

Measures of Variation

We've spent some time with measures of center, and now we turn to another kind of numerical statistic — measures of variation, spread, dispersion, or variability. The two most common such statistics are the range and the standard deviation.

The **range** of a set of data is the difference between the highest and lowest values.

$$\text{RANGE} = \text{MAX} - \text{MIN}$$

Since the range is highly impacted by an outlier, there is another measure of variability which is more commonly used. **Standard deviation** is an average of differences between data values and the mean and is often used along with the mean to summarize a given data set.

As a reminder, with the mean, whether the entire population or a sample is involved is really an insignificant point. Both calculations involve the total of the data values divided by the number of data values.

Population Mean	Sample Mean
$\mu = \frac{\Sigma X}{N}$	$\bar{X} = \frac{\Sigma X}{n}$

However, with the standard deviation, the population and sample standard deviation formulas differ slightly. As you decide which formula to use, think about whether you have a complete set of data or data from the entire population of interest. For example, we use the population standard deviation for the number of home runs Henry Aaron hit through every season of his career or for the ages of all students at our college.

On the other hand, if the data set is incomplete or the data was gathered from a sample or subset of the population, we use the sample standard deviation formula. For example, use the sample standard deviation for a current or active sports player or for weights of babies born in a given time period at Tift Regional Medical Center.

Population Standard Deviation (σ)	Sample Standard Deviation (s)
$\sigma = \sqrt{\frac{\Sigma(X - \mu)^2}{N}} \quad \text{OR}$ $\sigma = \sqrt{\frac{\Sigma X^2}{N} - \left(\frac{\Sigma X}{N}\right)^2}$	$s = \sqrt{\frac{\Sigma(X - \bar{X})^2}{n - 1}} \quad \text{OR}$ $s = \sqrt{\frac{\Sigma X^2 - (\Sigma X)^2/n}{n - 1}}$

Note: The major difference between the population and sample calculations is the $n - 1$ in the sample formula. Think of it as a compensation for the fact that \bar{X} is only an estimate of μ . The first formulas are more “definitional”, showing averaging of differences between data values and the mean. In each case, the second formula is best (and easiest) to use because it avoids the rounding issue; these two “shortcut formulas” just require two sums — the sum of the data values (ΣX) and the sum of the squares of the data values (ΣX^2).

Technology notes:

In a TI graphing calculator, after the data values have been entered into a list, you can calculate 1-variable statistics, and the TI calculator uses the symbol “ σ_x ” for the population standard deviation, “ s_x ” for the sample standard deviation, “minX” for the minimum value, “maxX” for the maximum value, and you can then calculate the range.

The computer spreadsheet software Excel uses the formula structure “=STDEV(first cell:last cell)” for the sample standard deviation, “=STDEVP(first cell:last cell)” for the population standard deviation, and “=MAX(first cell:last cell)–MIN(first cell:last cell)” for the range.

Examples:

- (1) Michael Jordan played 15 seasons of his NBA career (1984-1998, 2001-2003), and ended each regular season with these points per game (PPG) averages. Ignore the two seasons in which he only played in 17 and 18 games (notice the strikethroughs); the rest of the seasons, he played at least 60 games of the 82-game season. Compute the range and standard deviation for his scoring averages through these 13 typical seasons in his career.

28.2 ~~22.7~~ 37.1 35.0 32.5 33.6 31.5 30.1 32.6 ~~26.9~~ 30.4
29.6 28.7 22.9 20.0



Solutions: Range = $37.1 - 20.0 = 17.1$ PPG

Standard deviation (population):

$$\Sigma X = 28.2 + 22.7 + 37.1 + \dots + 20.0 = 392.2$$

$$\Sigma X^2 = 28.2^2 + 22.7^2 + 37.1^2 + \dots + 20.0^2 = 12,091.3$$

$$\sigma = \sqrt{\frac{12091.3}{13} - \left(\frac{392.2}{13}\right)^2} \quad \mathbf{B} \quad 4.5 \text{ PPG}$$

- (2) In an experiment on the effect of a drug on reaction time, a patient is asked to depress a button whenever a light flashes. The reaction time (in milliseconds) for 15 trials is:

102 135 123 143 134 130 122 123 103 98 93 112
117 105 128

Compute the range and standard deviation for these reaction times.

Solutions: Range = $143 - 93 = 50$ ms

Standard deviation: $\Sigma X = 102 + 135 + 123 + \dots + 128 = 1,768$

(for a sample) $\Sigma X^2 = 102^2 + 135^2 + 123^2 + \dots + 128^2 = 211,580$

Substituting these two calculated values, we have

$$s = \sqrt{\frac{211,580 - 1,768^2/15}{15 - 1}} = \sqrt{\frac{(211,580 - 1,768^2/15)}{14}} \quad \mathbf{B} \quad 15.1 \text{ ms}$$

- (3) The following data represent the numbers of years of teaching experience of a random sample of 10 faculty members from a state college

7 5 10 26 17 13 4 4 12 16

Compute the range and standard deviation for the faculty experience at this college.

Solutions: Range = $26 - 4 = 22$ years

Standard deviation (for a sample):

$$\Sigma X = 7 + 5 + \dots + 16 = 114$$

$$\Sigma X^2 = 7^2 + 5^2 + \dots + 16^2 = 1,740$$

$$s = \sqrt{\frac{1740 - 114^2/10}{9}} = \sqrt{\frac{(1,740 - 114^2/10)}{9}} \quad \mathbf{B} \quad 7.0 \text{ years}$$

Exercises In general, unless otherwise indicated, round mean and standard deviation results to one more decimal place than the given data values.

1. For the following golf scores, compute the mean and standard deviation. Then compare the players' performances using statistical language. By the way, Andy and Barney have played golf together before these 5 rounds, and they have played since these 5 rounds.

Andy:	78	83	81	84	80
Barney:	82	92	74	100	78

2. The number of consecutive hours a light bulb will last before it burns out is tested. The number of hours for 24 bulbs is shown below. Compute the range and standard deviation for the life of a light bulb.



402	405	409	389	456	423	432	440	425	436	423	430
421	423	408	411	434	448	392	400	439	430	412	409

3. For the following new home sale prices for 10 homes built in Tift County recently, compute the range and standard deviation. Here, round to the nearest dollar.

\$42,000	\$82,000	\$100,000	\$76,500	\$95,000
\$67,000	\$85,000	\$102,500	\$115,000	\$310,000

4. Here are Ted Williams' home run totals by season (1939-1960):

31	23	37	36	0*	0*	0*	38	32	25	43
28	30	1**	13**	29	28	24	38	26	10	29
				*Active service, WWII		**Active service, Korean War				



Compute the range and standard deviation for his home runs (excluding the 5 service year outliers).

5. A random sample of 10 College Algebra and 10 Math for Elementary Teachers students completed a survey form on several demographics including age. The ages of the 20 participants are listed below. Calculate the mean and standard deviation for each data set. Then briefly compare the courses in terms of a typical student's age and the spread or variability of the students' ages (using these samples).

College Algebra	20	19	19	21	25	19	20	22	34	29
Math for Teachers	26	24	24	19	27	25	32	29	27	22

6. Here are test scores in BIOL 2107 for students on the same 5 tests. Compute each student's mean, range, and standard deviation. Since this may be a sample of test results or a complete data set, calculate both σ and s . Compare the students using statistical language.

Rack 75, 82, 98, 90, 79 Shack 63, 95, 85, 76, 100 Benny 78, 82, 80, 83, 97

7. Refer to the following cattle weights:

748 lb	485 lb	807 lb	1,023 lb	761 lb
765 lb	934 lb	579 lb	865 lb	1,064 lb



Compute the range and standard deviation for these weights.

8. A policeman records the speeds to the nearest mile per hour of vehicles as they observe 30 minutes of rural highway traffic.

44	54	50	46	45	49	55	44	42	55	51	52
54	44	60	44	59	41	44	47	51	42	43	

Calculate the range and standard deviation of speeds for the drivers.

9. A teacher conducts a sample and records the following household sizes for 9 of their students: 5, 4, 3, 4, 2, 4, 5, 5, 5. Calculate the range and standard deviation for these household sizes.

10. A teacher conducts a sample and records the following numbers of credits enrolled (in the current semester) for 9 of their students: 12, 15, 12, 12, 13, 12, 12, 18, 15. Calculate the range and standard deviation for this data.

11. A dietician obtains the amounts of sugar (in grams) from 1 gram from each of 16 different cereals, including Cheerios, Corn Flakes, Fruit Loops, Trix, and 12 other brands. Find the range and standard deviation of these values.

0.03 0.24 0.30 0.47 0.43 0.07 0.47 0.13
 0.44 0.39 0.48 0.17 0.13 0.09 0.45 0.43

12. We've gathered demographic data from all 25 international students at Eastpark College. Use a random or systematic technique to sample 5 students, and then compute the mean and standard deviation for

(a) GPA

(b) IQ

(c) distance traveled to campus (in miles)

Student #	Gender	Class Rank	GPA	Miles traveled to school	IQ	Major Field
1	M	Fr	1.4	1	104	Bio
2	M	Fr	2.3	2	95	Ed
3	M	So	2.7	6	108	Psy
4	F	So	3.2	7	119	Eng
5	F	Gr	3.8	12	114	Ed
6	M	Jr	4.0	13	91	Psy
7	F	Jr	3.0	2	106	Eng
8	M	Jr	3.3	6	100	Bio
9	F	Se	2.7	9	102	Eng
10	F	So	2.3	5	99	Ed
11	M	Se	1.6	18	100	Bus
12	M	Gr	3.2	7	105	Psy
13	F	Gr	3.8	3	103	Bus
14	F	Se	3.1	5	97	Eng
15	F	Jr	2.7	5	106	Bio
16	F	Fr	1.4	4	114	Bus
17	M	So	3.6	17	102	Ed
18	M	Fr	2.2	1	101	Psy
19	F	So	4.0	7	108	Bus
20	M	Gr	2.1	4	97	Ed
21	F	Fr	2.0	3	113	Bio
22	F	So	3.6	4	104	Bio
23	F	Gr	3.3	16	110	Eng
24	F	Se	2.5	4	99	Psy
25	M	So	3.0	5	96	Psy

13. Find the range and standard deviation for the following reaction times (measured in milliseconds). Each measurement is rounded to the nearest hundredth of a second.

19 20 17 21 21 21 19 18 19 19
 17 17 15 17 18 17 18 18 18 17