

1. According to the Census Bureau's 2010 Current Population Survey, the mean and median 2009 income of people at least 25 years old who had a bachelor's degree but no higher degree were \$46,931 and \$58,762.

Which of these numbers is the mean and which is the median? Explain your reasoning.

2. Student engineers learn that, although handbooks give the strength of a material as a single number, in fact the strength varies from piece to piece. A vital lesson in all fields of study is that "variation is everywhere." Here are data from a typical student laboratory exercise: the load in pounds needed to pull apart pieces of Douglas fir 4 inches long and 1.5 inches square.

33190	32320	23040	24050	31860
33020	30930	30170	32590	32030
32720	31300	26520	30460	33650
28730	33280	32700	32340	31920

(a) Give the five-number summary of the distribution of breaking strengths.

(b) The stemplot shows that the distribution is skewed to the left. Does the five-number summary show the skew?

3. The table below gives the distribution of the weight at birth for all babies born in the United States in 2008.

Weight(grams)	Count	Weight(grams)	Count
Less than 500	6, 581	3,000 to 3,499	1, 663, 512
500 to 999	23, 292	3,500 to 3,999	1, 120, 642
1,000 to 1,499	31, 900	4,000 to 4,499	280, 270
1,500 to 1,999	67, 140	4,500 to 4,999	39, 109
2,000 to 2,499	218, 296	5,000 to 5,499	4, 443
2,500 to 2,999	788, 148		

(a) For comparison with other years and with other countries, we prefer a histogram of the percents in each weight class rather than the counts. Explain why.

(b) How many babies were there?

(c) Make a histogram of the distribution, using percents on the vertical scale.

(d) What are the locations of the median and quartiles in the ordered list of all birth weights? In which weight classes do the median and quartiles fall?

4. A report says that "the median credit card debt of American households is zero." We know that many households have large amounts of credit card debt. In fact, the mean household credit card debt is close to \$8000. Explain how the median debt can nonetheless be zero.

5. This is a standard deviation contest. You must choose four numbers from the whole numbers 0 to 10, with repeats allowed.

(a) Choose four numbers that have the smallest possible standard deviation.

(b) Choose four numbers that have the largest possible standard deviation.

(c) Is more than one choice possible in either (a) or (b)? Explain.

6. Create a set of 5 positive numbers (repeats allowed) that have median 7 and mean 10. What thought process did you use to create your numbers?

7. Adolescent obesity is a serious health risk affecting more than 5 million young people in the United States alone. Laparoscopic adjustable gastric banding has the potential to provide a safe and effective treatment. Fifty adolescents between 14 and 18 years old with a body mass index (BMI) higher than 35 were recruited from the Melbourne, Australia, community for the study. Twenty-five were randomly selected to undergo gastric banding, and the remaining twenty-five were assigned to a supervised lifestyle intervention program involving diet, exercise, and behavior modification. All subjects were followed for two years. Here are the weight losses in kilograms for the subjects who completed the study:

Gastric Banding

35.6	81.4	57.6	32.8	31.0	37.6
36.5	-5.4	27.9	49.0	64.8	39.0
43.0	33.9	29.7	20.2	15.2	41.7
53.4	13.4	24.8	19.4	32.3	22.0

Lifestyle Intervention

6.0	-17.0	2.0	-3.0	1.4	4.0
20.6	11.6	15.5	-4.6	15.8	34.6
6.0	-3.1	-4.3	-16.7	-1.8	-12.8

- (a) In the context of this study, what do the negative values in the data set mean?
- (b) Give a graphical comparison of the weight loss distribution for both groups using side-by-side boxplots. Provide appropriate numerical summaries for the two distributions and identify any high outliers in either group.
- (c) What can you say about the effects of gastric banding versus lifestyle intervention on weight loss for the subjects in this study?

8. Calculate by hand the standard deviation of the data set: 3.49, 1.64, and 4.73. (To receive credit you must show detailed and well written calculations of both the mean and standard deviation.)

9. We asked the students in a large first-year college class how many minutes they studied on a typical weeknight. Here are the responses of random samples of 30 women and 30 men from the class:

Women					Men				
270	150	180	360	180	120	120	30	45	200
120	180	120	240	170	90	90	30	120	75
150	120	180	180	150	150	90	60	240	300
200	150	180	120	240	240	60	150	60	30
120	60	120	180	180	30	230	120	95	150
90	240	180	115	120	0	200	120	120	18

- (a) Calculate \bar{x} and s for each group.
- (b) Generally speaking, what kinds of distributions are best summarized by \bar{x} and s ? Do you think these summary measures are appropriate in this case?
- (c) One student in each group claimed to study at least 300 minutes (five hours) per night. Remove these observations and recalculate \bar{x} and s for each group. How much does removing these outliers change the value of \bar{x} and s ?