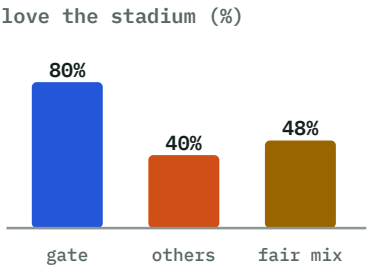
Who was asked? And how many?

A poll is only as good as **who** was asked: a big sample of the wrong people stays wrong. Then **how many**: even a fair, random poll wobbles by chance, about `100 ÷ √n` points. Question 7 of the ten to ask a number.

1 The stadium-gate poll

who stood there?

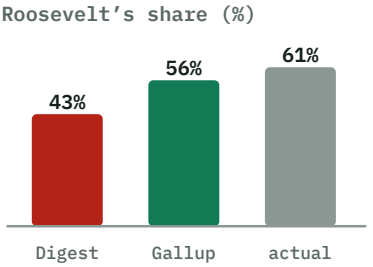
- **Coralia Gazette** (the toy island's newspaper): "**80%** of Coralians love the new stadium!" · 5,000 asked · all at the **stadium gate** → all fans
- Coralia: fans 2 in 10 → 80% love it · others 8 in 10 → 40% · fair mix → `(2 × 80 + 8 × 40) ÷ 10 = 48%`
- **Sample** = the people actually asked, used to learn about everyone · **random sample** = everyone has the same chance of being picked
- **Selection bias** = the sample leaves some kinds of people out, or picks others more often → it does not look like everyone



2 Big but biased

1936

- **The Literary Digest** (American magazine), **1936** election · ≈ 10 million **ballots** (voting cards to fill in and send back) · ≈ **2.4 million** back → **Landon** (challenger) **57%**, Roosevelt (president in office) 43%
- Names from **phone and car lists** → mostly better-off families in 1936 → poorer voters left out · only ≈ 1 ballot in 4 came back
- **George Gallup** (American pollster, 1901-1984) · ≈ **50,000** chosen to match the country's mix → Roosevelt ≈ 56% · result: Roosevelt ≈ **61%**
- ⇒ **who** is asked matters more than **how many** · 20,000 gym members still speak only for gym members



3 Margin of error

chance alone

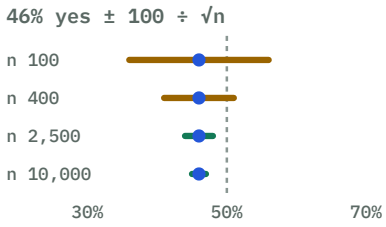
- **Margin of error** = how far a random poll is likely to be from the truth by chance alone · rule of thumb `margin ≈ 100 ÷ √n` percentage points · ± = plus or minus · n = people asked · √n = the number that times itself gives n (√400 = 20)
- **19 polls out of 20** land within the margin of the truth
- Halve the margin → **4 ×** the people (100 → 400) · accuracy gets costlier → polls stop near **1,000** → ≈ **± 3**

n	√n	margin
25	5	± 20
100	10	± 10
400	20	± 5
1,600	40	± 2.5
2,500	50	± 2
10,000	100	± 1

4 Reading a poll

result ± margin

- **Too close to call** → lead < ≈ 2 × margin · 1,000 asked, Mango 52 / Lime 48, ± 3 → 49-55 and 45-51 **overlap**
- "Fewer than half" → only if the **whole bar** is under 50% · 46% yes: 400 → 41-51 × · **2,500 → 44-48** ✓
- A poll that "moved": 48% → 51%, each ± 3 → 45-51 and 48-54 overlap → **within the margin** → may be chance · "about half, no clear change"
- Too few: 16 at random → ± 25 · 9 → ≈ ± 33 · 25 → ± 20



5 Surveys everywhere

who is missing?

- US unemployment rate (Macro Lesson 3) = a **survey**, not a count · **Current Population Survey** · Census Bureau (government statistics office) for the Bureau of Labor Statistics (publishes the rate) · ≈ **60,000** households a month → it has a margin too
- **Survivorship bias** (Finance Lesson 8) = judging a group only by the members that survived · funds that closed never enter the average → average looks **better**
- Online vote: "Vote now" button → only those who care most click → they lean one way → result misses, **however many click** · fix → 1,000 picked at random

Wrong people (selection bias)	Too few (huge margin)	Fair (random, enough)
stadium gate · magazine's own readers · click-to-vote · radio call-in · gym members · first 500 at a mall at 10 a.m.	16 or 9 picked at random	1,000 random by phone and online · 60,000 random homes visited · full address or student list

6 So what?

question 7 of 10

- **Ten questions to ask a number** · 1 · Where does the axis start, and what is its scale? · 2 · Average of what: the mean or the median? · 3 · Percent of what? Percent or percentage points? · 4 · Real or nominal: did prices change too? · 5 · What else changed? Could something else explain both? · 6 · Compared with whom? Where is the control group? · **7 · Who was asked, and who was not?**
- A good poll, in order: the question → pick people **at random** → ask **enough** → report the result **with its margin**
- Gazette rewritten → "A reporter at the stadium gate found that 80% of the fans she met love the new stadium. A random poll of 1,000 Coralians found 48%, give or take 3 points." · next → Lesson 8, what does the trend say?

Watch out: a big sample does not fix a biased one (2.4 million beat 50,000 for size, not accuracy) · size shrinks chance, not bias · always read a poll as result ± margin · a lead or a move smaller than the margin is not news