



Lesson 10: Summarizing Bivariate Categorical Data with Relative Frequencies

Student Outcomes

- Students summarize data on two categorical variables collected from a sample using a two-way frequency table.
- Given a two-way frequency table, students construct a relative frequency table and interpret relative frequencies.

Lesson Notes

Students continue the work with bivariate data from Lesson 9. Similar to Lesson 9, this lesson uses the superpower data to address the statistical question, “Do high school males have different preferences for superhero powers than high school females?” The two-way frequency table from Lesson 9 is used to develop a relative frequency table that will allow students to compare the responses of males and females. However, the statistical question is still not clearly answered. As students complete the exercises in this lesson, they begin to see the need for conditional relative frequencies, which are introduced in Lesson 11. Students also begin to understand how conditional summaries will be used to answer the statistical question.

This lesson revisits several topics that were developed in Grade 8. Students familiar with two-way tables will move through these questions within the suggested time frame. Students not as familiar with this work may need more time to complete the exercises.

Classwork

This lesson expands on your work with two-way tables from Lesson 9.

Exploratory Challenge 1 (5 minutes): Extending the Frequency Table to a Relative Frequency Table

Exploratory Challenge 1: Extending the Frequency Table to a Relative Frequency Table

Determining the number of students in each cell presents the first step in organizing bivariate categorical data. Another way of analyzing the data in the table is to calculate the *relative frequency* for each cell. Relative frequencies relate each frequency count to the total number of observations. For each cell in this table, the relative frequency of a cell is found by dividing the frequency of that cell by the total number of responses.

Consider the two-way frequency table from the previous lesson.

Two-Way Frequency Table:

	To Fly	Freeze Time	Invisibility	Super Strength	Telepathy	Total
Females	49	60	48	1	70	228
Males	51	71	27	25	48	222
Total	100	131	75	26	118	450

The relative frequency table would be found by dividing each of the above cell values by 450. For example, the relative frequency of females selecting “to fly” is $\frac{49}{450}$, or approximately 0.109, to the nearest thousandth. A few of the other relative frequencies to the nearest thousandth are shown in the following relative frequency table:

	To Fly	Freeze Time	Invisibility	Super Strength	Telepathy	Total
Females	$\frac{49}{450} \approx 0.109$					$\frac{228}{450} \approx 0.507$
Males			$\frac{27}{450} \approx 0.060$			
Total		$\frac{131}{450} \approx 0.291$			$\frac{118}{450} \approx 0.262$	

Read through the Exploratory Challenge as a class.

Ask students to look at the relative frequency 49 out of 450 students, which is the first cell highlighted in the two-way frequency table. This cell represents females selecting the superpower “to fly.” Indicate to students that the relative frequency of this cell is a fraction, a decimal, or a percentage. The highlighted cell expresses the frequency as 0.109, to the nearest thousandth. The value could also be interpreted as approximately 10.9%, or about 11% of the students surveyed were females whose favorite superpower is the ability “to fly.” Work with students on the other examples of relative frequencies that appear in the two-way table. Point out that approximately 0.06, or 6% of the students surveyed were males whose favorite superpower is “invisibility.” Indicate that 131 of the 450 students, or 0.291, approximately 29%, selected “freeze time” as their favorite superpower. Then, discuss the following:

- A relative frequency is found by dividing the frequency count by the total number of observations.
- A relative frequency can be expressed as a decimal or a percentage or a fraction.
- Explore the following question: How would the relative frequencies look if males and females had the same opinions about their favorite superpowers? This question is a prelude to Lesson 11. Use the question to indicate that because the number of females and the number of males are not the same, relative frequencies are difficult to interpret to answer this question. Some students may begin to indicate that a comparison to the total number of males or to the total number of females is more appropriate. This observation is addressed in the next lesson.

Exercises 1–7 (10 minutes)

Allow students to work independently on Exercises 1–7. Then, discuss and confirm answers as a class. Sample responses are indicated.

Exercises 1–7

1. Calculate the remaining relative frequencies in the table below. Write the value in the table as a decimal rounded to the nearest thousandth or as a percent.

Two-Way Frequency Table:

	To Fly	Freeze Time	Invisibility	Super Strength	Telepathy	Total
Females	0.109 10.9%	0.133 13.3%	0.107 10.7%	0.002 0.2%	0.156 15.6%	0.507 50.7%
Males	0.113 11.3%	0.158 15.8%	0.060 6.0%	0.056 5.6%	0.107 10.7%	0.493 49.3%
Total	0.222 22.2%	0.291 29.1%	0.167 16.7%	0.058 5.8%	0.262 26.2%	1.00 100%

2. Based on previous work with frequency tables, which cells in this table would represent the joint relative frequencies?

The joint relative frequency cells are represented by the unshaded cells within the body of the table.

3. Which cells in the relative frequency table would represent the marginal relative frequencies?

The marginal relative frequency cells are represented by the shaded cells in the total row and total column.

4. What is the joint relative frequency for females who selected “invisibility” as their favorite superpower?

The joint relative frequency for females who selected “invisibility” as their favorite superpower is 0.107, or approximately 10.7%. This indicates that approximately 11% of the students sampled were female who selected “invisibility” as their favorite superpower.

5. What is the marginal relative frequency for “Freeze Time”? Interpret the meaning of this value.

29.1%, or approximately 29% of the total number of people surveyed, selected “freeze time” as their favorite superpower.

6. What is the difference in the joint relative frequencies for males and for females who selected “to fly” as their favorite superpower?

The difference in the relative frequencies is 0.004, or 0.4%, or $0.113 - 0.109$.

7. Is there a noticeable difference between the genders and their favorite superpowers?

Yes. The most noticeable differences are in the following superpowers: invisibility, super strength, and telepathy.

Exploratory Challenge 2 (10 minutes): Interpreting Data

Exploratory Challenge 2: Interpreting Data

Interest in superheroes continues at Rufus King High School. The students who analyzed the data in the previous lesson decided to create a comic strip for the school website that involves a superhero. They thought the summaries developed from the data would be helpful in designing the comic strip.

Only one power will be given to the superhero. A debate arose as to what power the school's superhero would possess. Students used the two-way frequency table and the relative frequency table to continue the discussion. Take another look at those tables.

Scott initially indicated that the character created should have "super strength" as the special power. This suggestion was not well received by the other students planning this project. In particular, Jill argued, "Well, if you don't want to ignore more than half of the readers, then I suggest 'telepathy' is the better power for our character."

Read through the Exploratory Challenge as a class.

Then, discuss:

- Why would several students argue against Scott's initial suggestion of choosing super strength?
 - *0.2% of the responses were females who preferred this power.*
- Why might Scott have made this suggestion?
 - *Answers will vary. A reasonable answer could be personal preference.*
- Why would Jill say to use telepathy?
 - *It has the largest frequency count for females.*
- Do you think there is a difference in the superpowers selected by males and those selected by females?
 - *There are indications that males and females have differences in what superpowers they select. Less than 1% of the sample was females who selected "to fly," while over 5% of the sample was males who selected "to fly." The large difference suggests more males than females selected this superpower. However, point out to students that comparing the relative frequencies does not answer our question. Students may struggle to understand this question. (That is okay because the use of the relative frequency tables is limited.) Indicate that as we move into the next lesson, a direct comparison of these cells using conditional relative frequencies will provide a better indication of the gender differences in selection of superpowers.*

Exercises 8–10 (10 minutes)

Allow students to work in pairs. Then, discuss and confirm answers as a class.

Exercises 8–10

Scott acknowledged that "super strength" was probably not the best choice based on the data. "The data indicate that "freeze time" is the most popular power for a superhero," continued Scott. Jill, however, still did not agree with Scott that this was a good choice. She argued that "telepathy" was a better choice.

8. How do the data support Scott's claim? Why do you think he selected "freeze time" as the special power for the comic strip superhero?

"Freeze time" was the most popular choice among male students and among all students surveyed.

9. How do the data support Jill's claim? Why do you think she selected "telepathy" as the special power for the comic strip superhero?

Jill's claim was based on the observation that "telepathy" was the most popular choice of female students.

10. Of the two special powers "freeze time" and "telepathy," select one and justify why you think it is a better choice based on the data.

Answers will vary. The two categories each represent 26% to 29% of the sample. More females selected "telepathy," while more males selected "freeze time."

Closing (5 minutes)

Lesson Summary

- **Categorical data** are data that take on values that are categories rather than numbers. Examples include male or female for the categorical variable of gender or the five superpower categories for the categorical variable of superpower qualities.
- A **two-way frequency table** is used to summarize bivariate categorical data.
- A **relative frequency** compares a frequency count to the total number of observations. It can be written as a decimal or percent. A two-way table summarizing the relative frequencies of each cell is called a **relative frequency table**.
- The marginal cells in a two-way relative frequency table are called the **marginal relative frequencies**, while the joint cells are called the **joint relative frequencies**.

Exit Ticket (5 minutes)

Name _____

Date _____

Lesson 10: Summarizing Bivariate Categorical Data with Relative Frequencies

Exit Ticket

Juniors and seniors were asked if they plan to attend college immediately after graduation, seek full-time employment, or choose some other option. A random sample of 100 students was selected from those who completed the survey. Scott started to calculate the relative frequencies to the nearest thousandth.

	Plan to Attend College	Plan to Seek Full-Time Employment	Other Options	Totals
Seniors	$\frac{25}{100} = 0.250$	$\frac{10}{100} = 0.100$		
Juniors				$\frac{45}{100} = 0.450$
Totals	$\frac{60}{100} = 0.600$	$\frac{15}{100} = 0.150$	$\frac{25}{100} = 0.250$	$\frac{100}{100} = 1.000$

- Complete the calculations of the relative frequencies for each of the blank cells. Round your answers to the nearest thousandth.
- A school website article indicated that “A Vast Majority of Students from our School Plan to Attend College.” Do you agree or disagree with that article? Explain why you agree or disagree.
- Do you think juniors and seniors differ regarding after graduation options? Explain.

Exit Ticket Sample Solutions

Juniors and seniors were asked if they plan to attend college immediately after graduation, seek full-time employment, or choose some other option. A random sample of 100 students was selected from those who completed the survey. Scott started to calculate the relative frequencies to the nearest thousandth.

- Complete the calculations of the relative frequencies for each of the blank cells. Round your answers to the nearest thousandth.

The following is a completed chart:

	Plan to Attend College	Plan to Seek Full-Time Employment	Other Options	Totals
Seniors	$\frac{25}{100} = 0.250$	$\frac{10}{100} = 0.100$	$\frac{20}{100} = 0.200$	$\frac{55}{100} = 0.550$
Juniors	$\frac{35}{100} = 0.350$	$\frac{5}{100} = 0.050$	$\frac{5}{100} = 0.050$	$\frac{45}{100} = 0.450$
Totals	$\frac{60}{100} = 0.600$	$\frac{15}{100} = 0.150$	$\frac{25}{100} = 0.250$	$\frac{100}{100} = 1.000$

- A school website article indicated that “A Vast Majority of Students from our School Plan to Attend College.” Do you agree or disagree with that article? Explain why you agree or disagree.

Most of the students in the school plan to attend college after graduation (60% of sample). That is a clear majority. It is interesting, however, that a larger number of junior students plan to attend college, yet there were few juniors in the sample.

- Do you think juniors and seniors differ regarding after graduation options? Explain.

There are some interesting differences. One difference is noted above concerning college. Another is the selection of other options. A very small percentage of the students selecting that option were juniors.

Problem Set Sample Solutions

- Consider the Rufus King High School data from the previous lesson regarding after-school activities:

	Intramural Basketball	Chess Club	Jazz Band	Not Involved	Total
Males	20	2	8	10	40
Females	20	10	10	20	60
Total	40	12	18	30	100

Calculate the relative frequencies for each of the cells to the nearest thousandth. Place the relative frequencies in the cells of the following table. (The first cell has been completed as an example.)

	Intramural Basketball	Chess Club	Jazz Band	Not Involved	Total
Males	$\frac{20}{100} = 0.200$	$\frac{2}{100} = 0.020$	$\frac{8}{100} = 0.0800$	$\frac{10}{100} = 0.100$	$\frac{40}{100} = 0.400$
Females	$\frac{20}{100} = 0.200$	$\frac{10}{100} = 0.100$	$\frac{10}{100} = 0.100$	$\frac{20}{100} = 0.200$	$\frac{60}{100} = 0.600$
Total	$\frac{40}{100} = 0.400$	$\frac{12}{100} = 0.120$	$\frac{18}{100} = 0.180$	$\frac{30}{100} = 0.300$	$\frac{100}{100} = 1.000$

2. Based on your relative frequency table, what is the relative frequency of students who indicated they play basketball?

0. 40, or 40%

3. Based on your table, what is the relative frequency of males who play basketball?

0. 20, or 20%

4. If a student were randomly selected from the students at the school, do you think the student selected would be a male or a female?

The selection could be a female because 0. 60, or 60%, of the students in the sample were females.

5. If a student were selected at random from school, do you think this student would be involved in an after-school program? Explain your answer.

Yes. 70% of the students participate in an after-school program.

6. Why might someone question whether or not the students who completed the survey were randomly selected? If the students completing the survey were randomly selected, what do the marginal relative frequencies possibly tell you about the school? Explain your answer.

A person may question the sample based on the assumption that you would anticipate 50% of the students would be female and 50% would be male. The marginal relative frequencies indicate that the sample was 60% female and 40% male. Possibly, the school has more females than males.

7. Why might females think they are more involved in after-school activities than males? Explain your answer.

More females were involved in the after-school activities. 40 females were involved in the after-school activities, while 30 males were involved. (Note: Indicate to students that this will be explored in the next lesson.)