1.6: Levels of Measurement

Learning Objectives

- Define and distinguish among nominal, ordinal, interval, and ratio levels of measurement
- Give examples of errors that can be made by failing to understand the proper use of measurement levels

Types of Measurement

When we have a particular topic of interest and want to further our understanding, we need to collect appropriate data through some sort of measurement or observation. Exactly how the measurement is carried out depends on the variable of interest. To measure the time taken to respond to a stimulus, we might use a stop watch. Stop watches are of no use, of course, when it comes to measuring someone's attitude towards a political candidate. A rating scale is more appropriate in this case (with labels like "very favorable," "somewhat favorable," etc.). For a variable such as "favorite color," we can simply note the color-word (like "red") that the subject offers. Although measurements can differ in many ways, they can be classified using a few fundamental categories. These categories are called levels of measurement because each one is contained in the previous one. Just like we have categories of animal, mammal, canine, and dog, each referring to a differing level of specificity, so too do we have categories of nominal, ordinal, interval, and ratio. We'll start with the broadest category, nominal, and work our way down to the most specific category, ratio.

Nominal Scale

When we simply name or categorize responses, we are measuring on a nominal scale. Gender, handedness, favorite color, and religion are examples of variables measured on a nominal scale. The essential point about nominal scales is that they categorize without implying any ordering among the responses. For example, when classifying people according to their favorite color, there is no measurement sense in which green is placed "ahead of" blue. Responses are merely categorized. Nominally scaled data embody the lowest level of measurement.

Ordinal Scale

A researcher wishing to measure consumers' satisfaction with their microwave ovens might ask them to specify their feelings as either "very dissatisfied," "somewhat dissatisfied," "somewhat satisfied," or "very satisfied." The items in this scale are ordered, ranging from least to most satisfied. This is what distinguishes ordinal from nominal scales. Unlike nominal scales, ordinal scales allow comparisons because there is a meaningful order to the measurement values. For example, our satisfaction ordering makes it meaningful to assert that one person is more satisfied than another with their microwave ovens.

On the other hand, ordinal scales fail to capture important information that will be present in the subsequent levels we examine. In particular, the difference between two levels of an ordinal scale cannot be assumed to be the same as the difference between two other levels. In our satisfaction scale, for example, the difference between the responses "very dissatisfied" and "somewhat dissatisfied" is probably not equivalent to the difference between "somewhat dissatisfied" and "somewhat satisfied." Nothing in our measurement procedure allows us to determine whether the two differences reflect the same difference in psychological satisfaction. Indeed, even if we had two pairs of observations each with "very dissatisfied" and "somewhat dissatisfied" ratings, we could not determine whether the differences in these ratings are truly the same.

What if the researcher had measured satisfaction by asking consumers to indicate their level of satisfaction by choosing a number from one to four? Would the difference between the responses of one and two necessarily reflect the same difference in satisfaction as the difference between the responses two and three? The answer is No. Changing the response format to numbers does not change the meaning of the scale. We still are in no position to assert that the qualitative difference between 1 and 2 (for example) is the same as 3 and 4.

Text Exercise 1.6.1

Classify the following variables based on their nominal or ordinal level of measurement. Explain.

1. Eye color

Answer

Since there is no inherent, meaningful order to eye color (as normally denoted green, blue, hazel, and brown), eye color is nominal.

2. BMI weight type (underweight, healthy, overweight, obese, severely obese)

Answer

There is an inherent, meaningful order to BMI weight types, each subsequent value (as listed) indicates an increasing BMI. Hence BMI weight type is ordinal.

3. Shirt size (S, M, L, XL)

Answer

There is an inherent, meaningful order to shirt sizes, each subsequent value indicates a larger or smaller shirt (depending on the ordering). Hence shirt size is ordinal.

4. Phone number

Answer

Phone numbers are granted based on availability and hence have no meaningful order. Thus, phone number is nominal.

Interval Scale

Notice that qualitative data is nominal or ordinal. If we have quantitative data; that is, data with meaningful differences in values, then we have some important distinctions to make. Interval scales are numerical scales in which intervals have the same interpretation throughout. As an example, consider the Fahrenheit scale of temperature. The difference between 30 degrees and 40 degrees represents the same temperature difference as the difference between 80 degrees and 90 degrees. This is because each 10-degree interval has the same physical meaning (in terms of the kinetic energy of molecules).

To differentiate interval scale from our next level of measurement, we note that on interval scales a measurement of 0 does not represent the absence of some quantity. On interval data, 0 is artificially or arbitrarily defined and is not intrinsically meaningful to the quantity being measured. The Fahrenheit scale illustrates the issue. Zero degrees Fahrenheit does not represent the complete absence of temperature (the absence of any molecular kinetic energy). In reality, the label 0 is applied to its temperature for quite accidental reasons connected to the history of temperature measurement. This is important because if 0 doesn't mean "nothing," then it is not meaningful to divide nor multiply. For example, there is no sense in which the ratio of 40 to 20 degrees Fahrenheit is the same as the ratio of 100 to 50 degrees Fahrenheit; no interesting physical property is preserved across the two ratios. For this reason, it does not make sense to say that 80 degrees Fahrenheit is "twice as hot" as 40 degrees Fahrenheit.

One way that we mark the passage of time is by using the designation of year. Citizens of the United States remember the year 1776 as the year the Declaration of Independence was signed. Jews remember the year 70 as the year the temple was destroyed. The first recorded Olympic games occurred in 776BC. The differences between these years are meaningful; the Declaration of Independence occurred 1706 years after the Romans razed the temple and 2552 years after the first recorded Olympic games. So the year designation of time, must be at least of interval level. While there is a year 0, it does not represent the absence of passage of time; therefore, the year is measured on the interval level.

Ratio Scale

As the highest level of measurement, ratio scales allow the most varied statistical analyses. The ratio scale of measurement is an interval scale with the additional property that its zero position indicates the absence of the quantity being measured. Like a nominal scale, the ratio scale provides a name or category for each object (the numbers serve as labels). Like an ordinal scale, the objects are ordered (in terms of the ordering of the numbers). Like an interval scale, the same difference between values on the scale has the same meaning; however, with the ratio scale, the same ratio between values on the scale also carries the same meaning. This is the ratio scale's defining characteristic.

The Fahrenheit scale for temperature has an arbitrary zero point and is, therefore, not a ratio scale. However, 0 on the Kelvin scale is absolute zero (total absence of kinetic energy), making this a ratio scale. For example, if one temperature is twice as high as another as measured on the Kelvin scale, then that temperature has twice the kinetic energy of the other temperature. Therefore, it does make sense to say 80 K is twice as hot as 40 K.

Another example of a ratio scale is the amount of money you have in your pocket right now (25 cents, 55 cents, etc.). Money is measured on a ratio scale because, in addition to having the properties of an interval scale, it has a true zero point: if you have zero money, this implies the absence of money. Since money has a true zero point, we can say that someone with 50 cents has twice as much money as someone with 25 cents.

Ratio scales are very common. Most quantities of scientific interest tend to be ratio: distance, speed, weight, mass, pressure, volume, area, energy, population; these are all variables measured on ratio scales.

Text Exercise 1.6.2

Define a variable on the ratio scale that can take on both negative and positive numbers.

Answer

Note answers may vary. Since ratio scales have "zeros" that carry meaning relative to what we are actually measuring, we might think that ratio scales cannot change signs. This conclusion might be intuitive, but unfortunately the conclusion is false. Vector measurements, measurements that include a direction, can have both positive and negative values while remaining on the ratio scale. Consider displacement, velocity, and acceleration to name a few. Electrical charge is an example of a non-vector, ratio scale which can be negative or positive. Money in an account is another example, as negative values represent debt, but 0 still means no money in the account.

Consequences of Levels of Measurement

Why are we so interested in the type of scale that measures a variable? The crux of the matter is the relationship between the variable's level of measurement and the statistics that can be meaningfully computed with that variable. For example, consider a hypothetical study in which 5 children are asked to choose their favorite color from blue, red, yellow, green, and purple. The researcher codes the results as follows:

Table 1.6.1: Guide for Encoding Colors as Numbers

Color	Code
Blue	1
Red	2
Yellow	3
Green	4
Purple	5

This means that if a child said her favorite color was "Red," then the choice was coded as 2, if the child said her favorite color was "Purple," then the response was coded as 5, and so forth. Consider the following hypothetical data:

Table 1.6.2: Favorite Colors and Code from Sample of 5 Children

Subject	Color	Code
1	Blue	1
2	Blue	1
3	Green	4
4	Green	4
5	Purple	5

Each code is a number, so nothing prevents us from computing the average code assigned to the children. The average happens to be 3, but you can see that it would be senseless to conclude that the average favorite color is yellow (the color with a code of 3). Such nonsense arises because favorite color is a nominal scale, and taking the average of its numerical labels is like counting the number of letters in the name of a snake to see how long the beast is.

In a similar fashion, does it make sense to compute the average of numbers measured on an ordinal scale? Different fields of study might answer this question differently; after all, reviews online often give an average rating from 1 to 5 stars and surveys rating 0 (strongly disagree) to 10 (strongly agree) are often summarized with an average rating. Suppose an individual ran 4 races and placed 1^{st}, 2^{nd}, 3^{rd}, and 4^{th}. We could say the average placement was 2.5^{th}, but what does that mean? Suppose another individual placed 3^{rd}, 3^{rd}, 1^{st}, and 3^{rd} in the same races to also have an average placement of 2.5^{th}. Is it meaningful to say these two runners are, on average, equally fast? We cannot conclude that. Perhaps the second runner lost the first two races by a large margin, but the second two races were neck and neck. The problem is that we only know who was faster, we don't know by how much. We, therefore, recommend avoiding such a practice and considering results based on averaged ordinal data with a careful eye.

To conclude such a discussion, once we attain the levels of interval and ratio, computing the average of our data is well-defined and meaningful. As we move forward, always be aware of what descriptive statistics are possible given the consequences of the level of measurement at hand.
