Median and Size of Data Location

EGI SAFITRI, S.MAT., M.SI

Median

The median is the middle value of a dataset, or the value that divides the data into two equal parts.

Steps to determine the median:

1. **Sort the data** from the smallest to the largest.
2. **Determine the position** of the median. $$L_{Me} = \frac{n+1}{2}$$
3. **Find the median value:**
   - If the number of data points is **odd**:
     The median is the value in the exact middle position.
     $$Me = X_{L_{Me}}$$
   - If the number of data points is **even**:
     The median is the **average of the two middle values.**
     $$Me = \frac{1}{2}\left[X_{\frac{1}{2}n} + X_{\frac{1}{2}n+1} \right]$$

Median

The **median** is the middle value in an ordered dataset, dividing it into two equal parts.

Steps to Determine the Median:

1. **Sort** the data in ascending order.
2. **Count** the number of data points n.
3. Use the following **formulas** based on whether n is odd or even:

Median

1. If the number of data points is odd (n is odd):

$$Me = x_{\frac{n+1}{2}}$$

Where:

- Me = Median
- x = value at the $\frac{n+1}{2}$-th position

2. If the number of data points is even (n is even):

$$Me = \frac{x_{\frac{n}{2}} + x_{\frac{n}{2}+1}}{2}$$

Where:

- $x_{\frac{n}{2}}$ = the value at the $\frac{n}{2}$-th position
- $x_{\frac{n}{2}+1}$ = the value at the next position

Median of Ungrouped (Single) Data

If a set of data from observations is given as 5, 8, 10, 4, 10, 7, 12. **Find the Median**

Sort the data: 4, 5, 7, 8, 10, 10, 12

$$L_{Me} = \frac{7+1}{2} = 4$$

$$Me = X_4 = 8$$

So, the **median is 8.**

Median

Grouped Data

$$Me = b + p\left(\frac{\frac{n}{2} - F}{f}\right)$$

Where:

- b = lower boundary of the median class
- p = class width
- n = total frequency
- F = cumulative frequency before the median class
- f = frequency of the median class

Example

Class Interval	f	F
31 – 40	2	2
41 – 50	3	5
51 – 60	5	10
61 – 70	14	24
71 – 80	24	48
81 – 90	20	68
91 - 100	12	80
Total	80	

From the table:

- $n = 80,$
- $b = 70.5,$
- $p = 10,$
- $F = 24,$
- $f = 24$

$$Me = 70{,}5 + 10\left(\frac{\frac{80}{2} - 24}{24}\right) = 77{,}167$$

Example

Another Example – Grouped Data

Skor	Frekuensi
45 – 49	5
50 – 54	10
55 – 59	8
60 – 64	12
65 – 69	6
70 – 74	4
75 – 79	5
Jumlah	50

Example

Skor	*f*	*Frekuensi Kumulatif*
45 – 49	5	5
50 – 54	10	15
55 - 59	8	23
60 – 64	12	35
65 – 69	6	41
70 – 74	4	45
75 – 79	5	50
Jumlah	**50**	

Median class is the class that contains the $\frac{n}{2}$-th data = 25th data point (since $n = 50$).

Class: 60–64

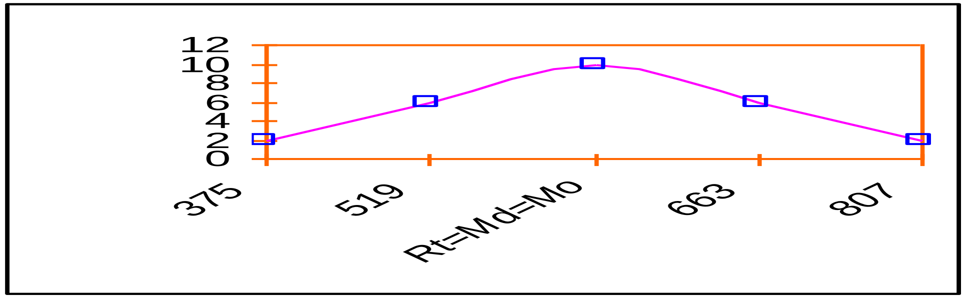
- $Tb = 59.5$ (lower boundary)
- $p = 5$
- $n = 50$, so $\frac{n}{2} = 25$
- $F = 23$ (cumulative frequency before median class)
- $f = 12$

$$Me = 59.5 + 5\left(\frac{25 - 23}{12}\right) = 60.33$$

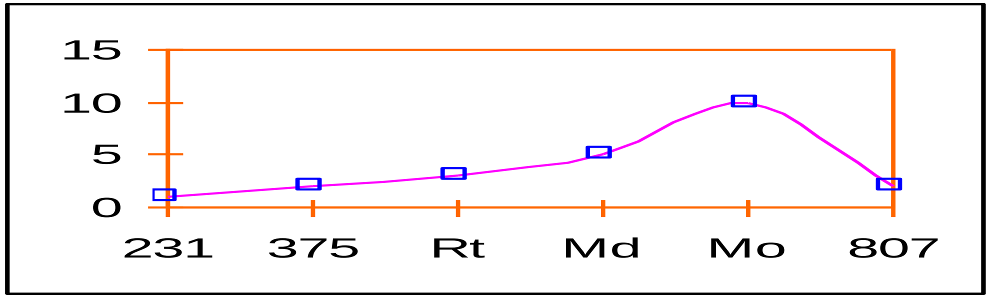
So, **the median is 60.33.**

Mean-Median-Modus Relationship

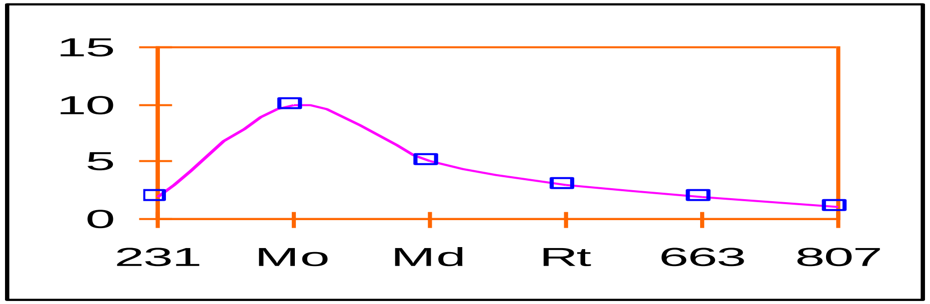
1. Symmetrical curveX= Md= Mo



3. Right skewed curve X < Md < Mo



2. Left skewed curve Mo < Md < X



A. Definition of Quartiles (Q)

Quartiles are values that divide a data set into four equal parts, after the data has been sorted from the smallest to the largest.

B. Quartiles of Ungrouped Data

Suppose we have data $x_1, x_2, x_3, \ldots, x_n$, with $x_1 < x_2 < x_3 < \ldots < x_n$, then the quartile is calculated as:

$$Q_i = \text{the value at position } \frac{i}{4}(n+1)$$

Explanation:

- Q_i = the i-th quartile
- n = total number of data points

$$Q_1 = \text{value at position } \frac{1}{4}(n+1) \quad \text{(lower quartile)}$$

$$Q_2 = \text{value at position } \frac{2}{4}(n+1) \quad \text{(median)}$$

$$Q_3 = \text{value at position } \frac{3}{4}(n+1) \quad \text{(upper quartile)}$$

Quartiles

Quartiles are values that divide an ordered dataset into **four equal parts.**

Steps to determine quartiles:

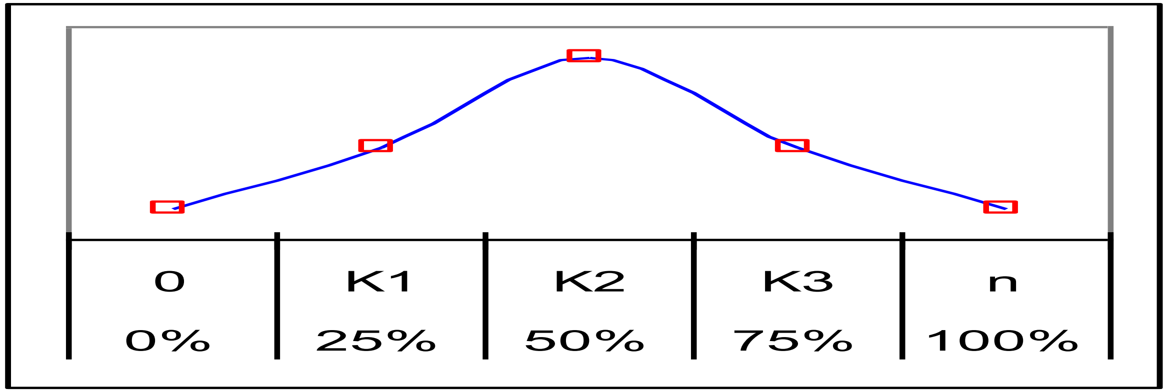
1. Sort the data from smallest to largest
2. Determine the position of the quartile
3. Calculate the quartile value using:

$$LK_i = \frac{i(n+1)}{2} = a,b$$

$$K_i = X_a + 0,b\left[X_{a+1} - X_a\right]$$

Quartile Locations

Ungrouped Data		Grouped Data
K1	= [1(n + 1)]/4	1n/4
K2	= [2(n + 1)]/4	2n/4
K3	= [3(n + 1)]/4	3n/4



Example

Given data:

7, 5, 8, 7, 9, 6, 6, 6, 8, 5, 9, 8, 6, 7, 9

Sorted:

5, 5, 6, 6, 6, 7, 7, 7, 8, 8, 8, 9, 9, 9, 9

$n = 15$

- Q_1 (1st quartile):

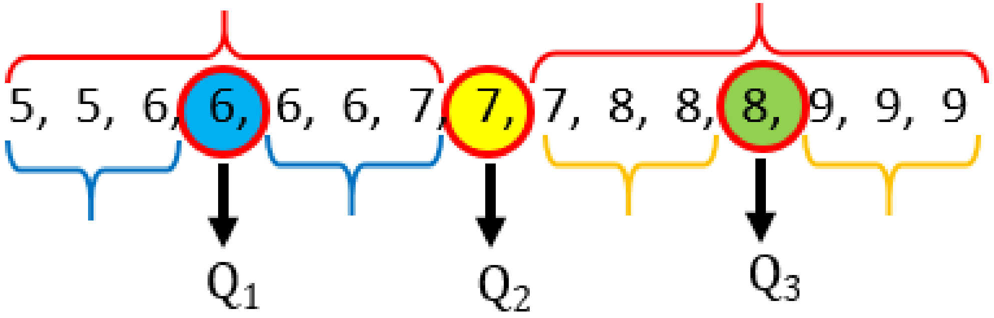
$$Q_1 = \text{value at } \frac{1}{4}(15+1) = \text{value at position } 4 = 6$$

- Q_2 (Median):

$$Q_2 = \text{value at } \frac{1}{2}(15+1) = \text{value at position } 8 = 7$$

- Q_3 (3rd quartile):

$$Q_3 = \text{value at } \frac{3}{4}(15+1) = \text{value at position } 12 = 8$$



Example

Suppose a sample has the following data: 78, 82, 66, 57, 97, 64, 56, 92, 94, 86, 52, 60, 70. Determine: a) K1 and b) K3

- Short the data:

  52, 56, 57, 60, 64, 66, 70, 78, 82, 86, 92, 94, 97

$$LK_1 = \frac{1(13+1)}{4} = 3{,}5 \qquad K_1 = X_3 + 0{,}5[X_4 - X_3] = 57 + 0{,}5(60 - 57) = 58{,}5$$

$$LK_3 = \frac{3(13+1)}{4} = 10{,}5 \qquad K_3 = X_{10} + 0{,}5[X_{11} - X_{10}] = 86 + 0{,}5(92 - 86) = 89$$

Quartile

Grouped Data

Steps to determine quartiles in grouped data:

1. Determine the quartile position: $LK_i = \frac{i(n+1)}{2}$

2. Find the value of the quartile:

$$K_i = b + p\left(\frac{\frac{i \cdot n}{4} - F}{f}\right)$$

Explanation:

- b = lower boundary of the quartile class
- p = class width
- F = cumulative frequency before the quartile class
- f = frequency of the quartile class
- $n = \sum_{i=1}^{n} f_i$ = total frequency

Example

Class Interval	f	F
31 – 40 41 – 50 51 – 60 61 – 70 71 – 80 81 – 90 91 - 100	2 3 5 14 24 20 12	2 5 10 24 48 68 80
Total	80	

$$LK_3 = \frac{3(80+1)}{4} = 60,75$$

Based on the adjacent table, we have:

- $b = 80.5$
- $p = 10$
- $F = 48$
- $f = 20$

$$K_3 = 80,5 + 10\left(\frac{\frac{3\times 80}{4} - 48}{20}\right) = 86,5$$

Exercise

The following data represents the survival to death, measured to the nearest tenth of a minute, of a random sample of 60 flies that have been sprayed with a new chemical in a laboratory experiment.

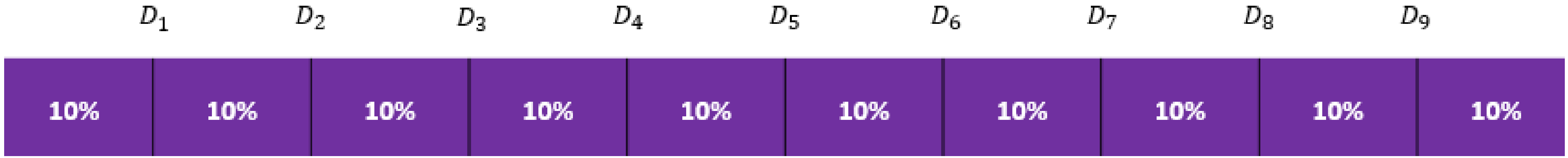
2.4	1.6	3.2	4.6	0.4	1.8	2.7	1.7	5.3	1.2
0.7	2.9	3.5	0.9	2.1	2.4	0.4	3.9	6.3	2.5
3.9	2.6	1.8	3.4	2.3	1.3	2.8	1.1	0.2	2.1
2.8	3.7	3.1	1.5	2.3	2.6	3.5	5.9	2.0	1.2
2.8	3.7	3.1	1.5	2.3	2.6	3.5	5.9	2.0	1.2
1.3	2.1	0.3	2.5	4.3	1.8	1.4	2.0	1.9	1.7

Using both single data and group data, determine:

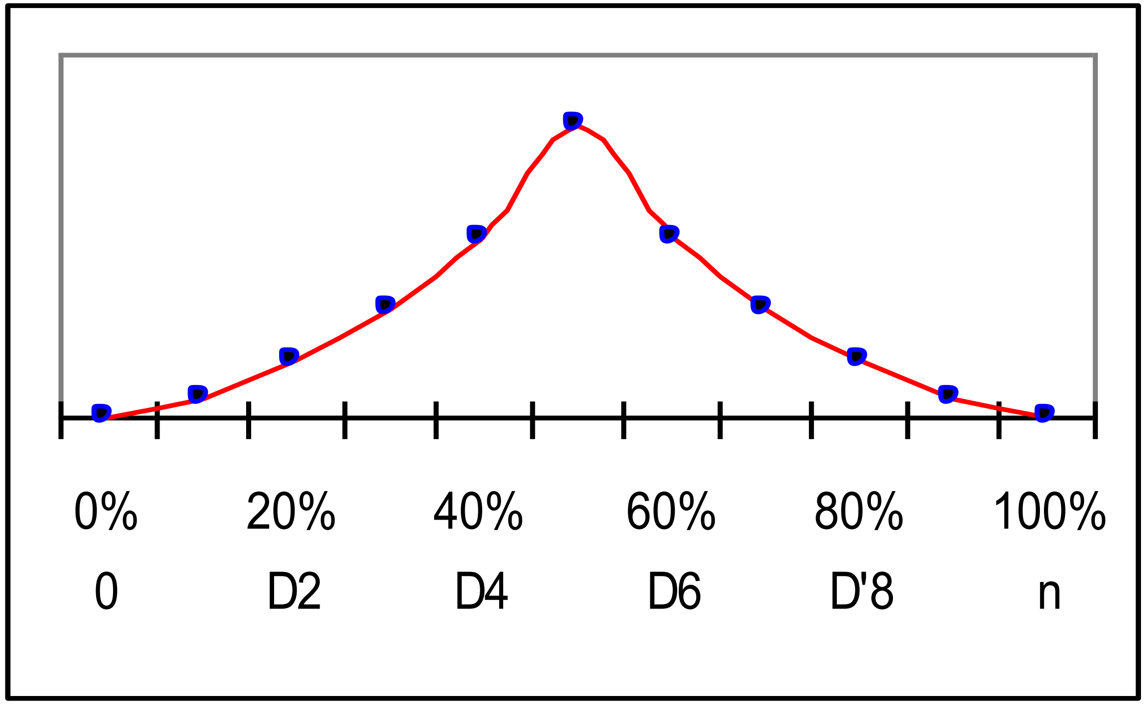
1. Median
2. Mode
3. Quartile Q1 and Q3

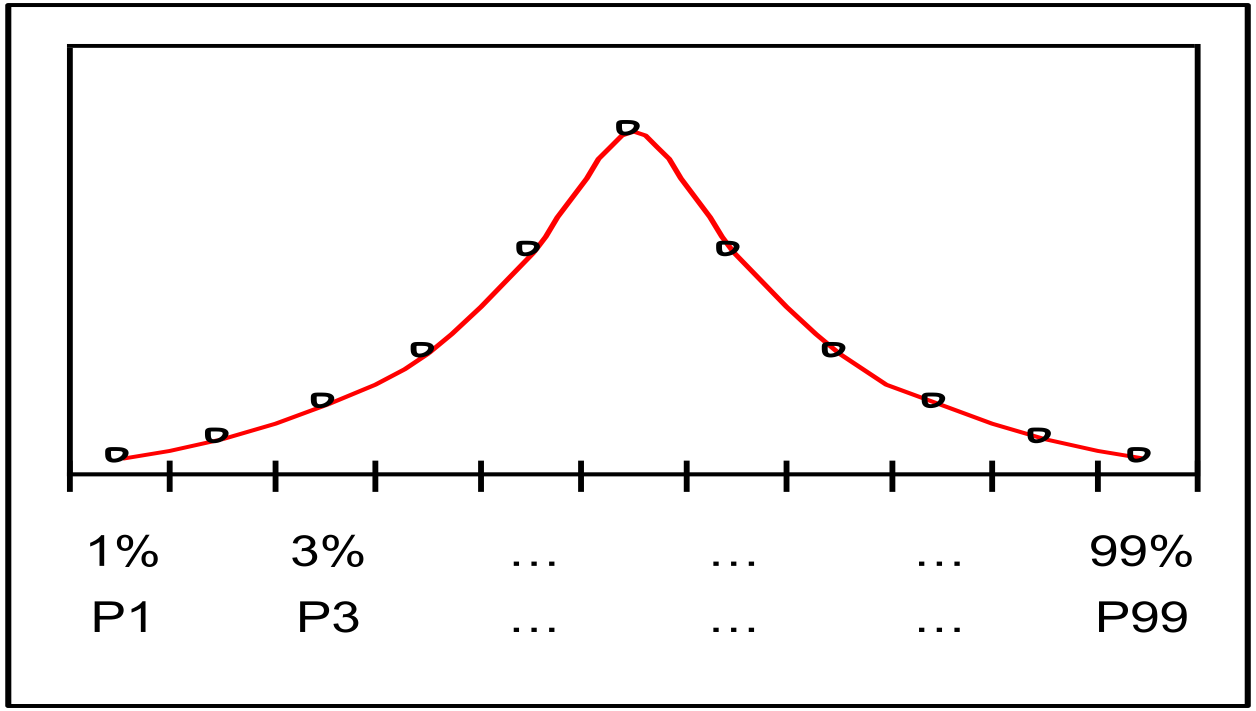
Deciles and Percentiles

Decile is a statistical term that divides grouped data into ten equal parts.
There are nine decile values that divide the dataset into ten equal portions.
Each part represents **1/10 or 10%** of the data group.
The 5th decile (D_5) is the **median** of the dataset divided by deciles.

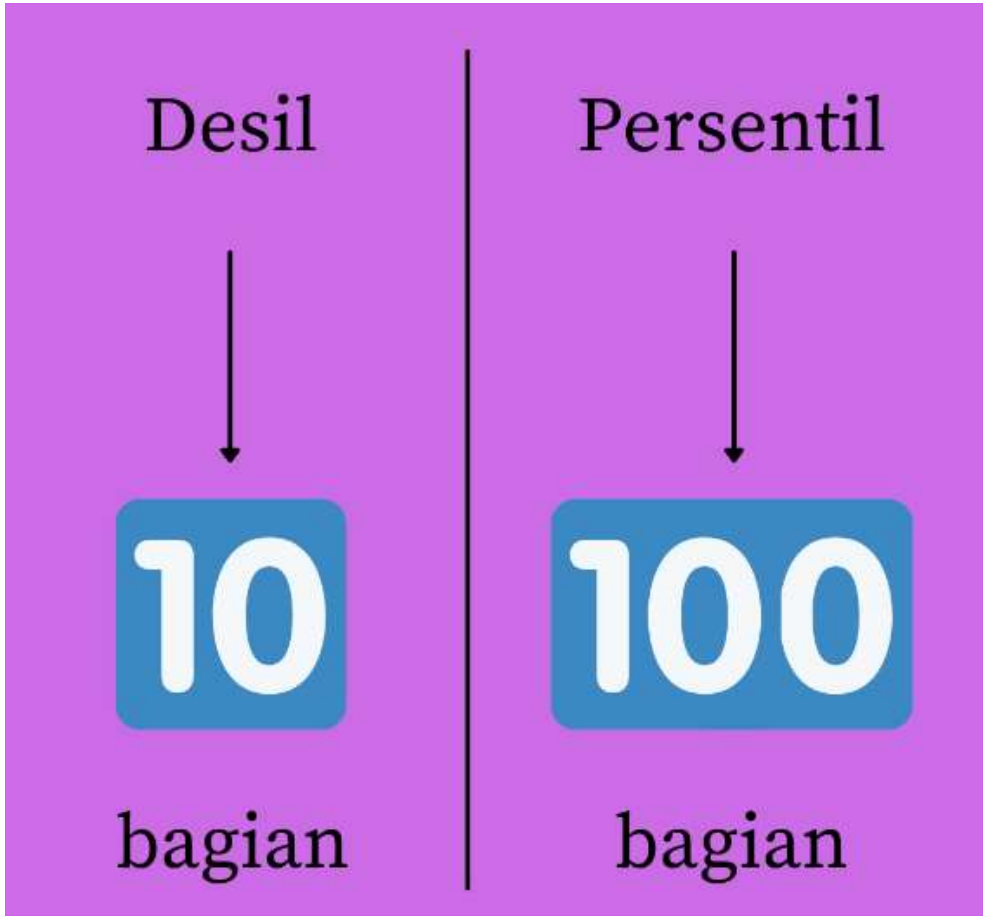


Decile and Percentile location chart





Decile vs Percentile



Ungrouped data decile

🧮 Formula for Deciles in Ungrouped Data

To find the i-th decile in ungrouped data:

$$D_i = x_{\frac{i}{10}(n+1)}$$

$$D_i = \frac{i(n+1)}{10} = a, b$$

$$D_i = x_a + 0, b[x_{a+1} - x_a]$$

Where:

- D_i = the i-th decile
- n = number of data points

Example

Find 3 Deciles from the Following Data

50, 15, 35, 59, 99, 96, 45, 70, 75, 28, 80, 41, 65, 85, 21, 15, 16, 41, 65, 85, 98, 15, 54, 99, 38, 83, 72, 17, 99, 23, 16, 45, 20, 55, 96, 99, 23, 30, 99, 72, 32, 17, 40, 82, 59, 67, 88, 43, 15, 75, 99, 95, 78, 63, 37, 53, 49, 99, 71, 98, 99, 85, 60, 98, 92, 97, 95, 55, 89, 94, 97, 84, 93, 90, 89

Solution

The number of data is 75 (n=75), then use the Single Data Decile formula. In the problem above, only 3 deciles are requested. The deciles chosen are for example Decile 1, Decile 3 and Decile 8.

$$D_1 = x_{\frac{1}{10}(n+1)} = x_{\frac{1}{10}(75+1)} = x_{\frac{1}{10}(76)} = x_{7,6}$$

$$D_3 = x_{\frac{3}{10}(n+1)} = x_{\frac{3}{10}(75+1)} = x_{\frac{3}{10}(76)} = x_{22,8}$$

$$D_8 = x_{\frac{8}{10}(n+1)} = x_{\frac{8}{10}(75+1)} = x_{\frac{8}{10}(76)} = x_{60,8}$$

From the above calculation, Decile 1 is the 7th.6 data, Decile 3 is the 22.8 data and Decile 8 is the 60.8 data.

Sorted Data (Ascending Order):

- 15, 15, 15, 15, 16, 16, 17, 17, 20, 21, 23, 23, 28, 30, 32, 35, 37, 38, 40 41, 41, 43, 45, 45, 49, 50, 53, 54, 55, 55, 59, 59, 60, 63, 65, 65, 67, 70, 71, 72, 72, 75, 75, 78, 80, 82, 83, 84, 85, 85, 85, 88, 89, 89, 90, 92, 93, 94, 95, 95, 96, 96, 97, 97, 98, 98, 98, 99, 99, 99, 99, 99, 99, 99, 99

$$\begin{aligned} D_1 &= x_{7,6} \\ &= x_7 + 0{,}6(x_8 - x_7) \\ &= 17 + 0{,}6(17 - 17) \\ &= 17 \end{aligned}$$

$$\begin{aligned} D_3 &= x_{22,8} \\ &= x_{22} + 0{,}8(x_{23} - x_{22}) \\ &= 43 + 0{,}8(45 - 43) \\ &= 43 + 1{,}6 \\ &= 44{,}6 \end{aligned}$$

$$\begin{aligned} D_8 &= x_{60,8} \\ &= x_{60} + 0{,}8(x_{61} - x_{60}) \\ &= 95 + 0{,}8(96 - 95) \\ &= 95 + 0{,}8 \\ &= 95{,}8 \end{aligned}$$

Percentiles

📘 Percentile Formula for Ungrouped Data

$$P_i = \text{value at position } \frac{i(n+1)}{100}$$

Where:

- P_i = the i-th percentile
- n = number of data points

Percentiles

📗 Percentile Formula for Grouped Data

$$P_i = Tb + \left(\frac{\frac{i}{100}n - f_k}{f_i} \right) p$$

Where:

- P_i = the i-th percentile
- Tb = lower boundary of the percentile class
- n = total number of data
- f_k = cumulative frequency before the percentile class
- f_i = frequency of the percentile class
- p = class width

Example

Feel free to search for examples of percentile problems independently.