Unit 3, Lesson 2

Relative Frequency Tables

Let's find relative frequencies of categorical data.

2.1 Notice and Wonder: Teacher Degrees

Several adults in a school building were asked about their highest degree completed and whether they were a teacher.

What do you notice? What do you wonder?

	teacher	not a teacher
associate degree	4%	16%
bachelor's degree	52%	64%
master's degree or higher	44%	20%

2.2 City Cat, Country Cat

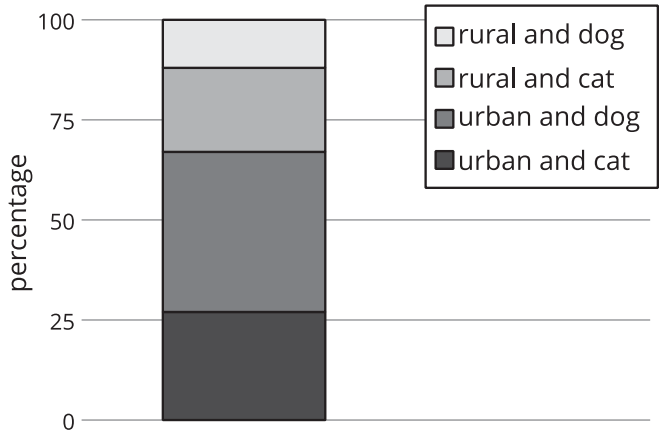
200 people were asked if they prefer dogs or cats and whether they live in a rural or urban setting.

The actual values collected from the survey are in the first table.

	urban	rural	total
cat	54	42	96
dog	80	24	104
total	134	66	200

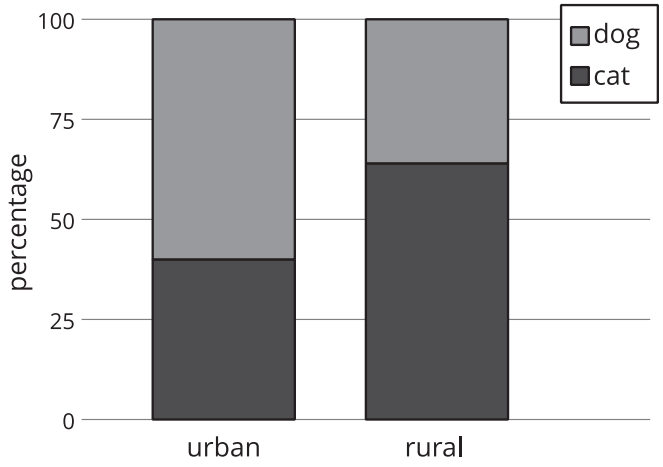
The next table shows what percentage of the 200 total people are represented by each combination of categories. The segmented bar graph represents the same information graphically.

	urban	rural
cat	27%	21%
dog	40%	12%



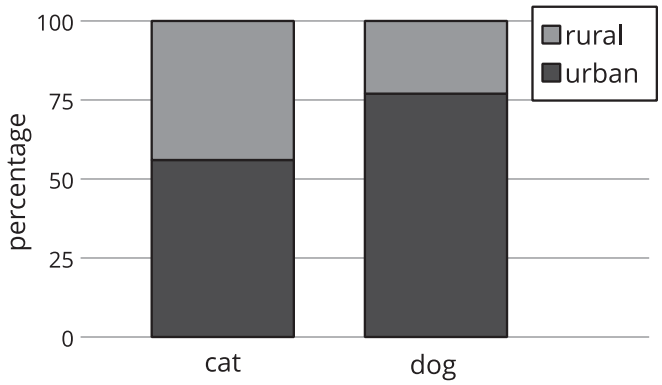
The next table is a column relative frequency table and shows the percentage of people represented in each column that have a certain pet preference. The segmented bar graph represents the same information graphically.

	urban	rural
cat	40%	64%
dog	60%	36%



The last table is a row relative frequency table and shows the percentage of people represented in each row that live in a certain area. The segmented bar graph represents the same information graphically.

	urban	rural
cat	56%	44%
dog	77%	23%



1. For each relative frequency table, select a percentage, and explain how numbers from the original table were used to get the percentage.

2. What percentage of those surveyed live in an urban area and prefer dogs?

3. Among the people surveyed who prefer dogs, what percentage of them live in an urban setting?

4. What percentage of people surveyed who live in an urban setting prefer dogs?

5. How many of the people responded that they prefer dogs and live in an urban setting?

6. Among the people surveyed, are there more people who prefer dogs or cats?

7. Your pet food company has access to a billboard in a rural setting. Would you recommend advertising dog food or cat food on this billboard? Which table did you use to make this decision? Explain your reasoning.

2.3 Analyzing a Study with Two Treatments

In an experiment to test the effectiveness of vitamin C on the length of colds, two groups of people with colds are given a pill to take once a day. The pill for one of the groups contains 1,000 mg of vitamin C, while the other group takes a placebo pill. The researchers record the results in a table.

	group A	group B
cold lasts less than a week	16	27
cold lasts a week or more	17	53

1. First, the researchers want to know what percentage (to the nearest whole percent) of people are in each combination of categories. 14% of all the participants had a cold that lasted less than a week and were in group A. What percentage of all the participants had a cold that lasted less than a week and were in group B? Complete the rest of the relative frequency table with the corresponding percentages.

	group A	group B
cold lasts less than a week	14% $\left(\frac{16}{16+27+17+53} \approx 0.14\right)$	
cold lasts a week or more		

2. Next, the researchers notice that among participants who had colds that lasted less than a week, 37% were in group A. Among participants who had colds that lasted a week or more, what percentage were in group B? Complete the table with the corresponding percentages.

	group A	group B
cold lasts less than a week	37% $\left(\frac{16}{16+27} \approx 0.37\right)$	
cold lasts a week or more		

3. Finally, the researchers notice that among the participants in group A, 48% had colds that lasted less than one week. Among the participants in group B, how many had colds that lasted a week or more? Complete the table with the corresponding percentages.

	group A	group B
cold lasts less than a week	48% $\left(\frac{16}{16+17} \approx 0.48\right)$	
cold lasts a week or more		

4. To understand the results, the researchers want to know the following: Among people whose colds lasted less than a week, what percentage are in group B? Explain your reasoning.

5. If the researchers believe that vitamin C has a small effect on the length of a cold, which group most likely got the pills containing vitamin C? Explain your reasoning.

Are you ready for more?

A teacher surveyed a group of 25 8th graders and a group of 20 12th graders who indicated that they know a computer programming language. Python and Scratch are programming languages. The results from the 8th-grade survey are displayed in the following two-way table.

	I know Python the best	I know Scratch the best	I know a different programming language the best
I have been taught a programming language at school	8	6	1
I have not been taught a programming language at school	1	7	2

The results from the 12th-grade survey are displayed in the two-way table.

	I know Python the best	I know Scratch the best	I know a different programming language the best
I have been taught a programming language at school	25%	35%	0%
I have not been taught a programming language at school	30%	5%	5%

1. Which programming language do a majority of the surveyed 8th graders know best?

2. Which programming language do a majority of the surveyed 12th graders know best?

3. How many of the surveyed 12th graders report that they were taught a programming language at school?

4. What percentage of the surveyed 8th graders report that they know Python the best and were not taught a programming language at school?

5. With the way the information is given in the tables, why is it difficult to decide if 12th graders or 8th graders use Python more?

Lesson 2 Summary

Converting two-way tables to *relative frequency tables* can help reveal patterns in paired categorical variables. **Relative frequency tables** are created by dividing the value in each cell in a two-way table by the total number of responses in the entire table, or the total responses in a row or a column. Depending on what patterns are important, different types of relative frequency tables are used. To examine how individual combinations of the categorical variables relate to the whole group, divide each value in a two-way table by the total number of responses in the entire table to find the relative frequency.

For example, this two-way table displays the condition of a certain textbook and its price for 120 of the books at a college bookstore.

	$10 or less	more than $10 but less than $30	$30 or more
new	3	9	27
used	33	36	12

A two-way relative frequency table is created by dividing each number in the two-way table by 120 because there are 120 values (3+9+27+33+36+12) in this data set. The resulting two-way relative frequency table can be represented using fractions, decimals, or percents. so inexpensive and that 10% of the books in the bookstore are both expensive and in used condition.

	$10 or less	more than $10 but less than $30	$30 or more
new	0.025	0.075	0.225
used	0.275	0.300	0.100

This two-way relative frequency table allows you to see what proportion of the total is represented by each number in the two-way table. The number 33 in the original two-way table represents the number of used books that also sell for $10 or less, which is 27.5% of all the books in the data set. Using this two-way relative frequency table, we can see that there are very few (2.5%) new books that are al

In other situations, it makes sense to examine row or column proportions in a relative frequency table. For example, to convert the original two-way table to a column relative frequency table using column proportions, divide each value by the sum of the column.

	$10 or less	more than $10 but less than $30	$30 or more
new	0.083	0.2	0.692
used	0.917	0.8	0.308

This shows that about 91.7% ($\frac{33}{3+33} \approx 0.917$) of the books that are sold for $10 or less are in used condition. Notice that each column of this column relative frequency table reveals the proportions of the books in each price category that are in each condition and that the relative frequencies in each column sum to 1. In particular, this shows that most of the inexpensive and moderately priced books are used, and most of the expensive books are new.