IM 9–12 MATH v.1

Unit 1
Sequences and Functions

ALGEBRA 1

Lesson 15

Comparing Data Sets

Illustrative Mathematics®

Kendall Hunt

Learning Goal

Let's compare statistics for data sets.

Algebra 1

Illustrative Mathematics®

Kendall Hunt

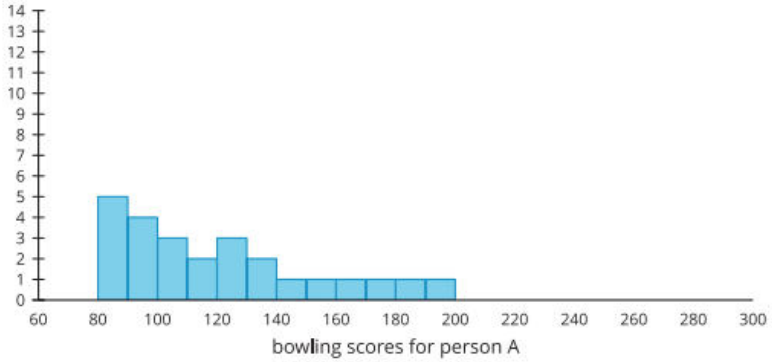
Bowling Partners

Warm-up

Each histogram shows the bowling scores for the last 25 games played by each person. Choose 2 of these people to join your bowling team. Explain your reasoning.

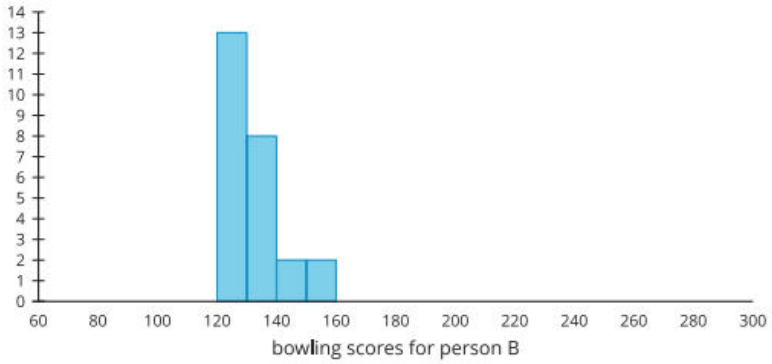
Person A

- mean: 118.96
- median: 111
- standard deviation: 32.96
- interquartile range: 44



Person B

- mean: 131.08
- median: 129
- standard deviation: 8.64
- interquartile range: 8

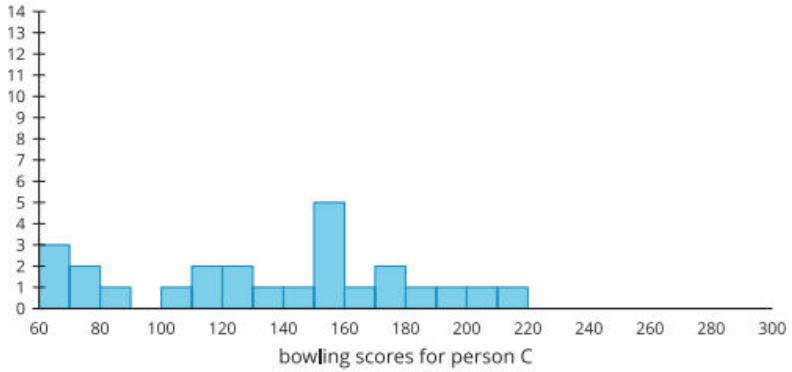


Bowling Partners

Warm-up

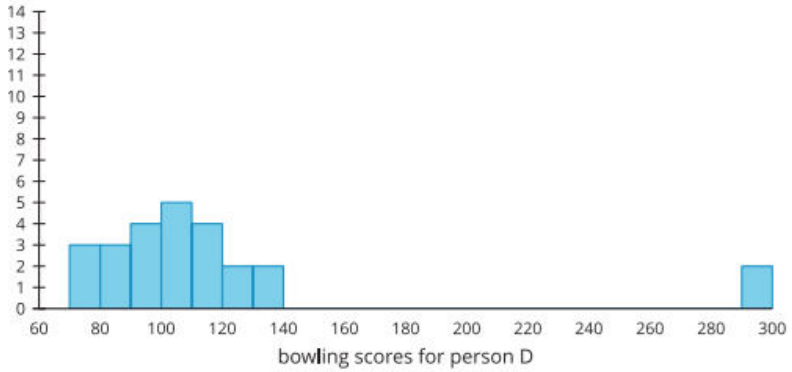
Person C

- mean: 133.92
- median: 145
- standard deviation: 45.04
- interquartile range: 74



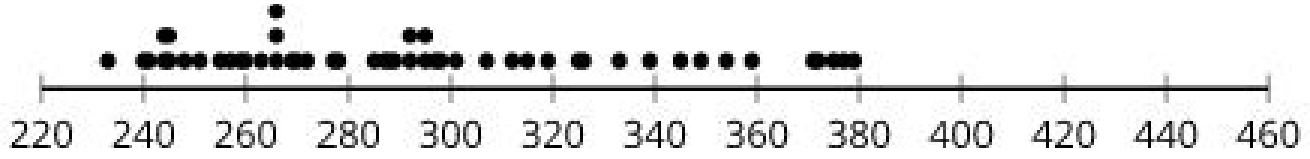
Person D

- mean: 116.56
- median: 103
- standard deviation: 56.22
- interquartile range: 31.5

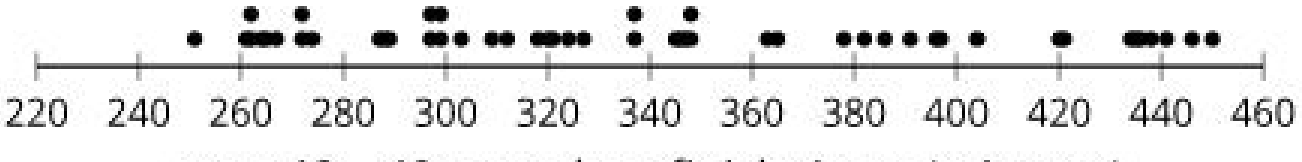


Comparing Marathon Times

All of the marathon runners from each of two different age groups have their finishing times represented in the dot plot.



ages 30 – 39 marathon finish times (minutes)



ages 40 – 49 marathon finish times (minutes)

1. Which age group tends to take longer to run the marathon? Explain your reasoning.
2. Which age group has more variable finish times? Explain your reasoning

Comparing Marathon Times

- What measure of center and measure of variability are most appropriate to use with the distributions? Explain your reasoning.
- Which points show values that are most likely to be outliers?
- Based on the displayed information, are there any outliers in these data sets?

Values for the Data Sets:

Ages 30–39

- Mean: 294.5 minutes
- Standard deviation: 41.84 minutes
- Median: 288.5 minutes
- IQR: 65 minutes
- Q1: 260 minutes
- Q3: 325 minutes

Ages 40–49

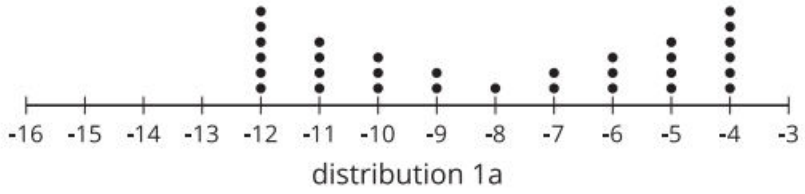
- Mean: 340.6 minutes
- Standard deviation: 59.93 minutes
- Median: 332 minutes
- IQR: 92 minutes
- Q1: 289 minutes
- Q3: 381 minutes

Comparing Measure

For each group of data sets,

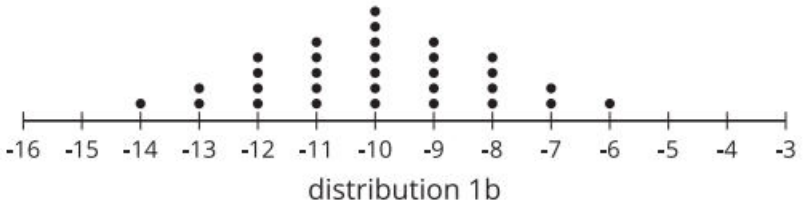
- Determine the best measure of center and measure of variability to use based on the shape of the distribution.
- Determine which set has the greatest measure of center.
- Determine which set has the greatest measure of variability.
- Be prepared to explain your reasoning.

1a



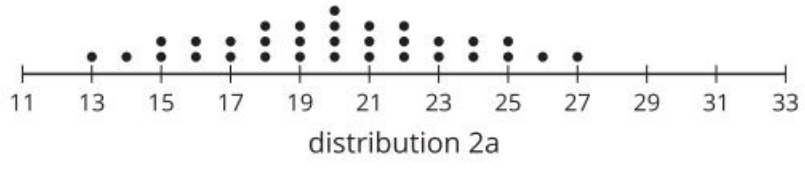
distribution 1a

1b



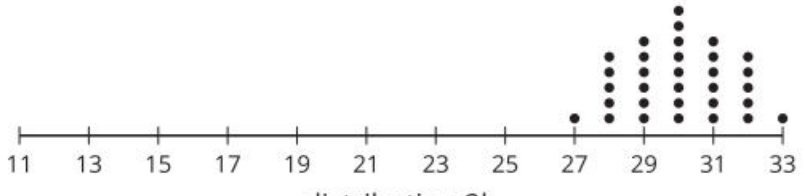
distribution 1b

2a



distribution 2a

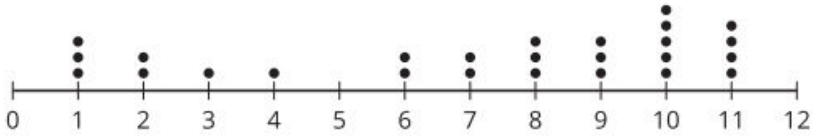
2b



distribution 2b

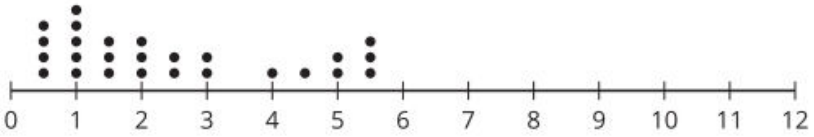
Comparing Measure

3a



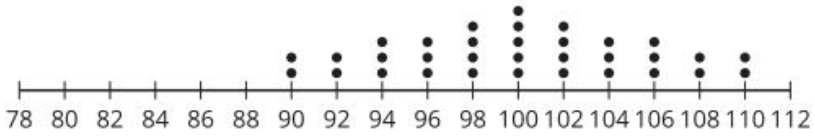
distribution 3a

3b



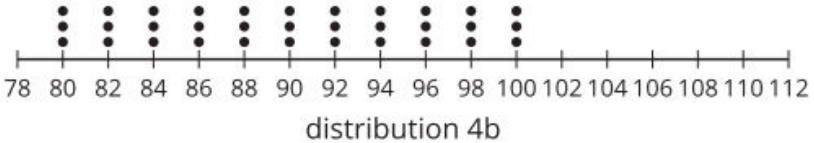
distribution 3b

4a



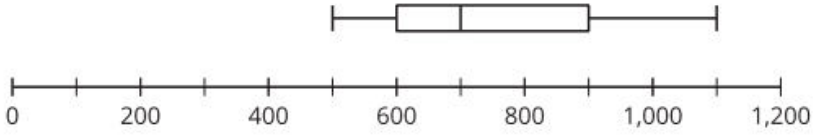
distribution 4a

4b



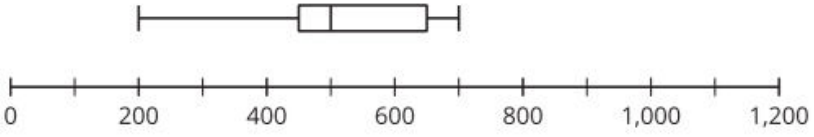
distribution 4b

5a



distribution 5a

5b



distribution 5b

Comparing Measure

6a

A political podcast has mostly reviews that either love the podcast or hate it.

6b

A cooking podcast has reviews that neither hate nor love the podcast.

7a

Stress testing concrete from site A has all 12 samples break at 450 pounds per square inch (psi).

7b

Stress testing concrete from site B has samples break every 10 psi starting at 450 psi until the last core is broken at 560 psi.

7c

Stress testing concrete from site C has 6 samples break at 430 psi and the other 6 break at 460 psi.

Comparing Measure

- What were some ways you handled the last two problems.
- Describe any difficulties you experienced and how you resolved them.
- How did you decide which data set showed greater variability

Comparing Data Sets

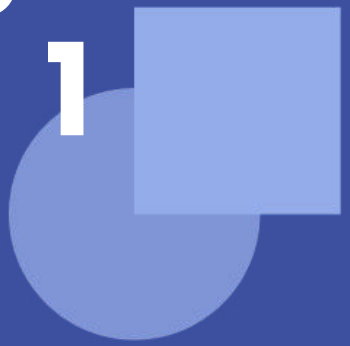
Lesson Synthesis

- How do you compare the measures of variability for a data set?
- How do you estimate variability when looking at data displays?
- How do you determine which measure of center to use for a data set?
- Why is the median the preferred measure of center for skewed distributions?
- Why is the mean the preferred measure of center for symmetric distributions?

I can compare and contrast situations using measures of center and measures of variability.

Learning Targets

Algebra 1



Comparing Mascots

Cool-down

A new pet food company wants to sell their product online and use social media to promote themselves. To determine whether to use a dog or a cat as their mascot, they research the number of people who click on links with an image of a dog or a cat.

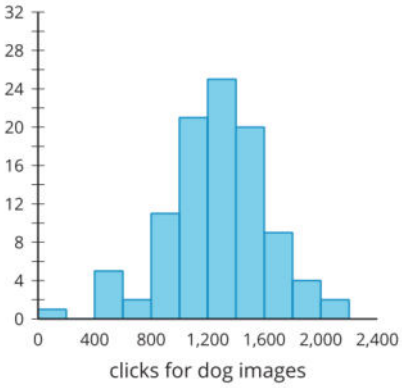
1. Based on the shape of the distributions, what measure of center and measure of variability would you use to compare the distributions? Explain your reasoning.
2. Based on the data shown here, should the company use a dog or cat mascot? Explain your reasoning.

mean: 1,263.5 clicks

median: 1,282 clicks

standard deviation: 357.4 clicks

IQR: 409 clicks

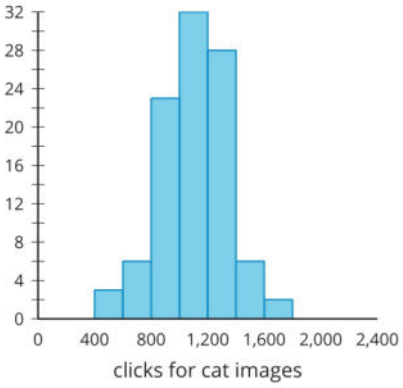


mean: 1,105.4 clicks

median: 1,125.5 clicks

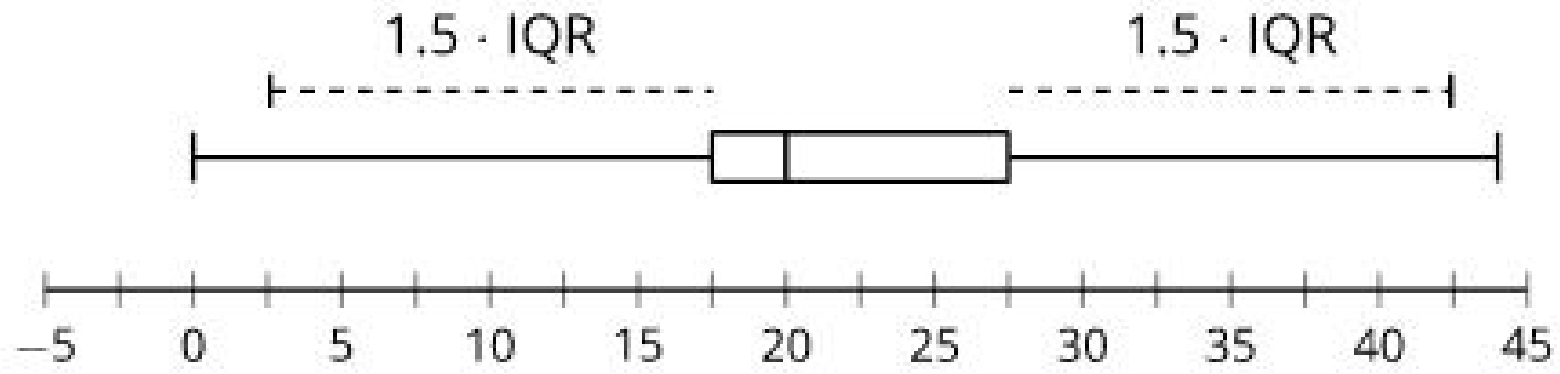
standard deviation: 239.3 clicks

IQR: 312.5 clicks



outlier

A data value that is unusual in that it differs quite a bit from the other values in the data set. In the box plot shown, the minimum, 0, and the maximum, 44, are both outliers.



standard deviation

A measure of the variability, or spread, of a distribution, calculated by a method similar to the method for calculating the MAD (mean absolute deviation). The exact method is studied in more advanced courses.

statistic

A quantity that is calculated from sample data, such as mean, median, or MAD (mean absolute deviation).