

# What Does Average Really Mean? Making Sense of Statistics

By Karen J. DeAngelis and Steven Ayers



**T**he recent shift toward greater accountability has put many educational leaders in a position where they are expected to collect and use increasing amounts of data to inform their decision making. Yet, because many programs that prepare administrators, including school business officials, do not require a statistics course or a course that is more theoretical than practical, administrators tend to have only limited knowledge of how to analyze, interpret, and use their data.

We're here to help. This article provides a user-friendly primer on basic statistics and how you can use them to make sense of your data.

Statistics can be broken into two broad categories: descriptive and inferential. Descriptive statistics are used to describe or summarize a collection of data.

You almost certainly have experience with descriptive statistics, such as the mean that is often used for reporting attendance data and personnel salaries. Inferential statistics, in contrast, are most often used when you want to generalize beyond the data that you have. For example, your district has results from a curricular intervention that was piloted in two elementary schools and needs to determine whether to expand the intervention to the rest of its elementary classrooms. You would use inferential statistics in your decision-making process.

The type of statistics you use depends on the data you have and the types of questions you want answered. Given the breadth of this topic, we focus herein on descriptive statistics and leave inferential statistics for another time.

Suppose your superintendent asks you for information regarding teacher salaries in your district. Chances are she is not interested in seeing the salary of every teacher. Rather, she wants a small amount of information that will enable her to judge how well teachers are compensated. A couple of summary statistics and a simple chart would likely provide the superintendent with the information she needs. Descriptive statistics can come in handy in this situation.

### Measures of Central Tendency

A measure of central tendency provides a single number that is most representative of your data points. It is an indicator of what value is “typical.” There are three measures of central tendency: mean, median, and mode. Which measure you use largely depends on the types of data you want to describe.

**Mean.** Your initial reaction to the superintendent’s request might be to provide her with the mean salary of teachers in your district. This makes sense given that the mean is perhaps the most well-known and often-used measure of central tendency. As Salkind (2008) explains, the mean of a distribution of data is like a fulcrum on a seesaw; it indicates the centermost point of your data. The mean is easy to calculate and is generally considered the most precise measure of central tendency because it takes into account all the data points in your collection of data (Gravetter and Wallnau 2008).

You must recognize, however, that the mean is not always a good representation of a set of data. For example, if your data contain outliers (i.e., data points that are much higher or much lower in value than the majority of your data points), those outliers will influence the mean. As a result, the mean will be higher or lower than what is typical of the majority of values in your data set and will poorly represent your data.

It is for this reason that the mean is not generally used to report home sale prices or managerial salaries (the sales price of mansions and the salaries of CEOs tend to be outliers). With teacher salary data, you could have a problem with outliers if you have a situation in which most teachers in your district are relatively new to the profession, but a few are on the verge of retirement after spending 30 years in the classroom.

Alternatively, you may be asked to report on types of data for which a mean is neither very informative nor appropriate for use. As a case in point, student performance on New York State assessment exams is reported on what is called an ordinal scale, ranging from 1 (below proficient) to 4 (advanced). The numbers 1 through 4 are used to represent rank-ordered achievement cate-

gories, but they have no real numerical meaning aside from the ranking. It is entirely possible to calculate a mean for ordinal data, but it will provide only limited information (what does a mean of 2.7 really tell you?).

## Together, a measure of central tendency and a measure of dispersion provide valuable information about the characteristics of a set of data.

Nominal data are similar to ordinal data, except nominal categories have no rank order and may not even be represented numerically. Examples of nominal data include gender, race/ethnicity, item responses on multiple-choice tests, and courses taken by high school students. The mean is not at all appropriate for this type of information.

**Median.** If you order your data points by value from lowest to highest, the value or score that falls exactly in the middle of the list is called the median. Thus, the median identifies the midpoint of a data distribution based on the number of values or scores involved. The median is typically used in place of the mean when a data set includes outliers. That is because the numerical value of the outliers has no influence on the median.

The median works well with ordinal data, too; once you know the median, you know (by definition) that half the scores in your distribution are at or below the median and half are at or above it. Finally, the median, like the mean, is inappropriate for use with nominal data because such data cannot be ordered.

**Mode.** The mode is the value or score that occurs most frequently in your data. The mode can be used with any types of data, including nominal data. When appropriate to use, the mean and median tend to be preferred to the mode because they convey better-representative information. Nonetheless, the mode is very useful for describing nominal data, like item responses on multiple-choice tests (e.g., B was the modal response on question 1).

Unfortunately, measures of central tendency by themselves provide an incomplete picture of what a set of data looks like. This is because two data sets can have the same result for a measure of central tendency, but they can be quite different for the range of values or scores that comprise them. Figure 1 provides an example using teacher salary data for two districts, both of which have a mean teacher salary of \$50,000. In district A, individual teacher salaries range from \$22,000 to \$78,000. In district B, salaries range from \$42,000 to \$58,000. There is much more spread in the salaries in district A. As a result, the mean by itself is not as good a representation of the salaries in district A as it is for district B.

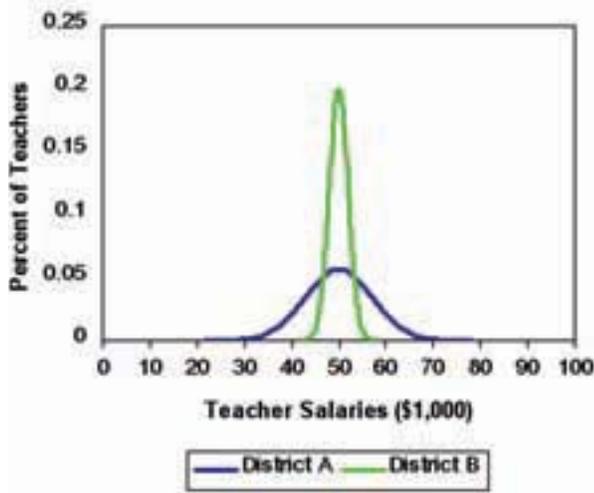


Figure 1. Distributions of teacher salaries.

### Measures of Dispersion or Variability

A measure of dispersion is used to capture how different or spread out the values in a data set are from one another. We present three common measures of dispersion here. It is important to note that there are no measures of dispersion that are appropriate to use with nominal data.

**Range.** The range is the simplest measure of dispersion. It is the difference between the largest and smallest values or scores in your data set. The ranges for the teacher salary data shown in figure 1 are \$56,000 and \$16,000 for the distributions of districts A and B, respectively. This indicates that the salaries of the lowest- and highest-paid teachers in district B differ by only \$16,000, but they differ by \$56,000 in district A.

A major drawback of the range is that it is determined by just two values in a data set. As a result, it is very sensitive to outliers and fails to account for how spread or clustered all the other values are. We can use our salary example to illustrate this latter problem. With a mean of \$50,000 and a range of \$16,000, it could be that half the teachers in district B earned \$42,000 and the other half earned \$58,000. Looking at figure 1, that is obviously not the case. However, we would be unable to determine that knowing just the mean and the range.

**Interquartile range.** The interquartile range tends to be preferred to the range as a measure of dispersion because it is based on only the middle 50% of a distribution. By disregarding the lowest 25% and highest 25% of values, the interquartile range is much less sensitive to extreme values than the range. This measure of dispersion still does not capture the degree of spread or clustering of the rest of the values in the data set. Nonetheless, the interquartile range is often used when the median is used as a measure of central tendency.

**Standard deviation.** The standard deviation is the most commonly used measure of dispersion. In very

simplistic terms, it indicates the average distance of all the values or scores from the mean of the distribution. The more each value differs from the mean, the larger the standard deviation will be. In our teacher salary example, district A's salaries have a standard deviation of \$7,000, 3.5 times larger than district B's standard deviation of \$2,000.

When the shape of a distribution of data is approximately normal (i.e., bell-shaped) like those shown in figure 1, knowing the standard deviation provides you with a wealth of information about the spread of values in your data set. More specifically, about 68% of the values or scores in a normal distribution lie within  $\pm 1$  standard deviation of the mean (roughly 34% fall within one standard deviation below the mean and another 34% fall within one standard deviation above the mean). About 95% of the values or scores fall within  $\pm 2$  standard deviations of the mean (Gravetter and Wallnau 2008). Applying this to our salary example, 68% of the teachers in district B earned salaries between \$48,000 and \$52,000, whereas 95% earned salaries between \$46,000 and \$54,000.

Given its relationship with the mean, the standard deviation is appropriate to use when the mean is used as the measure of central tendency.

### Your Turn . . .

Together, a measure of central tendency and a measure of dispersion provide valuable information about the characteristics of a set of data. There are additional descriptive tools not covered in this article, such as frequency tables and charts, that we know you would find useful as well. Statistical and spreadsheet programs, like SPSS and Excel, now make it easy to calculate such statistics. It is up to you, however, to determine what statistics to use to best represent your data. We hope the basic information provided here will prompt you to seek additional information that will help you make good use of your data.

### References

- Gravetter, F. J., and L. B. Wallnau. 2008. *Essentials of statistics for the behavioral sciences*. 6th ed. Belmont, Calif.: Thomson Wadsworth.
- Salkind, N. J. 2008. *Statistics for people who (think they) hate statistics*. 3rd ed. Thousand Oaks, Calif.: SAGE Publications, Inc.

---

**Karen J. DeAngelis** is an assistant professor in the Department of Educational Leadership at the University of Rochester in New York. Email: [kdeangelis@warner.rochester.edu](mailto:kdeangelis@warner.rochester.edu).

**Steven Ayers** is the assistant superintendent for business in the Hilton Central School District in New York. Email: [SAYERS@hilton.k12.ny.us](mailto:SAYERS@hilton.k12.ny.us).