



## Using Statistics

Brought to you by the Purdue University Online Writing Lab at <http://owl.english.purdue.edu>  
Graphics for this handout were designed by [Michelle Hansard](#).

---

For more information about citing sources in your research paper, see our workshop on writing a research paper at <http://owl.english.purdue.edu/workshops/hypertext/ResearchW/index.html>.

Statistics are often tossed around as if they could speak for themselves. For example, advertisers claim "Ivory soap is 99% pure." (Pure what?) Or a researcher may claim that "the average American today watches 5.3 hours of TV per day." (What does "average" mean?) All facts must be interpreted and presented in your argument; this handout presents five guidelines designed to help you use statistics responsibly.

[Define "Average"](#) || [Focus on Relevant Details](#) || [Give the Full Picture](#) || [Keep Visual Aids in Proportion](#) || [Give Bases of All Percentages](#)

---

## Define "Average"

There are basically three kinds of averages commonly used in statistics. These are defined as follows and illustrated in the examples given below.

### Mean (Arithmetic Mean)

a weighted average between extremes; a total of all numbers included divided by the quantity of numbers represented

### Median

- a. the number midway through an odd set of numbers For example, in the set (22, 27, 31, 32, 38, 40, 44), the median = 32 because there are 3 scores below this number and 3 above it.
- b. a value half way between two middle numbers in an even set. For example, in the set (22, 27, 31, 32, 35, 38, 40, 44), the median = 33.5 because that is halfway between the middle two numbers.

### Mode

a number or value that occurs most frequently in a series.

Note: if two or more values occur with equal frequency, the mode is the mean of those values.



this is a graph that shows examples of mean, median and mode

Always let the reader know which average you are using.

## When "average" is left undefined, the following confusion can result.

*"The average employee in Zoro Company already makes \$24,090, so there is no need to give raises this year," says the company president.*

*"The average employee in Zoro Company makes \$20,000. That is a decent salary," says the employment interviewer for*

Zoro Company.

"The average employee in Zoro Company makes only \$10,000 a year. That is a disgracefully low wage," says the union leader.

All three of these people are telling the truth. The problem is that each of them is failing to identify which average is being used.

Here is the list of employees and their salaries at Zoro Company:

Employee position	Salary	Number of Employees	Total amount of money paid to those employees
President	\$100,000	1 employee	\$100,000
Senior Executives	\$25,000	3 employees	\$75,000
Junior Executrices	\$20,000	2 employees	\$40,000
Office Clerks	\$10,000	5 employees	\$50,000
<b>Totals</b>	<b>\$155,000</b>	<b>11 employees</b>	<b>\$265,000</b>

The computation of possible averages for Zoro Company is as follows:

**Mean** = \$265,000 (total wages) / 11 (total employees) = \$24,090

**Median** = (middle number in list of eleven employees) = \$20,000

**Mode** = (number occurring most frequently) = \$10,000

Notice how the averages can vary tremendously. Only when the average is defined can the facts be correctly interpreted and properly understood.

[back to the top](#)

---

## Focus on Relevant Details

Selecting the significant facts is very important to persuasive writing. In this selection process, relevance to the decision or circumstance at hand is crucial. For example, if you were asked to perform a study to determine the feasibility of installing a word processor in a secretarial typing pool, you would probably begin your work by taking a random sample of the work load during a typical day. You might discover the following work load for letters being typed each hour:

Hour Observed	Number of Letters Typed
9-10:00	75
11-12:00	150
1-2:00	300
3-4:00	200
5-6:00	150
<b>Total for five hours</b>	<b>875</b>

This random sample will help you see what the typical work load requirements are throughout the day.

From the work load observations you can generate four significant figures:

**Mean** = 875 (total letters typed) / 5 (total hours observed) = 175

**Median** = (visual center of numbers in ranked order) = 150

**Mode** = (number appearing most often in hourly totals) = 150

**Peak** = (highest output during the day) = 300

If the break-even point for buying a word processor is 160 letters per hour, you could argue that the mean (175 letters per hour) is the most important figure, and thereby present evidence in favor of buying the word processor.

However, if the decision to buy a word processor has already been made and the only question is how much capacity this new machine must have, then the most relevant fact is the peak load requirement of 300 letters per hour.

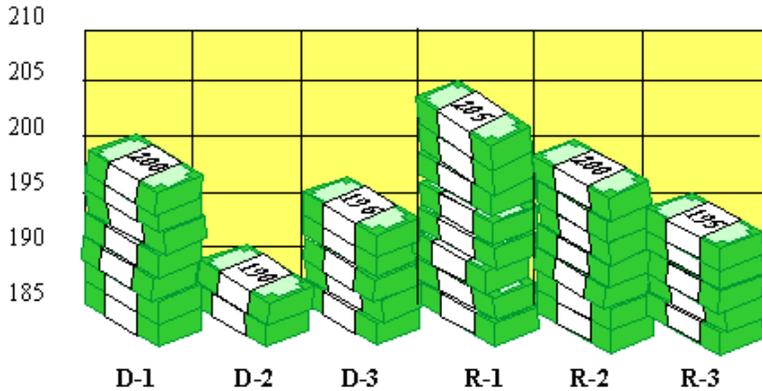
To return to an earlier example, in the Zoro Company illustration presented in the above section, a reader may justly assume that the president's salary is not relevant to the average employee's wage, thus making the mean only \$16,500. This more conservative figure reflects the typical employee's wages more accurately than does the earlier figure of \$24,090. On the other hand, a potential executive may simply choose to focus on the wages of people in his own category, thus making the figures for junior and senior executives the only relevant ones.

Whatever the circumstances, selecting relevant facts and figures is the mark of wisdom. Choose carefully and you will add to the persuasive qualities of your argument. (Should you be in doubt about the relevance of certain facts to your argument, you may include these facts in a footnote or an appendix.)

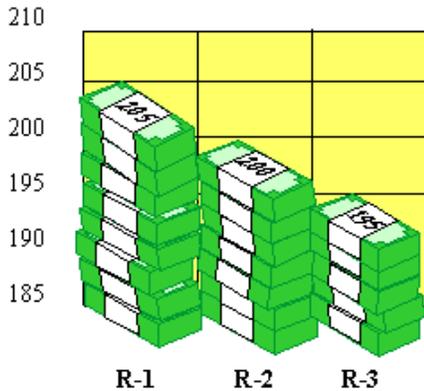
## Give the Full Picture

To earn the trust of readers, give them the complete information needed to draw their own conclusions. Being open about all the facts will add to your credibility and reinforce your argument. While emphasizing certain facts to support your conclusions is always in order, hiding important facts never is.

Election campaigns frequently involve hiding facts during debates on the economy, and statistics are often tossed around very freely. For the sake of discussion, assume the following bar graph depicts the Gross National Product (GNP) for three years preceding and three years following a change of presidency from the Democratic to the Republican party.

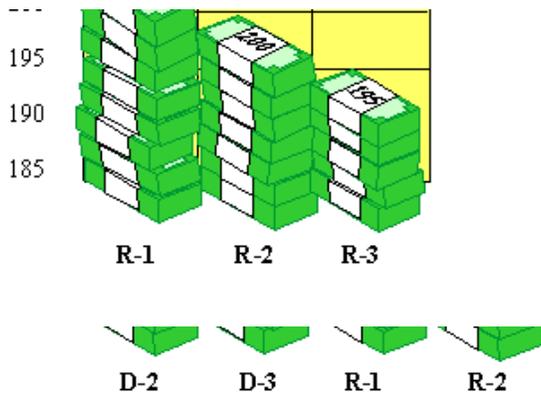


A Democrat arguing in favor of his party might choose to present just the corner of the graph that shows a decline in GNP under the Republican administration.



**Notice how the graph makes economic disaster seem imminent under the Republican party's rule.**

A Republican arguing in favor of his party might choose to present just that part of the graph which emphasizes how much the Republicans have done to improve the economy.



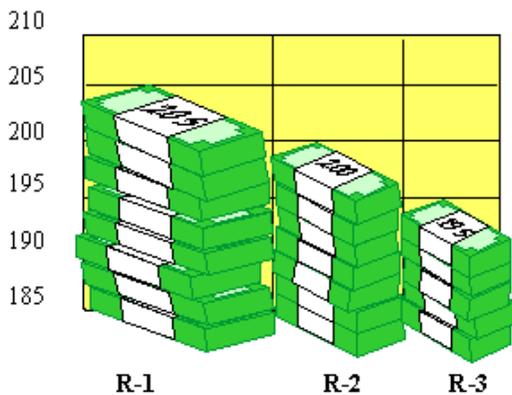
**Notice how this graph makes the Republicans seem like heroes who have brought the nation from poverty to prosperity.**

To be fair to all parties concerned, graphs should begin at zero and present enough information to allow readers to determine objectively their own conclusions.

[back to the top](#)

## Keep Visual Aids in Proportion

To present statistical information responsibly, you should keep visual aids in comparisons in proportion, both vertically and horizontally. For example, an illustration of decreased GNP could easily distort the changes as follows:



While this graph is reasonably proportional vertically, the horizontal changes make the GNP appear to have decreased by 75% in just two years. The visual effect thus distorts the truth and consequently erodes the credibility of the writer.

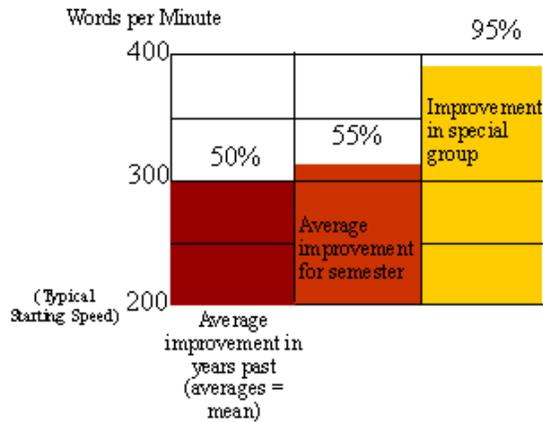
Whether you choose to use bar graphs, pie graphs, or unit graphs (in which an X or another symbol represents an increment, such as \$100), be sure to keep them in proportion so that your readers can draw their own conclusions from the facts and verify your assertions.

[back to the top](#)

## Give Bases of All Percentages

Because percentages are always derived from a specific base, they are meaningless until associated with a base. For example, to claim that customers can save 20% on purchases does not say much since that 20% savings may well be made after the products being sold have been marked up an additional 15% of the original sales price.

To cite another example, studies of performance changes often use percentages to indicate improvements or declines. The following graph describes the improvement of a special group of speed reading students in comparison to average speed reading students of the same semester and of previous years.



From this illustration several interesting percentages can be generated. The average students over the years have improved by 50% over their initial reading speed. The average students for the given semester measured improved by 55% over their initial speed, and the students in the special group improved by 95% over their initial speed of 200 words per minute. These three percentages are all measured from a common base and therefore represent fairly the performance comparisons of the students involved.

One can, however, easily exaggerate the performance of the special students. Though they performed only 40% (95% - 55%) above the students in their own semester (which offers the fairest base for comparison), they also performed 45% above the average student of previous years. (This is only a slight exaggeration.) One can also say that the special group of speed reading students did nine times better than the other students of their semester in exceeding the old average of 50% improvement.

### Note these computations:

Average students of the given semester exceed the old average improvement by 5% (55% - 50%).

Students in the special reading program exceeded the old average by 45% (95% - 50%), which is nine times better than the 5% improvement of the other students of the same semester.

Consequently, one can say that the special students did 900% better than the other speed reading students of the same semester. This percentage is dazzling, but would be misleading or meaningless if unexplained.

As with all statistics, percentages should be presented in context so that the readers can draw their own conclusions as you emphasize facts important to your argument.

[back to the top](#)

The following information must remain intact on every handout printed for distribution.

This page is located at [http://owl.english.purdue.edu/handouts/print/research/r\\_stats.html](http://owl.english.purdue.edu/handouts/print/research/r_stats.html)

Copyright ©1995-2001 by OWL at Purdue University and Purdue University. All rights reserved.

Use of this site, including printing and distributing our handouts, constitutes acceptance of our terms and conditions of fair use, available at <http://owl.english.purdue.edu/lab/fairuse.html>.

To contact OWL, please visit our contact information page at

<http://owl.english.purdue.edu/lab/contact.html> to find the right person to call or email.