Using Sample Data to Compare the Means of Two or More Populations

Suppose that Brett randomly sampled 12 tenth-grade girls and boys in his school district and asked them for the number of minutes per day that they text. The data and summary measures follow.

Gender	Number of Minutes of Texting													Mean	MAD
Girls	98	104	95	101	98	107	86	92	96	107	88	95		97.3	5.3
Boys	66	72	65	60	78	82	63	56	85	79	68	77		70.9	7.9

1. Draw dot plots for the two data sets using the same numerical scales. Discuss the amount of overlap between the two dot plots that you drew and what it may mean in the context of the problem.

2. Compare the variability in the two data sets using the MAD. Interpret the result in the context of the problem.

3. From 1 and 2, does the difference in the two means appear to be meaningful? Explain.

1. A school is trying to decide which reading program to purchase.
   a. How many MADs separate the mean reading comprehension score for a standard program ($\text{mean} = 67.8$, $\text{MAD} = 4.6$, $n = 24$) and an activity-based program ($\text{mean} = 70.3$, $\text{MAD} = 4.5$, $n = 27$)?
   b. What recommendation would you make based on this result?

2. Does a football filled with helium go farther than one filled with air? Two identical footballs were used: one filled with helium and one filled with air to the same pressure. Matt was chosen from your team to do the kicking. You did not tell Matt which ball he was kicking. The data (yards) follow.

Air	25	23	28	29	27	32	24	26	22	27	31	24	33	26	24	28	30
Helium	24	19	25	25	22	24	28	31	22	26	24	23	22	21	21	23	25

	Mean	MAD
Air		
Helium		

   a. Calculate the difference between the sample mean distance for the football filled with air and for the one filled with helium.
   b. On the same scale, draw dot plots of the two distributions, and discuss the variability in each distribution.
   c. Calculate the MAD for each distribution. Based on the MADs, compare the variability in each distribution. Is the variability about the same? Interpret the MADs in the context of the problem.
   d. Based on your calculations, is the difference in mean distance meaningful? Part of your reasoning should involve the number of MADs that separate the sample means. Note that if the MADs differ, use the larger one in determining how many MADs separate the two means.

3. Suppose that your classmates were debating about whether going to college is really worth it. Based on the following data of annual salaries (rounded to the nearest thousands of dollars) for college graduates and high school graduates with no college experience, does it appear that going to college is indeed worth the effort? The data are from people in their second year of employment.

College Grad	41	67	53	48	45	60	59	55	52	52	50	59	44	49	52
High School Grad	23	33	36	29	25	43	42	38	27	25	33	41	29	33	35

a. Calculate the difference between the sample mean salary for college graduates and for high school graduates.

b. On the same scale, draw dot plots of the two distributions, and discuss the variability in each distribution.

c. Calculate the MAD for each distribution. Based on the MADs, compare the variability in each distribution. Is the variability about the same? Interpret the MADs in the context of the problem.

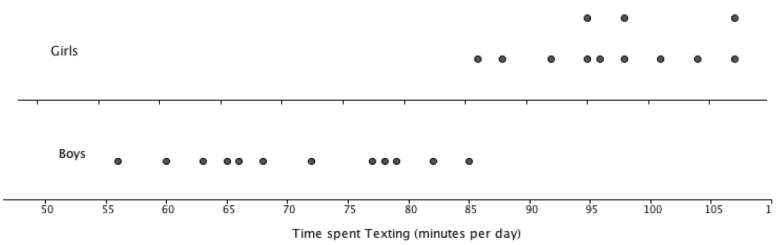
d. Based on your calculations, is going to college worth the effort? Part of your reasoning should involve the number of MADs that separate the sample means.

Suppose that Brett randomly sampled 12 tenth-grade girls and boys in his school district and asked them for the number of minutes per day that they text. The data and summary measures follow.

Gender	Number of Minutes of Texting													Mean	MAD
Girls	98	104	95	101	98	107	86	92	96	107	88	95		97.3	5.3
Boys	66	72	65	60	78	82	63	56	85	79	68	77		70.9	7.9

1. Draw dot plots for the two data sets using the same numerical scales. Discuss the amount of overlap between the two dot plots that you drew and what it may mean in the context of the problem.

   There is no overlap between the two data sets. This indicates that the sample means probably differ, with girls texting more than boys on average. The girls' data set is a little more compact than the boys, indicating that their measure of variability is smaller.



2. Compare the variability in the two data sets using the MAD. Interpret the result in the context of the problem.

   The MAD for the boys' number of minutes spent texting is 7.9 min., which is higher than that for the girls, which is 5.3 min. This is not surprising as seen in the dot plots. The typical deviation from the mean of 70.9 is about 7.9 min. for boys. The typical deviation from the mean of 97.3 is about 5.3 min. for girls.

3. From 1 and 2, does the difference in the two means appear to be meaningful?

   The difference in means is $97.3 - 70.9 = 26.4$ min. Using the larger MAD of 7.9 min., the means are separated by $\frac{26.4}{7.9} = 3.3$ MADs. Looking at the dot plots, it certainly seems as though a separation of more than 3 MADs is meaningful.

1. A school is trying to decide which reading program to purchase.

   a. How many MADs separate the mean reading comprehension score for a standard program ($\text{mean} = 67.8$, $\text{MAD} = 4.6$, $n = 24$) and an activity-based program ($\text{mean} = 70.3$, $\text{MAD} = 4.5$, $n = 27$)?

      The number of MADs that separate the sample mean reading comprehension score for a standard program and an activity-based program is $\frac{70.3 - 67.8}{4.6} = 0.54$, about half a MAD.

   b. What recommendation would you make based on this result?

      The number of MADs that separate the programs is not large enough to indicate that one program is better than the other program based on mean scores. There is no noticeable difference in the two programs.

2. Does a football filled with helium go farther than one filled with air? Two identical footballs were used: one filled with helium and one filled with air to the same pressure. Matt was chosen from the team to do the kicking. Matt did not know which ball he was kicking. The data (in yards) follow.

Air	25	23	28	29	27	32	24	26	22	27	31	24	33	26	24	28	30
Helium	24	19	25	25	22	24	28	31	22	26	24	23	22	21	21	23	25

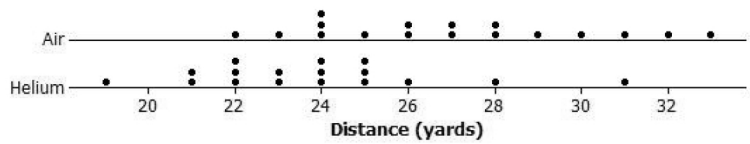
	Mean	MAD
Air	27.0	2.59
Helium	23.8	2.07

a. Calculate the difference between the sample mean distance for the football filled with air and for the one filled with helium.

The 17 air-filled balls had a mean of 27 yd. compared to 23.8 yd. for the 17 helium-filled balls, a difference of 3.2 yd.

b. On the same scale, draw dot plots of the two distributions, and discuss the variability in each distribution.

Based on the dot plots, it looks like the variability in the two distributions is about the same.



c. Calculate the MAD for each distribution. Based on the MADs, compare the variability in each distribution. Is the variability about the same? Interpret the MADs in the context of the problem.

The MAD is 2.59 yd. for the air-filled balls and 2.07 yd. for the helium-filled balls. The typical deviation from the mean of 27.0 is about 2.59 yd. for the air-filled balls. The typical deviation from the mean of 23.8 is about 2.07 yd. for the helium-filled balls. There is a slight difference in variability.

d. Based on your calculations, is the difference in mean distance meaningful? Part of your reasoning should involve the number of MADs that separate the sample means. Note that if the MADs differ, use the larger one in determining how many MADs separate the two means.

There is a separation of $\frac{3.2}{2.59} = 1.2$ *MADs. There is no meaningful distance between the means.*

3. **Suppose that your classmates were debating about whether going to college is really worth it. Based on the following data of annual salaries (rounded to the nearest thousand dollars) for college graduates and high school graduates with no college experience, does it appear that going to college is indeed worth the effort? The data are from people in their second year of employment.**

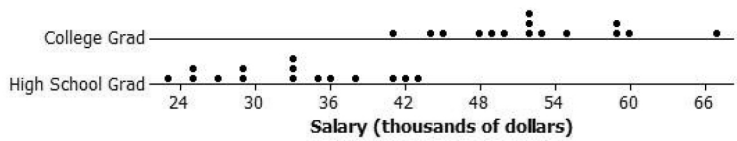
College Grad	41	67	53	48	45	60	59	55	52	52	50	59	44	49	52
High School Grad	23	33	36	29	25	43	42	38	27	25	33	41	29	33	35

a. **Calculate the difference between the sample mean salary for college graduates and for high school graduates.**

The 15 *college graduates had a mean salary of* $\$52,400$*, compared to* $\$32,800$ *for the* 15 *high school graduates, a difference of* $\$19,600$*.*

b. **On the same scale, draw dot plots of the two distributions, and discuss the variability in each distribution.**

Based on the dot plots, the variability of the two distributions appears to be about the same.



c. **Calculate the MAD for each distribution. Based on the MADs, compare the variability in each distribution. Is the variability about the same? Interpret the MADs in the context of the problem.**

The MAD is 5.15 *for college graduates and* 5.17 *for high school graduates. The typical deviation from the mean of* 52.4 *is about* 5.15 *(or* $\$5,150$*) for college graduates. The typical deviation from the mean of* 32.8 *is about* 5.17 *(*$\$5,170$*) for high school graduates. The variability in the two distributions is nearly the same.*

d. **Based on your calculations, is going to college worth the effort? Part of your reasoning should involve the number of MADs that separate the sample means.**

There is a separation of $\frac{19.6}{5.17} = 3.79$ *MADs. There is a meaningful difference between the population means. Going to college is worth the effort.*