



Two-way tables

Two categories, four boxes

- Today: categorical data, and why counts mislead.

The table

- Rows: has a pet / no pet.
- Columns: does chores / does not.
- Every student in exactly one box.

Check the totals first

- Rows add across. Columns add down.
- Both routes must reach the same grand total.

A table that does not add up cannot be interpreted.

The trap

- 42 pet owners do chores. 30 non-owners do.
- So pet owners do more chores?

There are 60 pet owners and only 40 non-owners. Bigger group, bigger count.

Divide by the group

- $42 \div 60 = 70\%$.
- $30 \div 40 = 75\%$.

The students WITHOUT pets do more chores. The counts had it backwards.

Say the denominator out loud

- 42 out of 60 pet owners: 70%.
- 42 out of 72 chore-doers: 58%.
- 42 out of 100 students: 42%.

One cell, three true percentages. “70%” alone says nothing.

How big is the gap?

- 70% against 75% is 5 points.
- That is two students out of forty.

Report the rate *WITH* its denominator so a reader can judge it.

Your turn

- Find the row relative frequencies.
- Which group does more chores — and by how much?

What did you divide by, and why?