How to find outliers?

Ankita Chaudhari

Steps Involved:

1. **Sort your data from low to high.**
2. **Identify the first quartile (Q1), the median, and the third quartile (Q3).**
3. **Calculate your IQR = Q3 – Q1**
4. **Calculate your upper fence = Q3 + (1.5 * IQR)**
5. **Calculate your lower fence = Q1 – (1.5 * IQR)**
6. **Use your fences to highlight any outliers, all values that fall outside your fences.**

Example:

STEP-1:

Dataset=[1,2,2,2,3,3,4,5,5,5,6,6,6,6,7,8,8,9,27]

STEP-2:

Find value of Q1:

Formula: value = (percentile/100)*(n+1)

Where Q1 = 25%

Value = (25 /100)* 20

Value = 5 index

Value = 3 = Q1

Similarly for Q3:

Where Q3 = 75%

Value = (75/100)*20

Value = 15 index

value = 7 = Q3

STEP-3:

IQR = Q3 -Q1 = 7- 4 = 4

IQR = 4

STEP-4:

. upper fence = Q3 + (1.5 * IQR)

=7 +(1.5 * 4)

=7 +6 = 13

upper fence = 13

STEP-5 :

- lower fence = Q1 – (1.5 * IQR)

  = 3 - (1.5* 4)

  =3 – 6 = -3

lower fence = -3

[lower fence , upper fence] = [-3 , 13]

[lower fence , upper fence] = [-3 , 13]

Check dataset .we observe 27 is outside of [-3 ,13]

Therefore 27 is outlier

Dataset=[1,2,2,2,3,3,4,5,5,5,6,6,6,6,7,8,8,9,27]

Removing outlier, we get,

Dataset=[1,2,2,2,3,3,4,5,5,5,6,6,6,6,7,8,8,9]

Removing outlier, we get,

Dataset=[1,2,2,2,3,3,4,5,5,5,6,6,6,6,7,8,8,9]

WE GET FIVE NUMBER SUMMARY:

1) Minimum =1

2) Q1 = 3

3) Median = 5 (n is odd ,we can take average)

4) Q3 = 7

5) Maximum = 9