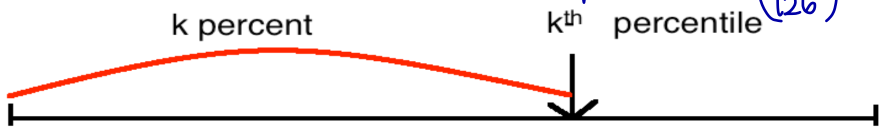
Learning Goal: Measures of Spread (One-variable statistics)

PERCENTILE – divide the data into 100 intervals that have equal numbers of values. It is a measure that indicates what percent of the given population scored at or below the measure.

89^{th} percentile $\left(\frac{70}{126}\right)$

89% of the data less than or equal to $\frac{70}{126}$



$1^{st} \rightarrow 99^{th}$

To calculate a particular percentile:

1) Multiply the total number of values in the data set by the percentile which will give you the index.
2) Order all of the values in the data set in ascending order (least to greatest)
3) If the index is a whole number, count the values in the data set from least to greatest until you reach the index, then take the index and the next greatest number and find the average.
4) If the index is not a whole number, round the number up, then count the values in the data set from least to greatest until you reach the index.
5) If the index is at .5, take the mean value of the raw data above and below in the data set from least to greatest.

Textbook page 145:

35	47	57	62	64	67	72	76	83	90
38	50	58	62	65	68	72	78	84	91
41	51	58	62	65	68	73	79	86	92
44	53	59	63	66	69	74	81	86	94
45	53	60	63	67	69	75	82	87	96
45	56	62	64	67	70	75	82	88	98

Percentile Rank: (R) to find the raw score in the data set that represent the percentile

$$R = \frac{p}{100}(n+1)$$

$$p = percentile; n = size\ of\ the\ population$$

Percentile (p): to determine the percentile of a raw score

$$p = 100\left(\frac{L + 0.5E}{n}\right)$$

$$L = number\ of\ data\ less\ than\ the\ data\ point$$

$$E = number\ of\ data\ equal\ to\ the\ data\ point$$

a) If a datum scored at 50^{th} percentile, what was its raw score?

$R = \frac{50}{100}(60+1)$

index $= 30.5^{th}$ term

$= \frac{30^{th} + 31^{st}}{2}$

$= \frac{67+67}{2}$

$= 67$

$\therefore 50^{th}$ percentile is 67.

b) What is the 90^{th} percentile for this data?

$R = \frac{90}{100}(60+1)$

index $= 54.9^{th}$ term

round up

$= 55^{th}$ term

$= 90$

$\therefore 90^{th}$ percentile is 90.

c) Does the score of 75 place it at the 75^{th} percentile?

$L = 40$ (40 numbers less than score of 75)

$E = 2$ (2 numbers are equal to 75)

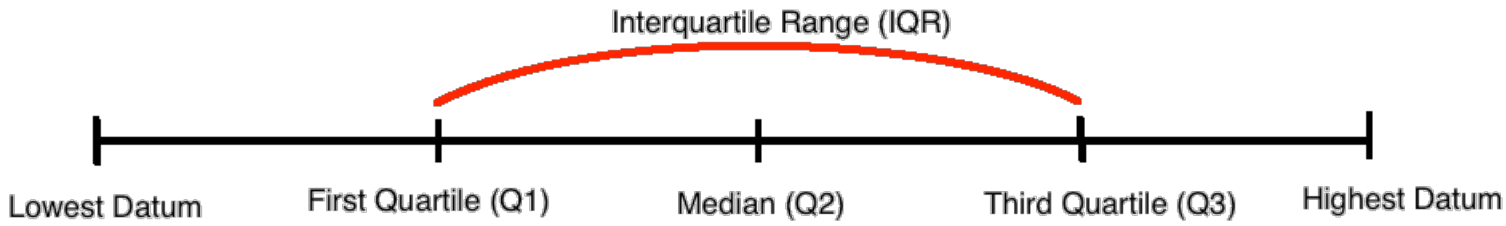
$P = 100\left(\frac{40 + 0.5(2)}{60}\right)$

$= 68\frac{1}{3}$ round up

$= 69^{th}$ percentile

$\therefore$ No, 75 score is at the 69^{th} percentile.

Quartile and Interquartile Range (IQR):

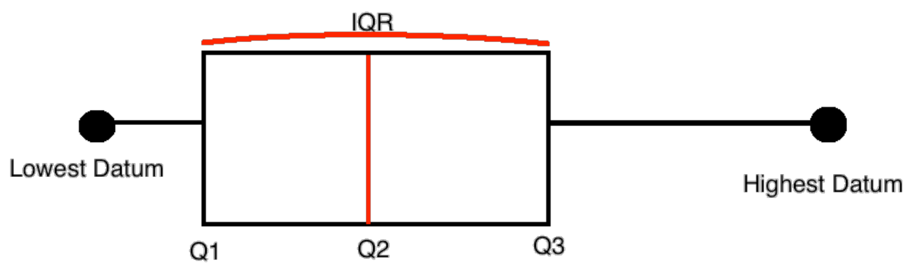


odd numbers of data:
$\underbrace{1, 1, 4}_{1^{st}\text{ half}}, 7, \underbrace{9, 11, 11}_{2^{nd}\text{ half}}$ $(n=7)$
median $Q2 = 7$
$Q1$ = median of the 1^{st} half = 1
$Q3$ = median of the 2^{nd} half = 11

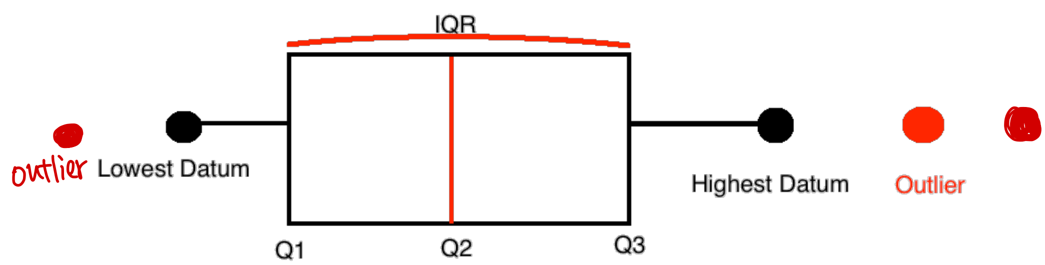
even numbers of data:
$\underbrace{1, 1, 4}_{1^{st}\text{ half}}, \underbrace{7, 9, 11}_{2^{nd}\text{ half}}$ $(n=6)$
median $Q2 = \frac{4+7}{2} = 5.5$
$Q1$ = median of 1^{st} half = 1
$Q3$ = median of 2^{nd} half = 9

$$Q2 = Median$$
$$Q1 = Median\ of\ the\ first\ half\ of\ the\ data$$
$$Q3 = Median\ of\ the\ second\ half\ of\ the\ data$$
$$IQR = Q3 - Q1$$
$$SemiIQR = SIQR = \frac{IQR}{2}$$

Box-And-Whisker Plot:



Modified Box-And-Whisker Plot: identify outliers



possible to have more than one outliers

To determine the value(s) of an outlier:

$$1.5 \times IQR$$

If a datum is less than $Q1 - (1.5 \times IQR)$, it is considered as an outlier.
If a datum is greater than $Q3 + (1.5 \times IQR)$, it is considered as an outlier. $\{-5, 0, \cdots\cdots, 45, 70, 75\}$

eg. $Q1 = 25$ $Q2 = 30$ $Q3 = 44$

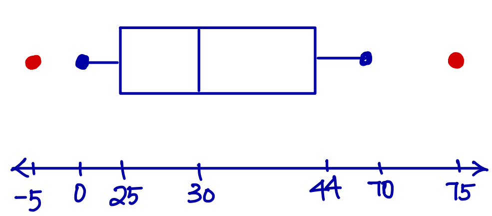
$IQR = 44 - 25$
$= 19$

$(1.5)(IQR) = 1.5(19)$
$= 28.5$

$Q1 - 28.5$
$= 25 - 28.5$
$= -3.5$

$Q3 + 28.5$
$= 44 + 28.5$
$= 72.5$

Modified Box-and-Whisker Plot:



Example #1: A survey of movie goers at a screening of "Rocky Horror Picture Show" asked how many times they have seen the movie. The results for 20 respondents were:

3 4 2 8 10 5 1 15 5 16 6 3 4 9 12 3 30 2 10 7

Step #1 rearrange data:

5th 6th ... 5th 6th

1 2 2 3 3 3 4 4 5 5 6 7 8 9 10 10 12 15 16 30

1st half ... 10th 11th ... 2nd half

a) Find the median, Q1, and Q3.

$$Q2 = \frac{n+1}{2} = \frac{20+1}{2} = 10.5^{th} = \frac{5+6}{2}$$

$$\boxed{Q2 = 5.5}$$

$$Q1 = \frac{10+1}{2} = 5.5^{th} = \frac{5^{th}+6^{th}}{2} = \frac{3+3}{2}$$

$$\boxed{Q1 = 3}$$

$$Q3 = \frac{10+10}{2}$$

$$\boxed{Q3 = 10}$$

b) Calculate IQR and SIQR.

$$IQR = Q3 - Q1 = 10 - 3 = 7$$

$$SIQR = \frac{IQR}{2} = \frac{7}{2} = 3.5$$

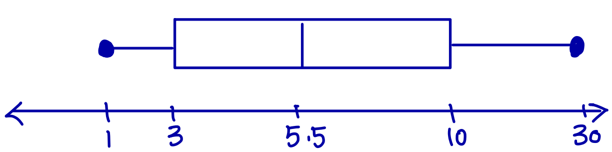
c) Are there any outliers? Explain.

$$Q1 - 1.5(IQR) = 3 - 1.5(7) = -7.5$$

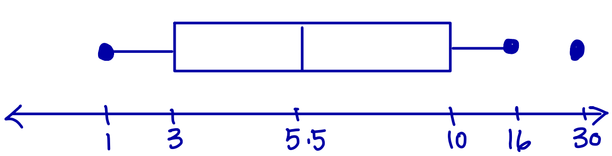
$$Q3 + 1.5(IQR) = 10 + 1.5(7) = 20.5$$

∴ the outlier is 30 times.

d) Make a box-and-whisker plot of the results.



e) Make a modified box-and-whisker plot of the results accounting for any outliers. How do the quartiles compare?



Using the previous example, find the score of:

a) 20^{th} percentile
b) 65^{th} percentile
c) 95^{th} percentile

a) 20^{th} percentile

$$\frac{20}{100}(20+1) = 4.2^{th} \doteq 5^{th}$$

∴ 20^{th} percentile is 3 times.

b) 65^{th} percentile

$$\frac{65}{100}(20+1) = 13.65^{th} = 14^{th}$$

∴ 65^{th} percentile is 9 times.

c) 95^{th} percentile

$$\frac{95}{100}(20+1) = 19.95^{th} = 20^{th}$$

∴ 95^{th} percentile is 30 times.