

KEY CONCEPT OVERVIEW

This topic extends the concept of a relationship between variables to **bivariate categorical data**. Students organize bivariate categorical data in a two-way table. They calculate **row** and **column relative frequencies**, decide whether there is an **association** by examining the differences (or similarities), and interpret the frequencies in the context of problems. Students discover that when two categorical variables have an association, knowing the value of one variable can help them predict the value of the other variable.

You can expect to see homework that asks your child to do the following:

- Organize data in a two-way table.
- Calculate relative frequencies.
- Determine whether there is an association in the data.

SAMPLE PROBLEMS (From Lesson 14)

Below is a two-way table of row relative frequencies for preferred movie types based on gender.

	Movie Preference			
	Action	Drama	Science Fiction	Comedy
Female	0.25	0.325	0.033	0.392
Male	0.625	0.0125	0.2	0.1625

1. If you randomly select a female participant, would you predict that her favorite type of movie is action? If not, what would you predict and why?

I would not predict that a female participant's favorite type of movie is action. I would predict that a female participant is more likely to prefer comedy since it has the greatest row relative frequency in the female row.

2. Is there an association between the variables of gender and movie preference? Explain your answer.

Yes, there is an association because the row relative frequencies are not the same in each row in the table. If I know a participant's gender, I can use the highest row relative frequency to predict that participant's movie preference.

Additional sample problems with detailed answer steps are found in the *Eureka Math Homework Helpers* books. Learn more at GreatMinds.org.

HOW YOU CAN HELP AT HOME

You can help at home in many ways. Here are some tips to help you get started.

- Make a two-way table for your child that includes only **frequency** values. (An online search for two-way tables produces many options.) Have her use the data to calculate the **cell relative frequencies**. Using the same original data, she then calculates the row relative frequencies and the column relative frequencies. (See Terms for descriptions on determining these values.) Have a discussion about patterns she sees in the data and whether she thinks there is an association between the variables.
- Guide your child to practice converting fractions to decimals or **percents**. For example, given the fraction $\frac{15}{37}$, have him divide the numerator by the denominator to convert to a decimal that is approximately 0.41. Now he can convert the decimal into a percent: $0.41 = 41\%$. Other examples include the following:
 $\frac{5}{8} = 0.625 = 62.5\%$, and $\frac{7}{12}$ is approximately 0.583, or 58.3%.

TERMS

Association: A relationship between the two variables of a bivariate data set. If there is an association between two categories (such as gender and movie preference), knowing a characteristic of one category allows one to make a prediction about the other category. (See Sample Problems.)

Bivariate categorical data: Data on two different categories, such as age and favorite social media site.

Categorical data: Data that can be divided into groups, such as race, gender, age, or favorite ice cream flavor.

Frequency: The number of times the data category occurs in a data set. For example, if 13 students choose chocolate ice cream as their favorite, the frequency is 13, and that number is placed in the cell for chocolate in a two-way frequency table.

Percent: One part in every hundred. One out of 100 is written as $\frac{1}{100}$ and 1%. Percentages can be used to describe relative frequency. For example, if 13 out of 25 students prefer chocolate ice cream, then $\frac{13}{25}$, or 52%, of the students prefer chocolate ice cream.

Proportion: A comparison of a part with the whole. Relative frequency is an example of a proportion.

Relative frequency: A fraction, decimal, or percent that represents how often a specific outcome occurs. To calculate the relative frequency, divide the target outcome by the total number of possible outcomes. For example, if a team won 9 out of 15 games last season, then the relative frequency of winning was $\frac{9}{15}$, 0.6, or 60%.

Cell relative frequency: A cell frequency divided by the table total (i.e., the total number of observations).

Column relative frequency: A cell frequency divided by the column total.

Row relative frequency: A cell frequency divided by the row total.