

Two-Way Frequency Tables: A New 8th Grade Common Core Standard

Presented by:

Chase Orton

chase@mathandteaching.org

The Center for Mathematics and Teaching

www.mathandteaching.org

Session Poll Code: 11092

Date: _____

WARM UP

Here's some data to ponder before we get started.

	Lived	Died				
Male	367	1364				
Female	344	126				

What statements can you make about this data?

What questions would you like to ask?

I invite you to share your thoughts with your elbow partners.

SOME STRUCTURED WORK SPACE

The table on the right is an example of a two-way frequency table that shows data broken into two categorical variables: mortality and gender.

	Lived	Died	Total
Male	367	1364	1731
Female	344	126	470
Total	711	1490	2201

What do each of the five cells labeled "total" represent?

Another way to examine this data is to create relative frequency tables, like this one:

	Lived	Died	Total
Male ($n = \underline{\hspace{2cm}}$)			
Female ($n = \underline{\hspace{2cm}}$)			

And this one:

	Lived ($n = \underline{\hspace{2cm}}$)	Died ($n = \underline{\hspace{2cm}}$)
Male		
Female		
Total		

QUESTIONS TO DISCUSS

What might be good titles for these tables?

Which conclusions are easier to see using these relative frequency tables?

Which conclusions are easier to see in the frequency table that contains the raw data?

Which table is "better"?

THE STANDARD THAT WE ARE EXPLORING

Read the standard below.

As you read, underline key vocabulary that is new to you.

8.SP.4: Understand that patterns of association can also be seen in bivariate categorical data by displaying frequencies and relative frequencies in a two-way table. Construct and interpret a two-way table summarizing data on two categorical variables collected from the same subjects. Use relative frequencies calculated for rows or columns to describe possible association between the two variables. *For example, collect data from students in your class on whether or not they have a curfew on school nights and whether or not they have assigned chores at home. Is there evidence that those who have a curfew also tend to have chores?*

What questions do you have about standard 8.SP.4?

SOME ADDITIONAL DATA IF YOU'D LIKE TO EXPLORE MORE

Here's some other data about the Titanic you could explore if you wanted to geek out more. The statistics vary slightly from source to source so these numbers may deviate slightly from the data we explored earlier in the lesson.

Adults	Survivors		Non-Survivors	
	Male	Female	Male	Female
1st Class	57	140	118	4
2nd Class	14	80	154	13
3rd Class	75	76	387	89
Crew	192	20	670	3

Children	Survivors		Non-Survivors	
	Male	Female	Male	Female
1st Class	5	1	0	0
2nd Class	11	13	0	0
3rd Class	13	14	35	17
Crew	0	0	0	0

Below are some examples of frequency tables we could create.

	Survivors	Non-Survivors	Total
Passenger	499	817	1316
Crew	212	673	885
Total	711	1490	2201

	Survivors	Non-Survivors	Total
1 st class	197	122	319
2 nd class	94	167	261
3 rd class	151	476	627
Total	442	765	1207

What might be good titles for these tables?

What conclusions could you draw from the data in the tables above?

What kind of relative frequency tables would you need to construct in order to see if those conclusions might be true?

Here is some space for you to use if you'd like to create more tables.



A large empty grid consisting of 12 columns and 12 rows of squares, intended for creating tables.

**SIMPSON'S PARADOX:
A CONCEPTUAL EXTENSION FOR YOU TO PONDER**

In 1999 at the University of Calizona at Los Phoenix (UCLP), 393 men applied to the graduate school and 294 were admitted for an admission rate of 74.8%. The same year 444 women applied and 135 were admitted for an admission rate of 30.4%. To UCLP administrators this strongly suggested sex bias favoring men in graduate admissions. To track down the source of the bias, administrators ordered the individual graduate programs at UCLP to report their admission rates for men and women in 1999. This was a simple task in that UCLP has graduate programs in only four fields: English, Physics, Psychology, and Materials Science.

	Men			Women		
	Accepted	Applied	Rate	Accepted	Applied	Rate
English	12	43	27.9%	48	130	36.9%
Physics	119	124	96.0%	8	8	100.0%
Psychology	7	60	11.7%	59	285	20.7%
Materials Scien	156	166	94.0%	20	21	95.2%
	294	393	74.8%	135	444	30.4%

The baffled administration at UCLP is at a loss to explain the situation. Every department admits women at a higher rate than men, but overall the university admits men at a much higher rate than women. This counterintuitive situation is a concrete example of Simpson's Paradox. This paradox asserts that event A may be positively relevant to event B in every block of some partition of the population and yet be negatively relevant in the population as a whole (e.g., being female is positively relevant to being admitted in every program yet negatively relevant to being admitted overall).

In our example the source of the problem is clear. The two departments that have low admission rates overall also have high numbers of applications from women and low numbers from men. In the two departments with high admission rates overall the situation is reversed. Thus a large percentage of men are admitted and a large percentage of women are not, even though every department admits women at a higher rate than men.

Here are some other scenarios in which Simpson's Paradox might arise:

–Baseball player Willie may have a higher batting average than player Hank every year of their careers, and yet Hank may have the higher lifetime batting average.

–Cancer treatment A may produce a higher recovery rate than treatment B at every hospital in which both are tested, and yet treatment B may have the higher overall recovery rate when the data are combined.

CATEGORICAL DATA

Summary (Ready)

We will use two-way tables to display the frequencies and relative frequencies of categorical data. We will examine patterns of association in bivariate categorical data. We will draw conclusions about possible associations.

Goals (Set)

- Construct two-way frequency and relative frequency tables using bivariate data.
- Interpret two-way frequency tables.
- Conduct a survey.

Warmup (Go)

Numerical data is data consisting of numbers.

Below are some data about Ms. Robertson's 8th grade math class.

• Number of boys.....	12
• Number of girls.....	18
• Number of students who ate breakfast this morning.....	20
• Number of students who passed the math test today.....	25

Use the data to answer the questions below.

1. How many students are in the class? $n =$ _____

2. What percent of the students in the class are boys? _____

What percent of the students in the class are girls? _____

Why must the sum of these percents equal 100%?

3. What percent of the students ate breakfast this morning? _____

What percent of the students passed the math test today? _____

Why is it possible that the sum of these percents doesn't equal 100%?

WHAT IS CATEGORICAL DATA?

Categorical data is data sorted into categories, such as colors, ranges of measurements, or other attributes of the data. Generally there are only finitely many categories.

1. For each topic below, list categories for organizing collected data. Then create a possible survey question that you could use to collect data for each category.

Topic	Possible Categories	Possible Survey Question(s)
Example: Hair color	Black, brown, blonde, red, other	What color is your hair? OR Do you have dark hair?
Music		
Art		
Fruit		
Ecosystems		

2. If you were reading a table that had categorical data labeled “swimming,” “football,” and “basketball,” what might be the larger topic that is being explored?

3. Choose a topic that you could collect data about and list possible categories related to that topic.

Topic: _____

Categories:

CATEGORICAL QUESTIONS

Categorical survey questions are used to collect categorical data. These questions usually have word responses. For example, “Do you own at least one dog?” is a categorical survey question because the answer is either “yes” or “no.”

Numerical survey questions are used to collect numerical data. Numerical data sometimes comes from counting. It can sometimes come from measurements. These questions usually have number responses. For example, “How many dogs do you own?” is a numerical survey question because the answer is a number like zero or five.

State whether each question in the table below is a categorical or numerical survey question.

Question	Categorical (C) or Numerical (N)?
1. Are you male or female?	
2. What is your favorite color?	
3. How many dress shirts do you own?	
4. Are you an only child?	
5. How many siblings do you have?	
6. What were the class scores on the test?	
7. What types of birds did we observe today?	
8. What are the average family incomes of different cities in my state?	

9. State whether the question is categorical or numerical. Then rewrite the question in the other form.

“How many video games do you own?” Type: _____

Rewritten in other form:

10. Create a categorical question and then rewrite it as a numerical question.

BIVARIATE CATEGORICAL DATA

Bivariate data is data that has two variables based on the same population.

The results of an 8th grade survey about favorite color are below.

	Pink	Red	Blue	Purple	TOTAL
Boys	2	10	10	3	
Girls	8	2	3	7	
TOTAL					

1. Complete the table by finding totals for the rows and the columns.
2. How many students total were surveyed? $n =$ _____
3. How many girls were surveyed? _____
4. What percent of girls preferred pink? _____
5. How many boys were surveyed? _____
6. What percent of boys preferred pink? _____
7. How many students preferred pink? _____
8. What percent of the students who preferred pink are girls? _____
9. What percent of the students who preferred pink are boys? _____
10. Compare the **questions** from problems 4 and 8. How are they different?
11. Compare your **answers** from problems 4 and 8. Why are they different?
12. The table above is called a two-way table. Explain what you think a two-way table is in your own words.

CHORES AND CURFEWS

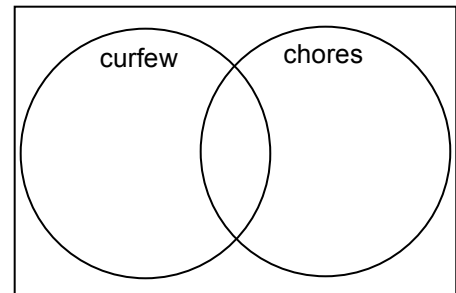
Ten different 8th graders were asked the following questions:

- Do you have a curfew?
- Are you assigned chores at home?

Data was collected on their responses and recorded in the table below.

	Students A through J									
	A	B	C	D	E	F	G	H	I	J
Curfew	Yes	No	No	Yes	Yes	No	Yes	No	No	Yes
Chores	Yes	No	Yes	Yes	No	No	Yes	Yes	No	Yes

1. Use the sections in the Venn diagram to record the number of students in each category. Be sure to include the number of students with neither chores nor curfew outside of the circles.



2. How many students were surveyed? _____
This is called the number of observations (n) or sample size for the survey.

3. How many students surveyed had chores? _____

4. How many students surveyed had a curfew? _____

5. What percent of students had both chores and a curfew? _____

6. What percent of students had neither chores nor a curfew? _____

7. What percent of students who had chores also had a curfew? _____

8. What percent of students who had a curfew also had chores? _____

9. Madhav thinks that most students have both a curfew and chores. Does the data support Madhav's claim? Explain.

TWO-WAY FREQUENCY TABLES

A frequency table is a table that lists items and the number of times they occur in a data set.

A two-way frequency table is useful for displaying bivariate categorical data.

1. Use the data on the previous page to complete the table.

	Students with Curfew	Students with No Curfew	TOTAL
Students with Chores			
Students with No Chores			
TOTAL			

Based on this table:

2. Circle the number that indicates the sample size for the whole survey.

What percent of all students had chores? _____

3. Draw a square around the number that indicates the total number of students who had chores.

What percent of students who had chores also had a curfew? _____

4. Draw a triangle around the number that indicates the total number of students who did not have a curfew.

What percent of the students who did not have a curfew also did not have chores?

5. Draw a parallelogram around the number that indicates the total number of students who had a curfew.

What percent of students who had a curfew also had chores? _____

6. Raji thinks that most students who had chores were more likely to have a curfew. Does the data support Raji's claim? Explain.

RELATIVE FREQUENCY TABLES

A frequency table with raw data can be used to create a relative frequency table that contains percents. The strength of any association or relationship between two variables can be easier to describe using percents.

- Here is one way to create a relative frequency table about curfews and chores. Use data from the previous page to complete the table.

	Table 1: Curfew and Chores		
	Curfew	No Curfew	TOTAL
Chores ($n = \underline{\quad}$)			100%
No Chores ($n = 4$)	$\frac{1}{4} = 25\%$		100%

- Here is another way to create a relative frequency table about curfews and chores. Use data from the previous page to complete this table.

	Table 2: Curfew and Chores	
	Curfew ($n = 5$)	No Curfew ($n = \underline{\quad}$)
Chores		
No Chores	$\frac{1}{5} = 20\%$	
TOTAL	100%	100%

- Circle the percent of students with both curfews and chores in each table. Why is this percent different in the two tables?
- Raji thought that students who had chores were more likely to have a curfew as well. Do these relative frequency tables support her claim? ___ Explain.
- Barbara states, "Of the students who were surveyed, students without curfews are less likely to have chores compared to those with a curfew."

Draw squares around the percents that show the association between not having a curfew and not having chores. Then determine if Barbara's reasoning is correct.

TASK: A SURVEY

1. As a class, conduct a survey using the questions below.

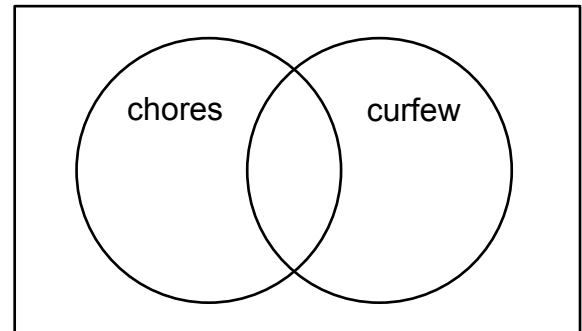
- Do you have chores at home?
- Do you have a curfew?

Record your data in the table below.

Student	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Chores? Y/N															
Curfew? Y/N															

Student	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
Chores? Y/N															
Curfew? Y/N															

2. Compile your results in the Venn diagram.



3. On a separate sheet of paper, use the Venn diagram to create a frequency table.

Then use the frequency table to create two relative frequency tables. You may use the tables in the student pages as a guide to help you.

4. On a separate sheet of paper, analyze your data by answering the following questions:

- How do your results compare to the “Curfew and Chores” data you examined in class?
- Is there an association between having a curfew and having chores for students in your class? Explain.
- What other categorical variables might be associated with having a curfew or having chores? How could you find out?

CATEGORICAL DATA

Summary

Students use two-way tables to display the frequencies and relative frequencies of categorical data. Students examine patterns of association in bivariate categorical data. Students draw conclusions about possible associations.

Goals

- Construct two-way frequency and relative frequency tables using bivariate data.
- Interpret two-way frequency tables.
- Conduct a survey.

PREVIEW / WARMUP

Whole Class

Page 0
Word Bank

Lesson 10.1 contains challenging and subtle concepts, and may require more direct instruction and guidance than other lessons in this program.

Page 1
Categorical Data

- Introduce the goals and standards for the lesson. Discuss important vocabulary as relevant. Students complete the warmup activity.

Why must the percentages of boys and girls in the class total 100%? Because all (100%) of the students are either boys or girls.

Why doesn't the sum of the percentage of students who ate breakfast and the percentage students who passed the math test have to total 100%? Because it is possible for some students to have breakfast and also pass the math test. Students can be counted in both categories.

- At the conclusion of the warmup exercise, make it explicit to students that the prior knowledge they used will be extended to new knowledge in this packet.

INTRODUCE 1 / EXPLORE 1

Whole Class/
Partners

Page 2
What is Categorical
Data?

- Explain the meaning of categorical data, and use hair color as an example.

If we wanted to survey people about music preferences, what are some categories we could use to organize our data? Answers will vary, but some categories could include rock, rap, country, etc.

What are some survey questions we might ask? Answers will vary, but some questions might be: What is your favorite music? What do you like to listen to on the radio? Where do you listen to music the most?

- Students complete the rest of the table with a partner. Then ask students to share responses with the rest of the class.

Is the question “Do you like art?” a categorical question? Yes. ***What would the categories be?*** “Yes.” “No.” Some of the categorical questions later in the lesson have “Yes/No” responses.

- Students answer the open-ended questions below the table before sharing responses.

Demographics (statistics relating to populations of people) could be an interesting topic for discussion to promote critical thinking for all students. ***“How are people organized into categories?”*** may be a worthwhile question to ask your students depending on their maturity and social awareness. Such a question invites a discussion about the challenges and consequences of organizing people by race, gender, age, socioeconomic status, sexual orientation, etc. Students encounter demographic data in future math and non-math classes. Use your discretion and assess the culture of your classroom before entering into such a discussion. See TN5 for further discussion on this topic.

- Students distinguish between categorical and numerical survey questions.

[Problem 9] Is the question categorical or numerical? Numerical is the obvious answer, but one could think of the answer in categories (0, 1, 2, 3, 4, etc.... or 0-2, 3-5, 6-8, etc.).

How would you write the question as a categorical question? Possibilities include, “Do you own any video games?” or “What kind of video games do you own?” (e.g sports, puzzles, role playing games etc.)

- [Problem 10] Have students share their examples of categorical and numerical questions, and critique the clarity of the questions.

INTRODUCE 1 / EXPLORE 1 (Continued)

Whole Class/
Individuals/Partners

Page 4
Bivariate Categorical
Data

- Review the meaning of categorical and bivariate data. Ask some clarifying questions about the table. Check to see that students can total the rows and columns correctly.

What two topics are measured and displayed in this table? Gender in the two rows and favorite color in the four columns.

What categories are used to describe gender? “Boy” and “girl.”

What categories are used to describe the favorite color? “Pink,” “red,” “blue,” and “purple.”

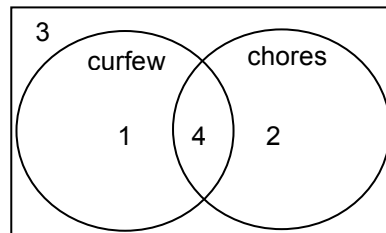
How many students were surveyed in total? Both the “total” column and row add to 45. Remind students that we use “ n ” to represent the total population in a survey.

- Students finish the rest of the page individually or collaboratively. Support as needed.
- Students may find it difficult to identify the correct denominator when finding percents. The following prompts and problems are meant to help students become aware of how they need to read carefully to determine which number to use as a denominator.

How are questions from problems 4 and 8 different? In problem 4, students find the percent of girls who prefer pink. In problem 8, students find the percent of “pink favorites” who are girls.

How are the answers from problems 4 and 8 different? In problem 4, students find the percent out of the total girls ($8/10$). In problem 8, they find percent out of the total students who like pink ($2/10$)

- Ask similar questions to help students compare problems 5 and 9.
- Students share their understanding of two-way frequency tables.
- Introduce the survey questions and help students complete the Venn diagram using the data in the table. This data is used on the following page



How many students were surveyed? 10. **How many had chores?** 6. **How many had curfew?** 5. **What percent of students had both chores and a curfew?** $4/10 = 40\%$.

What percent of students had neither chores nor a curfew? $3/10 = 30\%$.

What percent of students who had chores also had a curfew? $4/6$ (about 67%).

What percent of students who had curfew also had chores? $4/5 = 80\%$.

Does the data support Madhav’s claim that most students have both a curfew and chores? No, only 40% of students have both curfew and chores. However, some students might suggest (and correctly so) that there appears to be an association between the two variables and that students with a curfew are more likely to also have chores and vice versa.

Page 5
Chores and Curfews

INTRODUCE 2 / EXPLORE 2

Whole Class/
Partners

Page 6
Two-Way Frequency
Tables

- Introduce two-way frequency tables as a tool for displaying bivariate categorical data. Help students fill in the table using the Venn diagram on the previous page. As students do problems, be sure they understand that each number in the “total” columns and rows can be used to represent the sample size for different subgroups. Some questions are repeats from the previous page for emphasis.

	Students with Curfew	Students with No Curfew	TOTAL
Students with Chores	4	2	6
Students with No Chores	1	3	4
TOTAL	5	5	10

What is the sample size of the whole survey? 10.

See TN7 for more teaching strategies to help students.

What is the total number of students who had chores? 6.

What percent of all students had chores? $6/10 = 60\%$

What percent of student who had chores also had a curfew? $4/6$ is about 67%.

What is the total number of students who did not have a curfew? 5. **What percent of the students who did not have a curfew also did not have chores?** $3/5 = 60\%$.

What is the total number of students who had curfew? 5. **What percent of students who had curfew also had chores?** $4/5 = 80\%$.

- Students analyze Raji’s claim and assess its validity using evidence in the table.

[Problem 6] Does the data support Raji’s claim that most students who had chores had a curfew? Yes. **Explain.** 67% (4 out of 6) of students who had chores had curfews.

How is Raji’s claim different from Madhav’s on the previous page? Madhav uses all the students in his denominator. Raji is analyzing using subsets of the sample size in her denominators.

Why might there be a relationship between students having chores and having a curfew? What might it say about parenting behavior? Answers will vary.

Do you think this relationship might exist if we conducted this survey in class? Answers will vary.

INTRODUCE 2 / EXPLORE 2 (Continued)

Whole Class/
Partners

Page 7
Relative Frequency
Tables

- Explain that it can be easier to see if inferences like Raji's are accurate by creating relative frequency tables. Students calculate the percentages for each cell in the relative frequency tables. Identifying the correct denominator (n) to use for each row (in Table 1) and each column (Table 2) is critical. The numerator can be found in the appropriate field in the frequency table.

	Curfew	No Curfew	TOTAL
Chores ($n = 6$)	$4/6 = 67\%$	$2/6 = 33\%$	100%
No Chores ($n = 4$)	$1/4 = 25\%$	$3/4 = 75\%$	100%

	Curfew ($n = 5$)	No Curfew ($n = 5$)
Chores	$4/5 = 80\%$	$2/5 = 40\%$
No Chores	$1/5 = 20\%$	$3/5 = 60\%$
TOTAL	100%	100%

Circle the percent that compares curfews and chores in each table. Why are the two percents that compare curfews and chores different in each table? They have different sample sizes. In Table 1, the sample size is 6 (the total students with chores). In Table 2, the sample size is 5 (the total students with curfews).

- Students explain why both Raji's and Barbara's inferences are correct and cite percents from the table as evidence. **Does Barbara's argument support or refute Raji's argument?** Answers will vary. But together, their arguments seem to support the idea that there is a relationship between these categorical variables.

SUMMARIZE

Whole Class

- Formatively assess student understanding of key vocabulary. Students can create a word bank on a classroom wall that lists definitions and examples of categorical and numerical data, categorical data questions, bivariate data, frequency tables, and relative frequency tables. Use examples from the lesson to assist in the discussion.

What is bivariate data? Data that has two variables. **What is categorical data?** Data sorted into categories. **What do you think bivariate categorical data means?** Bivariate categorical consists of two attributes that are collected from each member of the sample population such as hair color and eye color of students at a school.

How are relative frequency tables a useful tool for analyzing bivariate data? They help us see possible associations between variables.

CLOSURE

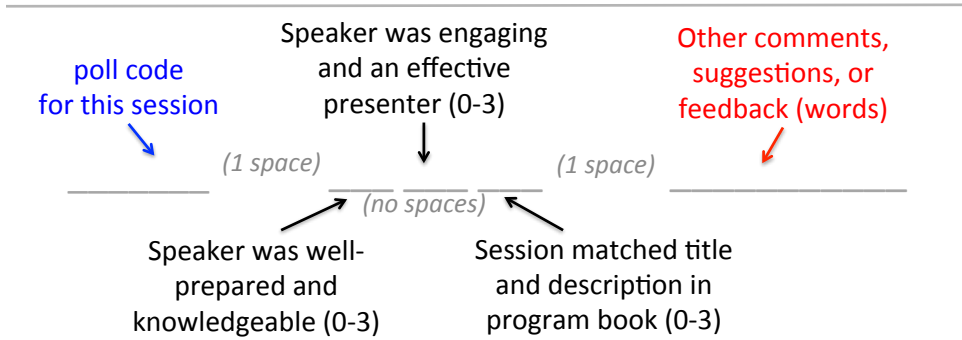
Whole Class

- Review the goals, standards, and vocabulary for the lesson.

POLL CODE: 11092

Strongly Disagree	Disagree	Agree	Strongly Agree
0	1	2	3

Send your text message to this Phone Number: 37607



Example: 38102 323 Inspiring, good content

Non-Example: 38102 3 2 3 Inspiring, good content

Non-Example: 38102 3-2-3Inspiring, good content

