I'm not robot

reCAPTCHA

Continue

Finding mean median and mode in a frequency table

When given a data set, it is possible to construct a frequency table in order to make the data easier to analyse. Example: 66 people were asked about how many bathrooms they had in their house. 30 people had 1 bathroom, 21 people had 2, 5 people had 3 and 7 people had 4 bathrooms. The rest had 5 or more. Use this information to construct a frequency table. We write the number of bathrooms in our left column, then frequency in the right. The only calculation that has to be done is $30+21+5+7 = 63$ As the question tells us there are 66 people in total, this must mean there are 3 people who have 5 or more bathrooms. The table therefore looks like this. Below is a frequency table of data based on a survey where 89 women were asked what their shoe size was. Calculate the mean, median, and mode of the data. Mode: Simply identify the shoe size with the highest frequency: 5. Median: The median is the $\frac{89 + 1}{2} = 45\text{th}$ term. We must find the 45th term from the bottom when in order. $5 + 12 + 18 = 35$, so the 35th person is the last one with size 5 feet. $5 + 12 + 18 + 19 = 54$, so the 54th person is the last one with size 5.5 feet. The 45th must therefore fall into the size 5.5 category, thus the median is 5.5. Mean: To calculate the mean, we need to multiply shoe size by the frequency to get a new row, as shown below; Next we add up this row to find the total shoe size. We divide total shoe size by the number of people. $\text{Mean} = \frac{\text{Total shoe size}}{\text{Total number of women}}} = \frac{512}{89} = 5.8$ (1 dp) a) The number of bathrooms with the highest frequency is the 1 bathroom category, so the mode is 1. To find the median, we need to find the middle value(s). In order to find the middle value(s), we need to find how many values there are in total. The number of values in total is the sum of all the frequencies: There are $30+21+5+7+3=66$ values in total. Since the number of values is an even number, this means that there is no single middle value, so we will need to locate the two middle values. To find the middle values we need to use the formula $\frac{n + 1}{2}$ where n represents the total number of values: $\frac{66 + 1}{2} = 33.5$ This means that the median is halfway between the 33rd and the 34th value. If we go back to the frequency table, we can see that the first 30 values are in the 1 bathroom category, and the following 23 values are in the 2 bathroom category. Therefore values 33 and 34 are in the 2 bathroom category. Since the 33rd and the 34th values are identical, then the median is simply 2 bathrooms. b) It is not possible to calculate the mean due to the fact that there is a category of '5 bathrooms or more'. We do not know exactly how many bathrooms people have who are in this category (they could have 5, they could have 500!). a) Working out the mode is the easy part. Which category was the most common (has the highest frequency)? 1 goal per game is therefore the mode. b) To find the median, we need to find how many values there are in total. The number of values in total is the sum of all the frequencies: There are $7+14+13+8+3+4+1=50$ values in total. Since the number of values is an even number, this means that there is no single middle value, so we will need to locate the two middle values. To find the middle values we need to use the formula $\frac{n + 1}{2}$ where n represents the total number of values: $\frac{50 + 1}{2} = 25.5$ This means that the median is halfway between the 25th and the 26th value. If we go back to the frequency table, we can see that the first 7 values are in the 0 goals category, and the following 14 values are in the 1 goal category. This means that the first 21 values fall in the 0 goal or the 1 goal category. The following 13 values fall into the 2 goal category, so values 25 and 26 must be in this category. Since the 25th and the 26th values are identical, then the median is simply 2 goals. c) The mean is the total number of goals divided by the total number of games. In this question, the frequency represents the total number of games which is 50 (which we had already calculated from the previous question). To work out the total number of goals, we need to multiply the number of goals by the frequency (if the team scored 5 goals on 4 occasions, then the team scored 20 goals in these 4 matches combined): 7times 0 goals = 0 goals 14times 1 goal = 14 goals 13times 2 goals = 26 goals 8times 3 goals = 24 goals 3times 4 goals = 12 goals 4times 5 goals = 20 goals 1times 6 goals = 6 goals Now that we know how many goals were scored in each category, we can work out the total number of goals scored: $\text{Total number of goals scored} = 0+14+26+24+12+20+6=102$ If the team scored 102 goals in 50 games, then the mean number of goals scored can be calculated as follows: 102 goals $\div$ 50 games = 2 goals (to the nearest goal) a) We know that a total of 240 divers were surveyed. This means that the total of the frequency column is 240. Therefore, if we subtract all the known values from 240, we can work out the value of x and y combined: $240-15-76-32-9=108$ divers Therefore $x + y = 108$ divers We have been told that the ratio of x to y is 7 : 5. This means that x is $\frac{7}{12}$ of the total. (We are dealing in twelfths here because the sum of the ratio is 12.) We can calculate the value of x as follows: $\frac{7}{12}\times\, 108$ divers = 63 divers We can calculate the value of y as follows: $\frac{5}{12}\times 108$ divers = 45 divers b) The modal number of shark encounters is the category with the highest frequency. This is clearly the 2 shark encounters category since 76 divers fall into this category, more than any other. c) To find the median, we need to find how many values there are in total. Since we have been told that there are 240 divers, we do not need to calculate this. Since the number of values is an even number, this means that there is no single middle value, so we will need to locate the two middle values. To find the middle values we need to use the formula $\frac{n + 1}{2}$ where n represents the total number of values: $\frac{240 + 1}{2} = 120.5$ This means that the median is halfway between the 120th and the 121st value. If we go back to the frequency table, we can see that the first 9 values are in the 0 shark encounters category, and the following 32 values are in the 1 shark encounter category. This means that the first 41 values fall in the 0 or the 1 shark encounter categories. The following 76 values fall into the 2 shark encounters category, so the first 117 values fall in the 0 or 1 or 2 shark encounter categories. The following 63 values fall in the 3 shark encounters category, so values 120 and 121 must be in this category. Since the 120th and the 121st values are identical, then the median is simply 3 shark encounters. d) The mean is the total number of shark encounters divided by the total number of divers (240). To work out the total number of shark encounters by the frequency: 9times 0 shark encounters = 0 shark encounters 32times 1 shark encounters = 32 shark encounters 76times 2 shark encounters = 152 shark encounters 63times 3 shark encounters = 189 shark encounters 45times 4 shark encounters = 180 shark encounters 15times 5 shark encounters = 75 shark encounters Now that we know how many shark encounters there are in each category, we can work out the total number of shark encounters: $\text{Total number of shark encounters} = 0+32+152+189+180+75=628$ If 240 divers had a total of 628 shark encounters, then the mean number of shark encounters can be calculated as follows: 628 shark encounters $\div$ 240 divers = 3 shark encounters (to the nearest whole number) In order to continue enjoying our site, we ask that you confirm your identity as a human. Thank you very much for your cooperation. We at Cuemath believe that Math is a life skill. Our Math Experts focus on the "Why" behind the "What." Students can explore from a huge range of interactive worksheets, visuals, simulations, practice tests, and more to understand a concept in depth. Book a FREE trial class today! and experience Cuemath's LIVE Online Class with your child. Introduction to Mean, Median And Mode We come across data every day. We find them in newspapers, articles, in our bank statements, mobile and electricity bills. The list is endless; they are present all around us. Now the question arises if we can figure out some important features of the data by considering only certain representatives of the data. This is possible by using measures of central tendency or averages. A measure of central tendency describes a set of data by identifying the central position in the data set as a single value. We can think of it as a tendency of data to cluster around a middle value. In statistics, the three most common measures of central tendencies are Mean, Median and Mode. Choosing the best measure of central tendency depends on the type of data we have. Let's begin by understanding the meaning of each of these terms. Mean The arithmetic mean of a given data is the sum of all observations divided by the number of observations. For example: A cricketer's scores in five ODI matches are as follows: 12, 34, 45, 50, 24 To find his average score in a match, we calculate the arithmetic mean of data using the mean formula: $\text{Mean} = \dfrac{\text{Sum of all observations}}{\text{Number of observations}}$ $\begin{align}\text{Mean} \&= \dfrac{12+34+45+50+24}{5}\\\text{Mean} \&= \dfrac{165}{5}\\ \&=33\end{align}$ Mean is denoted by $\bar{x}$ (pronounced as x bar) Let the mean be $\x_{1}\, \x_{2}\, \x_{3}\, ...\x_{n}$ be n observations, then the mean is: 1. $(\x_{1}+k)\, (\x_{2}+k)\, (\x_{3})+k)\, ...(\x_{n})+k)$? 2. $(\x_{1}\-k)\, (\x_{2}\-k)\, (\x_{3})\-k)\, ...(\x_{n})\-k)$? 3. $k(\x_{1})\, k(\x_{2})\, k(\x_{3})\, ...k(\x_{n})$? Types of Data Data can be present in raw form or tabular form. Let's find the mean in both cases. 1. Raw Data Let $\x_{1}\, \x_{2}\, \x_{3}\, ...\x_{n}$ be n observations. We can find the arithmetic mean using the mean formula. $\begin{align} \text{Mean}\, \bar{x}=\dfrac{\x_1+\x_2+...\x_n}{n}\end{align}$ Example 1 If the heights of 5 people are 142 cm, 150 cm, 149 cm, 156 cm, and 153 cm. Find the mean height. Mean height $\begin{align} \&=\bar{x} \\ \&=\dfrac{142+150+149+156+153}{5} \\ \&=\dfrac{750}{5} \\ \&=150 \end{align}$ Thus, the mean height is 150 cm. 2. Frequency Distribution (Tabular) Form When the data is present in tabular form, we use the following formula: $\begin{align}\text{Mean}\, \bar{x}=\dfrac{\x_1f_1 + x_2f_2+...x_nf_n}{f_1 + f_2+...f_n}\end{align}$ Consider the following example. Example 1 Find the mean of the following distribution: x 4 6 9 10 15 f 5 10 10 7 8 Solution Calculation table for arithmetic mean: xi fi xifi 4 5 20 6 10 60 9 10 90 10 7 70 15 8 120 $\\ \sum f_i=40\\ \sum x_i f_i=360\\ \begin{align} \therefore \text{Mean} \&=\bar{x} \\ \&=\dfrac{360}{40} \\ \&=9 \end{align}$ Example 2 Here is an example where the data is in the form of class intervals. The following table indicates the data on the number of patients visiting a hospital in a month. Find the average number of days visiting the hospital in a day. Number of patients Number of days visiting hospital 0-10 2 10-20 6 20-30 9 30-40 7 40-50 4 50-60 2 Solution In this case, we find the classmark (also called as mid-point of a class) for each class. Note: $\text{Class mark } =\dfrac{\text{lower limit + upper limit}}{2}$ Let $\x_{1}\, \x_{2}\, \x_{3})\, ...\x_{n}$ be the class marks of the respective classes. Hence, we get the following table: Class mark (xi) frequency (fi) xifi 5 2 10 15 6 90 25 9 225 35 7 245 45 4 180 55 2 110 Total $\\sum f_i=30\\ \\sum f_xi_i=860\\ \begin{align} \therefore \text{Mean} \&=1 + \dfrac{\dfrac{n}{2}-c}{f}\times h\\ \&= \dfrac{860}{30} \\ \&=28.67 \end{align}$ Median The value of the middlemost observation, obtained after arranging the data in ascending order, is called the median of the data. For example, consider the data: 4, 4, 6, 3, 2. Let's arrange this data in ascending order: 2, 3, 4, 4, 6. There are 5 observations. Thus, median = middle value i.e. 4 We can see here: 2, 3, 4, 4, 6 (Thus, 4 is the median) Case 1: Ungrouped Data Step 1: Arrange the data in ascending or descending order. Step 2: Let the total number of observations be n. To find the median, we need to consider if n is even or odd. If \(n\) is odd, then use the formula: $\text{ Median} = \dfrac{(n+1)}{2})^{th}$ {observation}\\ Example 1 Let's consider the data: 56, 67, 54, 34, 78, 43, 23. What is the median? Solution Arranging in ascending order, we get: 23, 34, 43, 54, 56, 67, 78. Here, n (no.of observations) = 7 So, $\begin{align} \dfrac{7+1}{2}\&=4 \\\therefore Median \&= 4^{th}\text { observation}\end{align}$ If \(n\) is even, then use the formula: $\text{ Median} = \dfrac{ \dfrac{n}{2}^{th}\text{ obs.}+ (\dfrac{n}{2}+1)^{th}\text{ obs.}}{2}$ Example 2 Let's consider the data: 50, 67, 24, 34, 78, 43. What is the median? Solution Arranging in ascending order, we get: 24, 34, 43, 50, 67, 78. Here, n (no.of observations) = 6 $\\dfrac{6}{2}=3$ Using the median formula, $\begin{align} \text { Median } \&= \dfrac{ 3^{rd}\text{ obs.}+ 4^{th}\text{ obs.}}{2}\end{align}$ Case 2: Grouped Data When the data is continuous and in the form of a frequency distribution, the median is found as shown below: Step 1: Find the median class. Let n = total number of observations i.e. $\\sum f_i$ Note: Median Class is the class where $\dfrac{n}{2}$ lies. Step 2: Use the following formula to find the median. $\text{ Median} = l + \dfrac{(\dfrac{n}{2}-c\{f})}{\times h}$ where, $\(l= \)$ lower limit of median class $\(c= \)$ cumulative frequency of the class preceding the median class $\(f= \)$ frequency of the median class $\(h= \)$ class size Let's consider the following example to understand this better. Example 1 Find the median marks for the following distribution: Classes 0-10 10-20 20-30 30-40 40-50 Frequency 2 12 22 8 6 Solution We need to calculate the cumulative frequencies to find the median. Calculation table: Classes Number of students Cumulative frequency 0-10 2 2 10-20 12 2 + 12 = 14 20-30 22 14 + 22 = 36 30-40 8 36 + 8 = 44 40-50 6 44 + 6 = 50 N=50 \\ \dfrac{N}{2} =15=p\\ \sum {f_i}=25=p\\ \\sum {f_i}_1\\= \(530+20p\) \begin{align} \text{ Mean} \&=1 + \dfrac{ \dfrac{n}{2}-c}{f})\times h \\\&=20+\dfrac{25-14}{22} \times 10 \\\&= 20 + \dfrac{11}{22} \times 10 \\\&=20+5= 25 \end{align} \\therefore\) Median = 25 Mode The value which appears most often in the given data i.e. the observation with the highest frequency is called a mode of data. Case 1: Ungrouped Data For ungrouped data, we just need to identify the observation which occurs maximum times. Mode = Observation with maximum frequency The example discussed above has only 1 mode, so it is unimodal. Case 2: Grouped Data When the data is continuous, the mode can be found using the following steps: Step 1: Find modal class i.e. the class with maximum frequency. Step 2: Find mode using the following formula: $\text{ Mode} = l + \dfrac{ f_m-f_1}{2f_m-f_1-f_2}\times h)$ where, $\(l= \)$ lower limit of modal class, $\(f_m= \)$ frequency of modal class, $\(f_1= \)$ frequency of class preceding modal class, $\(f_2= \)$ frequency of class succeeding modal class, $\(h= \)$ class width Consider the following example to understand the formula. Example 1 Find the mode of the given data: Marks Obtained 0-20 20-40 40-60 60-80 80-100 Number of students 5 10 12 6 3 Solution The highest frequency \(=\) 12, so the modal class is 40-60. \(l= \) lower limit of modal class = 40 \(f_m= \) frequency of modal class =12 \(f_1=\)frequency of class preceding modal class = 10 \(f_2=\)frequency of class succeeding modal class = 6 \(h=\)class width \(=\) 20 Using the mode formula, $\begin{align} \text{ Mode} \&= l + [\dfrac{ f_m-f_1}{2f_m-f_1-f_2}]\times h\\\&= 40+[\dfrac{12-10}{2 \times 12 - 10-6}]\times 20\\\&= 40+[\dfrac{2}{8}]\times 20\\\&= 45 \end{align}\) Let's try finding mean, median, and mode in the case mentioned below. Instructions: Click on the green board and observe the mean, median, and mode of the numbers appearing on the screen. Verify the answer by calculating using the formula discussed above. Help your child score higher with Cuemath's proprietary FREE Diagnostic Test. Get access to detailed reports, customized learning plans, and a FREE counseling session. Attempt the test now. Relation Between Mean, Median and Mode The three measures of central values i.e. mean, median, and mode are closely connected by the following relations (called an empirical relationship). \(\text{2Mean + Mode =3Median} \) For instance, if we are asked to calculate the mean, median, and mode of continuous grouped data, then we can calculate mean and median using the formulae as discussed in the previous sections and then find mode using the empirical relation. For example, we have a data whose median \(=\) 65 and median \(=\) 61.6. Then, we can find the mean using the above relation. $\begin{align} 2 \text {Mean}+ \text {Mode} \&=3 \text { Median } \\ \therefore 2 \operatorname{Mean} \&=3 \times 61.6-65 \\\therefore 2 \operatorname{Mean} \&=119.8 \\\Rightarrow \text {Mean} \&=\dfrac{119.8}{2} \\\Rightarrow \text {Mean} \&= 59.9 \end{align}$ Solved Examples on Mean, Median and Mode If the mean of the following data is 20.6, find the missing frequency (p). x 10 15 20 25 35 f 3 10 p 7 5 Solution: Let us make the calculation table for this : xi fi xifi 10 3 30 15 10 150 20 20 p 20 x p = 20p 25 7 175 35 5 175 Total: \\sum {f i}=25+p\\ \\sum {f xi_i}=\(530+20p\) \begin{align} \text{ Mean} \&=\dfrac{530+ 20p}{25+p} \\ 530 + 20p\&=515+20.6p \\15 \&= 0.6p \\\p \&=\dfrac{15}{0.6}\\\&=25 \end{align}\) Thus, the missing frequency (p) =25 The mean of 5 numbers is 18 If one number is excluded, their mean is 16 Find the excluded number. Solution: Given, n = 5, \(\bar{x}\) = 18 \(\therefore\) \(\bar{x}\) = \(\dfrac{\sum {x_i}}{n}\) \\sum {x_i}\= \(5\times 18\) = 90 Thus, the total of 5 numbers = 90 Let the excluded number be \(a\). Therefore, total of 4 numbers\(=90-a)\\ \\begin{align} \text {Mean of 4 numbers} \&=\dfrac{90-a}{4} \\ 16\&= \dfrac{90-a}{4} \\ 90-a\&= 64 \\\end{align}\\ \\therefore \text{a} = 26 \) A survey on the heights (in cm) of 50 girls of class X was conducted at a school and the following data was obtained: Height (in cm) 120-130 130-140 140-150 150-160 160-170 Total Number of girls 2 8 12 20 8 50 Solution \begin{align} \text {Maximum class} \&= 150-160\end{align}\) {as it has maximum frequency} \(\begin{align}\text{ Modal class} \&= 150-160\end{align}\) as l= 150, h= 10, f_m =20, f_1 =12,f_2 =8 \\ Mode\&= l +\dfrac{ f_m-f_1}{2f_m-f_1-f_2}\times h\\ \\therefore \text {Mode}\&=154 \end{align}\) To find the median, we need cumulative frequencies. Consider the table: Class Intervals No. of girls (fi) Cumulative frequency (c) 120-130 2 2 130-140 8 2+8=10 140-150 12 f= f1 10+12=22 (c) 150-160 20 fm 22+20=42 160-170 8 f2 42+8=50 (n) \begin{align} n \&= 50 \\ \dfrac {n}{2} &=25 \\\therefore \text{ Median class} \&= 150-160\\\ l =150, c= 22, f\&=20, h= 10\\\ \text{ Median} \&=l + \dfrac{\dfrac{n}{2}-c}{f} \times h \\\&=150+\dfrac{25-22}{20}\times 10 \\\&= 150+ 1.5 \\\&= 151.5 \end{align}\) Mode = 154, Median= 151.5 Heena and Harry received their answer sheets for their mid-term exams. There were five questions, each carrying ten marks. Their scores were as follows: Question Number 1 2 3 4 5 Heena's score 10 8 9 8 7 Harry's score 4 7 10 10 10 Both of them got into an argument. Heena said that she had a better average score so she was a better performer. Whereas Harry argued that he had scored 10 out of 10 most number of times so he was a better performer. Help them decide who performed better. Difference between Mean and Average The term average is frequently used in everyday life to denote a value that is typical for a group of quantities. Average rainfall in a month or the average age of employees of an organization is a typical examples. We might read an article stating "People spend an average of 2 hours every day on social media. " We understand from the use of the term average that not everyone is spending 2 hours a day on social media but some spend more and some less. However, we can understand from the term average that 2 hours is a good indicator of the amount of time spent on social media per day. Most people use average and mean interchangeably even though they are not the same. Average is the value that indicates what is most likely to be expected. They help to summarise large data into a single value. An average tends to lie centrally with the values of the observations arranged in ascending order of magnitude. So, we call an average measure of the central tendency of the data. Averages are of different types. What we refer to as mean i.e. the arithmetic mean is one of the averages. Mean is called the mathematical average whereas median and mode are positional averages. Difference Between Mean and Median Mean is known as the mathematical average whereas the median is known as the positional average. To understand the difference between the two, consider the following example. A department of an organization has 5 employees which include a supervisor and four executives. The executives draw a salary of ₹10,000 per month while the supervisor gets ₹40,000. Mean, \(\begin{align} \&=\dfrac{10000+10000+10000+10000+40000}{5}\\\&=\dfrac{80000}{5}=16000\end{align}\) Thus, the mean salary is ₹16,000. To find the median, we consider the ascending order: 10000, 10000, 10000, 10000, 40000. \(n=5, so\dfrac{n+1}{2}=3\) . Thus, the median is the 3rd observation. \((Median = 10000)\) Thus, the median is ₹10,000 per month. Now let us compare the two measures of central tendencies. We can observe that the mean salary of ₹16,000 does not give even an estimated salary of any of the employees whereas the median salary represents the data more effectively. One of the weaknesses of mean is that it gets affected by extreme values. Look at the following graph to understand how extreme values affect mean and median: So, mean is to be used when we don't have extremes in the data. If we have extreme points, then the median gives a better estimation. Here's a quick summary of the differences between the two. Mean Vs Median Mean Median Definition Average of given data (Mathematical Average) Central value of data (Positional Average) Calculation Add all values and divide by total number of observations Arrange data in ascending / descending order and divide by total number of observations Greatly affected by extreme points Doesn't get affected by extreme points Practice Questions Here are a few activities for you to practice. Select/Type your answer and click the "Check Answer" button to see the result. IMO (International Maths Olympiad) is a competitive exam in Mathematics conducted annually for school students. It encourages children to develop their math solving skills from a competition perspective. You can download the FREE grade-wise sample papers from below: To know more about the Maths Olympiad you can click here Frequently Asked Questions (FAQs) It is the value that appears maximum number of times in data. It is best to put the numbers in order and then, count the number of times each value occurs. The value with the highest frequency will be the modal value of the data. A mode represents the most common value. On a bar chart, the mode is the highest bar. It is used with categorical data such as most sold T-shirts size. Median is the value of the middlemost observation, obtained after arranging the data in ascending order. To find the same, we need to consider two cases. \(\text {Median= } (\dfrac {n+1}{2})^{ th}\text { observation}\), \(\text {when n is odd }\) \(\text {Median= } \dfrac {\dfrac {n}{2} ^{th}\text {obs.}+ (\dfrac{n}{2}+1)^{ th}\text {obs.}}{2}\), \(\text {when n is even }\) For grouped data, the median is obtained using the median formula: \(\text {Median= } l + [\dfrac {\dfrac{n}{2}-c}{f}]\times h\) No, mean and median are not the same. Mean is the average of the given set of numbers. We need to add the numbers up then divide their sum by the number of observations. For finding the mode, we find whether any number appears more than once. The number which appears most is the mode. If there are other numbers that repeat to the same level, there may be more than one mode. A set could be bimodal or trimodal. But the mean of a given data is unique.