Organizing Data with Frequency Tables

Inquire: Easy Interpretation Without Calculation

Overview

Much of the work done in organizing data requires formulas and calculations; however, visual models can simplify that. Frequency tables are an efficient way of both presenting data and providing analysis and interpretation quickly and accurately. When tables include relative frequencies and cumulative relative frequencies, relationships and percentages are apparent and require little actual calculation. This lesson will provide information that will allow you not only to construct frequency tables, but also to confidently interpret the tables you meet along the way. By becoming familiar with relative frequencies, cumulative relative frequencies, and levels of measurement, you will find easy interpretation essentially without calculation.

Big Question: In the study of descriptive statistics, why is it valuable to understand relative frequencies, cumulative relative frequencies, and levels of measurement?

Watch: What’s in Your Pocket?

The way a set of data is measured is called its level of measurement. Not every operation can be used with every set of data, so you need to be familiar with them all. If your data includes a list of colors, you might create a frequency table listing the times each color is selected. In this case, give yellow a frequency of 10 and blue a frequency of 7. This hardly means that yellow is “of greater value” than blue. This set of data is nominal and cannot be ordered or ranked.

How about a satisfaction chart from drivers of a particular brand of car? The frequency table shows Car A with a satisfaction rate of 65%, and Car B with 35%. If we order the cars in their rank of customer satisfaction, Car A is ranked higher than Car B. However, the differences in their ranks cannot be measured. We only know that one is higher than the other. If three runners finish a race, we can order them as one, two, and three, but this rank doesn’t tell us about the differences in their times. This set of data is ordinal. It can be ordered, but nothing can be calculated.

A third type, and a higher classification of measurement, is the interval data scale. Suppose we record the ages of 50 people in a mall. We could chart each year of age separately or create intervals. Either way, in this table, each interval will have the same interpretation. However, in an interval scale, there is no zero. There are no people in the mall who are zero years old.

The last and highest level of measurement, the ratio scale, combines qualities of the others and adds something new. Let’s look at how much money you have in your pocket right now. Create a frequency
table and survey as many other people as you want. With this type of data, you can calculate and interpret the values in many ways. You can order the amounts from greatest to least, and the differences will have meaning. If you have $10 and a friend has $40, he has four times as much money as you. The intervals have the same interpretation in every situation, and we have an absolute zero. Hopefully, that’s not what is in your pocket, but it is a possibility. There is no “less than zero,” but there is a zero.

Read: Organizing Data with Frequency Tables

Overview

If you planned well during the collection of your data, you probably have some ideas of how you want to classify and organize your work. The methods of organization you choose should highlight the data to make it easily understood, useful, and meaningful. This lesson will give you more information about levels of measurement and frequencies to help in making these decisions.

Levels of Measurement

The way a set of data is measured is called its **level of measurement**. Data can be classified into four levels of measurement. They are:

- **Nominal**
- **Ordinal**
- **Interval**
- **Ratio**

Data measured using a **nominal scale** is qualitative. Categories, colors, names, labels, and favorite foods are examples of nominal data. Nominal data are not ordered. For example, trying to classify people according to their favorite food does not make sense. Putting pizza first and sushi second is not meaningful.

Smartphone companies are another example of nominal data. The data are the names of the companies, but there is no order of these brands, even though people may have personal preferences. Nominal data cannot be used in calculations.

Data measured using an **ordinal scale** is similar to nominal scale data, but there is one big difference. The ordinal data can be ordered. An example of ordinal data is a list of the top five national parks in the United States. The top five parks can be ranked from one to five, but we cannot measure differences between the data.

Another example of the ordinal scale is a survey in which the responses are “excellent,” “good,” “satisfactory,” and “unsatisfactory.” These responses are ordered from the most desired response to the least desired; however, the differences between two pieces of data cannot be measured. Like the nominal data, ordinal data cannot be used in calculations.

Data that is measured using the **interval scale** is similar to ordinal data because it has a definite ordering, but this time, there is a difference between data. The differences between interval data can be measured even though the data does not have a starting point.

Temperature scales are measured by using the interval scale. In both Celsius and Fahrenheit, 40° is equal to 100° minus 60°. Although there is a zero in each scale, zero is not the absolute lowest temperature. Temperatures like -10° F and -15° C exist and are colder than zero.
Interval data can be used in calculations, but one type of comparison cannot be done. 80° is not four times as hot as 20°, even though the number 80 is four times the number 20. There is no meaning to the ratio of 80 to 20 (or four to one) in this interval scale.

Data measured using the **ratio scale** gives the most information. Ratio scale data is like interval scale data, but it has a zero point, and ratios can be calculated.

For example, four multiple choice exam scores are 80, 68, 20, and 92. The data can be put in order from lowest to highest: 20, 68, 80, 92, and the differences between the data have meaning. The score 92 is more than the score 68 by 24 points. Ratios can be calculated. The smallest score is zero. The score of 80 is four times better than the score of 20.

**Frequency Tables**

In data collection, there is often repetition of particular values. The number of times a particular data value is repeated is known as **frequency**. The table constructed based on the data values with corresponding frequency is a **frequency table**. These tables can be used to describe the number of occurrences of values, identify trends, compare data with other tables, and calculate probability.

20 students were asked how many hours they worked per day. Their responses, in hours, were as follows: 5, 6, 3, 3, 2, 4, 7, 5, 2, 3, 5, 6, 5, 4, 4, 3, 5, 2, 5, 3. The following table lists the different data values in ascending order and their frequencies.

<table>
<thead>
<tr>
<th>Data Value</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
</tr>
</tbody>
</table>

According to the table, there are three students who work two hours, five students who work three hours, and so on. The sum of the values in the frequency column, 20, represents the total number of students included in the sample. The table is far easier to understand and analyze than the list of raw data.

**Relative Frequency**

A **relative frequency** is the ratio of the number of times a value of the data occurs in the set of all outcomes to the total number of outcomes. To find the relative frequencies, divide each frequency by the total number of students in the sample — in this case, 20. Relative frequencies can be written as fractions, percentages, or decimals.
<table>
<thead>
<tr>
<th>Data Value</th>
<th>Frequency</th>
<th>Relative Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>3</td>
<td>3/20 or 0.15</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
<td>5/20 or 0.25</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>3/20 or 0.15</td>
</tr>
<tr>
<td>5</td>
<td>6</td>
<td>6/20 or 0.30</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>2/20 or 0.10</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>1/20 or 0.05</td>
</tr>
</tbody>
</table>

The sum of the values in the relative frequency column is 20/20, or 1.00.

**Cumulative Relative Frequency**

*Cumulative relative frequency* is the accumulation of the previous relative frequencies. To find the cumulative relative frequencies, add all the previous relative frequencies to the relative frequency for the current row.

<table>
<thead>
<tr>
<th>Data Value</th>
<th>Frequency</th>
<th>Relative Frequency</th>
<th>Cumulative Relative Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>3</td>
<td>3/20 or 0.15</td>
<td>0.15</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
<td>5/20 or 0.25</td>
<td>0.15 + 0.25 = 0.40</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>3/20 or 0.15</td>
<td>0.40 + 0.15 = 0.55</td>
</tr>
<tr>
<td>5</td>
<td>6</td>
<td>6/20 or 0.30</td>
<td>0.55 + 0.30 = 0.85</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>2/20 or 0.10</td>
<td>0.85 + 0.10 = 0.95</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>1/20 or 0.05</td>
<td>0.95 + 0.05 = 1.00</td>
</tr>
</tbody>
</table>

The last entry of the cumulative relative frequency column is one, indicating that 100 percent of the data has been accumulated. Because of rounding, the relative frequency column may not always equal one, and the last entry in the cumulative relative frequency column may not be one. However, they each should be close to one.

**Reflect Poll: Levels of Measurement**

In thinking about this lesson, how do you think it can help you to know about levels of measurement?

- To decide how to interpret data
- To know what operations I can do with data
- To make sure I use data correctly
- To score well on my tests
Expand: More Practice with Frequency Tables

Overview

Frequency tables can be useful for describing the number of occurrences of a particular value, and are one of the most basic tools for visually representing data. They provide a quick interpretation of the distribution of data. They are easy to analyze and can display large data sets in a small space. When tables include relative frequency and cumulative relative frequency, even more analysis is possible.

Practice One

The following table shows the amount, in inches, of annual rainfall in a sample of towns.

<table>
<thead>
<tr>
<th>Rainfall (Inches)</th>
<th>Frequency</th>
<th>Relative Frequency</th>
<th>Cumulative Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.95 - 4.97</td>
<td>6</td>
<td>6/50 or 0.12</td>
<td>0.12</td>
</tr>
<tr>
<td>4.97 - 6.99</td>
<td>7</td>
<td>7/50 or 0.14</td>
<td>0.12 + 0.14 = 0.26</td>
</tr>
<tr>
<td>6.99 - 9.01</td>
<td>15</td>
<td>15/50 or 0.30</td>
<td>0.26 + 0.30 = 0.56</td>
</tr>
<tr>
<td>9.01 - 11.03</td>
<td>8</td>
<td>8/50 or 0.16</td>
<td>0.56 + 0.16 = 0.72</td>
</tr>
<tr>
<td>11.03 - 13.05</td>
<td>9</td>
<td>9/50 or 0.18</td>
<td>0.72 + 0.18 = 0.90</td>
</tr>
<tr>
<td>13.05 - 15.07</td>
<td>5</td>
<td>5/50 or 0.10</td>
<td>0.90 + 0.10 = 1.00</td>
</tr>
</tbody>
</table>

In this sample, there are 6 towns whose annual rainfall amounts fall within the interval 2.95 - 4.97 inches, 7 towns whose amounts fall within the interval 4.97 - 6.99 inches, 15 towns whose mounts fall within the interval 6.99 - 9.01 inches, 8 towns whose rainfall is within the interval 9.01 - 11.03 inches, 9 whose rainfall is within the interval 11.03 - 13.05 inches, and 5 towns whose rainfall is within the interval 13.05 - 15.07. All annual rainfall amounts fall between the endpoints of an interval and not at the endpoints.

From this table, how can you find the percentage of rainfall that is less than 9.01 inches?

Solution: The first three entries are for towns with less than 9.01 inches of rainfall annually. We know this because it is given that the rainfall is between the endpoints of an interval and not at the endpoints. The answer can be found in two different ways. First, add the three relative frequencies to arrive at 0.56 or 56%. A second way to find the answer is to look at the cumulative relative frequencies on the third row, 56%. This is an example of how a frequency table can provide information without the necessity of lengthy calculations.

Lesson Toolbox

Additional Resources and Readings

A video clearly explaining the four levels of measurement in the first seven minutes

- Link to resource: https://youtu.be/B0ABvLa__u88
A video explaining the four levels of measurement
  ● Link to resource: https://youtu.be/klgFMJppfcY

A video covering how to make a relative frequency distribution chart
  ● Link to resource: https://youtu.be/7jUIt39tUBM

A video showing how to solve problems that require reading frequency tables
  ● Link to resource: https://www.youtube.com/watch?v=MVTizqEX99g

Lesson Glossary

cumulative relative frequency: the sum of the relative frequencies for all values that are less than or equal to the given value
frequency: the number of times the data value occurs
frequency table: an arrangement of collected data values in ascending order of magnitude with their corresponding frequencies
level of measurement: a classification that describes the nature of information within the values assigned to variables
nominal scale: the lowest of the four levels of measurement; categorical and allows no ordering or computation (Example: names of students)
ordinal scale: the second of the four levels of measurement; attributes can be ordered (Example: satisfaction scales)
interval scale: the third of the four levels of measurement; order is meaningful and intervals have the same interpretation; there is no zero (Example: temperature scales)
ratio scale: the fourth of the four levels of measurement; includes qualities of the other levels and has a fixed zero (Example: amount of money in your pocket)
relative frequency: the ratio of the number of times a value of data occurs in the set of all outcomes to the total number of outcomes

Check Your Knowledge

1. Which of the four levels of measurement can be calculated?
   a. nominal
   b. ordinal
   c. interval
   d. ratio

2. __________ is the ratio of the number of times a value occurs in the outcomes to the total number of outcomes.
   a. Relative frequency
   b. Cumulative frequency
   c. Ordinal scale data
   d. Frequency

3. Which level of measurement would be appropriate for measuring temperature?
   a. nominal
   b. ordinal
   c. interval
   d. ratio
Answer Key:

Citations

Lesson Content:
Authored and curated by Kathryn R. Price, M.Ed, NBCT for The TEL Library. CC BY NC SA 4.0

Adapted Content:
Title: Introductory Business Statistics – 1.3 Levels of Measurement. Openstax. License CC BY 4.0. Link to resource: https://cnx.org/contents/tWu56V64@34.114:B2ZsQUfk@13/Levels-of-Measurement

Title: Introductory Statistics – 1.3 Frequency, Frequency Tables, and Levels of Measurement. Openstax. License CC BY 4.0. Link to resource: https://cnx.org/contents/MBiUQmmY@22.8:P71MkpUV@15/Frequency-Frequency-Tables-and-Levels-of-Measurement