LEARNING OBJECTIVE: We will summarize bivariate categorical data in two-way tables (Lesson 65)

CONCEPT DEVELOPMENT:

We have been looking at relationships between numerical variables. Now, we will look at relationships between categorical variables

TEMPERATURE WEATHER FORECAST sunny, cloudy...

Numerical Variables: Variables that represent data that is measured in numbers.
Example: Size of a house. Miles per hour.

Categorical Variables: Variables that represent data evaluated using specific categories or descriptions.
Example: Favorite ice cream flavor, gender, age range.

Univariate Categorical Data: Data on one variable that is categorical.

One-Way Frequency Table: A way to organize and present univariate categorical data. *Number of occurrences.*
Example: 30 students were asked their favorite ice cream flavor.

Ice Cream Flavor	Chocolate	Vanilla	Cookie Dough	Total
Number of Students	15	9	6	30

Number of occurrences

Relative Frequency: A description of the frequency of the occurrences of each categorical data in relation to the whole. A **proportion** measured by the following fraction: $\frac{frequency}{total}$.
Example:

Ice Cream Flavor	Chocolate	Vanilla	Cookie Dough	Total
Number of Students	.50 $\frac{15}{30}$	.30 $\frac{9}{30}$	.20 $\frac{6}{30}$	1.00

Bivariate Categorical Data: Data on two variables that is categorical. This data is easily organized and summarized in a **two way frequency table**.
Example: The number of boys and girls who like a specific flavor of ice cream.

		Favorite Ice Cream Flavor			
		Chocolate	Vanilla	Cookie Dough	Total
Gender	Male	7 $\frac{7}{30}$	8	2	17
	Female	8 $\frac{8}{30}$	1	4	13
	Total	15	9	6	30

Guided Practice:

Steps for Summarizing Bivariate Data
1. Select the variables you would like to evaluate.
2. Enter in the values for each of the data.
3. Figure out relevant relative frequencies.
4. Answer any questions regarding the data.

Below is a one way table that reports data collected on how we normally get to school. Answer all questions below.

More than 1 drive.

Mode of Transportation to School	Walk	Skateboard/ Scooter	Bike	In a car	TOTAL
Number of students	18 $\frac{18}{87}$ .21	1 KAIA .01	21 .24	47 .54	87 1.00

1. What is the relative frequency for each of the different ways to get to school?

Below is a two-way table that reports data collected on gender and cell phone ownership. Answer all questions below.

Columns

		Gender		
		Male	Female	TOTAL
Cell Phone Owner?	Yes	36 $\frac{36}{87}$ .41	35 .40	71 .82
	No	11 .13	5 .06	16 .18
	TOTAL	47 .54	40 $\frac{40}{87}$ .46	87 1.00

Rows

1. What is the relative frequency for each of the cells above?

FREQUENCY   PERCENT

2. Of the girls, what percent have cell phones?

40   35 $\frac{35}{40}$ = .88   88%

COLUMN RELATIVE FREQUENCY

3. Of cell phone owners, what percent are girls?

71   49% $\frac{35}{71}$ .49

ROW RELATIVE FREQUENCY

Below is a two way table that reports data collected. Answer all questions below

		Mode of Transportation to School				
		Walk	Skateboard/ Scooter	Bike	In a car	Total
Gender	Male	11 .13	0 0	10 .11	26 .30	47 .54
	Female	7 .08	1 .01	11 .13	21 .24	40 .46
	Total	18 .21	1 .01	21 .24	47 .54	87

1. What is the relative frequency for each cell in the table above?

2. What percent of the girls ride a skateboard or a scooter to school?

$2\frac{1}{2}\%$ .03 (FREQ)   ROW RELATIVE FREQ

3. Of those who reported they walked to school, what percent were boys?

boy walkers → $\frac{11}{18}$ = 61%   .61   COLUMN RELATIVE FREQ
All walkers →

4. If a student is selected at random, how would you predict they would get to school?

IN A CAR!

5. If the randomly selected student was one who walked, do you think they are a boy or a girl?

More boys walk to school than girls
.61   .39

Below is a two-way table that reports data collected on gender and the snacks that students like to eat. Answer all questions below.
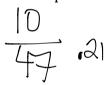
		Candy Bar	Baked Goods	Salty	Spicy	Healthy	Total
				Favorite Snack			
Gender	Male	9	10	15	5	8	47 .54
	Female	2	13	14	1	10	40 .46
	Total	11 .13	23 .26	29 .33	6 .07	18 .21	87

1. List the relative frequencies for each cell in the table.

Lots of salty kids, only 1 spicy girl.   Only 2 sweet girls
Girls are healthier than boys.

2. What is the difference in how you'd determine the proportion of male students who prefer baked goods and the proportion of students who are male AND prefer baked goods? Explain this in words.

$\frac{10}{23}$ $\frac{10}{47}$ .21 $\frac{10}{87}$ ≈ .11

3. What proportion of the female students prefer healthy foods?

$\frac{10}{40}$ .25   ROW FREQUENCY

4. What proportion of the students who prefer spicy snacks are male?

$\frac{5}{6}$ ≈ .83

CATEGORICAL DATA

Below is a two way table that reports data collected on how people get to school and the snacks they like to eat. Answer all questions below.

		Favorite Snack					
		Candy Bar	**Baked Goods**	**Salty**	**Spicy**	**Healthy**	**Total**
Mode of Transportation	**Walk**	2	3	7	0	6 .33	**18**
	Skateboard/ Scooter	0	0	1	0	0	**1**
	Bike	1 .05	11	5	0	4	**21**
	In a car	8	9	16	6	8	**47**
	Total	**11** .13	**23**	**29**	**6**	**18** .20	**87**

1. Write down three interesting observations that you would be willing to share with the class.
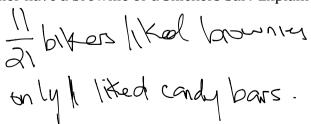
• Only 1 biker likes candy bars
• All spicy people in cars.
• Bikers ♡ cookies.
• More health nuts in car than walk.

2. What is the proportion of students that bike to school and like salty snacks?

$$\Rightarrow \frac{freq.}{total} = \frac{5}{87} = .06$$

3. A student is selected at random. They biked to school this morning. Would they rather have a brownie or a Snickers bar? Explain your thoughts.

$\frac{11}{21}$ bikers liked brownies

only 1 liked candy bars.

Below is a two-way table that reports data collected on sports we like to watch and sports we like to play. Answer the questions below.

		Favorite Sport to WATCH					
		Baseball	**Basketball**	**Football**	**Hockey**	**Soccer**	**Total**
Favorite Sport to PLAY	**Baseball**	13 .72	0	2 .11	2 .11	1 .06	**18**
	Basketball	4 $\frac{4}{21}$ .19	10 $\frac{10}{21}$ = .48	3 .14	0	4 .19	**21**
	Football	0	1 .20	4 .80	0	0	**5**
	Hockey	0	1 .25	1 .25	2 .50	0	**4**
	Soccer	5 .13	5 .13	6 .15	3 .08	20 .51	**39**
	Total	**22**	**17**	**16**	**7**	**25**	**87**

1. Which sport is the most popular to play?

Soccer $\frac{39}{87} = .45$

2. Which sport is the most popular to watch?

Soccer $\frac{25}{87} = .29$

3a. Of those who said soccer was their favorite sport to PLAY, what percent also said it was their favorite sport to WATCH? $\frac{20}{39} = .51$

3b. Of those who said soccer was their favorite sport to WATCH, what percent also said is was their favorite sport to PLAY?

$\frac{20}{25} = .80$

3c. What conclusions can you draw from this?

If you watch soccer, you're more likely to play soccer.

4. Create row relative frequencies for the favorite sport to PLAY.

Independent Practice:

If time, independent practice will be to create their own two-way table with the data collected.

Activating Prior Knowledge:

We understand percentages and proportions.

| A bag of M&Ms has 300 candies in it. 70 are red, 30 are blue, 60 are green, 50 are brown and the rest are yellow. What is the percentage of yellow M&Ms in the bag? $\frac{90 \div 3}{300 \div 3} = .3 = 30\%$ | If you randomly selected 20 M&Ms from the bag, how many would you expect to be blue? 300 total, 30 are blue. That's 10%. So... 10% of 20 would be blue too.... and that's 2 M&Ms $\frac{30}{300} = \frac{x}{20}$ |

Closure:

Why can't you graph these relationships on a coordinate plane?

Notes:

This maps to lesson 13 from Grade 8, Module 6.