Stat Support Activity: Quartiles and IQR

Notes

- The <u>quartiles</u> are three numbers that divide the data set into four parts (four quarters) when the data is put in order.

- The <u>1st Quartile</u> (Q_1) is the divider where approximately 25% of the numbers in the quantitative data are lower than Q_1. For this reason, the 1st quartile is also called the 25th Percentile (P_{25}).

- The <u>2nd Quartile</u> (Q_2) is also called the median. It is the divider where approximately 50% of the numbers in the quantitative data are lower than the median (Q_2). For this reason, the 2nd quartile is also called the 50th Percentile (P_{50}).

- The <u>3rd Quartile</u> (Q_3) is the divider where approximately 75% of the numbers in the quantitative data are lower than Q_3. For this reason, the 3rd quartile is also called the 75th Percentile (P_{75}).

- Think of the quartiles as three dividers that quarter the data. Approximately 25% of the numbers in data will be below Q_1. Approximately 25% of the numbers in the data will be between Q_1 and Q_2 (median). Approximately 25% of the numbers in the data will be between Q_2 (median) and Q_3. Approximately 25% of the data will be higher than Q_3.

- <u>Typical Values</u>: For non-normal or skewed data, typical values fall between the 1st Quartile (Q_1) and the 3rd Quartile (Q_3).

- <u>To calculate the three Quartiles</u>:

    o   Put the data in order from lowest to highest.

    o   Calculate the median (Q_2). If there is an odd amount of numbers in the data (sample size "n" is odd), the median will be the number in the exact middle. If there is an even number of values in the data (sample size "n" is even), the median will be half-way in-between the two numbers in the middle.

    o   If the median is a number in the data set (sample size "n" is odd), then the bottom half of the data and the top half of the data will include the median. If the median is <u>not</u> a number in the data set (sample size "n" is even), then the bottom half of the data and the top half of the data will <u>not</u> include the median.

    o   To calculate the 1st Quartile (Q_1), find the median (center) of the bottom half of the data in order.

    o   To calculate the 3rd Quartile (Q_3), find the median (center) of the top half of the data in order.

Stat Support Activity: Quartiles and IQR

- Interquartile Range (IQR):

    - The Interquartile Range (IQR) is the most accurate measure of spread for non-normal or skewed data.

    - IQR measures the distance between Q_1 and Q_3 (the middle 50% of the numbers in the data).

    - IQR is the maximum distance that typical values are from each other in non-normal or skewed quantitative data.

    - To calculate IQR, subtract the 3rd Quartile minus the 1st Quartile. ($IQR = Q_3 - Q_1$)

Problems

1.

The following data represents the lengths of various rocks in centimeters (cm).

15.2, 3.4, 1.7, 2.5, 2.9, 4.8, 5.6, 6.3, 1.4, 3.3, 6.7, 3.9

a) How many numbers are in the sample data (sample size "n")?
b) If you divide the sample size "n" by 4, what would you get? (This is how many numbers should be in each of the four groups separated by the quartiles.)
c) Put the numbers in order for smallest to largest.
d) Find the median average (Q_2). If sample size "n" is odd find the number in the middle. If the sample size "n" is even, add the two numbers in the middle and divide by 2.
e) Fill in the blank to complete the following sentence:
   "The average size of the rocks in the data set was _____ cm."
f) Is the median a number in the data set or was it in-between two numbers in the data set?
g) Now find the 1st Quartile (Q_1). If the median (Q_2) is a number in the data set, it should be included in the bottom half. If the median (Q_2) is not a number in the data set, it should not be included in the bottom half. To find Q_1, find the median (center) of the bottom half of the data.
h) Fill in the blank to complete the following sentence: "Approximately 25% of the rocks had a length below _____ cm."
i) Now find the 3rd Quartile (Q_3). If the median (Q_2) is a number in the data set, it should be included in the top half. If the median (Q_2) is not a number in the data set, it should not be included in the top half. To find Q_3, find the median (center) of the top half of the data.
j) Fill in the blank to complete the following sentence: "Approximately 75% of the rocks had a length below _____ cm."
k) Fill in the blanks to complete the following sentence: "Typical rocks in the data set have a length between _____ cm and _____ cm."

Stat Support Activity: Quartiles and IQR

l) How many numbers in the data set were between Q_1 and Q_3? (Do not include Q_1 and Q_3 in the count since they are not numbers in the data set.) What percent of the data is in between Q_1 and Q_3 exclusively?

m) Find the Interquartile Range (IQR) with the formula $IQR = Q_3 - Q_1 = $?

n) Fill in the blank to complete the following sentence: "Typical rock lengths in the data set are within _____ cm of each other."

o) Look at the numbers in order. Label Q_1, Q_3, Median and IQR so you can see the four groups and the typical spread.

2.

The following data represents the amount of cash carried by adults at the mall in dollars.

20, 50, 35, 75, 10, 0, 400, 15, 100, 80, 40, 60, 30, 5, 20, 40, 30

a) How many numbers are in the sample data (sample size "n")?

b) If you divide the sample size "n" by 4, what would you get? Round up the answer to the next whole number. (This is how many numbers should be in each of the four groups separated by the quartiles.)

c) Put the numbers in order for smallest to largest.

d) Find the median average (Q_2). If sample size "n" is odd find the number in the middle. If the sample size "n" is even, add the two numbers in the middle and divide by 2.

e) Fill in the blank to complete the following sentence:
"The average amount of cash carried by people at the mall in the data set was $_____ ."

f) Is the median a number in the data set or was it in-between two numbers in the data set?

g) Now find the 1st Quartile (Q_1). If the median (Q_2) is a number in the data set, it should be included in the bottom half. If the median (Q_2) is not a number in the data set, it should not be included in the bottom half. To find Q_1, find the median (center) of the bottom half of the data.

h) Fill in the blank to complete the following sentence: "Approximately 25% of the amounts of cash carried by people at the mall in the data set were below $_____ ."

i) Now find the 3rd Quartile (Q_3). If the median (Q_2) is a number in the data set, it should be included in the top half. If the median (Q_2) is not a number in the data set, it should not be included in the top half. To find Q_3, find the median (center) of the top half of the data.

j) Fill in the blank to complete the following sentence: "Approximately 75% of the amounts of cash carried by people at the mall in the data set were below $_____ ."

k) Fill in the blanks to complete the following sentence: "Typical amounts of cash carried by people at the mall in the data set were between $_____ and $_____ ."

l) How many numbers in the data set were between Q_1 and Q_3? (Include Q_1 and Q_3 in the count since they are numbers in the data set.) What percent of the data is in between Q_1 and Q_3 inclusively?

m) Find the Interquartile Range (IQR) with the formula $IQR = Q_3 - Q_1 = $?

n) Fill in the blank to complete the following sentence: "Typical amounts of cash carried by people at the mall in the data set were within $_____ of each other."

o) Look at the numbers in order. Label Q_1, Q_3, Median and IQR so you can see the four groups and the typical spread.