

Name: Answer Key Date: _____

Learning Task: If the Shoe Fits!

MCC9-12.S.ID. 1 Represent data with plots on the real number line (dot plots, histograms, and box plots).

MCC9-12.S.ID. 2 Use statistics appropriate to the shape of the data distribution to compare center (median, mean) and spread (interquartile range, mean absolute deviation) of two or more different data sets.

MCC9-12.S.ID. 3 Interpret differences in shape, center, and spread in the context of the data sets, accounting for possible effects of extreme data points (outliers).

Welcome to CSI at School! Over the weekend, a student entered the school grounds without permission. Even though it appears that the culprit was just looking for a quiet place to study undisturbed by friends, school administrators are anxious to identify the offender and have asked for your help. The only available evidence is a suspicious footprint outside the library door.

After the incident, school administrators arranged for the data in the table below to be obtained from a random sample of this high school's students. The table shows the shoe print length (in cm), height (in inches), and gender for each individual in the sample.

Shoe Print Length	Height	Gender	Shoe Print Length	Height	Gender
24	71	F	24.5	68.5	F
32	74	M	22.5	59	F
27	65	F	29	74	M
26	64	F	24.5	61	F
25.5	64	F	25	66	F
30	65	M	37	72	M
31	71	M	27	67	F
29.5	67	M	32.5	70	M
29	72	F	27	66	F
25	63	F	27.5	65	F
27.5	72	F	25	62	F
25.5	64	F	31	69	M
27	67	F	32	72	M
31	69	M	27.4	67	F
26	64	F	30	71	M
27	67	F	25	67	F
28	67	F	26.5	65.5	F
26.5	64	F	30	70	F
22.5	61	F	31	66	F
			27.25	67	F

1. Explain why this study was an observational study and not an experiment.

Observational - The admin were taking note of length + height but not doing experiments w/ the students

2. Why do you think the school's administrators chose to collect data on a random sample of students from the school? What benefit might a random sample offer?

They can't sample everyone so they sampled a few. This could help them to narrow down the students if they found a correlation between height + length.

3. Suggest a graph that might be used to use to compare the shoe print length data distributions for females and males.

Box plot - can easily compare 2 data sets

4. Describe one advantage of using comparative box plots instead of comparative dot plots to display these data.

Box plots are able to give you more info (5 point summary). Dot plots can be time consuming.

F: 22.5, 22.5, 24, 24.5, 24.5, 25, 25, 25, 25, 25, 25.5, 25.5, 26, 26, 26.5, 26.5, 27, 27, 27, 27, 27, 27.25, 27.4, 27.5, 27.5, 28, 29, 30, 31

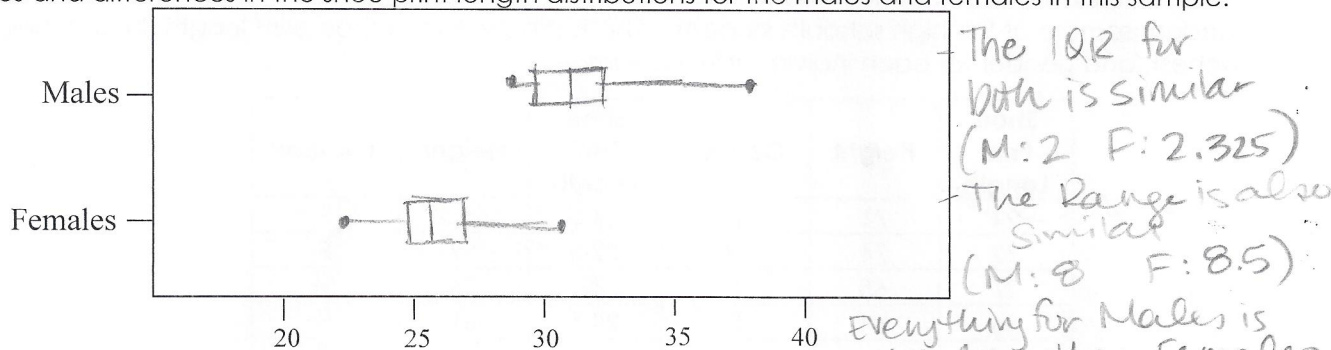
M: 29, 29.5, 30, 30, 31, 31, 31, 32, 32, 32.5, 37

5. For each gender calculate the five-number summary for the shoe print lengths. Additionally, for each gender, determine if there are any outlying shoe print length values.

	Minimum	Quartile 1 (Q1)	Median (Q2)	Quartile 3 (Q3)	Maximum
Male	29	30	31	32	37
Female	22.5	25	26.5	27.325	31

All 5 values calculated are larger for males than the females.

6. Construct comparative box plots for the shoe print lengths of males and females. Discuss the similarities and differences in the shoe print length distributions for the males and females in this sample.



7. For each gender calculate the mean shoe print length. What information does the mean shoe print length provide? Males: $\bar{x} = 31.4$ Females: $\bar{x} = 26.3$

8. The mean will give us an indication of a typical shoe print length. In addition to knowing a typical length we would also like to know how much variability to expect around this length. For each gender calculate the **Range**; **Interquartile Range**; and **Mean Absolute Deviation** of the shoe print lengths. Interpret each of the calculated values.

	Range	IQR	M.A.D.
Male	8	2	1.5
Female	8.5	2.325	1.5

MAD:

$$F: \frac{2(3.8) + 2.3 + 2(1.8) + 4(1.3) + 2(0.8) + 2(1.3) + 2(2) + 5(.7) + .95 + 1.1 + 2(1.2) + 1.7 + 2.7 + 3.7 + 4.7}{28} = 1.5$$

$$M: \frac{2.4 + 1.9 + 2(1.4) + 3(0.4) + 2(0.6) + 1.1 + 5.6}{11} = 1.47 \quad \text{M: } 1.5$$

9. If the length of a student's shoe print was 32 cm...

A. Would you think that the print was made by a male or a female?

Male

B. How sure are you that you are correct? Explain your reasoning. Use results from Questions 5 through 8 in your explanation.

It falls within the middle 50% of the male data + the max female was only 31

10. How would you answer Question 9 if the suspect's shoe print length was 27 cm?

Female - because the min male was 29 + 27 falls within the middle 50% of female data