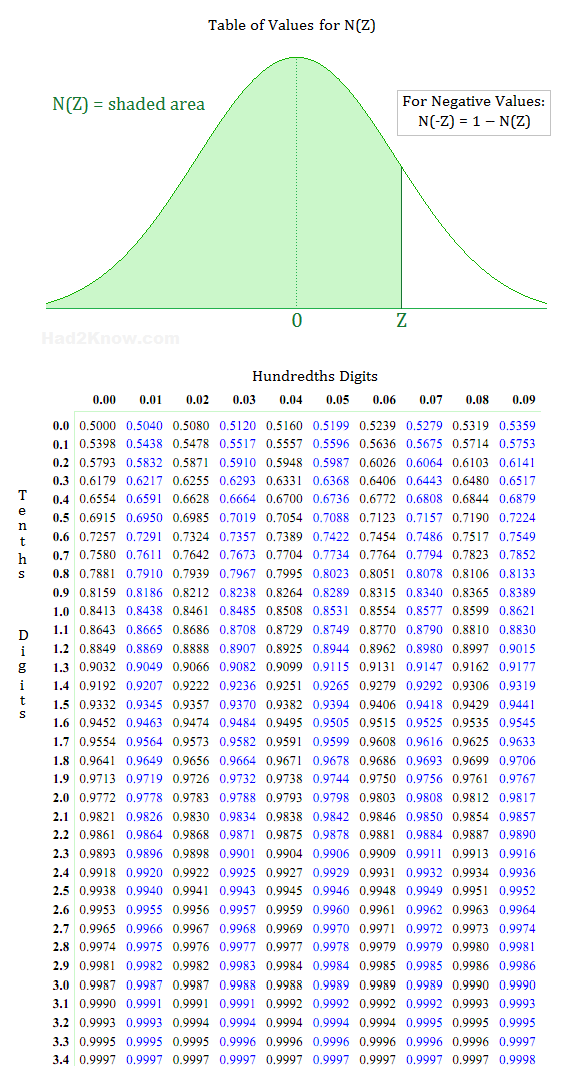
<http://www.youtube.com/watch?v=Oyjp6Ke8V9I> **Z Score Probability Table** <http://www.had2know.com/academics/normal-distribution-table-z-scores.html>



Data Analysis by Robert Niles Source: <http://www.robertniles.com/stats/dataanly.shtml>

You wouldn't buy a car or a house without asking some questions about it first. So don't go buying into someone else's data without asking questions, either. Okay, you're saying... but with data there are no tires to kick, no doors to slam, no basement walls to check for water damage. Just numbers, graphs and other scary statistical things that are causing you to have bad flashbacks to your last income tax return. What the heck can you ask about data? Plenty. Here are a few standard questions you should ask any human beings who slap a pile of data in front of you and ask you write about it. Where did the data come from? Always ask this one first. You always want to know who did the research that created the data you're going to write about. You'd be surprised — sometimes it turns out that the person who is feeding you a bunch of numbers can't tell you where they came from. That should be your first hint that you need to be very skeptical about what you are being told. Even if your data have an identifiable source, you still want to know what it is. You might have some extra questions about a medical study on the effects of secondhand smoking if you were to learn that it came from researchers employed by a tobacco company instead of from, say, a team of research physicians from a major medical school. You might question a study about water safety that came from a political interest group that had been lobbying Congress for a ban on pesticides. Just because a report comes from a group with a vested interest in its results doesn't guarantee the report is a sham. But you should always be skeptical when looking at research generated by people with a political agenda. At the very least, they have plenty of incentive NOT to tell you about data they found that contradict their organization's position. Which brings us to the next question: Have the data been peer-reviewed? Major studies that appear in journals like the New England Journal of Medicine undergo a process called "peer review" before they are published. That means that professionals — doctors, statisticians, etc. — have looked at the study before it was published and concluded that the study's authors followed the rules of good scientific research and didn't torture their data like a middle ages infidel to make the numbers conform to their conclusions.

Always ask if research was formally peer reviewed. If it was, you know that the data you'll be looking at are at least minimally reliable. And if it wasn't peer-reviewed, ask why. It might be that the research just wasn't interesting to enough people to warrant peer review. Or it could mean that the research had as much chance of standing up to professional scrutiny as a $500 mobile home has of standing up in a tornado. How were the data collected? This one is real important to ask, especially if the data were not peer-reviewed. If the data come from a survey, for example, you want to know that the people who responded to the survey were selected at random. How many times have you seen news reports based on call-in polls or website surveys? Those can be fun (see my notes on the ladder of engagement in Chapter 8), but they aren't news that other publications should be reporting. Why? These types of surveys simply reflect the views of what statisticians call a "self-selected sample." People who feel really passionately about one side or the other can flood the poll, skewing the results from what they would have been had you polled only a random sample of people in the community. When this happens in online surveys, it's called "Freeping" the poll, after the website FreeRepublic.com, whose readers have become notorious over the years for doing this sort of thing to polls on other websites. Another problem with data is "cherry-picking." This is the social-science equivalent of gerrymandering, where you draw up a legislative district so that all the people who are going to vote for your candidate are included in your district and everyone else is scattered among a bunch of other districts. Be on the lookout for cherry-picking, for example, in epidemiological (a fancy word for the study of disease that sometimes means: "We didn't go out and collect any data ourselves. We just used someone else's data and played 'connect the dots' with them in an attempt to find something interesting") studies looking at illnesses in areas surrounding toxic-waste dumps, power lines, high school cafeterias, etc. It is all too easy for a lazy researcher to draw the boundaries of the area he or she is looking at to include several extra cases of the illness in question and exclude many healthy individuals in the same area. When in doubt, plot the subjects of a study on map and look for yourself to see if the boundaries make sense.

Finally, be aware of numbers taken out of context. Again, data that are "cherry picked" to look interesting might mean something else entirely once they are placed in a different context. Consider the following example from Eric Meyer, a professional reporter who went on to teach the University of Illinois: "My personal favorite was a habit we use to have years ago, when I was working in Milwaukee. Whenever it snowed heavily, we'd call the sheriff's office, which was responsible for patrolling the freeways, and ask how many fender-benders had been reported that day. Inevitably, we'd have a lede that said something like, "A fierce winter storm dumped 8 inches of snow on Milwaukee, snarled rush-hour traffic and caused 28 fender-benders on county freeways" — until one day I dared to ask the sheriff's department how many fender-benders were reported on clear, sunny days. The answer — 48 — made me wonder whether in the future we'd run stories saying, "A fierce winter snowstorm prevented 20 fender-benders on county freeways today." There may or may not have been more accidents per mile traveled in the snow, but clearly there were fewer accidents when it snowed than when it did not." It is easy for people to go into brain-lock when they see a stack of papers loaded with numbers, spreadsheets and graphs. (And some sleazy sources are counting on it.) But your readers are depending upon you to make sense of that data for them. Use what you've learned on this page to look at data with a more critical attitude. (That's critical, not cynical. There is a great deal of excellent data out there.) The worst thing you can do as a writer is to pass along someone else's word about data without having any idea whether that person's worth believing or not.