Percentile and Quartiles

Percentage : {1, 2, 3, 4, 5}

Q: what is the percentage of even numbers?
A:

$$\% \text{ of even } \# = \frac{2}{5} = 0.4 = 40\%$$

∴ 40% of the # are even.

* **Percentile :**
A percentile is a value below which a certain percentage of observations lie.

95 Percentile → Means that a person got better marks than 95 % of the students.

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

Q: what is the percentile rank of 10?

$$\text{Percentile rank of 10} = \frac{\#\text{ of values below 10}}{n} \times 100$$

$$= \frac{16}{20} \times 100 = 80 \text{ Percentile.}$$

Interpret : • Percentile ranking of 10 → 80 Percentile
• 80% of values present in the dataset are smaller than 10.

Q: what value exists at percentile rank of 25?

$$\text{value} = \frac{\text{Percentile}}{100} \times (n+1)$$

$$= \frac{25}{100} \times (21) = 5.25 \Rightarrow \text{Index}$$

This index is not present. Present index is 5th and next value after 5th index is 6th index, i.e., value 5

$$\therefore \text{value} = \frac{5+5}{2} = 5.$$

Hence, value 5 exists at percentile rank of 25.

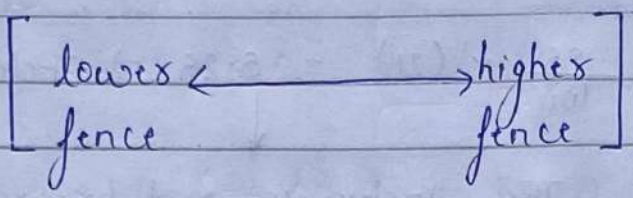
Q what if 5.25 index was present?
→ Index always be a whole number and if the index we got was 8th, then the value 6 will be value existing at percentile rank of 25.

* Quartiles : (one-fourth)

- 25% → 1st quartile → Q1
- 50% → 2nd quartile (Median) → Q2
- 75% → 3rd quartile → Q3.

IQR = Q3 − Q1
↳ Inter Quartile Range

Q: why we are learning this?
A: This will be used to identify the outliers.



This is a range (Border)
- Below lower fence, every observation is outlier.
- Above higher fence, every observation is outlier.

→ Lower fence :

lower fence = $Q1 - 1.5 \times IQR$

→ Higher fence :

higher fence = $Q3 + 1.5 \times IQR$

Eg: calculating outliers

$x : \{1, 2, 2, 2, 3, 3, 4, 5, 5, 5, 6, 6, 6, 6, 7, 8, 8, 9, 29\}$

$Q1\ (25\ \text{percentile}) = \frac{25}{100} \times (19+1)$

$= 5^{th}$ element

∴ value 3 exist at 25^{th} percentile.

$Q3\ (75\ \text{percentile}) = \frac{75}{100} \times (19+1)$

$= 15^{th}$ element

∴ Value 7 exist at 75^{th} percentile.

Now,

→ Lower Fence and Higher fence :

$IQR = Q3 - Q1$

$= 7 - 3 = 4.$

Lower Fence $= Q1 - 1.5 \times IQR$

$= 3 - 1.5 \times 4$

$= -3$

Higher Fence $= Q3 + 1.5 \times IQR$

$= 7 + 1.5 \times 4$

$= 13.$

Lower fence and Higher fence are in the range between

$[-3, 13]$

∴ Value 29 is an outlier.