

PERCENTILE



What is percentile? 🤔

In statistics, a percentile indicates how a particular score compares to others within the same dataset. Although there isn't a single universal definition, it is generally described as the percentage of data points that fall below a specific value. or A **percentile** is a way to show how your result compares to others.

Example 📖

Imagine you're in a class with 100 students who took a test. If your score is in the **90th percentile**, it means your score is better than **90% of the students**.

In simple terms: **A percentile tells you how many people scored less than you.** It's like your rank but expressed in percentages!

How to identify the percentile ? 🧮

Identifying the percentile generally follows two methods: one is the **rank method**, and another is the **linear interpolation method**. In this method, calculating the percentile through linear interpolation is the standard method. Let's dive deep into the linear interpolation method.

Steps

❑ Sort the dataset

❑ Calculate the position (index):

$$I = (p/100) * (n + 1)$$

where

I is the position or rank.

p is the percentile value.

n is the amount of data in dataset.

🔗 If the rank or position is a **whole number**, use that position directly; if it's a decimal, identify the **two adjacent values surrounding this position** and step 3.

📐 Interpolate: Calculate the percentile using:

$$P = L + (H - L) \times d$$

where:

L is the lower value (at the integer part of rank).

H is the upper value (at one position higher).

d is the decimal part of your calculated rank or position.

Practical Implementation in Statistics 📊

Dataset = {2, 2, 3, 4, 5, 5, 5, 6, 7, 8, 8, 8, 8, 8, 9, 9, 10, 11, 11, 12}

Question: find the 25 percentile Value? 🧐

☐ Already Sorted, move to the step 2.

$$\square I = (p/100) * (n + 1)$$

$$I = (25/100) * (20 + 1) = 5.25$$

📖 Note: 5.25 is the index Position of the dataset, the position is decimal so we want to see the adjacent values surrounding this Position.

Now the 5.25 is not the whole number it has Decimal value $d = 0.25$

L :-> 5th index of the dataset is $L = 5$ (Keep in mind that $I = 5.25$ is the index not a value so 5 is the 5th index)

H :-> H is assigned a value of 5. It is one position higher than L, meaning H corresponds to the element at index 6 (the position following index 5).

📐 Interpolate:

$$P = L + (H - L) \times d$$

$$= 5 + (5 - 5) \times 0.25$$

$$= 5$$

🗨 **Answer: The 25th percentile is 5, which means that 25% of the data values in the dataset are less than or equal to 5.**

For more further clarity we may see another example

Question: find the 75 percentile Value? 🧐

☐ Sort the dataset (Already Sorted).

$$\square I = (p/100)*(n + 1)$$

$$= (75/100)*(20+1) = 15.75$$

15 is the whole number $L = 9$ (15th index in the dataset is 9).

$$d = 0.75$$

$H = 9$ (16th Index of dataset, one position higher than L)

☐ Interpolate:

$$P = L + (H - L) \times d$$

$$= 9 + (9 - 9) \times 0.75$$

$$= 9.$$

🧐 **Answer: The 75th percentile is 9. So 75% of the data values in the dataset are less than or equal to 9, and the remaining 25% of the data values are greater than 9.**

In Conclusion 🌀

Percentiles are a powerful tool in statistics, helping us understand how individual scores compare to the larger dataset. By using the linear interpolation method, we can calculate percentiles with precision, as shown in the examples above.