

Transforming Data

Section 2.1 Day 2

You will receive your test later in the period.

LEARNING TARGET

✓ DESCRIBE the effect of adding, subtracting, multiplying by, or dividing by a constant on the shape, center, and variability of a distribution of data.

Translate Data to analyze it easier

Translate Data to analyze it easier

Dilate (compress or expand)

Exploration

(handout)

1) Make a dotplot of the following data, then calculate the following statistics using your calculator
 {0, 0, 0, 1, 2, 3}

Mean: Median:
 SD: IQR:

2) Add 5 to each of the values in #1, graph the data, and calculate the summary statistics.

Mean: Median:
 SD: IQR:

1) Make a dotplot of the following data, then calculate the following statistics using your calculator
 {0, 0, 0, 1, 2, 3}

Mean: 1 Median: 1.5
 SD: 1.26 IQR: 2

2) Add 5 to each of the values in #1, graph the data, and calculate the summary statistics.

Mean: 6 Median: 5.5
 SD: 1.26 IQR: 2

5 5 5 6 7 8

3) What is the effect of **adding or subtracting** a constant from each observation?

Shape:

Center:

Spread:

The Effect of Adding or Subtracting a Constant

Adding the same positive number **a** to (subtracting a from) each observation:

- Adds **a** to (subtracts a from) measures of center and location (mean, five-number summary, percentiles)
- Does not change measures of variability (range, IQR, standard deviation)
- Does not change the shape of the distribution

1) Make a dotplot of the following data, then calculate the following statistics using your calculator
 {0, 0, 0, 1, 2, 3}

Mean: 1 Median: 1.5
 SD: 1.26 IQR: 2

4) Double each of the values in #1, graph the data, and calculate the summary statistics.

Mean: 2 Median: 1
 SD: 2.53 IQR: 4

0 0 0 2 4 6

5) What is the effect of **multiplying or dividing** each observation by a constant?

Shape:

Center:

Spread:

Transforming converts the original observations from the original units of measurements to another scale. Transformations can affect the shape, center, and spread of a distribution.

The Effect of Multiplying or Dividing by a Constant

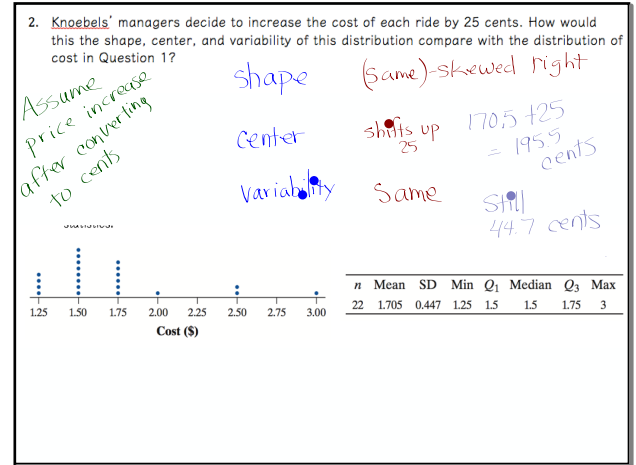
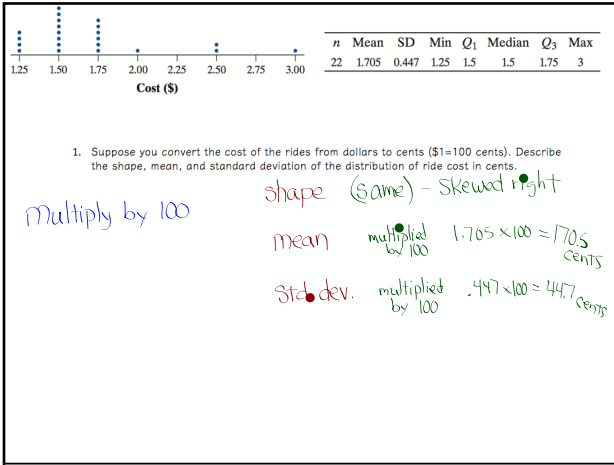
Multiplying (or dividing) each observation by the same positive number **b**:

- Multiplies (divides) measures of center and location by **b** (mean, five-number summary, percentiles) by **b**
- Multiplies (divides) measures of variability by **b** (range, IQR, standard deviation)
- Does not change the shape of the distribution

Check Your Understanding:

Knoebels Amusement Park in Elysburg, Pennsylvania, has earned acclaim for being an affordable, family-friendly entertainment venue. *Knoebels* does not charge for general admission or parking, but it does charge customers for each ride they take. How much do the rides cost at *Knoebels*? The figure shows a dot-plot of the cost for each of 22 rides in a recent year, along with summary statistics.

<i>n</i>	Mean	SD	Min	Q_1	Median	Q_3	Max
22	1.705	0.447	1.25	1.5	1.5	1.75	3



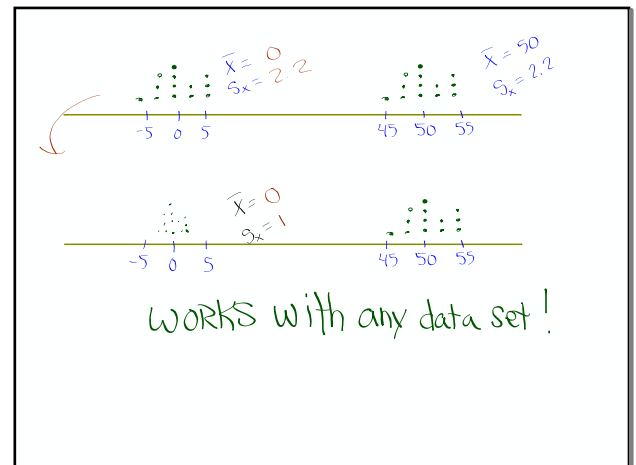
before we do #3 we're going to standardize every number in a data set using Excel

3. Now suppose you convert the increased costs from Question 2 to z-scores. What would be the shape, mean, and standard deviation of this distribution? Explain your answers.

A	B	C	D	E	F
	Data	Data +10	Data *2	Data - Mean	E / std dev
	20	30	40	-19	-1.0084153
	30	40	60	-9	-0.4776704
	35	45	70	-4	-0.212298
	40	50	80	1	0.05307449
	70	80	140	31	1.64530916
mean	39	49	78	0	0
st. deviation	18.8414437	18.8414	37.6829	18.8414437	1

Moral of the Story

Standardizing a distribution will produce a mean of 0 and a standard deviation of 1 for any distribution.



3. Now suppose you convert the increased costs from Question 2 to z-scores. What would be the shape, mean, and standard deviation of this distribution? Explain your answers.

shape same (skewed right)
 mean $\bar{x} = 0$
 std. deviation $s_x = 1$

Do You Use Uber ?

Taking an Uber ride in New York City has an initial fee of \$2.55 with an additional charge of \$1.75 per mile (we will ignore the small per minute waiting fee). In equation form,

$$\text{cost} = 2.55 + 1.75(\text{miles})$$

A local New York City resident records the number of miles for his first 25 rides with Uber. The mean distance of his rides is 5.6 miles with a standard deviation of 1.2 miles.

(a) Find the mean cost of the 25 trips.

(b) Calculate the standard deviation of the cost of the 25 trips. Interpret this value in context.

Do You Use Uber ?

Taking an Uber ride in New York City has an initial fee of \$2.55 with an additional charge of \$1.75 per mile (we will ignore the small per minute waiting fee). In equation form,

$$\text{cost} = 2.55 + 1.75(\text{miles})$$

A local New York City resident records the number of miles for his first 25 rides with Uber. The mean distance of his rides is 5.6 miles with a standard deviation of 1.2 miles.

(a) Find the mean cost of the 25 trips.

$$\text{Mean} = (1.75)(5.6) + 2.55 = \$12.35$$

(b) Calculate the standard deviation of the cost of the 25 trips. Interpret this value in context.

$$s_x = (1.75)(1.2) = \$2.10$$

• the costs of any Uber ride typically vary from \$12.35 by \$2.10

↑
 Adding or Subtracting won't change variability so we don't add

M/C 70%

$$\text{FR } \frac{10}{12} = 83\%$$

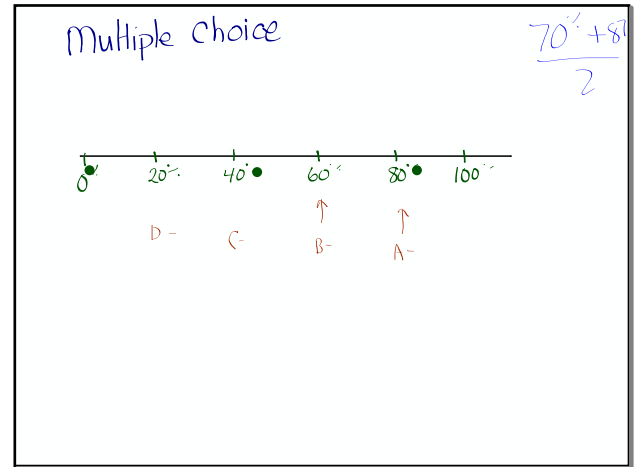
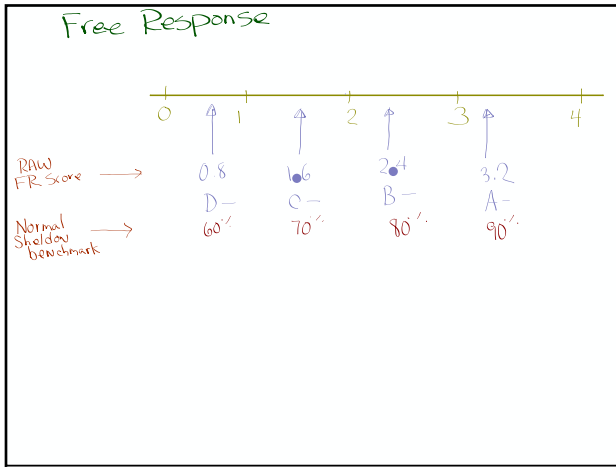
$$\rightarrow (.833)(4) = 3.33$$

BB

then see your tests

Free response [Feedback in terms of AP score Free response 0 to 4] but converted to our system

this test



Assignment

21 21, 25, 29, 31, 33-38