Chapter 1

The Fundamental Process of Measurement

Measurement is the process of assigning numbers to properties. This process is so basic that often we don't think much about it. However, failing to think about some fundamental aspects of measurement can turn out to be a costly mistake, for several reasons:

- We may attach unwarranted significance to aspects of the numbers which do not convey meaningful information.
- We may fail to simply data when we could easily do so, and consequently waste resources.
- We may manipulate our data in ways that destroy the information they contain.
- We may perform meaningless statistical operations on our data.

In this chapter, we review the basics of measurement theory in a bit more detail than you probably saw in your undergraduate course. Then we sketch some of the practical implications of this theory.

1.1 Levels of Measurement

Measurement theory deals with what happens when numbers are assigned to objects in an attempt to describe some property of the object. For example, suppose we are measuring the amount of time it takes several events to occur by assigning numbers to each event. Numbers have properties, and quantities have properties. These properties, which are closely related to the concept of levels of measurement, are summarized in Table 1.1.

For example, consider two numbers. These numbers have a "nominal," or "same-difference" property, namely, they are either the same or different. Now consider two amounts of time. These quantities also have a "nominal," or "same-difference" property.

Table 1.1: Some Identifiable Properties of Numbers and the Quantities They Measure

Property	**Description**
Nominal	Two numbers or attributes are either the same or different. For example, you and I may have the same number of children, or different numbers of children.
Ordinal	If two numbers or attributes are different, one is either greater then or less than the other. For example, if we are different ages, I am either older or younger than you.
Interval	If pairs of attributes or numbers A,B and C,D are compared, the difference between A and B and the difference between C and D have relative sizes.
Ratio	If two numbers or attributes are compared they can have a ratio. For example, 2 is twice as large as 1, and I may weigh twice as much as you.

We see that quantities have properties, and numbers have properties, and these properties can be very similar. When we assign numbers to quantities so that the properties of the numbers match the properties of the quantities, we have good measurement. When the properties of the numbers do not match the properties of the quantities they are assigned to, we have bad measurement.

There are several identifiable properties of numbers and quantities, and they form a hierarchy. To the extent that a property of the quantities is captured by a numerical assignment, we achieve a certain level of measurement. First, let's review the properties of numbers and quantities that we will be trying to match

1.1.1 Nominal Measurement

Two attributes are either the same or different. Two numbers are either the same or different. If numbers are assigned to a group of objects so that any two objects are given the same score if and only if they have the same value on the attribute being measured, we say the numbers have achieved a nominal level of measurement.

Mathematically, we express this in the following definition.

Definition 1.1 (Nominal Measurement) Let objects A and B have values $V(A)$ and $V(B)$ on the attribute being measured. Let $M(A)$ and $M(B)$ be the scores assigned to A and B by the measurement process. We say measure M achieves a *nominal level of measurement* when for all pairs of objects A and B,

$$M(A) = M(B) \text{ iff } V(A) = V(B)$$

In the preceding definition, the symbol "iff" stands for "if and only if."

Table 1.2: Levels of Measurement - An Example

Athlete	True Time	Watch V (Nominal)	Watch W (Ordinal)	Watch X (Ordinal)	Watch Y (Interval)	Watch Z (Ratio)
A	10	23	11	2	21	20
B	11	12	14	3	23	22
C	13	20	15	4	27	26
D	20	19	18	5	41	40
E	13	20	15	4	27	26
S	0	26	9	1	1	0

Example 1.1 (Levels of Measurement) The mathematical discussion of levels of measurement is rather abstract, and the following numerical example may help illuminate the various concepts. Suppose a group of athletes are trying out for the track team at the university, and they are all being timed in the 100 meter dash. Several coaches record times for each athlete. However, they use some very unusual stopwatches. The stopwatches are at several different levels of measurement. Table 1.2 presents the "true" times of the athletes in an arbitrary unit of time roughly (but not exactly) equal to 1 second. It also presents the times as recorded by the various stopwatches.(By the way, Athlete S is a mild mannered student named Clark, who is often seen ducking into phone booths, and occasionally wears strange-looking tee shirts with a large red "S" on the front!)

If we examine the "true times" for the athletes, and compare them with the times recorded by Stopwatch V, we see that the stopwatch has done an extraordinarily poor job of reflecting meaningful information about the athletes' performances.

However, Stopwatch V has captured the same-difference property in the true times. If you examine any pair of values recorded by Stopwatch V, you can see that the two measured values are equal if and only if the true times for those runners are equal. So Athletes C and E both receive a score of 20, and all other athletes receive unique, different values. Stopwatch V hasn't done very well, but is has recorded something, i.e., the nominal (same-difference) property, correctly. Consequently, we say it has achieved a nominal level of measurement.

1.1.2 Ordinal Measurement

If two objects A and B have a different value on some attribute, then in many cases they can be ordered. If so, A is either greater than B or less than B. Similarly, if two numbers are different, they have an ordering. If numbers are assigned to a group of objects so that the same-different property is captured and the ordering is captured as well, we say the numbers have achieved an *ordinal level of measurement.*

Definition 1.2 (Ordinal Measurement) Suppose J objects X_j have attribute values $V(X_j)$ and measured values $M(X_j)$ that satisfy the qualities of Nominal Measurement. We say the measure M achieves an *ordinal level of measurement* when the measured values also satisfy following property:

$$M(X_i) > M(X_j) \text{ iff } V(X_i) > V(X_j)$$

Numbers achieve an ordinal level of measurement if they capture the same-difference distinction in a set of attributes, and also reflect a proper ordering. Examine the numbers recorded by Stopwatch W. These values properly reflect the nominal (same-different) and ordinal (ordering) properties of the numbers. Two times recorded by the stopwatch are the same if and only if the corresponding true times are the same. One time is greater than another if and only if the true times have the same ordering. We say that Stopwatch W has achieved an ordinal level of measurement.

Notice that the relative differences between the numbers recorded by Stopwatch W do not convey correct information about the relative differences in the true times. For example, the difference in true times between Athletes B and C is twice as great as the difference between athletes A and B. This fact is not represented correctly in the numbers recorded by Stopwatch W. Hence, although the ordering of the numbers recorded by this stopwatch is correct, the differences between values are basically worthless, and add a misleading complexity to the data values.

To pursue this point, consider the numbers recorded by Stopwatch X. These numbers are also at the ordinal level of measurement. They capture the same-different property and the ordering just as well as the numbers yielded by Stopwatch W.

Does the simplicity of the numbers produced by Stopwatch X make you feel uncomfortable? It should not. These numbers capture as much meaningful information as those produced by Stopwatch W, and they eliminate needless complexity.

1.1.3 Interval Measurement

An important aspect of an attribute is the amount by which objects differ on that property. A measure can achieve an ordinal level of measurement, but not properly reflect the relative size of attribute differences, as we shall see in an example below. If, in addition to the nominal and ordinal properties mentioned above, the relative size of the differences in the numbers assigned to objects by a measure accurately reflects the relative differences in the attribute being measured, the numbers achieve an interval level of measurement. Formally, all that is required to achieve this property is that, throughout the scale generated by M, equal differences in a measure reflect equal differences in the attribute being measured. If this is true, then it can be shown that relative differences in the measure reflect relative differences in the attribute.

Definition 1.3 (Interval Measurement) Suppose objects X_j have attribute values $V(X_j)$ and measured values $M(X_j)$ that satisfy the qualities of ordinal measurement. We say the measure M achieves an interval level of measurement when

$$M(X_i) - M(X_j) = M(X_k) - M(X_l) \text{ iff } V(X_i) - V(X_j) = V(X_k) - V(X_l)$$

Now consider 3 objects, A, B, and C. Suppose that, on the attribute value, they are ordered and evenly spaced. (For example, $V(A) = 10$, $V(B) = 5$, $V(C) = 0$). Then $V(A) - V(B) = V(B) - V(C)$. Then the difference in the attribute between A and C is twice as great as the difference between A and B, since

$$\begin{aligned} 2[V(A) - V(B)] &= [V(A) - V(B)] + [V(A) - V(B)] \\ &= [V(A) - V(B)] + [V(B) - V(C)] \\ &= V(A) - V(C) \end{aligned}$$

And, from the definition of interval measurement, it immediately follows that $M(A) > M(B) > M(C)$, and $M(A) - M(B) = M(B) - M(C)$ This in turn implies that $M(A) - M(C) = 2(M(A) - M(B))$.We can extend this argument indefinitely, and so when numbers achieve an interval level of measurement, they capture the same-difference and ordering properties in a set of numbers, *and the relative size of the differences between values mirrors relative differences in the attribute being measured*.

Examine the numbers produced by Stopwatch Y. Clearly, they satisfy the properties of nominal and ordinal measurement, but they go further. Note that the difference in times for Athletes B and C is twice as great as the difference in times for Athletes A and B. Indeed, the relative differences for all pairs of values produced by Stopwatch Y match the relative differences for the true times. Consequently, Stopwatch Y has achieved the interval level of measurement.

1.1.4 Ratio Measurement

Stopwatch Y has achieved an interval level of measurement, but it has failed to reflect some important qualities of the true times. For one thing, the ratios of the numbers produced for any two athletes by Stopwatch Y are not the same as the correct ratios of true times. For example, the true times show that Athlete D took twice as long as Athlete A, but the times recorded by Stopwatch Y (41 and 21) are not in the ratio of 2 to 1. Indeed, all of the ratios produced by Stopwatch Y are incorrect, in that they do not match the ratios of true times.

Moreover, the zero point for Stopwatch Y also appears to be incorrect. Athlete S took no time at all, yet Stopwatch Y assigns the athlete a nonzero time. Numbers which reflect all the properties of interval measurement, but also have correct ratio information and a correct zero point are said to have achieved a *ratio level of measurement*. Stopwatch Y has not reached this level.

Definition 1.4 (Ratio Measurement) Suppose objects X_j have attribute values $V(X_j)$ and measured values $M(X_j)$ that satisfy the qualities of interval measurement. We say the measure M achieves a ratio level of measurement if

$$M(X_j) = 0 \text{ iff } V(X_j) = 0$$

and, as a consequence, for any objects X_i, X_j.

$$M(X_i)/M(X_j) = V(X_i)/V(X_j)$$

Clearly, stopwatch Z achieves a ratio scale of measurement.

1.1.5 Absolute Measurement

In some cases, one ratio scale is clearly more appropriate than any other scale, for overriding reasons of convenience. In that case, we say that numbers that use the one "best" scale have achieved an absolute level of measurement. For example, suppose we were recording data for the number of children given birth to by a group of women in Nashville. If 3 women had 0,1,2 children respectively, we could record 0,1,2 *or* 0,2,4 and still be at the interval level of measurement. (Do you see why?) But the former set of numbers is simply better, because the quantity being measured has a natural correspondence with the integers 0,1,2.

1.1.6 The Hierarchy of Measurement Scales

By now it is probably already obvious to you that measurement scales have properties that form a hierarchy. That is, the scales are ordered, and each scale possesses all the properties of the scales below it, plus additional virtues not present in the scales below it.

In the following section, we pursue the notion of permissible transforms, the ways that numbers at a particular measurement scale can be changed without destroying the information in them. We find that permissible transforms also form a hierarchy.

1.2 Permissible Transforms for Measurement Scales

At this point, students comparing the "real times" with those in the last column of Table 1.2 sense that there is something vaguely arbitrary about the numbers in the table. Would the discussion in Section 1.1 have differed substantively if the numbers in the "Real Times" and "Stopwatch Z" columns had been exchanged? In virtually every important respect, the answer is simply "No!" The entire discussion would have proceeded the same way. Both sets of numbers, then, are clearly at the ratio level of measurement.

It seems that either column of numbers could be considered the "True Times," and either could be considered ratio measures of the true times. Does that mean there is something arbitrary about these numbers? Indeed there is something arbitrary, just as the choice to measure height in inches or centimeters is, in an important sense, arbitrary. Many students, when encountering this example for the first time, prefer the times given

Table 1.3: Permissible Transforms for Measurement Scales

Scale	**Permissible Transform**
Absolute	$Y = X$
Ratio	$Y = aX$ for $a > 0$
Interval	$Y = aX + b$ for $a > 0$
Ordinal	$Y = f(x)$ with $f()$ monotonic strictly increasing

in the table as the "True times" over the numbers for Stopwatch Z. They seem more "natural."That is because, through familiarity, specific scales of measurement acquire greater meaning for us. The second is a specific unit of time that we have been using all our lives. The original "true times" in our example are in a unit approximately equal to seconds, whereas the times recorded by Stopwatch Z are closer to half a second. The former seem more natural. But if you think carefully about it, you will realize that familiarity is the only advantage for this particular scale. If, many centuries ago, originators of the watch had chosen a unit of time for the "second" half as long as the one we use now, we might prefer the times given by Stopwatch Z.

The fact that one ratio scale contains all the meaningful information in any other ratio scale means that preferences for one ratio scale over another are largely a matter of habit and personal taste. In many important respects, the numbers we assign to quantities are arbitrary. This raises an important question. "How can we transform a set of numbers that are at a particular level of measurement and be certain we are not destroying useful information?"

The answer is that each scale has its own class of "permissible transforms," or equations by which the numbers can be tranformed without lowering the scale of measurement. The permissible transforms for scales of measurement are shown in Table 1.3.

In the table, X stands for the original numbers, Y for the transformed values. The permissible transform for a given scale is a numerical operation that we are allowed to perform on the numbers without reducing the level of measurement. The scales form a hierarchy, and the permissible transforms are also a hierarchy, in that, as we move up the table, each permissible transform is a more restricted special case of the one below it. In general, if you perform a transformation on a list of numbers that is only permissible for a lower level of measurement, you will drop the level of measurement to this lower level. So, for example, if you have ratio data and you

We see that there is no permissible transform (other than the identity function) for numbers at an absolute scale of measurement. Numbers at an absolute level of measurement already have the only values they are permitted to have.

Numbers at a ratio scale will not drop below that level if you multiply them by any positive constant. (The constant must be positive to avoid reversing the order of the scale values. For numbers at an interval level of measurement, the permissible transform is the equation for any straight line with a positive slope a and y-intercept b. We will refer to such a transformation as a *positive linear transformation*. For efficiency, often we simply will refer to such a transformation a *linear transformation*. Unless explicitly stated otherwise, we assume throughout this text that linear transformations are order preserving, i.e., have a positive slope. Numbers at an ordinal level of measurement will retain their ordinal properties if they are transformed with any function that is order-preserving. A function is order-preserving if it is monotonic and strictly increasing. If you plot the function in a plane, the function line keeps going up from left to right.

Several comments are in order about the permissible transforms. Notice that they form a hierarchy. As you increase the level of measurement, the permissible transforms become increasingly specialized and hence increasingly restrictive. The permissible transform at each level is a special case of the transforms at lower levels. For example:

- The identity transform $Y = X$ is a special case of the multiplicative transform $Y = aX$, with $a = 1$.
- The multiplicative transform $Y = aX$ is a special case of the linear transform $Y = aX + b$, with $b = 0$, etc.

Example 1.2 (Reducing the Level of Measurement) Suppose you have numbers that are at a ratio scale of measurement, but you transform them with the function $Y = 2X + 1$. The numbers will drop to an interval level of measurement. If you re-examine the numbers in Table 1.2on page 3, you may verify that the interval scale of measurement for Stopwatch Y was generated with this transform from the "true times."

1.3 Levels of Measurement in Behavioral Data

One reason that measurement theory is difficult to apply in practice is that many data from behavioral sciences have an uncertain, or debatable level of measurement. If we are measuring maze-running times with an accurate clock, or norepinephrine levels with an accurate chemical assay technique, we can assert with confidence that our data are ratio measurement. If we measure performance by the rank a person obtained in a class, we can say that the data are ordinal.

Unfortunately, many data seem to hover somewhere between ordinal and interval measurement. For example, consider course grades of 80, 79, and 44 given to three individuals A, B, and C. If grades were merely ordinal, absolutely no significance could legitimately be attached to the fact that A and B are 35 times closer together than B and C.

However, most course grades are such that we would feel confident stating there is a greater difference between B and C than there is between A and B, so obviously course grades are better than ordinal. On the other hand, I would not feel confident saying that the difference in knowledge between B and C is exactly 35 times as great as the difference in knowledge between A and B. Taken as an indicator of knowledge, course grades appear to be "semi-interval," that is better than ordinal but not quite interval. How do you deal with such data?

In practice, people have taken two approaches.

- One group simply legislates the problem out of existence by declaring, in effect, that course grades are at an absolute scale of measurement. That is, your performance in my course is defined as the grade I give you!
- The more liberal analysts admit there is a problem, but conclude that there is less damage done by acting as though course grades are at an interval level of measurement than by acting as though they are ordinal. Consequently, data analytic procedures appropriate for interval data are applied.

1.3.1 Robustness in Statistical Analysis

The relationship between measurement theory and course grades is in an important sense typical of what happens all the time in statistics.

- Theory and related procedures are developed on the basis of idealistic assumptions.
- These assumptions are often not quite true.
- The question then becomes, "How much damage is done by assuming an assumption is true when it is actually false."

This is our first introduction to the notion of *robustness*. Robustness analysis is difficult to avoid in statistics, because more often than not the idealistic assumptions necessary to make statistical derivations tractable are not quite true even in the best of circumstances.

1.4 Levels of Measurement in Statistical Practice

Despite the problems posed by behavioral data, measurement theory has serious implications for statistical practice. Here are two.

- The discussion of permissible transforms highlights the fact that what you can do to a set of numbers without destroying information they contain depends on what information is there to be destroyed. You are free to do more to ordinal data than you are to ratio data.
- Certain statistical operations depend for their meaning on the scale of measurement of the data they are applied to. We must be particularly sensitive to whether a transform that is permissible for a set of data alters the interpretation of a statistic calculated on those data.

Table 1.4: Weights in Pounds for 3 Groups

	I	II	III
	150	160	170
	160	140	180
	170	180	190
Mean	160	160	180

Table 1.5: Mixed Measurement Scales (Weights in Boldface are in Kilograms)

	I	II	III
	150	160	170
	352	**308**	**396**
	170	180	190
Mean	224	216	252

Proposition 1.1 (Averaging Numbers from Different Scales) Unless all numbers averaged are on the same measurement scale, the sample mean need not even be on a nominal scale of measurement.

Proof. To prove the assertion, we simply need a counterexample to the assertion that the numbers *are* at a nominal level. Suppose you have three groups of three people each, and their weights in pounds, measured on an accurate scale, are as shown in Table 1.4. Group I and Group II both have the same average weight, 160 pounds, while Group III individuals average 180 pounds. These averages are on a ratio scale of measurement, the same as the weights of the individuals.

Now suppose the second individual in each group had chosen to report their weights in kilograms rather than pounds, and you had chosen to average the numbers as reported.

As you can see, if the numbers averaged are not all on the same scale of measurement, the resulting averages need not be even at a nominal level of measurement! Groups I and II have the same average weight, but the averages of the reported values need not be the same if the values are not all reported on the same scale. This completes the proof. ■

Clearly, averaging pounds and kilograms is like the proverbial "comparing apples and oranges." This point is easy for most people to grasp in the context of the above example. However, people frequently perform the equivalent of averaging pounds and kilograms in more complex contexts, as we see in the following example.

Example 1.3 (Measurement Theory and Grading Policy) Consider the following idealized version of something that happens quite frequently at universities. A course is "team taught" by two professors. One professor teaches the first half and assigns a grade

Table 1.6: Grading Practices for Two Hypothetical Professors

	Professor	
Performance	Nizegy	Meenie
Good	90	80
Average	85	65
Bad	80	50

Table 1.7: Performance of Three Students

Student	**Part 1**	**Part 2**
A	Good	Bad
B	Average	Average
C	Bad	Good

for that half of the course. Another professor takes over the second half, and assigns grades for that portion of the course. Grades from the first and second halves of the course are then averaged to provide a final grade for students.

There are several technical problems with this procedure, and we will consider only one very interesting issue here. The problem, in a nutshell, is that professors often have different grading standards, different exam construction philosophies, different teaching styles, that lead, in effect, to their grades being comparable to two different scales of measurement for the same quantity.

Suppose, for simplicity, that there are three levels of performance in the course, and that they are equally spaced along the continuum of "true performance." The three levels are "Bad," "Average," and "Good." Any set of 3 numbers that is evenly spaced and in the correct order will satisfy the definition of an interval scale of measurement for these performance levels. In Table 1.6, we consider 2 such interval scales, which are the scales of grading used by two different hypothetical professors.

We can see that Professor Nizegy gives easy grades, and does not distinguish very much between good and bad performance. On the other hand, Professor Meenie gives low grades, and the grades differ sharply between good and bad performance.

Suppose three students take a year long course that is "team taught" by Professor Nizegy and Professor Meenie. Nizegy teaches the first half, and Meenie teaches the second half. Imagine the situation shown in Table 1.7.

Over the course of the year, all three students were at an average level. For two students, bad performance in one half was balanced by good performance in the other, while one student performed at an average level all year. All three students should have gotten the same grade in the course. Now let's see how they actually fared under the team-teaching system. The grades obtained by the three students are shown in Table 1.8.

Table 1.8: Grades Obtained by the Three Students

Student	**Part 1**	**Part 2**	**Course**
A	90	50	70
B	85	65	75
C	80	80	80

Notice that the final grades obtained by the three students are not equal, even though their performance was. Indeed, student C got a grade a full 10 points higher than student A!

The problem is, in a theoretical sense, virtually identical to what happens when people average pounds and kilograms. Each professor grades on an interval scale of measurement, but the two professors have different interval scales. Consequently, for them to average their grades is the same as averaging pounds and kilograms, which are two different ratio scales of weight.

A natural question to ask is "Can anything be done about this problem?" In essence, we might do the same thing with course grades that we would do with body weights — change them all to the same scale of measurement before averaging. With body weights, we could convert all weights to pounds or all weights to kilograms before averaging. With course grades, we can do very much the same thing. It is called "scaling," and we will talk about its technical aspects in great detail in a subsequent module.

Note that this problem generalizes to many other situations. If averaging grades on different scales creates a problem within a single team-taught course, it also creates a problem when universities compute grade point averages across several courses. Isn't the validity of these gradepoint averages also compromised? In a sense, they are. Strictly speaking, grade point averages probably do not even achieve an ordinal level of measurement. That is, some students who performed better than others will receive a lower grade point average, because they took a disproportionate number of courses from professors like Dr. Meenie. The question again revolves around robustness. We know that the assumptions necessary to validly average numbers are violated by the grades on students' transcripts. Most universities ignore this, and average them anyway. That is why the intelligent consumer of such information does not place too much emphasis on minor differences in grade point average. Unfortunately, consumers of numerical information are not always sophisticated, and the "fallacy of misplaced precision" is commonplace in the analysis of data. Often people will apply analyses of enormous sophistication and precision to numbers that simply cannot support such precision.

The preceding critique also implies that averaging grades across different exams within a single course can also create problems. It is incumbent upon teachers to use rescaling technology to reduce such problems. However, we should realize they can never be eliminated entirely, and that minor differences in grade point average are undoubtedly meaningless.

1.5 Precision and Accuracy in Numerical Data

In this section we discuss two fundamentally different types of data, which we call *discrete* (or *categorical)* and *continuous*. There are some important and surprisingly subtle technical issues related to the difference between categorical and continuous data, and we will explore a few of them in detail.

1.5.1 Categorical Data

Some data naturally fall into discrete, discontinuous categories. Often, the number of categories is small as well. An example is outcome of a soccer match, which for each team is either a win, loss, or tie. Such data can typically be recorded easily, with no error. Moreover, numerical calculations involving such data can also be error-free. For example, we can readily calculate, without any error, the average number of wins for a baseball team over the last ten years.

1.5.2 Continuous Data and Round-Off Error

Some quantities that we measure are in principle *continuous*, i.e., can take on any value within certain limits. Body weight is such a quantity. Amount of time needed to run a mile is another. In some cases, quantities that are continuous in principle are reported as categorical data. For example, we are all familiar with questionnaires that ask us to rate our feelings about some topic on a “5 point” scale. For example, you might be asked “How confident are you that nuclear war will not take place again before 2010?” You might then be presented with 5 alternatives, “Very Confident,” “Moderately Confident,” “Uncertain,” “Moderately Unconfident,” “Very Unconfident.” Typically, the instructions on the questionnaire will ask you to choose the alternative that “best” matches your feelings. None of these descriptors may fit your actual feelings, in which case you might engage in some kind of internal “round-off” process in order to select the best alternative.

In this case, you have reduced a continuous quantity to a categorical one. When continuous real number values are reported, they generally are only approximated. As well, most computer data analytic systems maintain only an approximation of a real number. Usually, the approximation is far better than we actually need for practical purposes. When a real number is reported, we can talk about two aspects of the reported value, *precision* and *accuracy*. Informally, precision refers to how precisely a value is stated, while accuracy refers to how close a number is to the *correct* value. A number can be very precise, yet very inaccurate. For example, if a person actually weighs 200.3 pounds, and his weight is listed as 156.19989544 pounds, the value would be stated very *precisely*, but also very *inaccurately*.

Definition 1.5 (Precision) The precision of an approximate real number X is the total number of significant decimal digits in X.

Table 1.9: Rounding Numbers to a Specified Level of Precision

Original Number	**Level of Precision**	**Rounded Value**
2.798765	2	2.80
29.4972	3	29.497
16.25	1	16.3
1242.1184	2	1242.12

Definition 1.6 (Accuracy) The accuracy of an approximate real number X is the absolute difference between X and the quantity it is measuring.

In one sense, continuous data are always reported as categorical data, since it is impossible to record many numbers to their full level of accuracy with the decimal system. In this text, we will always round off the last digit up if the next digit (to its right) is greater than or equal to 5, and down if the next digit is less than 5. Table 1.9 gives examples of numbers rounded off to a specified level of precision.

1.5.3 Nominal and Real Limits

When a set of numerical quantities is reported, we can speak of the *Nominal Limits* and *Real Limits* for the set.

Definition 1.7 (Nominal Limits) When a group of numbers is reported, the *nominal limits* for the group are the lowest and highest reported values.

Definition 1.8 (Real Limits) When a group of numbers is reported, the *lower and upper real limits* for the group are, respectvely, the lowest value that the lower nominal limit could have been rounded up from, and the highest value that the upper nominal limit could have been rounded down from.

Example 1.4 (Nominal and Real Limits) Suppose three football players weigh 198, 212, and 235 pounds on a scale that accurately rounds weight to the nearest whole pound. What are the nominal and real limits for this group of 3 numbers? The nominal limits are 198 and 235. However, the lower nominal limit, 198, could in principle represent a value as low as 197.5 pounds, so the lower real limit is 197.5. The upper nominal limit of 235 could have been rounded down from any value below 235.5, so the (non-inclusive) upper real limit is 235.5.

Many psychology students encounter the notion of nominal and real limits in connection with some elementary calculations on frequency distributions, particularly in connection with "estimation of percentiles from grouped frequency data," a topic we shall discuss later. At this point, I want to take a rather lengthy digression to illustrate some points which are truly fundamental to the study of statistics. The material is at a very low level mathematically, yet illustrates the surprising subtleties one encounters even in the most mundane areas of inquiry.

The preceding two definitions seem simple and straightforward. Yet, with virtually any idea one encounters in statistics, it is possible to blunder seriously (and obviously) if one tries to apply the idea mechanically. The danger is especially great if one is given the mechanical rule without being told the idea that produced it. For example, consider the following definition.

Definition 1.9 (Inclusive Range) For a set of numbers X_i, the *inclusive range* is the difference between the upper and lower real limits, i.e., the largest possible difference between the highest and lowest numbers underlying the reported values.

This is an extremely simple notion, yet it is considered too complex by some textbook writers, who instead report the following "simplified" definition.

Definition 1.10 (Inclusive Range B) For a set of numbers X_i, the *inclusive range* is

$$X_{\max} - X_{\min} + 1$$

where $X_{\max}$ is the highest nominal value, and $X_{\min}$ is the lowest nominal value.

For example, Glass and Hopkins, (1996, p. 16, footnote 3), report such a definition.

> Some statisticians define the range as $X_{\max} - X_{\min} + 1$, so that it extends from the upper real limit of $X_{\max}$ (i.e., $X_{\max} + .5$), to the lower real limit of $X_{\min}$ (i.e., $X_{\min} - .5$). (Glass and Hopkins, 1996, p. 16)

This "simplified" definition is easy to use. Indeed, it is the only definition for the *inclusive range* given in most textbooks. Yet, it is often inconsistent with the original definition, and can lead to gross errors, as we see in the next example.

Example 1.5 (Weighing Trucks) Suppose 5 trucks stop by a roadside weigh station, where the scales report a weight rounded to the nearest 100 pounds. The trucks' weights are reported as 14200, 16400, 19100, 22400, and 22700. What is the inclusive range for these data?

The answer would seem to be straightforward. The lightest truck is reported to be 14200 pounds. It could really weigh as little as 14150 pounds, however, because the scale rounds to the nearest 100 pounds.. The heaviest truck is reported as 22700 pounds, but because the weights are rounded to the nearest 100 pounds, its actual weight could be any value between 22650 and 22750 pounds. So the upper real limit for the highest values is 22750. Consequently, the inclusive range is $22750 - 14150 = 8600$. Note, however, that the "simplified" definition (Inclusive Range B) gives the result $22700 - 14200 + 1 = 8501$. The discrepancy results from the fact that the "simplified" formula implicitly assumes that the level of round-off is always to the nearest whole unit. If the scales weighed trucks to the nearest pound, the formula would give the correct answer. In many cases, this assumption is clearly inappropriate, and consequently so is the "simplified" formula.

One might consider replacing the definition with an improved one. Instead of $X_{\max} - X_{\min} + 1$ use the new definition $X_{\max} - X_{\min} + L$ where L is the "level of rounding off." So, for example, if trucks are weighed to the nearest hundred pounds, $L = 100$ and we obtain the correct answer. Indeed, this "improved simplified formula" makes far more sense than the one it replaces. Moreover, the inclusion of the variable L instead of the constant 1 starts to hint at the conceptual foundation of the formula.

One might ask how could a formula that is so clearly inadequate be printed in so many places. It seems so obvious that an equally simple and obviously superior formula is available! It is obvious now, because you understand the rationale on which the original, inadequate, "simplified" formula was based. Many people who were taught this formula were never told the rationale on which it was based. The very straightforward lesson here is that, if you do not understand why you are doing something, it is easy to apply a formula in a situation where it is simply incorrect. The idea behind "inclusive" range is that it calculates a range of values which, if there is no reporting error, is wide enough to cover all real values. That simple enough idea has been replaced in the minds of many students (and teachers) of statistics with a formula, $X_{\max} - X_{\min} + 1$. To many math phobics, formulas are more impressive than the ideas underlying them. They are more "mathematical," more concrete, more easily memorized. Ironically, their use can be anxiety reducing.

But our lesson is not over! Actually, the new and improved "simplified" formula $X_{\max} - X_{\min} + L$ does not always work either, as we see in the next example.

Example 1.6 (Percentage Grades) Suppose I give a multiple-choice exam with a large number of questions, and I round the grades to the nearest whole percent. The grades for 5 students are 86,72,100,59,68. What is the inclusive range for these data?

Either of the two "simplified" formulas gives $100 - 59 + 1 = 42$. However, this result is not correct. The actual answer is 41.5. Do you see why? Ponder this a minute before reading the next paragraph.

The reason why the answer is 41.5 is that the maximum nominal value of 100 percent has an upper real limit of 100, not 100.5. It is not possible to have more than 100 percent right on an exam, so a nominal value of 100 could stand for a real value of at most 100. On the other hand, the lowest nominal value of 59 could have been rounded up from

Table 1.10: Number of Votes Received by the Winning Candidates — 1990 Connecticut Congressional Elections

Name	**Number of Votes**
Kennelly	$126,343$
Gedjenson	$103,945$
DeLauro	$89,742$
Shays	$106,396$
Franks	$93,041$
Johnson	$138,758$

a value as low as 58.5, so the inclusive range is $100 - 58.5 = 41.5$. Here is another example with a similar message.

Example 1.7 (Inclusive Range with Discrete Data) In the 1990 U.S. Congressional elections, there were 6 representatives elected in the state of Connecticut. Their names, and the number of votes they obtained, are given in Table 1.10. What is the inclusive range for the number of votes in these data? The answer is $138,758-89,742 = 49,016$. Since these data are not continuous, the real limits are identical to the nominal limits. One cannot obtain half a vote. Assuming no error in the vote count, the range of values can be computed exactly here. Consequently, only the second "simplified formula" is appropriate. In this case, $L = 0$ since there is absolutely no round-off error.

Even in a comparatively simple situation, formulas are not enough. To produce intelligent answers across a reasonable variety of circumstances, you have to understand the ideas that generated the formulas. We have illustrated this point here in a context where, admittedly, the consequences of error are relatively trivial. However the lesson holds across the whole range of statistical practice, often in cases where negative consequences of a "cookbook" approach can be far from trivial. As you pursue the study of statistics, try to approach the material with a critical mindset. In particular, always try to master the ideas that underly the formulas. In some cases we will cover later, you will discover that, if you know the underlying ideas, you actually don't need the formulas.

1.6 Exercises

1. Suppose you are measuring weights of a group of individuals on a scale that measures in pounds. Unfortunately, following the installation of a new rubber mat on the scale, it was not "zeroed" properly, and so each person's weight has a pound added on to it. Assuming that the scale is otherwise in working condition, what level of measurement will the recorded weights achieve?

2. Suppose every watch in North America suddenly started running exactly twice as fast as before. What level of measurement would time be be measured at by these watches?

3. Suppose you are planning to study amount of weekly alcohol consumption by teenagers as a function of socioeconomic status of their parents. What level of measurement can you expect to achieve in measuring your independent and dependent variables?

4. Suppose you have a list of numbers that are at the ratio level of measurement, and you cube each number (i.e., raise it to the third power). What level of measurement will your numbers now achieve?

5. Suppose you have a list of numbers that are at the interval level of measurement, and you square each number. What level of measurement will your numbers now achieve?

6. Suppose you have numbers that are at an interval level of measurement, and you modify them by subtracting 3 from all of the numbers. What level of measurement will they now achieve?