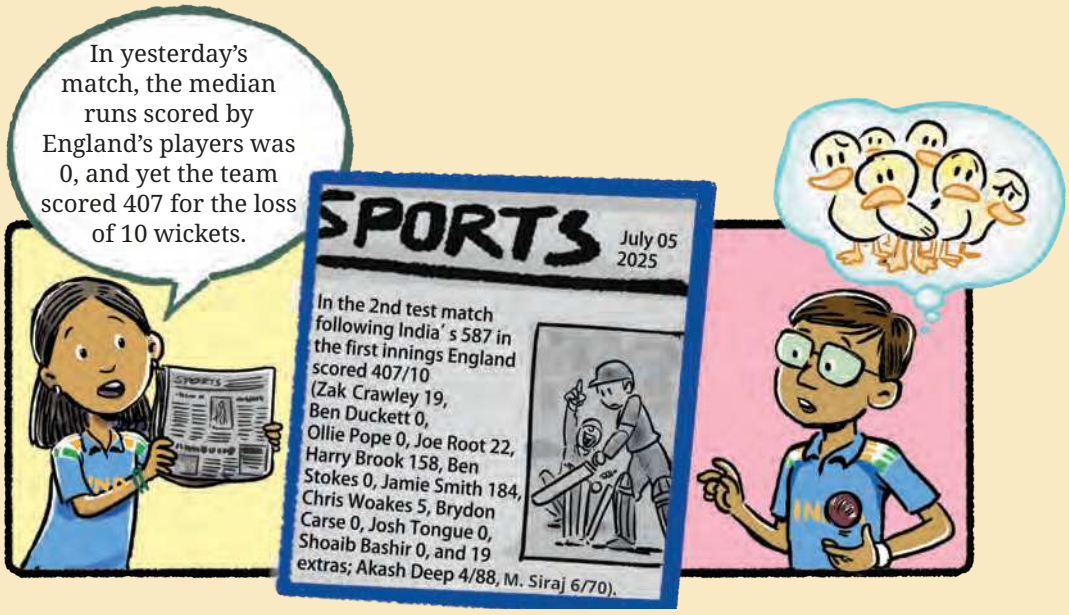
Discuss how well both the groups fared at this activity. Describe and compare the variability in data and their central tendency.

Math Talk



Zero Median Runs Scored!

In a cricket match, can a team's median runs scored by a player be 0 but the team's total score be 407/10?



Accounting for the extras, the average runs scored by a player in this innings is (407 – 19) ÷ 11 = 35.27.

Zero vs. No value

Suppose a player scores 57, 13, 0, 84, —, 51, 27 in a series. Notice that the player played Match 3 and scored 0 runs whereas the player did not play Match 5. So, we consider the total number of matches to be 6 and not 7. We calculate their average runs scored per match as
(57 + 13 + 0 + 84 + 51 + 27) ÷ 6.

Sita has a mango tree in her backyard. The number of mangoes the tree gave every month over the last year, from January to December, is 0, 0, 8, 24, 41, 16, 5, 0, 0, 0, 0, 0 respectively. If we want to find the mean or median number of mangoes per month, it would be appropriate to consider only the (summer) months when mangoes are expected to grow.

A Mean Foot

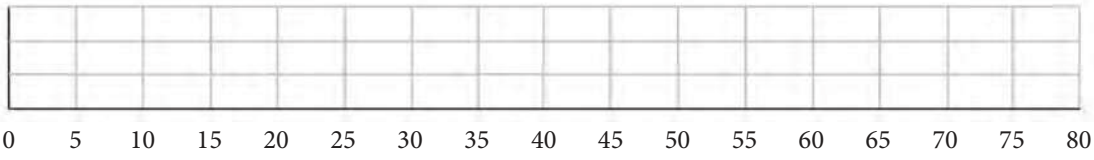
In the early 1500s in Europe, the basic unit of land measurement was the rod, defined as 16 feet long. At that time, a foot meant the length of a human foot! But foot sizes vary, so whose foot could they measure? To solve this, 16 adult males were asked to stand in a line, toe to heel, and the length of that line was considered the 16-foot rod. After the rod was determined, it was split into 16 equal sections, each representing the measure of a single foot. In essence, this was the arithmetic mean of the 16 individual feet, even though the term 'mean' was not mentioned anywhere.



Jacob Kobel's depiction of the determination of 1 foot

Figure it Out

1. Find the median of onion prices in Yahapur and Wahapur.
2. Sanskruti asked her class how many domestic animals and pets each had at home. Some of the students were absent. The data values are 0, 1, 0, 4, 8, 0, 0, 2, 1, 1, 5, 3, 4, 0, 0, —, 10, 25, 2, — , 2, 4. Find the mean and median. How would you describe this data?
3. Rintu takes care of a date-palm tree farm in Habra. The heights of the trees (in feet) in his farm are given as: 50, 45, 43, 52, 61, 63, 46, 55, 60, 55, 59, 56, 56, 49, 54, 65, 66, 51, 44, 58, 60, 54, 52, 57, 61, 62, 60, 60, 67. Fill the dot plot, and mark the mean and median. How would you describe the heights of these palm trees? Can you think of quicker ways to find the mean? How many trees are shorter than the average height?

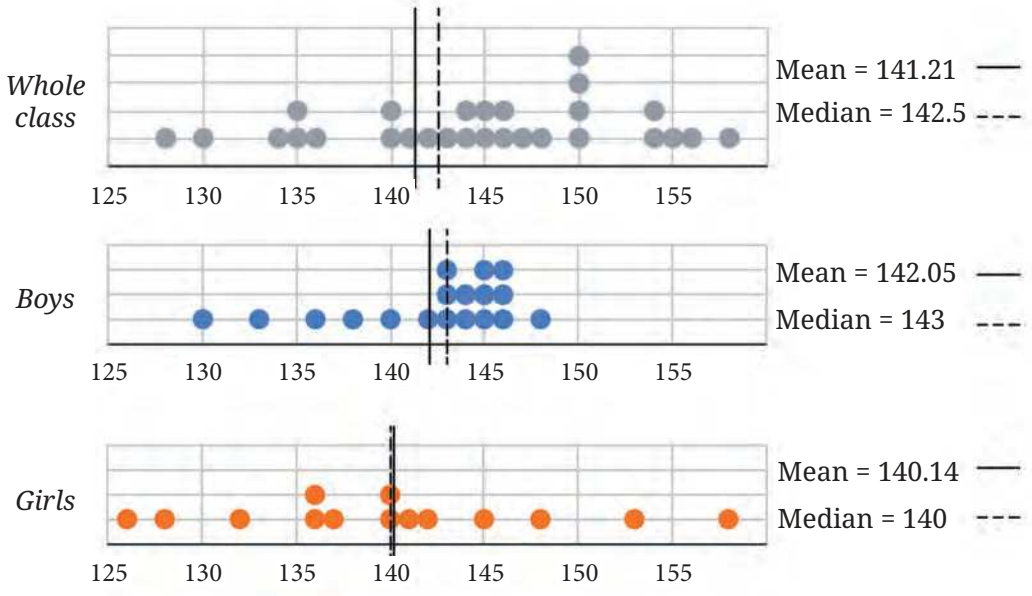


4. The daily water usage from a tap was measured. The usage in liters for the first few days are: 5.6, 8, 3.09, 12.9, 6.5, 12.1, 11.3, 20.5, 7.4.
   (a) Can the mean or median daily usage lie between 25 and 30? Justify your claim using the meaning of mean and median.
   (b) Can the mean or median be lesser than the minimum value or greater than the maximum value in a data?
5. The weights of a few newborn babies are given in kgs. Fill the dot plot provided below. Analyse and compare this data.

Boys	3.5	4.1	2.6	3.2	3.4	3.8
Girls	4.0	3.1	3.4	3.7	2.5	3.4



6. The dot plots of heights of another section of Grade 5 students of the same school are shown below. Can you share your observations? What can we infer from the dot plots and the central tendency measures?



? Compare the heights of the two sections. Share your observations.

7. The weights of some sumo wrestlers and ballet dancers are: Sumo wrestlers: 295.2 kg, 250.7 kg, 234.1 kg, 221.0 kg, 200.9 kg. Ballet dancers: 40.3 kg, 37.6 kg, 38.8 kg, 45.5 kg, 44.1 kg, 48.2 kg. Approximately how many times heavier is a sumo wrestler compared to a ballet dancer?



Math Talk

5.3 Visualising Data

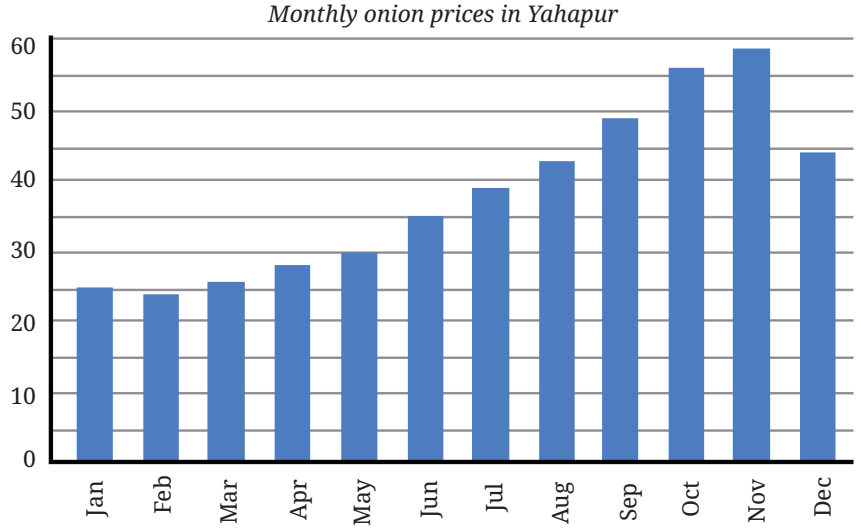
We can often understand data more clearly if it is presented as a picture. This is called **data visualisation**. Last year, we saw how to visualise data using graphs. Let us explore visualisation further.

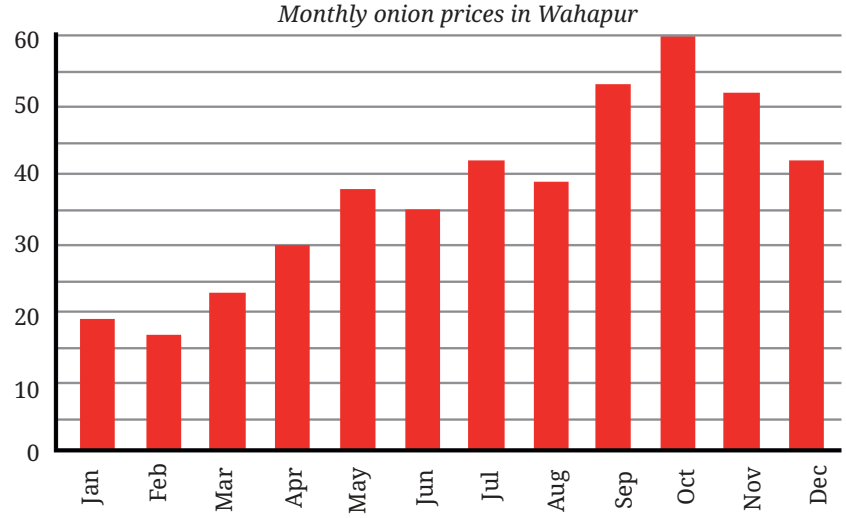
Clubbing the Columns

Earlier, we looked at the monthly onion prices in Yahapur and Wahapur.

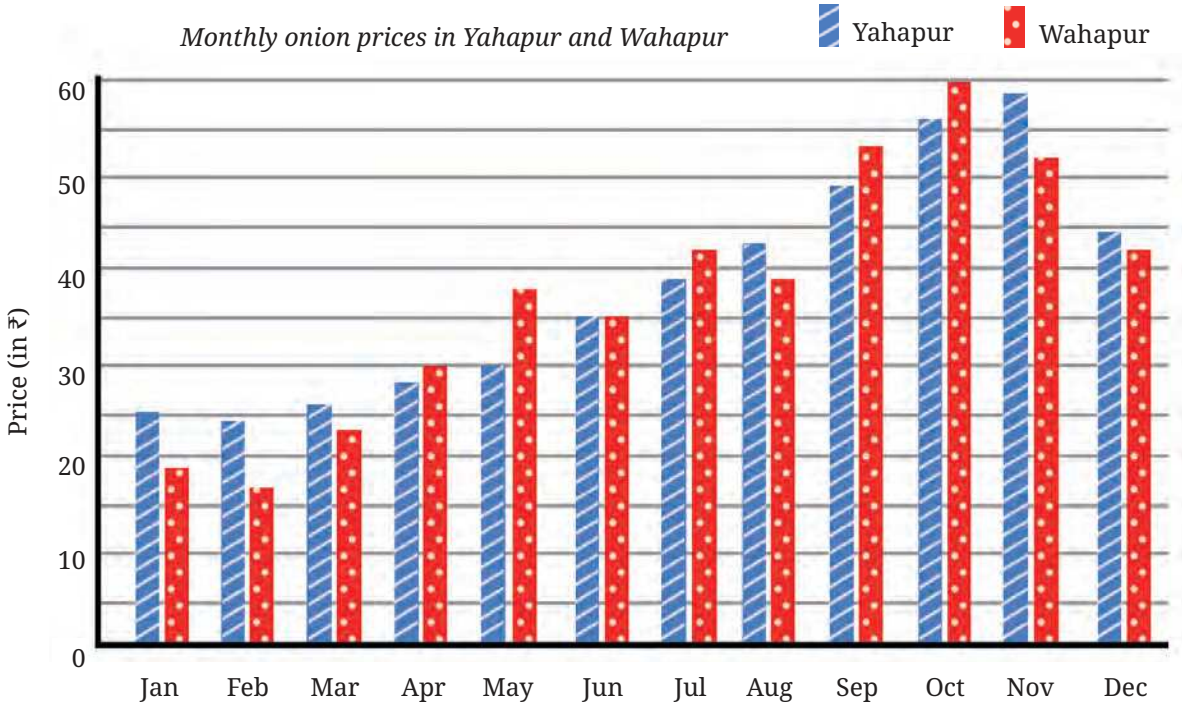
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Yahapur	25	24	26	28	30	35	39	43	49	56	59	44
Wahapur	19	17	23	30	38	35	42	39	53	60	52	42

Two column graphs for this data are given below.





The two graphs can also be combined into a single graph. We just draw the bars side by side! Verify if the data in the table matches the graph below.



We use different colours to clearly separate the data from the two places. This is called a **clustered column graph**. Since it has two columns in each cluster, we also call it a **double column graph**.

? What is the scale used in this graph?

The relative heights of the bars tell us where onions are costlier in each month. We can also visually estimate the difference by referring to the markings along the vertical line.

? Is it now easier to compare month-wise prices in both places?

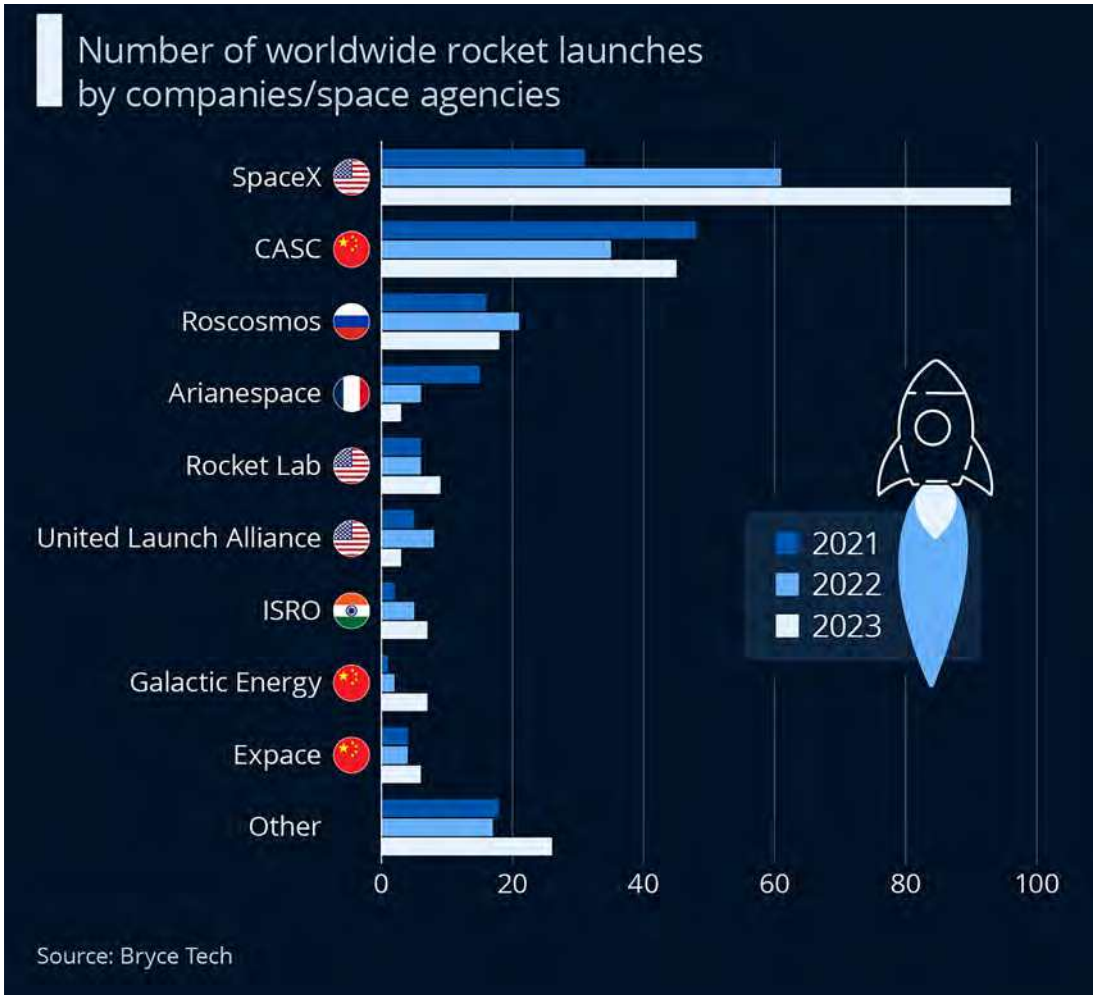
The dots and slanted lines within the bars help people who find difficulty in distinguishing colours. It is also useful when things are printed in greyscale (black-and-white).

10...9...8...7...6...5...4...3...2...1...Take Off!

You might have heard about scientific probes (like Chandrayaan-3 launched in 2023 by ISRO or the Voyager-1 launched in 1977 by NASA), observational satellites (like Aryabhata launched in 1975 by ISRO or Sputnik-1 launched in 1957 by the Soviet Space program), or about human spaceflights to the International Space Station. All space missions are

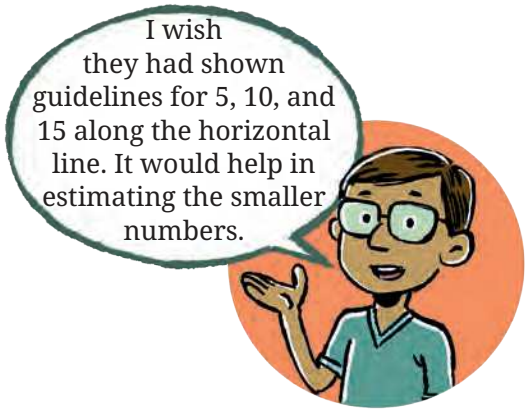
launched using rockets. Look at the graph below showing the number of worldwide rocket launches by different organisations.

? Share your observations (you may take the teacher's help to identify the countries these organisations belong to).



Source: https://www.statista.com/chart/29410/number-of-worldwide-rocket-launches-by-companies-and-space-agencies/

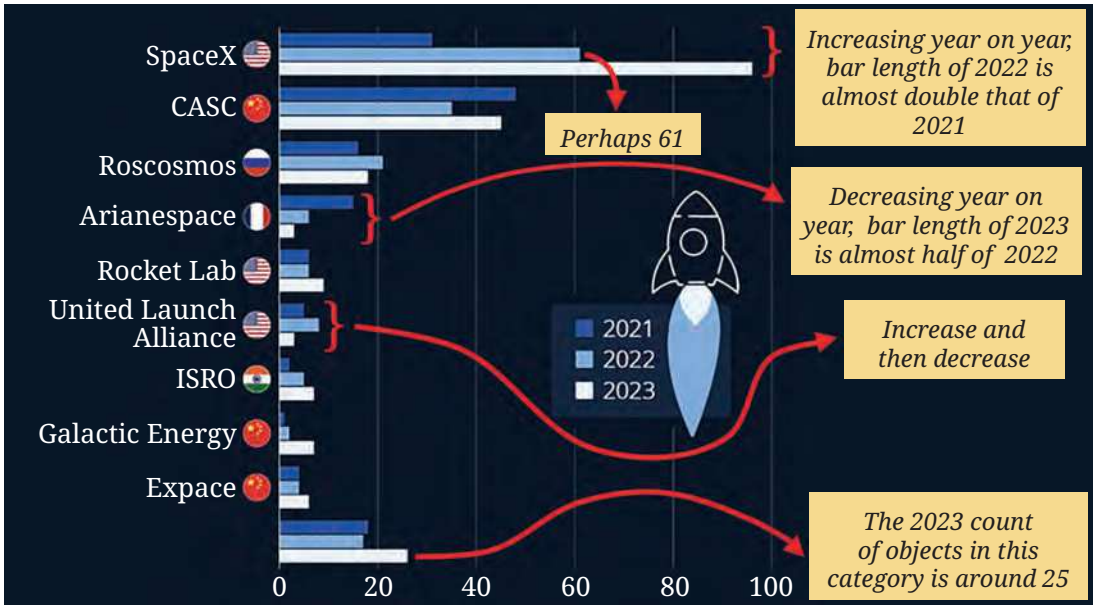
Often there is a lot of information in graphs and it may be difficult to understand. We can follow a 2-step process to simplify making sense of the data in graphs.



Step 1: Identify what is given

? Notice how the graph is organised, what scale is used, and what patterns the data shows.

- For each organisation, the numbers of rocket launches for the years 2021, 2022, and 2023 are shown as three adjacent bars. The scale used is 1 unit length = 20 rockets. Notice the numbers at the bottom.
- The 'Others' category indicates multiple organisations worldwide that are clubbed together to keep the graph short.
- Note that in the double bar graph of onion prices, the months are shown in order, i.e., January to December, to observe the change over time, whereas in this case, a change in the order of organisations does not affect the meaning.



Step 2: Infer from what is given

Analyse and interpret each of your observations.

- We can say that the USA, China, and Russia are the leading rocket launching countries in the given time period.
- SpaceX launched about twice the number of rockets in 2022 compared to 2021. And it launched about 35 more rockets in 2023 compared to 2022.
- The number of rockets launched by Arianespace decreased every year.
- United Launch Alliance launched more rockets in 2022 than in 2021. They launched fewer rockets in 2023 than in both the years 2022 and 2021.
- Other organisations launched about 25 rockets in 2023.

? Identify which of the following statements can be justified using this data.

(a) All organisations launched more rockets than the previous years.

(b) Only an organisation from the USA launched more than 50 rockets in a single year.

(c) The total number of rockets launched by France in all 3 years is less than 40.

(d) The average number of rockets launched by CASC in these 3 years is around 40.

(e) ISRO launched more rockets than Galactic Energy in these 3 years.

(f) Russia launched more than 60 rockets in these 3 years.

? List the organisations that have consistently launched more rockets every year.

? Estimate the total number of rockets launched worldwide in 2023.

(a) less than 200 (b) 200 to 400
(c) 400 to 600 (d) more than 600

We may have many questions after looking at the graph. We might wonder why USA launches so many more rockets than other countries. Or we might be eager to look at the data from previous and later years. What are you curious to know after looking at this graph?

Summer and Winter at the Same Time

The tables below show data related to weather in two cities in different countries. The numbers given are in hours. Can you guess what the data might be related to?

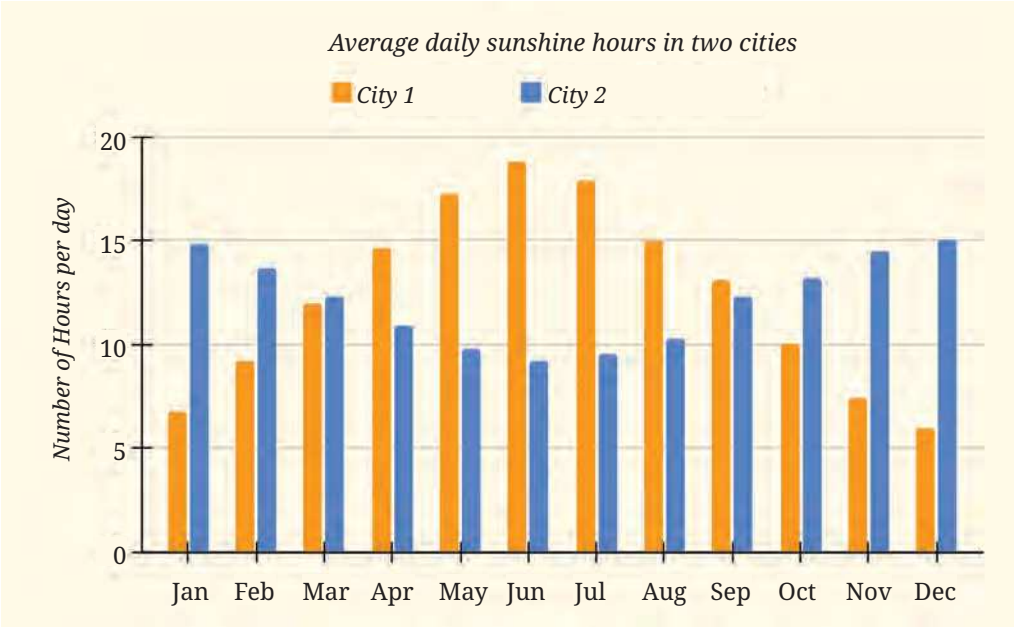
City 1

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
210	257	372	441	536	564	555	465	394	310	222	186

City 2

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
459	384	381	327	304	276	295	318	369	409	435	468

The data shows the monthly hours of daylight (i.e., the Sun is at least partly above the horizon) in these two cities over the year. Based on this data, a clustered bar graph showing the average daylight hours per day in each month is given below. This average is obtained by dividing the monthly daylight hours by the number of days in the month.



Let us follow the 2-step process to identify and interpret the information presented.

Step 1: Identify what is given

Notice how the graph is organised, what scale is used, and what patterns the data shows.

- The horizontal line shows the months of the year. The vertical line shows the average daylight hours per day, using the scale 1 unit = 5 hours. The month of June has the maximum value for City 1 and the minimum value for City 2.

Step 2: Infer from what is given

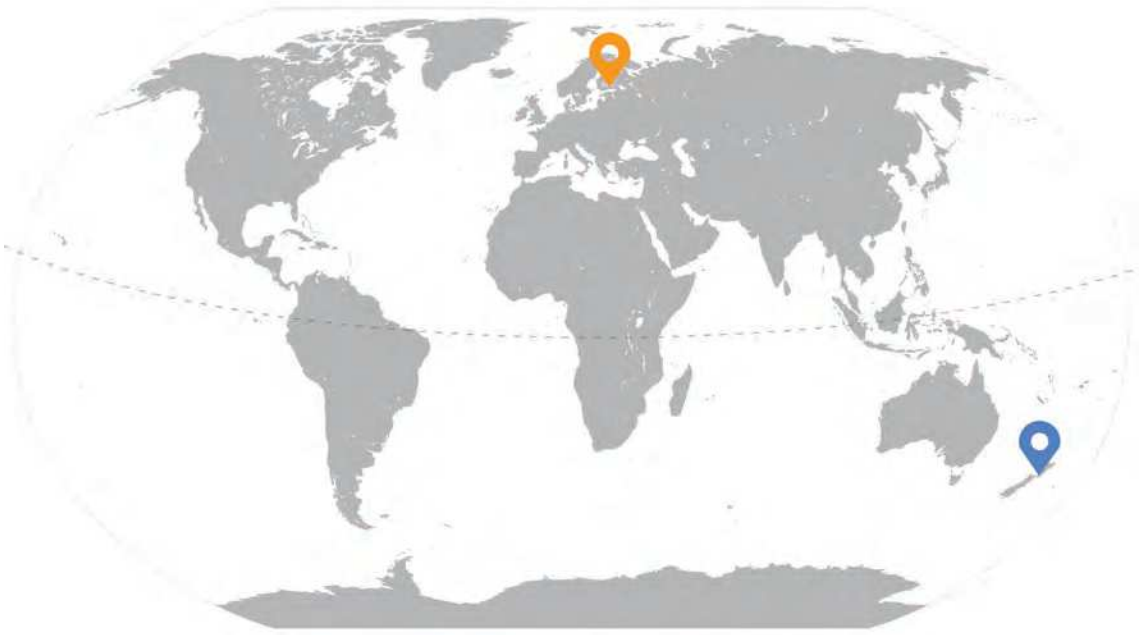
Analyse and interpret each of your observations. Share appropriate summary and conclusion statements.

- The average number of daylight hours per day in City 1 increases from January, reaching a maximum of about 17 – 18 hours in June. It then decreases, reaching a minimum of about 6 hours in December.

- The average number of daylight hours per day in City 2 decreases from January, reaching a minimum of about 9 hours in June. It then increases, reaching a maximum of about 15 hours in December.
- The maximum and minimum values in City 1 are more extreme than those of City 2. That is, the maximum number of daylight hours per day of City 1 is more than that of City 2, and the minimum number of daylight hours per day of City 1 is less than that of City 2.
- In June, City 1 experiences daylight for about $\frac{3}{4}$th of the full day (24 hours), whereas during December–January, it only experiences daylight for about $\frac{1}{4}$th of the full day.

? Does this give some idea of where these two cities are located?

City 1 and City 2 are located away from the Equator in the Northern and Southern hemispheres, respectively. City 1 is Helsinki, Finland, and City 2 is Wellington, New Zealand. These are also shown on the map. In June, the Northern Hemisphere is tilted towards the Sun, resulting in longer daylight hours; it is summertime here. Meanwhile, the Southern Hemisphere is tilted away from the Sun, leading to shorter days; it is winter time here. The inverted seasonal daylight pattern is due to the cities' location in opposite hemispheres. The large variation in the data is because they are away from the Equator.



? Is there anything more that you wish to explore?

Math Talk



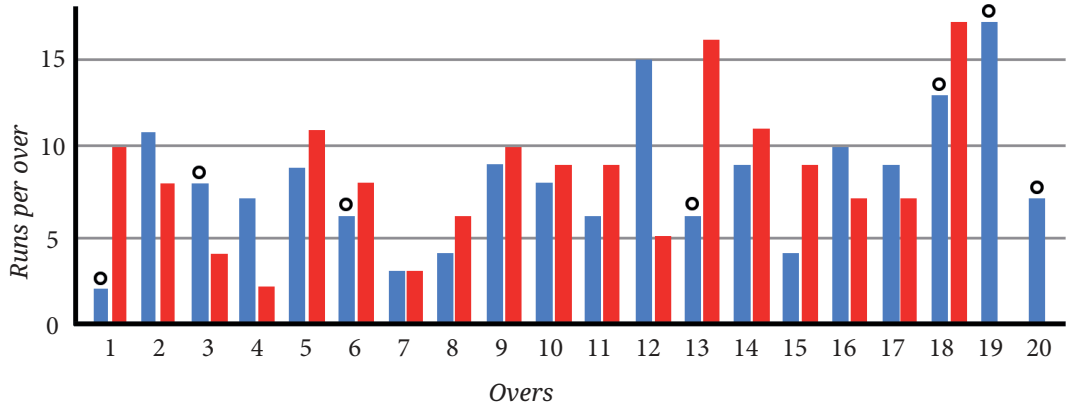


©timeanddate.com/Brendan Goodenough

The Midnight Sun at Andøya, Norway.

All it Takes is a Minute

Have you ever missed watching a cricket match? You can catch up in a minute by looking at a graph. You might have seen graphs like the following one.



The horizontal line lists the overs starting from 1, and the vertical line indicates the runs scored in each over. The graph shows the number of runs scored per over as a double bar graph — each bar corresponding to a team. Let us call them the blue team (denoted by blue) and the red team (denoted by red). The scale used for the runs per over is 1 unit = 5 runs. The circles shown on top of the bars indicate that a wicket fell in that over.

Answer the following questions based on the graph:

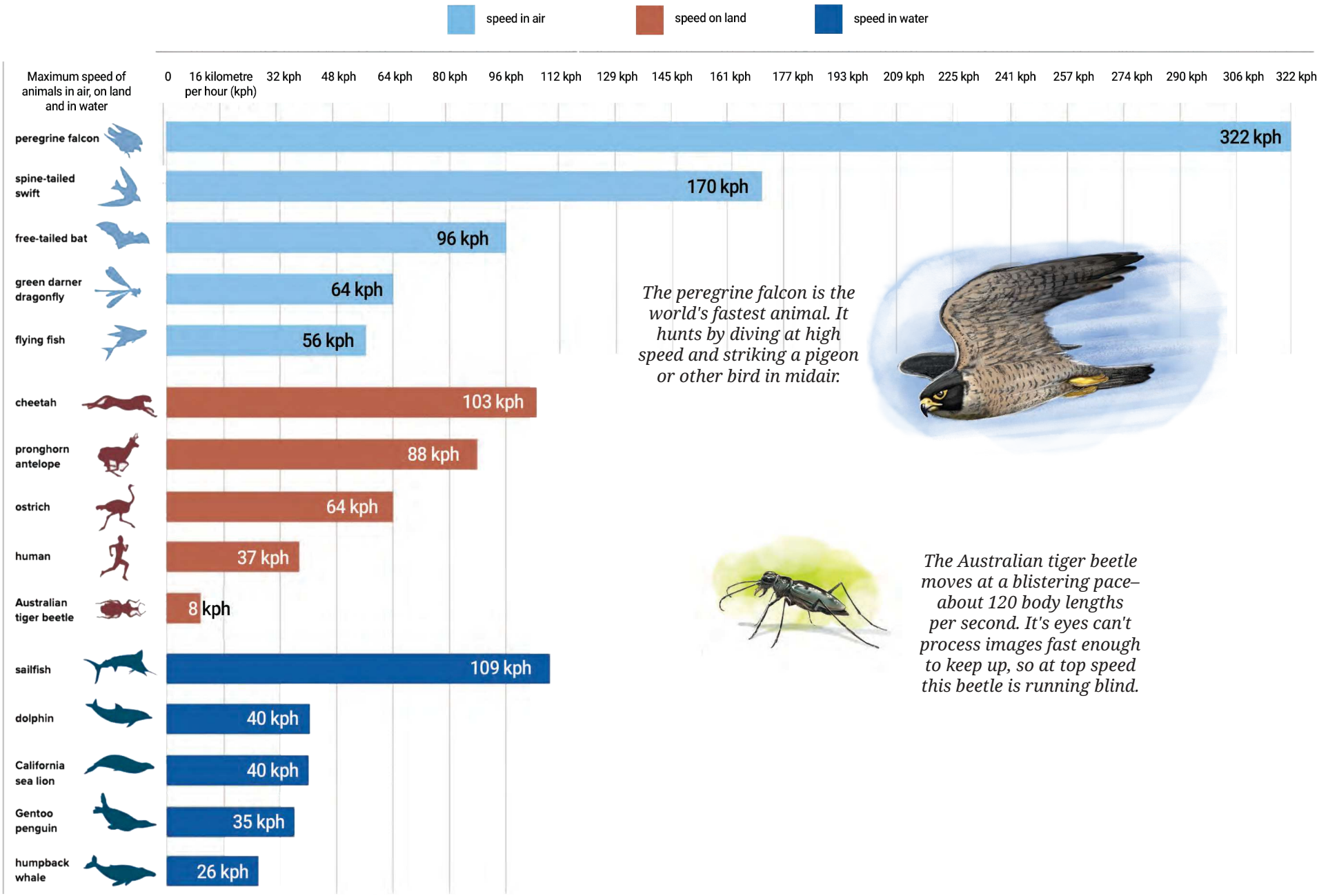
1. Can we tell who batted first? Who won the match?
2. How many runs did the blue team score in over 12?
3. In which over did the red team score the least number of runs?
4. Is it easy to tell the target set by the team batting first?

Figure it Out

1. The following infographic shows the speeds of a few animals in air, on land, and in water. Can we call this graph a bar graph?
   (a) What is the scale used in this graph?
   (b) What did you find interesting in this infographic? What do you want to explore further?
   (c) Identify a pair of creatures where one's speed is about twice that of the other.
   (d) Can we say that a sailfish is about 4 times faster than a humpback whale? Can we say that a sailfish is the fastest aquatic animal in the world?

Math Talk

How fast?



The peregrine falcon is the world's fastest animal. It hunts by diving at high speed and striking a pigeon or other bird in midair.

The Australian tiger beetle moves at a blistering pace– about 120 body lengths per second. It's eyes can't process images fast enough to keep up, so at top speed this beetle is running blind.

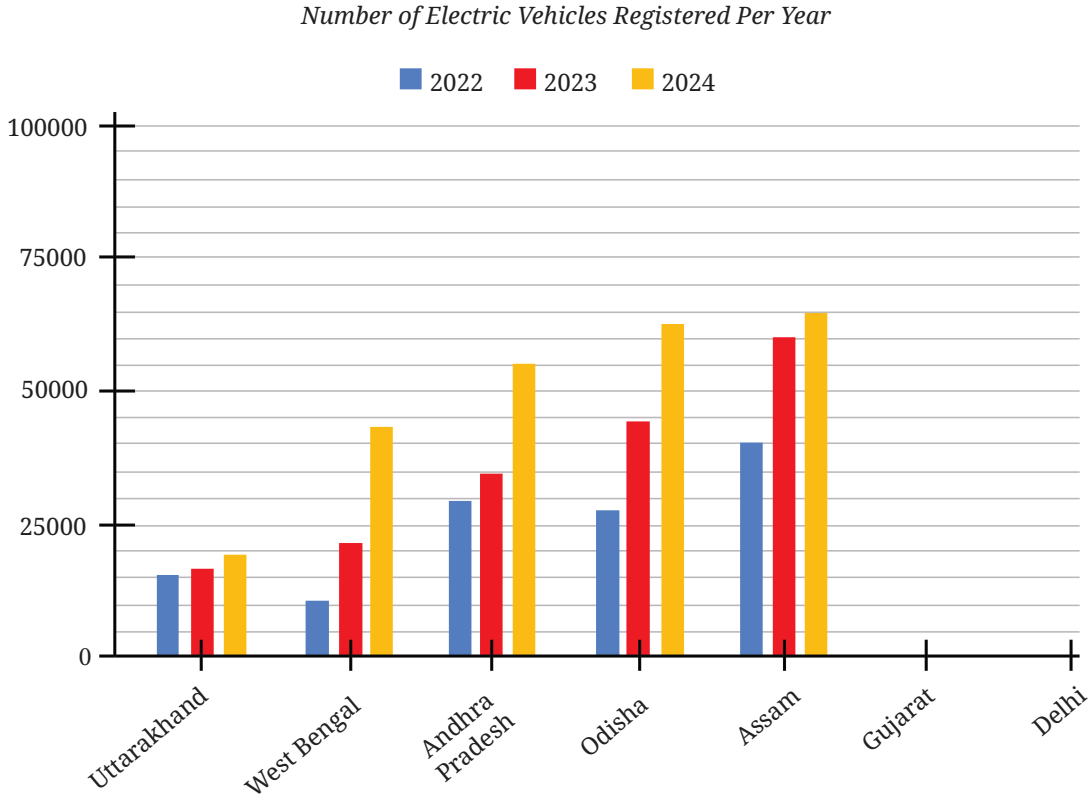
2. Preyashi asked her students 'If you were to get a super power to become aquatic (water-borne), aerial (air-borne), or spaceborne which one would you choose?'. The responses are shown below. Some chose none. Draw a double-bar graph comparing how both grades chose each option. Choose an appropriate scale.

Grade 5	w, a, a, a, w, n, s, a, n, w, a, a, a, a, a, w, w, s, a, a, n, w, a, a, n
Grade 9	n, w, s, a, s, w, s, s, a, a, w, s, s, a, s, a, n, w, s, s, a, w, a, w, a

3. The temperature variation over two days in different months in Jodhpur, Rajasthan, is given below. Draw a double-bar graph. Use the scale 1 unit = 4°C. Can you guess which two months these days might belong to?

	12 am	**3 am**	**6 am**	**9 am**	**12 pm**	**3 pm**	**6 pm**	**9 pm**
Day 1	20°C	18°C	16°C	20°C	26°C	34°C	30°C	24°C
Day 2	37°C	34°C	30°C	33°C	37°C	43°C	42°C	39°C

4. The following clustered-bar graph shows the number of electric vehicles registered in some states every year from 2022 to 2024.



(a) The data (rounded-off to thousands) for the states of Gujarat and Delhi are given in the table below. Mark the corresponding bars on the bar graph. (It is enough if you place the top of the bars between the two appropriate vertical guidelines.)

	2022	2023	2024
Gujarat	69000	89000	78000
Delhi	62000	74000	81000

(b) Notice how the graph is organised, what scale is used, and what patterns the data shows.

(c) How would you describe the change for various states between 2022 and 2024?

(d) Approximately how many more registrations did Assam get in 2023 compared to 2022?

(e) How many times more did the registrations in West Bengal increase from 2022 to 2024?

(f) Is this statement correct — 'There were very few new registrations in Uttarakhand in 2023 and 2024, as the increase in the bar lengths is minimal'?

5.4 Data Detective

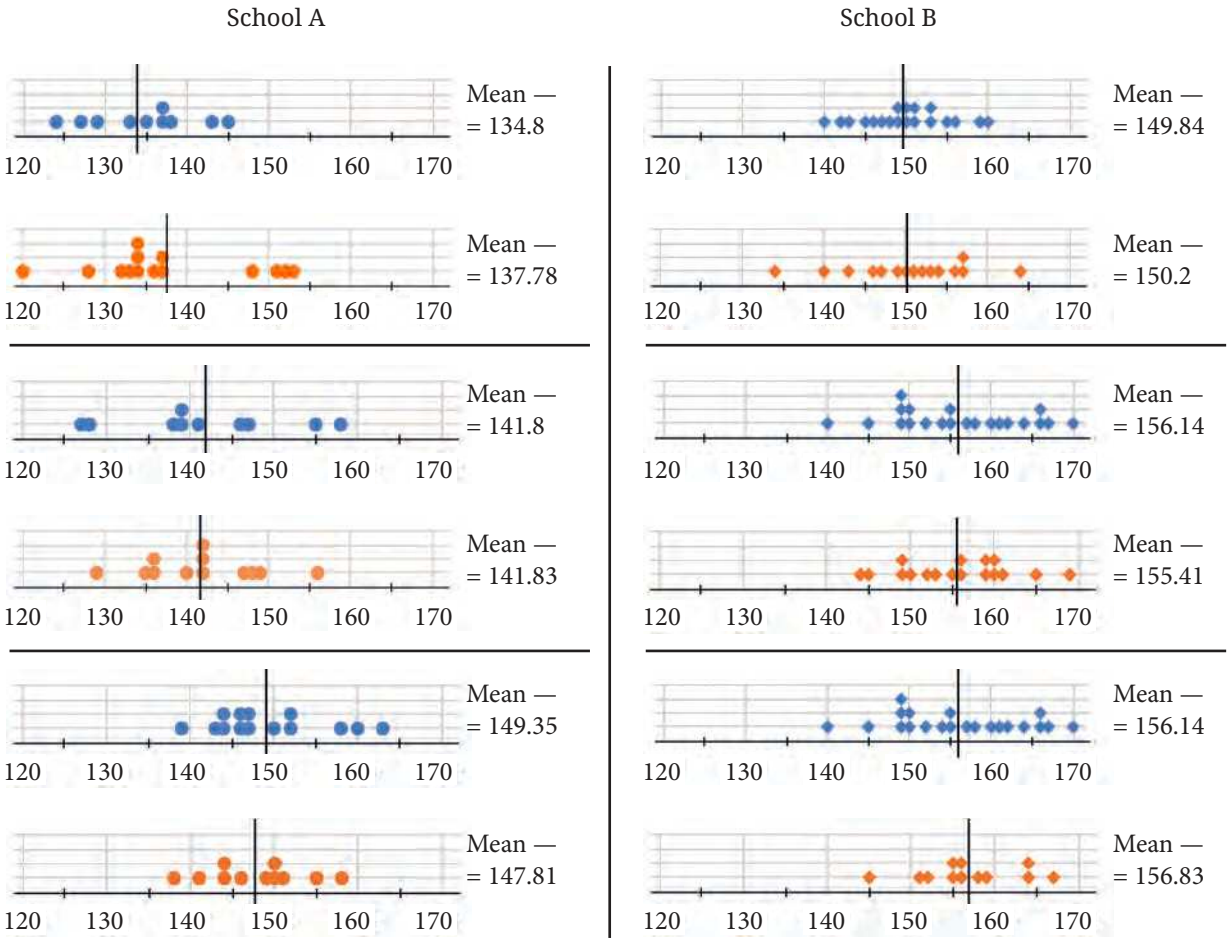
We put well-formed sentences one after the other to make a beautiful story. In the same way, well-organised and well-presented data can tell interesting stories, and can also expose new mysteries or help solve mysteries!

Telling Tall Tales

Earlier, we saw data of two Grade 5 classrooms with heights of boys and girls in each class. There, the average height of girls was more than boys in one class and vice versa in the other class.

? Following are the dot plots of heights of boys (in blue) and girls (in orange) of Grades 6, 7 and 8 (in that order) of two different schools. What do you notice? Share your observations.

Math Talk



Looking at this data, you might wonder:

"Why is there a considerable difference in heights in the same grades across these two schools?"

"Where are these schools located?"

"How tall are students in Grades 6 to 8 in my school?"

"What is the average height of all Grade 6 boys and girls?"

We see that men are taller than women in general. But what about the heights of boys and girls? Are boys taller than girls? Well, just by looking at the data of one or two schools, we cannot generalise for all children in our country, or around the world.

Let us look at some data (based on a survey) of the heights of boys and girls of different ages in India over time. The following table shows the average heights of boys and girls (in centimeters) across ages 5 to 19 in the years 1989, 1999, 2009, and 2019. In each year, the first column shows boys' heights and the second column shows girls' heights.

Age	1989		1999		2009		2019	
5	101.3	100	102.4	101.7	105.1	104	107.1	107.2
6	107.5	106	108.7	107.5	111	109.7	113.1	112.9
7	113	111.4	114.2	112.6	116.2	114.8	118.6	118
8	118.1	116.5	119.2	117.5	120.9	119.6	123.5	122.7
9	122.9	121.7	123.9	122.4	125.2	124.5	128.1	127.6
10	127.5	127.3	128.3	127.8	129.4	129.9	132.6	132.8
11	132.2	133.4	132.8	133.6	133.7	135.7	137	138.6
12	137.7	139	138	139.1	138.9	141.1	142.2	143.8
13	144.2	143.2	144.3	143.1	145.2	145.1	148.4	147.7
14	150.6	146.2	150.5	146.1	151.5	148	154.4	150.4
15	155.4	148.5	155.2	148.4	156.3	150.1	159	152.4
16	158.9	150.1	158.7	150.1	159.9	151.6	162.3	153.8
17	161.3	151.2	161.4	151.3	162.6	152.6	164.6	154.7
18	162.9	151.8	163.2	152.1	164.3	153	166	155.2
19	163.5	151.9	164.2	152.4	165.1	153	166.5	155.2

? Spend sufficient time observing the data presented in this table. Share your findings with the class.

Math Talk

These are some prompts for you to probe —

- Changes in the heights of boys or girls of a certain age from 1989 to 2019.
- The heights of boys vs. girls at different ages in a particular year.
- Changes in height between successive ages in boys and girls in 2019.

? Which of the following statements can be justified using the data?

1. The average heights of both boys and girls at every age increased from 1989 to 2019.
2. The average height of 13-year-old girls in 1989 is more than the average height of 14-year-old girls in 2009.
3. The average height of 15-year-old boys in 2019 is more than the average height of 16-year-old boys in 1989.
4. All girls aged 13 are taller than all girls aged 11.
5. Throughout the age period 5 to 19, the average boy's height is more than the average girl's height.
6. Boys keep growing even beyond age 19.

In 2019, between which two successive ages from 5 to 19 did boys grow the most? Between which two successive ages from 5 to 19 did girls grow most?

Math Talk

Suppose the average height of a newborn is 50 cm. Estimate the average height of young children of ages 1 to 4.

Based on the trend observed in the table, write your estimates of the heights of boys and girls for ages 5 to 19 in the year 2029.



Whenever you see data or some graph, look closely to know the story it has to say and the mysteries it may hold.

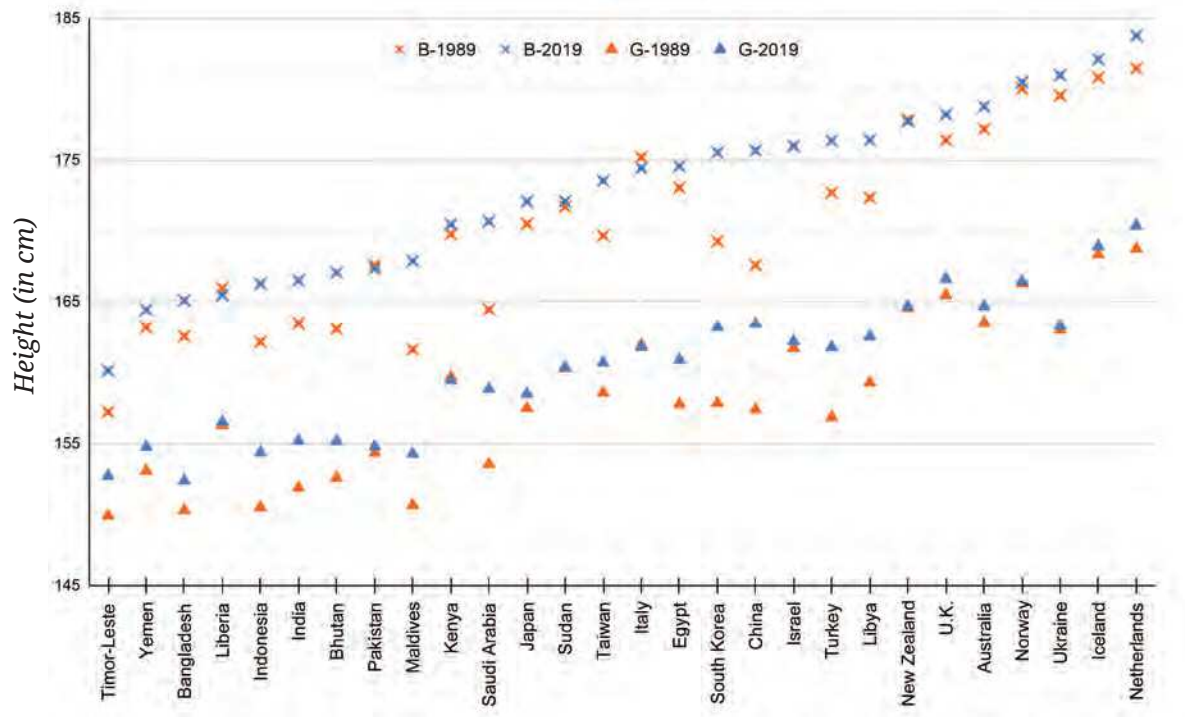
You may want to look at the data of weights. Or you might be curious to see if such patterns are present in other countries as well. You might also wonder if humans were much shorter a few centuries ago! Also, are people in some countries taller than others? The following visualisation shows the change in average heights of 19-year-old boys and girls of different countries from 1989 to 2019.

How is the graph organised? What information is presented?

What do you find interesting?

Math Talk

Notice that the vertical line starts from 145 cm — this helps give a closer (zoomed in) look at the heights.



A Mean Decision!

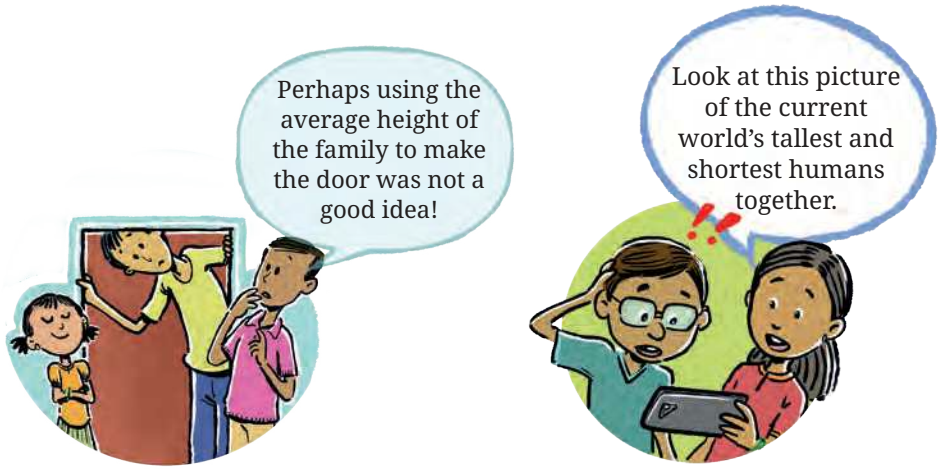
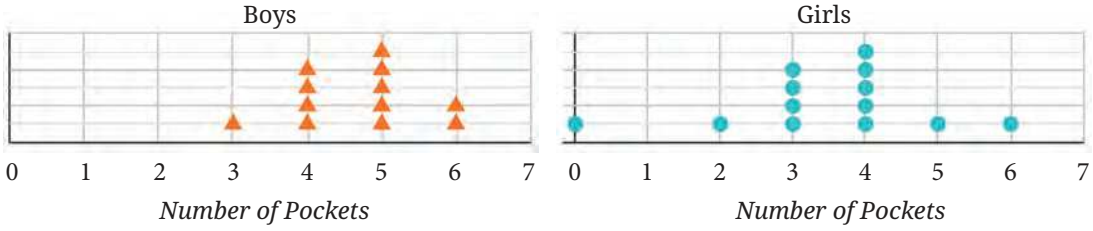


Figure it Out

1. The dot plots below show the distribution of the number of pockets on clothing for a group of boys and for a group of girls.



Based on the dot plots, which of the following statements are true?

(a) The data varies more for the boys than for the girls.

(b) The median number of pockets for the boys is more than that for the girls.

(c) The mean number of pockets for the girls is more than that for the boys.

(d) The maximum number of pockets for boys is greater than that for the girls.

2. The following table shows the points scored by each player in four games:

Player	Game 1	Game 2	Game 3	Game 4
A	14	16	10	10
B	0	8	6	4
C	8	11	Did not play	13

Now answer the following questions:

(a) Find the average number of points scored per game by A.

(b) To find the mean number of points scored per game by C, would you divide the total points by 3 or by 4? Why? What about B?

(c) Who is the best performer?

3. The marks (out of 100) obtained by a group of students in a General Knowledge quiz are 85, 76, 90, 85, 39, 48, 56, 95, 81 and 75. Another group's scores in the same quiz are 68, 59, 73, 86, 47, 79, 90, 93 and 86. Compare and describe both the groups performance using, mean and median.

4. Consider this data collected from a survey of a colony.

Favourite Sport	Cricket	Basket Ball	Swimming	Hockey	Athletics
Watching	1240	470	510	430	250
Participating	620	320	320	250	105

Choose an appropriate scale and draw a double-bar graph. Write down your observations.

5. Consider a group of 17 students with the following heights (in cm): 106, 110, 123, 125, 117, 120, 112, 115, 110, 120, 115, 102, 115, 115, 109, 115, 101. The sports teacher wants to divide the class into two groups so that each group has an equal number of students: one group has students with height less than a particular height and the other group has students with heights greater than the particular height. Suggest a way to do this. Can you guess the age of these students based on the tabular data in the 'Telling Tall Tales' section?

6. Describe the mean and median of heights of your class. You can visualise the heights on a dot plot.

7. There are two 7th grade sections at a school. Each section has 15 boys and 15 girls. In one section, the mean height of students is 154.2 cm. From this information, what must be true about the mean height of students in the other section?

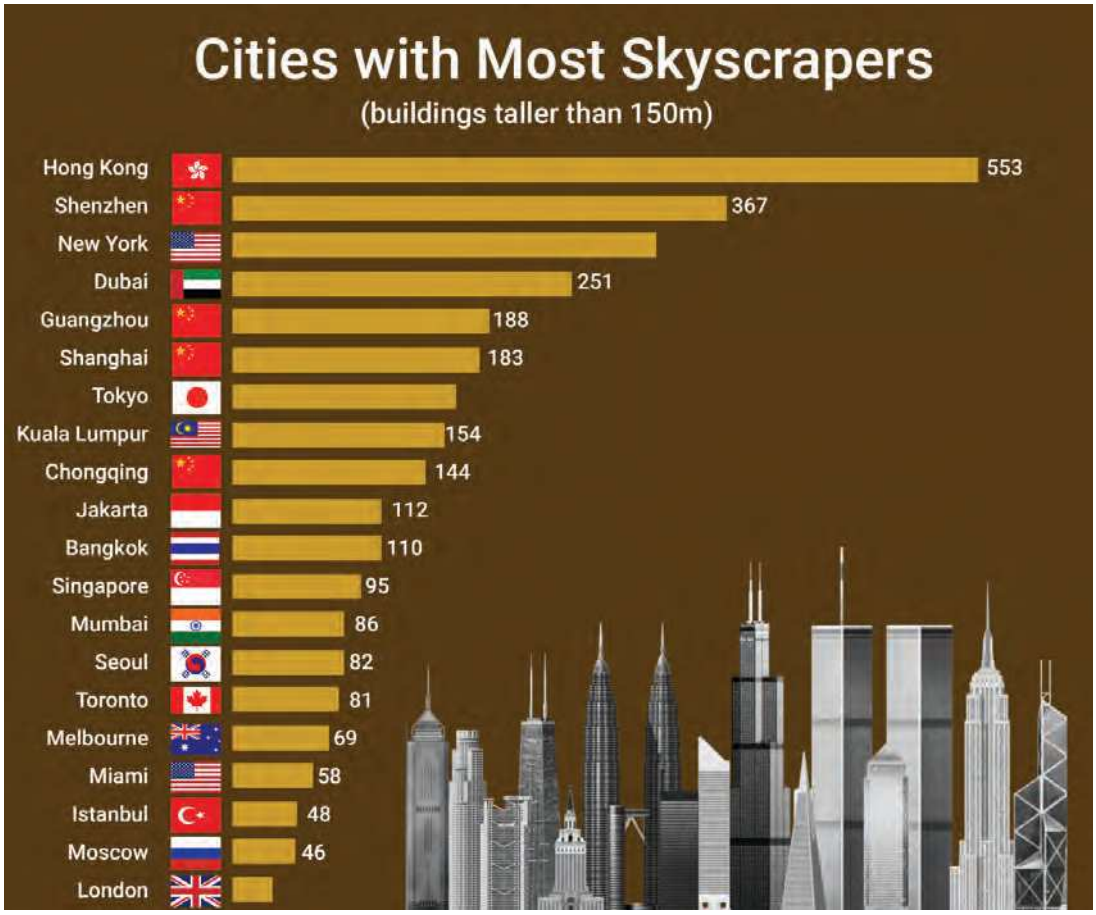
(a) The mean height of students in the other section is 154.2 cm.

(b) The mean height of students in the other section is less than 154.2 cm.

(c) The mean height of students in the other section is more than 154.2 cm.

(d) The mean height of students in the other section cannot be determined.

8. Standing tall in the storm.



(a) Write estimated values for the number of skyscrapers in New York, Tokyo, and London.

(b) Are the following statements valid?

(i) Only 12 cities have more skyscrapers than Mumbai.

(ii) Only 7 cities have fewer skyscrapers than Mumbai.

(iii) The tallest building in the world is in Hong Kong.

9. Estimate and then measure the objects listed in the following table. Draw a double bar graph based on the data. How accurate were your estimates? Find the average difference between the estimated and measured values.

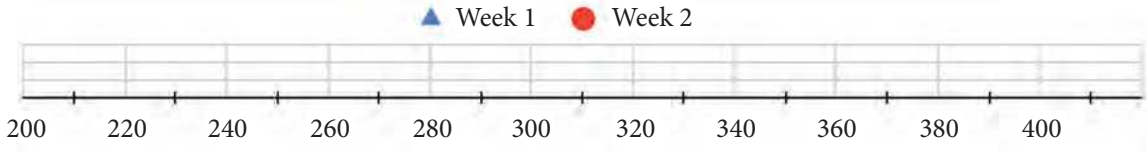
Object	Estimate (in cm)	Measure (in cm)	Positive Difference
Length of a pen			
Length of an eraser			
Length of your plam			
Length of your geometry box			
Length of your math notebook			

10. Aditi likes solving puzzles. She recently started attempting the 'Easy' level Sudoku puzzles. The time she took (in seconds) to solve these puzzles are — 410, 400, 370, 340, 360, 400, 320, 330, 310, 320, 290, 380, 280, 270, 230, 220, 240. The first nine values correspond to Week 1 and the rest to Week 2.



(a) Construct a dot plot below showing the data for both weeks.

(b) Describe the mean, median, and any observations you may have about the data.



11. **Individual Project:** Pick at least one of the following:

(a) How Long is a Sentence? Pick any two textbooks from different subjects. Choose any page with a lot of text from each book.

(i) Use a dot plot to describe how many words the sentences have on each page.

(ii) Compare the data of both the pages using mean and median.

(b) What is in a Name? Write down the names of all of your classmates. The following are some interesting things you can do with this data!

(i) Find the mean and median name length (number of letters in a name).

(ii) Visualise the data and describe its variability and central tendency.

(iii) Which starting letters are more popular? Which are less popular?

(iv) What is the median starting letter? What does this say about the number of names starting with the letters A – M and N – Z?

(v) Plot a double-bar graph showing the number of boys' names and girls' names that:

- start and end with vowels,
- start with vowels and end with consonants,
- start with consonants and end with vowels,
- start and end with consonants.

12. **Individual project (long term):** This requires collecting data over 2 weeks or more.
In and Out: Track how many times you step out of your house in a day. Do this for a month.

(i) Describe the variability and central tendency of this data. Make a dot plot.

(ii) Do you find anything interesting about this data? Share your observations.

(iii) You can ask any of your family members or friends to do this as well.



13. **Small-group project:** Pick at least one of the following. Make groups of 8 to 10. Collect data individually as needed. Put together everyone's data and do the appropriate analysis and visualisation.

(a) Our heights vs. our family's heights: Collect the heights of your family members.

(i) Make a dot plot showing heights of just your family members. Describe its variability and central tendency.

(ii) Make a double-bar graph showing each student's height next to their family's mean height.

(iii) Look at everyone's data and share your observations.

(b) Estimating time: Check the time and close your eyes. Open them when you think 1 minute has passed (no counting). Note down after how m-any seconds you opened your eyes. Collect this data for yourself and for your family members. Repeat this activity to estimate 3 minutes.

(i) Make two dot plots (for 1 minute and 3 minutes) showing estimates of just your family members.

(ii) Mark these on the respective dot plots. Describe its variability and central tendency.

(iii) Make a double bar graph showing each family's mean 1 minute estimate and mean 3 minute estimate.

(iv) Look at everyone's data and share your observations.

SUMMARY

- Dot plots help us get a quick glimpse of the variability of the data — minimum, maximum, range, and how data is clustered or spread out.
- The Arithmetic Mean = $\frac{\text{Sum of all the values in the data}}{\text{Number of values in the data}}$.
- The Median is the number in the middle of any sorted data. If there are an even number of values, then the median is the average of the two middle numbers.
- We can describe and compare data in several ways including by referring to the minimum, maximum, total, range, arithmetic mean, and median.
- We learnt how to read and make clustered bar graphs. These graphs can be used to compare and visualise values across categories and across time.
- Examining data can lead to new questions and directions to probe further.



Connect the Dots...

A number lock has a 3-digit code. Find the code using the hints below.

