Investigation 5

Variability and Associations in Categorical Data

Early roller coasters had wooden frames. Now, most roller coasters have steel frames, even though wood is still popular. A recent roller coaster census counted 174 wood-frame coasters, with 129 in North America.

How can people compare wood- and steel-frame roller coasters? Wood and steel are types of frames, not numbers. *Type of roller coaster* is what statisticians call a **categorical variable** that has values *wood* and *steel.*

A study comparing popularity of dogs would investigate the categorical variable *breed of dog* with values *standard poodle, Irish setter, German pointer, Chihuahua, whippet,* and so on.

Common Core State Standards

8.SP.A.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.

Also 8.EE.B.5, 8.EE.C.7, 8.F.A.3, S.ID.B.5 and S.ID.C.9

5.1 Wood or Steel? That's the Question

Relationships in Categorical Data

To plan a new amusement park, a team of coaster designers asked customers, "Do you prefer wood or steel frames in roller coasters?" The table shows the preferences by age group.

	Prefer Wood	Prefer Steel
Age ≤ 40 years	45	60
Age > 40 years	15	20

Does it look like younger and older riders have the same preferences in roller coaster type?

Problem 5.1

Study the roller coaster survey data by age of rider. Make a recommendation about the type of coaster that should be installed in the new park.

A Use the survey data. Is each statement *true* or *false*? Explain.

1. Younger riders are three times as likely as older riders to prefer wood-frame coasters.
2. Younger riders are three times as likely as older riders to prefer steel-frame coasters.
3. The number of riders who prefer wood-frame coasters is about three quarters of the number who prefer steel-frame coasters.
4. Younger riders are more likely than older riders to prefer steel-frame coasters.
5. Older riders are more likely than younger riders to prefer wood-frame coasters.

continued on the next page >

Problem 5.1 continued

B Suppose that a park installed one of each type of roller coaster. One day there were 210 riders over the age of 40 and 420 riders under the age of 40. Use the survey data from Question A.

1. How many riders would you expect on the wood-frame coaster and how many on the steel-frame coaster?
2. How would you expect those riders to be distributed by age and coaster type in the following table?

	Prefer Wood	Prefer Steel	Total
Age ≤ 40 years	■	■	420
Age > 40 years	■	■	210
Total	■	■	■

C If only one roller coaster type could be installed in the park, which would you recommend? Explain your choice.

ACE Homework starts on page 119.

5.2 Politics of Girls and Boys

Analyzing Data in Two-Way Tables

Every four years social studies teachers at the middle school hold a mock election. Each student registers as a Democrat, Independent, or Republican, all of which are categorical data values. Then the classes hold primary and final elections for President.

The table shows the student registrations in one class.

	Democrat	Independent	Republican
Boys	8	4	12
Girls	8	2	6

- Do you think boys and girls have different party preferences?
- What evidence could you give as support?

Problem 5.2

There are different ways to answer the question about political preferences of girls and boys in the sampled class.

A Use the table on the previous page. Do you think each statement is *true* or *false*? Justify your answers.

1. Girls and boys are equally likely to be Democrats.
2. Boys are more likely than girls to be Independents.
3. Boys are more likely than girls to be Republicans.
4. Girls are only half as likely as boys to be Republicans.

B Study the table of party choices and claims about differences between boys and girls. Notice that there are 24 boys and 16 girls in the class.

1. Copy and complete this extended table.

	Democrat	Independent	Republican	Totals
Boys	8	4	12	
Girls	8	2	6	
Totals				

2. Do the totals of political party choices change your answers to Question A? Explain your reasoning.

C One way to compare groups with unequal numbers of members is to compute percents.

1. Copy and complete the table below to show the fractions or percents of boys and girls with each preference.

	Democrat	Independent	Republican
Boys	$\frac{8}{24}=\frac{1}{3}=33\frac{1}{3}\%$		
Girls			

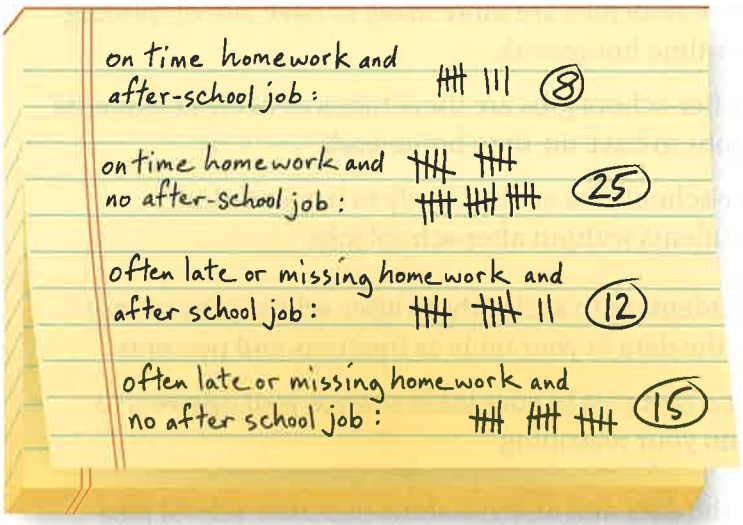
2. Do the percent calculations change your answers to Question A? Explain your reasoning.

ACE Homework starts on page 119.

5.3 After-School Jobs and Homework

Working Backward: Setting up a Two-Way Table

The teachers at the high school did a study to see whether students who had jobs after school were more or less likely to turn in homework on time than students who did not have after-school jobs. Each student was categorized as *usually on time* or *often late* with homework and as *having a job* or *not having a job*. Here are the results.



Is there evidence that students with after-school jobs are more likely to have late or missing homework than students without after-school jobs?

Problem 5.3

Use the information about the students to answer these questions.

A Make a table to display the data on students and after-school jobs.

B Use your table from Question A. Do you think each statement is *true* or *false*? Justify your answers.

1. Students without after-school jobs are more likely to have late or missing homework than students with after-school jobs.
2. Students with after-school jobs are more likely to have late or missing homework than on-time homework.
3. Students without after-school jobs are three times as likely as students with after-school jobs to have on-time homework.
4. Students with after-school jobs are less likely to have on-time homework than students without after-school jobs.

C

1. The numbers of students with and without after-school jobs are not the same. Rewrite the data in your table as fractions and percents.
2. Do the fractions and percents in your table change your answers to Question B? Explain your reasoning.

D If someone claims that the data and analysis show that after-school jobs cause students to have late or unfinished homework, what alternate explanations would you offer? What do you think could be the cause of late or unfinished homework other than after-school jobs?

ACE Homework starts on page 119.