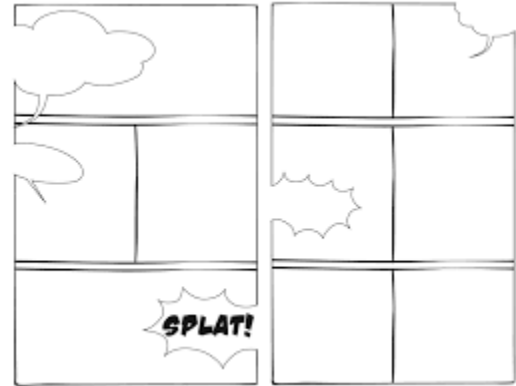


Lesson 10: Conditional Relative Frequencies and Association

Exploratory Challenge 1: Interpreting Relative Frequency 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.



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

Relative Frequency Table

divide each by 450

	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%

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.”

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.

1. 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 is the most popular among males.

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

Telepathy is the most popular among females.

3. 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.

Exploratory Challenge 2: Conditional Relative Frequencies

After further discussion, the students involved in designing the superhero comic strip decided that before any decision is made, a more careful look at the data on the special powers a superhero character could possess was needed.

There is an **association** between gender and superpower response if the superpower responses of males are not the same as the superpower responses of females. Examining each row of the table can help determine whether or not there is an association.

4. Do you think there is an association between gender and the superhero power chosen? Support your idea with evidence.

A **conditional relative frequency** compares a frequency count to the marginal total that represents the condition of interest.

For example, the condition of interest in the first row is females. The row conditional relative frequency of females responding invisibility as the favorite superpower is $\frac{48}{228}$, or approximately 0.211. This conditional relative frequency indicates that approximately 21.1% of females prefer invisibility as their favorite superpower. Similarly, $\frac{27}{222}$, or approximately 0.122 or 12.2%, of males prefer invisibility as their favorite superpower.

5. A. Based on the example given above, how is conditional relative frequency determined?

- B. How would we interpret the conditional relative frequency of 21.1% in the example above?

- C. How does this compare to the relative frequency for females responding invisibility as their favorite superpower? (See Exploratory Challenge 1)

6. Use the frequency counts from the table at the right to calculate the missing cells of row conditional relative frequencies. Round the answers to the nearest thousandth.

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

	To Fly	Freeze Time	Invisibility	Super Strength	Telepathy	Total
Females	$\frac{49}{228} \approx 0.215$	$\frac{60}{228} \approx 0.263$	$\frac{48}{228} \approx 0.211$	$\frac{1}{228} \approx 0.004$	$\frac{70}{228} \approx 0.307$	$\frac{228}{228} = 1$
Males	$\frac{51}{222} \approx 0.230$	0.320	0.122	0.113	0.216	$\frac{222}{222} = 1.000$
Total	$\frac{100}{450} \approx 0.222$	0.291	0.167	0.058	0.262	$\frac{450}{450} = 1$

7. Suppose that a student is selected at random from those who completed the survey. What do you think is the gender of the student selected? What would you predict for this student's response to the superpower question?

They are about equal. Freeze time is the most likely.

8. Suppose that a student is selected at random from those who completed the survey. If the selected student is male, what do you think was his response to the selection of a favorite superpower? Explain your answer.

Freeze time. $71/222$

9. Suppose that a student is selected at random from those who completed the survey. If the selected student is female, what do you think was her response to the selection of a favorite superpower? Explain your answer.

Telepathy. $70/228$

10. What superpower was selected by approximately one-third of the females? What superpower was selected by approximately one-third of the males? How did you determine each answer from the conditional relative frequency table?

Females → telepathy
Males → Freeze time

Lesson Summary

- A conditional relative frequency compares a frequency count to the marginal total that represents the *condition* of interest.
- The differences in conditional relative frequencies are used to assess whether or not there is an association between two categorical variables.

Homework Problem Set

Consider again the summary of data from the 100 randomly selected students in the Rufus King High School investigation of after-school activities and gender.

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

- Construct a row conditional relative frequency table for this data. Decimal values are given to the nearest thousandth.

	Intramural Basketball	Chess Club	Jazz Band	Not Involved	Total
Females					60
Males					40
Total					

- For what after-school activities do you think the row conditional relative frequencies for females and males are very different? What might explain why males or females select different activities?
- If John, a male student at Rufus King High School, completed the after-school survey, what would you predict was his response? Explain your answer.
- If Beth, a female student at Rufus King High School, completed the after-school survey, what would you predict was her response? Explain your answer.

5. Notice that 20 female students participate in intramural basketball and that 20 male students participate in intramural basketball. Is it accurate to say that females and males are equally involved in intramural basketball? Explain your answer.

Column conditional relative frequencies can also be computed by dividing each frequency in a frequency table by the corresponding column total to create a column conditional relative frequency table. Column conditional relative frequencies indicate the proportions, or relative frequencies, based on the column totals.

6. If you wanted to know the relative frequency of females surveyed who participated in chess club, would you use a row conditional relative frequency or a column conditional relative frequency?
7. If you wanted to know the relative frequency of band members surveyed who were female, would you use a row conditional relative frequency or a column conditional relative frequency?
8. For the superpower survey data, write a question that would be answered using a row conditional relative frequency.
9. For the superpower survey data, write a question that would be answered using a column conditional relative frequency.