[INTER QUARTILE RANGE(IQR)]

[LESSER RANGE & GRATER RANGE]

Ragu surya Prakash
Hope Artificial Intelligence Pvt Ltd | Learn Data Science, Machine Learning, Python, Deep Learning, NLP, TSA [Company address]

INTER QUARTILE RANGE[IQR]

About IQR (Inter quartile range)?

1] It represents the range of values within which the middle 50% of the data falls, making it less sensitive to outliers than the range

2] Calculated the difference between the third quartile (Q3) and the first quartile (Q1)

3]Quartile=Q1----25%-------Q2-----75%-----Q3-----MAX-----Q4

4]Calculate between of outlier range,50% data called central balance data.

Uses of 1.5rule in IQR:

Outlier Detection: Identifies data points that deviate significantly from the rest of the dataset.

Data Cleaning: Helps remove or adjust outliers to improve data accuracy.

Quality Control: Used in manufacturing, finance, etc., to spot anomalies indicating errors or fraud.

Robust Estimation: Provides reliable estimates despite outliers' influence.

Data Visualization: Outliers can be highlighted in visualizations for better understanding.

My understanding about 1.5rule:

[1.5 Rule]: Efficiently identifies outliers, data reset or replace, and find man error or fraud.

INTER QUARTILE RANGE[IQR]

Formula of IQR:

HEAR TWO TYPES OF OUTLIERS:

1] LESSER RANGE

2]GREATER RANGE

Calculate between of outlier range,50% data called central blance data.

Lesser than quartile range: [IQR=Q3-Q1]

[Q3=75% Central balance data Q1=25%]

Q1-1.5rule*IQR

Calculate between of outlier range,50% data called central blance data.

Greater than quartile range: [IQR=Q3+Q1]

[Q3=75% Central balance data Q1=25%]

Q3+1.5rule*IQR

Examples:

Data: [2,6,8,10,12,30]

Min	Q1	Median	Q3	max
2	6	10	12	30

Find IQR: IQR=Q3-Q1

12-6=6

Using 1.5rule:

6-1.5*6=-3

Lesser range of outlier=-3

INTER QUARTILE RANGE[IQR]

Find IQR: IQR=Q3+Q1

12+6=18

Using 1.5rule: Q3+1.5rule*IQR

12+1.5*18=39

Greater range of outlier=39

For the following 13 real estate prices calculate the IQR and determine if price are potential outliers pries are in dollars

[3]:
```
import pandas as pd
import numpy as np
```

[4]:
```
lists = [114950, 158000, 230500, 387000,389950, 479000,488800, 529000, 575000, 639000,659000,1095000,5500000]
```

[8]:
```
q1, q2, q3, q4 = np.percentile(lists, [25, 50, 75, 100])
```

[47]:
```
table = {"s/no": [1], 'Min': [min(lists)], 'Q1': [308150], 'Median': [488800.0], 'Q3': [649000], 'Max': [max(lists)]}
```

[48]:
```
dataset= pd.DataFrame(table)
dataset
```

[48]:

	s/no	Min	Q1	Median	Q3	Max
0	1	114950	308150	488800.0	649000	5500000

[90]:
```
#find median value
#why using int (int is unnecessary in this context,isn't required for removing the extra zero)
median=488800
median = (230500+ 387000) / 2
median= int(median)
print("q1_median:",median)
median=488800
median = (639000+ 659000) / 2
median= int(median)
print("q3_median:",median)
```

```
q1_median: 308750
q3_median: 649000
```

[95]:
```
IQR = 649000-308750
IQR
```

[95]:
```
340250
```

[96]:
```
Q1=1.5*340250
q1
```

[96]:
```
510375.0
```

[85]:
```
lesser=308750-510375
lesser
```

[85]:
```
-201625
```

[99]:
```
greater =649000+510375
greater
```

[99]:
```
1159375
```

INTER QUARTILE RANGE[IQR]

Q1=25% or 25^{th} percential value

25% of the price =308,750

Q3=75% or 75^{th} percential value

75% of the price =649,000

The real estate prices in the central area have increased by 50%. Now, calculate the Interquartile Range (IQR) using only the median formula.

Find the Interquartile Range (IQR), we can use the median formula directly, calculating the median of the first quartile (Q1) and the third quartile (Q3)

Find lesser range of outlier:

IQR=649,000-308,750
=340,250
(1.5) (IQR)= (1.5) (340,250)
=510,375
Q1-(1.5) (IQR)=308,750-510,375
=-201,625

In Python code or calculation:

```
median=488800

median = (230500+ 387000) / 2

median= int(median)

print ("q1_median:", median)
IQR = 649000-308750

Q1=1.5*340250

lesser=308750-510375
lesser range of outlier: -201,625
```

INTER QUARTILE RANGE[IQR]

The "Thumb rule" states that any value less than -201,625 is considered to be within the lesser range compared to other values.

Greater range in python calculations:

```
median=488800
median = (639000+ 659000) / 2
median= int(median)
print ("q3_median:", median)
greater =649000+510375
greater range=1,159,375
```

greater range of outlier: 1,159,375

The "Thumb rule" states that any value greater than 11,59,375 is considered to be within the greater range compared to other values.

Conclusion of this data:

The prices are considered potential outliers. If the value exceeds 5,500,000, it will be replaced with the original price of 1,159,375.

A) The interquartile range. Compare the two interquartile range.
b) Any outliers in eithers set.

The five number summary for the day and night classes is

	min	Q1	Median	Q3	Maximum
day	32	56	74.5	82.5	99
night	25.5	78	81	89	98

INTER QUARTILE RANGE[IQR]

Calculate the interquartile range (IQR) for each dataset:

“Lesser day”

IQR=Q3-Q1

82.5-56=26.5

IQR=26.5

The lesser outliers 1.5rule:

lesser outliers: Q1-1.5*IQR

56-1.5*26.5=16.25

lesser day=16.25

“Greater day:”

IQR=Q3-Q1

89-78=11

IQR=11

The greater outliers 1.5rule:

greater outliers: Q3+1.5*IQR

82.5+1.5*26.5=16.25

greater day=122.5

“NIGHT”

Lesser night=78-1.5*11=61.5

Greater night=89+1.5*11=105.5

"Day" category, the value of 99 may be a potential outlier is =122.5

"Night" category, there are no potential outliers

----------THE END----------