

GRADE 8 · STATISTICS & PROBABILITY · ANSWER KEY

TWO-WAY TABLES

Today's goal: I can understand that patterns of association can also be seen in bivariate categorical data by displaying frequencies and relative frequencies in a two-way table and use relative frequencies calculated for rows or columns to describe possible association between the two variables.



How to use this key. Problem numbers match the student sheets. Each answer shows the steps and the slip Mr DJ sees most, so a wrong answer tells you what to reteach. Pepper levels: **Mild** warm-up, **Medium** stretch, **Spicy** challenge, **Extra Spicy** big brain.

Practice

9 problems

Use: In-class practice after the deck. Item 6 is the honest one and the most grown-up idea on the sheet: the gap rests on two students out of forty, so it may be nothing — and that is why a rate should always be reported with its denominator. Item 8 asks children to CONSTRUCT a case, which is the strongest test of whether they see that the count and rate comparisons are independent; equal group sizes is the easiest route. Item 5's best evidence is that pet owners lead in both columns at once.

Timing: 15–20 min · pairs · calculator allowed · items 6, 8 and 9 on the board.

- 1 because a mis-copied or miscounted table will not add to the same grand total both ways
- 2 pet owners 30%; non-owners 25%
- 3 about 58.3%, out of the 72 students who do chores
- 4 each pair adds to 100%, because every student in a row is in exactly one of the two columns
- 5 because a bigger group produces bigger numbers in every column, so a larger count may only mean a larger group
- 6 two students out of the forty non-owners — so the gap is small and could easily be noise
- 7 about 83.3% against 75% — the association is stronger, and it now runs the other way
- 8 for example: 45 of 50 pet owners do chores (90%) and 30 of 50 non-owners do (60%) — a bigger count and a bigger rate
- 9 because a row or column rate removes the effect of group size and answers a question about one group, while a share of the whole table cannot compare groups at all

1 MILD Why check the totals

because a mis-copied or miscounted table will not add to the same grand total both ways

1. Rows across and columns down must reach the same total.
2. If they do not, at least one number is wrong, and any conclusion drawn from it is worthless.
3. It takes seconds and it is done before any interpretation.

Watch for: Said “to be neat”. It is an error check, and it is the cheapest one available.

2 MEDIUM Non-chose rates

pet owners 30%; non-owners 25%

1. 18 out of 60 = 30%.
2. 10 out of 40 = 25%.

Watch for: Divided by the grand total, giving 18% and 10%.

3 **MEDIUM** Share of chore-doers with a pet

about 58.3%, out of the 72 students who do chores

1. $42 \text{ out of } 72 = 0.583\dots$, about 58.3%.
2. The denominator is the CHORES column total, 72, because the question asks about chore-doers.
3. This is a different question from “what share of pet owners do chores”, which was 70%.

Watch for: Used 60 or 100. The group named in the question is the group you divide by.

4 **SPICY** The rows sum to 100%

each pair adds to 100%, because every student in a row is in exactly one of the two columns

1. Pet owners: $70\% + 30\% = 100\%$.
2. Non-owners: $75\% + 25\% = 100\%$.
3. The two columns cover every possibility with no overlap, so the shares of a row must account for all of it.
4. This is a useful check: if a pair of row percentages does not reach 100, something has been divided by the wrong total.

Watch for: Noticed the pattern without saying why. The reason — the categories are exhaustive and exclusive — is what makes it a reliable check rather than a coincidence.

5 **SPICY** Why counts compare size

because a bigger group produces bigger numbers in every column, so a larger count may only mean a larger group

1. There are 60 pet owners and 40 non-owners, so pet owners are 1.5 times as numerous.
2. They therefore have more chore-doers AND more non-chore-doers: 42 against 30, and 18 against 10.
3. A count mixes the group's size with its behaviour, so it cannot separate them.
4. Dividing by the row total removes the size, leaving only the behaviour — which is exactly what a comparison needs.

Watch for: Said counts are “not accurate”. They are accurate counts; the problem is that they answer a different question, and noticing that pet owners lead in BOTH columns is the clearest evidence.

6 **SPICY** How fragile the gap is

two students out of the forty non-owners — so the gap is small and could easily be noise

1. 5% of 40 is 2 students.
2. If two non-owners had said they do no chores, their rate would fall to 70% and the gap would vanish.
3. So the finding rests on two answers, which is not much.
4. This is why a rate should be reported with its denominator: 75% of 40 and 75% of 4000 are very different levels of evidence.

Watch for: Said the gap is small without quantifying it. Two students is the useful form — it turns “small” into something a reader can judge.

7 **SPICY** The other school

about 83.3% against 75% — the association is stronger, and it now runs the other way

1. $50 \text{ out of } 60 = 83.3\%$. $30 \text{ out of } 40 = 75\%$.
2. The gap is about 8.3 percentage points, larger than our 5.
3. And the direction has reversed: here it is the pet owners who do more chores.
4. So neither the size nor the direction of an association carries over between schools — each has to be read from its own table.

Watch for: Compared 50 with 30 again. The whole lesson is that the counts cannot be compared when the groups differ in size.

8

SPICY

A table where both agree

for example: 45 of 50 pet owners do chores (90%) and 30 of 50 non-owners do (60%) — a bigger count and a bigger rate

1. With equal group sizes, the count comparison and the rate comparison always agree.
2. With unequal groups they may agree or disagree, so the rates must be checked either way.
3. A valid answer needs the four cells and both totals, so the rates can be verified.
4. The lesson is not that counts are always wrong — it is that they are not to be TRUSTED without the rates.

Watch for: Gave a table where the counts and rates disagree, or one that does not add up. Building the example is what shows the two comparisons are independent.

9

EXTRA SPICY

Why rows or columns

because a row or column rate removes the effect of group size and answers a question about one group, while a share of the whole table cannot compare groups at all

1. Dividing by the grand total gives each cell's share of everybody: 42%, 18%, 30%, 10%.
2. Those add to 100% and describe the table, and they cannot tell you whether pet owners behave differently.
3. A row rate holds the group fixed and asks about behaviour within it, which is what an association question needs.
4. So the choice of denominator is the choice of question, and the standard names rows and columns because those are the denominators that compare groups.

Watch for: Said row percentages are “more useful”. The reason is that they hold the group fixed — grand-total shares describe the table rather than comparing its groups.

Homework

8 problems

Use: Homework after the practice sheet, and the last of Grade 8. Item 5 shows the counts error doing a DIFFERENT kind of damage from the exit ticket's: here it does not reverse the direction, it hides a 20-point gap behind a 3-person difference — so the lesson is not “counts reverse things” but “counts cannot be trusted”. Item 7 is worth marking generously: a child who calls the claim false has over-corrected, since it may well be true and simply unreadable. Item 8 asks what “no association” looks like in a table, which is the idea that ties all three Statistics skills together. **Timing:** 20–25 min · independent · calculator allowed · items 5, 4 and 7 first when marking.

1 walkers 60 (30 bike, 30 no bike); drivers 90 (27 bike, 63 no bike); bike owners 57; non-owners 93; total 150

2 walkers 50%; drivers 30%

3 about 52.6%, out of the 57 bike owners

4 girls, at 50% against 40% — a gap of 10 percentage points

5 walkers, at 50% against 30%. The counts 30 and 27 are close and would have suggested the two groups are similar

6 because the same count gives different percentages depending on the total, and a percentage of a small group is much weaker evidence than the same percentage of a large one

7 how many members there are in each category, and how many the club has in total — and ideally what the comparison group is

8 because association means knowing one variable tells you something about the other, and that can be true of categories as well as measurements — the row rates differing IS the association

1 **MEDIUM** The table

walkers 60 (30 bike, 30 no bike); drivers 90 (27 bike, 63 no bike); bike owners 57; non-owners 93; total 150

1. Walkers: 60 total, 30 with a bike, so 30 without.
2. Drivers: 90 total, 27 with a bike, so 63 without.
3. Bike column: $30 + 27 = 57$. No-bike column: $30 + 63 = 93$.
4. Check: $57 + 93 = 150$, and $60 + 90 = 150$. Both agree.

Watch for: Left cells blank rather than working them out by subtraction from the row totals, or reached different grand totals.

2 **MEDIUM** Two rates

walkers 50%; drivers 30%

1. 30 out of 60 = 50%.
2. 27 out of 90 = 30%.

Watch for: Divided by 150, giving 20% and 18%.

3 **MEDIUM** Share of bike owners who walk

about 52.6%, out of the 57 bike owners

1. 30 out of 57 = 0.526..., about 52.6%.
2. The denominator is the bike-owner column total, 57.

Watch for: Used 60 or 150. The question names bike owners, so 57 is the total.

4 **MEDIUM** Boys and girls choosing science

girls, at 50% against 40% — a gap of 10 percentage points

1. Boys: 80 out of 200 = 40%.
2. Girls: 60 out of 120 = 50%.
3. So girls chose science more, by 10 percentage points, despite the smaller count.
4. The boys' count is higher only because there are more boys.

Watch for: Compared 80 with 60 and answered boys. The group sizes differ, so the counts cannot be compared.

5 **SPICY** Which group, and would counts agree?

walkers, at 50% against 30%. The counts 30 and 27 are close and would have suggested the two groups are similar

1. Walkers own bikes at 50%, drivers at 30% — a large gap.
2. The raw counts, 30 and 27, differ by only 3, so they would suggest almost no difference.
3. The driver group is half again as large, which inflates its count relative to its rate.
4. So here the counts do not reverse the direction, they HIDE a large difference — which is the same error causing a different kind of damage.

Watch for: Said the counts would give the same answer. They point weakly in the same direction and conceal the size, which matters just as much as direction — a 20-point gap read as a 3-person difference.

6 **SPICY** Why a bare percentage is uninterpretable

because the same count gives different percentages depending on the total, and a percentage of a small group is much weaker evidence than the same percentage of a large one

1. 42 out of 60 is 70%, out of 72 is 58%, out of 100 is 42% — one cell, three true percentages.
2. So without the denominator you cannot tell which question is being answered.
3. And 75% of 4 people is three people, while 75% of 4000 is strong evidence — the same percentage, very different weight.
4. So a rate needs its denominator both to be understood and to be judged.

Watch for: Gave only the first reason. Both matter, and the small-sample point is the one that catches misleading claims in the world.

7 **SPICY** The club's claim

how many members there are in each category, and how many the club has in total — and ideally what the comparison group is

1. “More beginners than experts” is a comparison of two counts within one club, so it needs both counts.
2. With 11 beginners and 9 experts it is barely a pattern; with 90 and 10 it is a real one.
3. It would also help to know whether other clubs differ — most clubs having more beginners would make this unremarkable.
4. So the claim as stated is not false, it is unreadable.

Watch for: Said the claim is false or misleading. It may be perfectly true; the problem is that no one can judge it, and naming the missing numbers is the answer.

8 **EXTRA SPICY** Why tables show association

because association means knowing one variable tells you something about the other, and that can be true of categories as well as measurements — the row rates differing IS the association

1. A scatter plot shows association by a trend in the dots; a table shows it by the row rates differing.
2. If walkers own bikes at 50% and drivers at 30%, then knowing how someone travels tells you something about whether they own a bike.
3. If both rates were equal, the two variables would be unrelated — the table's version of no association.
4. So the underlying question is identical, and only the display changes with the type of data.

Watch for: Said tables are “just a different way of showing data”. True and it misses the answer: the equal-rates case is what “no association” looks like in a table, and that is what makes the two displays answer the same question.

Exit Ticket

4 problems

Use: End of the two-way tables lesson and the last Grade 8 exit ticket. Item 3 is the diagnostic and it is built so the wrong method gives a confidently wrong answer: the larger group has the larger count AND the smaller rate, so comparing counts reverses the direction rather than blurring it. A child who writes “she should have used percentages” has the method and not the finding — the two rates are the argument. Item 4 is the denominator habit, and both percentages are correct, which is the point. **Timing:** 5–8 min · independent · calculator allowed.

1 rows 60 and 40; columns 72 and 28; grand total 100

2 pet owners 70%; non-owners 75%

3 the groups are different sizes — 60 pet owners against 40 non-owners — and by rate it is the NON-owners who do more chores, 75% against 70%

4 they have different denominators: 70% of PET OWNERS do chores, and 58% of CHORE-DOERS have a pet

1 **MEDIUM** The table

rows 60 and 40; columns 72 and 28; grand total 100

1. Pet row: $42 + 18 = 60$. No-pet row: $30 + 10 = 40$.
2. Chores column: $42 + 30 = 72$. No-chores column: $18 + 10 = 28$.
3. Grand total: $60 + 40 = 100$, and $72 + 28 = 100$ — both routes agree, so the table is consistent.

Watch for: Omitted the totals, or reached different grand totals by the two routes. Adding both ways is the check, and it costs seconds.

2 **MEDIUM** Two rates

pet owners 70%; non-owners 75%

1. 42 out of 60 pet owners: $42 \div 60 = 0.7$, so 70%.
2. 30 out of 40 non-owners: $30 \div 40 = 0.75$, so 75%.
3. Each cell is divided by its own ROW total, because the question asks about a share of each group.

Watch for: Divided both by 100, giving 42% and 30% — the share of ALL students, which does not compare the two groups. The denominator has to be the group you are describing.

3 **SPICY** Ines's error

the groups are different sizes — 60 pet owners against 40 non-owners — and by rate it is the NON-owners who do more chores, 75% against 70%

1. 42 is bigger than 30, and there are also more pet owners to count.
2. The rates are $42/60 = 70\%$ and $30/40 = 75\%$.
3. So the pattern runs the opposite way to the one the counts suggested.
4. This is not a small correction: comparing counts reversed the DIRECTION, because the bigger group holds more of everything.

Watch for: Said she should have used percentages, without giving them. The two rates are the argument, and the important observation is that the answer reverses — counts here are not merely imprecise, they are misleading.

4 **SPICY** Two correct percentages

they have different denominators: 70% of PET OWNERS do chores, and 58% of CHORE-DOERS have a pet

1. $42 \div 60$ uses the pet row, so it answers “what share of pet owners do chores?”
2. $42 \div 72$ uses the chores column, so it answers “what share of chore-doers have a pet?”
3. Both divisions are correct and they answer different questions.
4. So a percentage is unreadable without its denominator — the same cell also gives 42% of all students.

Watch for: Said one of them is wrong. Both are right; what differs is the question. Naming the denominator out loud is the habit that prevents this.

Lesson slides

teaching notes & timing

Slide	Time	What to do
1	0:00	Categorical, not measurement — contrast with the scatter plots of the last two lessons. Same question (is there an association?), different display.
2	0:45	Build the table with the class. Insist every person lands in exactly one box.
3	1:30	Totals both ways. Cheap, and it catches a mis-copied table before any conclusion is drawn from it.
4	2:15	THE slide. Ask the question and take the vote. Most of the class will say pet owners, and they should be allowed to commit before the group sizes appear.

5	3:15	The reveal. 70% against 75%. Not a small correction — the direction reverses, which is what makes counts dangerous rather than merely imprecise here.
6	4:30	The denominator. Three divisions of the same cell, all correct arithmetic, answering three different questions. The habit: say “70% OF PET OWNERS”, never “70%”.
7	6:00	How big is 5 points? Two students out of forty. Worth being honest here: a small gap on small numbers may be nothing at all, and reporting the denominator is what lets a reader judge it.
9	7:30	The row rates and the comparison, then the exit ticket. Whiteboards: everyone tries Mild, then climbs. These are #1, #2, #4 on the practice sheet.
10		Fingers up: 3 = got it, 2 = almost, 1 = not yet. Pull the 1s for a quick reteach while everyone else starts the practice sheet.
11		Launch the practice sheet. Exit ticket in the last 5–8 minutes.