

Psych 10: Intro to Statistics

Lecture 5 · The mean as a model · Wed Oct 7, 2026

Prof. Judith Fan · Fall 2026

Reminders & Announcements

- Finish CourseKata **Chapters 4 and 5** by Friday.
- Milestone 2 is due Sunday, October 11 on Gradescope.
- Friday: **Case #2, *The Vault of Fallen Labels***. Check your discussion section's Slack channel for your time slot.

Today

The mean as a model

1. What is a model?
2. Modeling data with a single number
3. Where does “average” make sense?
4. Could it be chance?

Sixteen men out of church



Warm-up. In 1535, a German surveyor's manual gave this recipe for the rod, the unit used to measure fields. Stop sixteen men as they come out of church, tall ones and short ones, and have them line up their left feet toe to heel. That length is the rod, and one sixteenth of it is a foot. **Why sixteen men instead of one? And what kind of error would remain in a rod made from sixteen feet?**

Jacob Köbel, *Geometrei*, Frankfurt, 1535. The story is told in Stephen Stigler, *The Seven Pillars of Statistical Wisdom* (2016)

Sixteen feet is a mean

- Different feet vary—some are longer than others. When you lay sixteen feet end to end, the long ones and the short ones partly cancel out, and a sixteenth of the total comes out close to a typical foot. That is a mean!
- The new idea, in 1535, was to treat sixteen different men's feet as sixteen measurements of the same thing, **a foot**, and to treat the differences between the men as **random variation**.

1 What is a model?

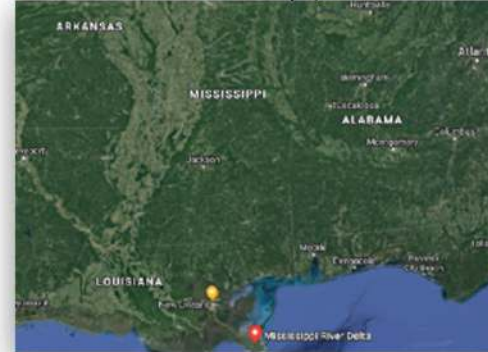
The foot became a standardized length—**a model derived from aggregating individual measurements**. A model is the same kind of thing: something simpler that we use in place of the real thing.

Models leave things out on purpose

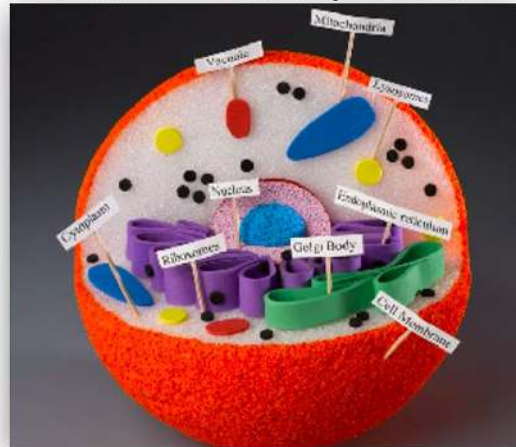
Mississippi River Basin Model



Actual Mississippi River Basin



Model of Eukaryotic Cell



Actual Image of Cell



On exactitude in science

... In that Empire, the Art of Cartography attained such Perfection that the map of a single Province occupied the entirety of a City, and the map of the Empire, the entirety of a Province. In time, those Unconscionable Maps no longer satisfied, and the Cartographers Guilds struck a Map of the Empire whose size was that of the Empire, and which coincided point for point with it.

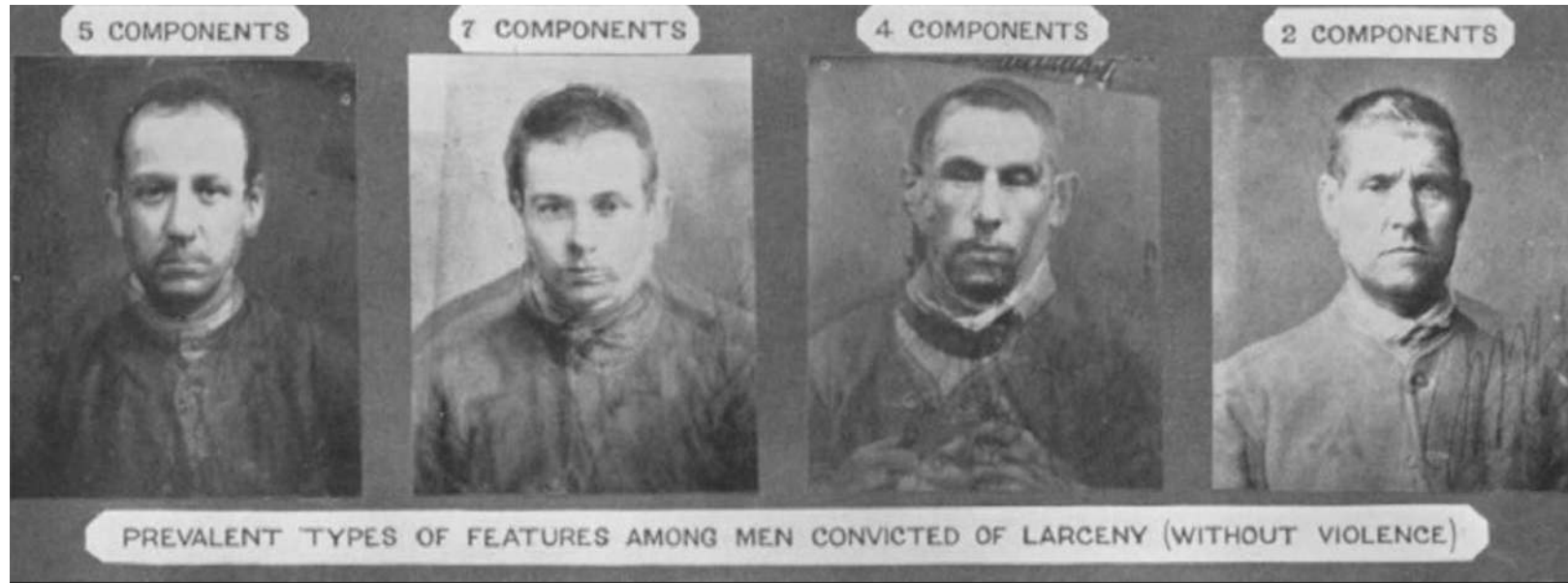
Jorge Luis Borges, "On Exactitude in Science" (1946)

With your neighbor

What are other examples of models you've encountered in your life?



What does an average face look like?



Here is a model made out of photographs. In 1878, Francis Galton exposed one photographic plate to several men's portraits, one after another, each for a fraction of the usual time. Each face left a faint image on the plate, and the result was an average of the faces: a composite portrait.

Composite portraits by Francis Galton, 1878-1883. The men were prisoners convicted of larceny.

What Galton expected, and what he got



- Galton was looking for the face of a “criminal type”. He expected the features the men had in common to show up dark in the composite, and the features particular to each man to come out faint.
- The composites came out ordinary. In his words, the individual peculiarities disappeared and the common humanity prevailed. The averaged faces were smoother and more regular than any of the faces that went into them, and there was no sign of a criminal type.

A composite is a model

- Each man's difference from the composite is his own error. A feature that only one man has, say a scar, shows up as a faint mark in the composite, because none of the other faces has anything there.
- Nobody looks exactly like the composite, and the composite doesn't look exactly like anybody. It is still the best single picture of the group.
- But it is only a picture of the group. Galton claimed more than that. He took the composites as evidence of an inherited criminal "type". The composite shows what the men's faces had in common. It says nothing about why they were in prison, and it doesn't make anyone's face the right face.

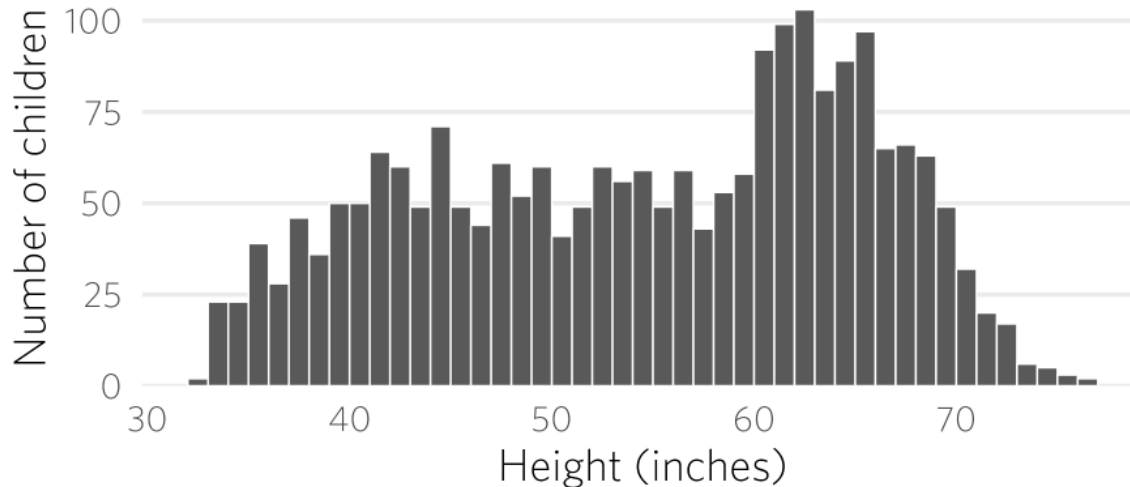
2 Modeling data with a single number

In 1535, one length was used in place of sixteen feet. Now we do the same thing with a column of data: we pick one number to use in place of all of the values. Which number is the best choice?

Predict a child's height

A child is picked at random from a national health survey. **What height would you predict, in inches?**

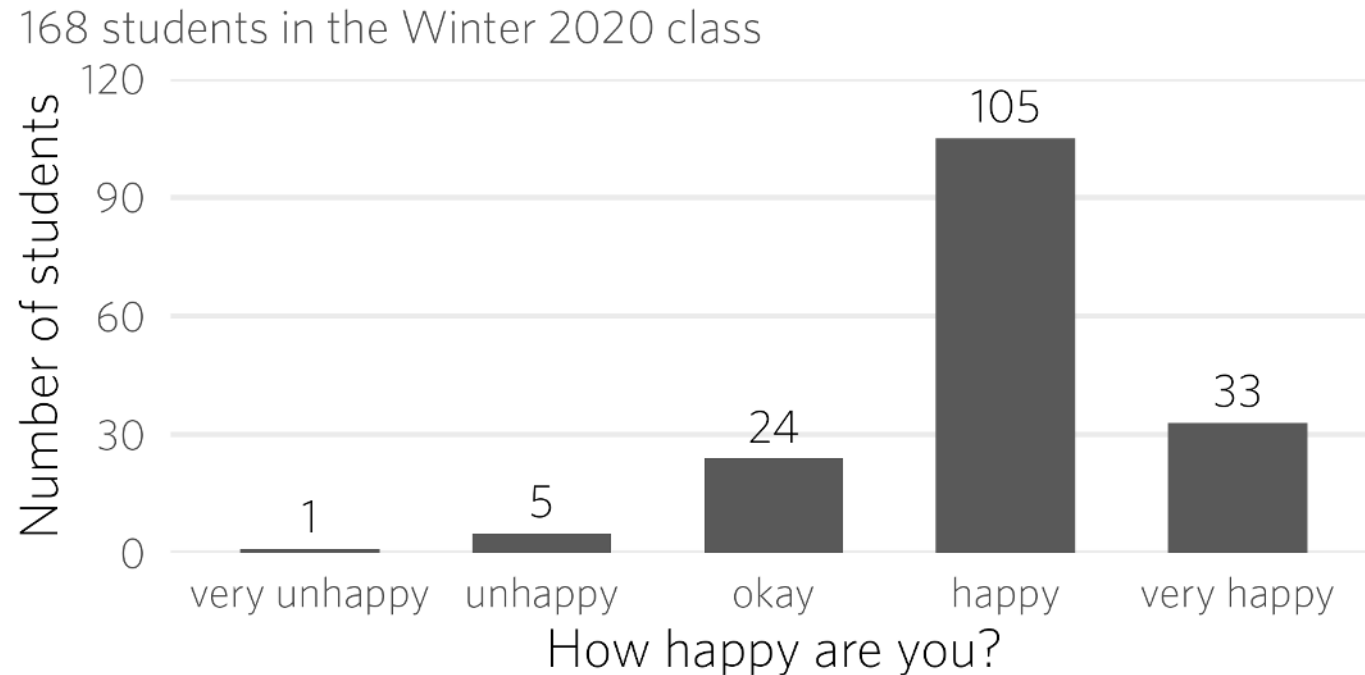
2,223 children, ages 2 to 17 (NHANES)



Which number should it be: the most common height, the median height, or the mean height? Let's try the most common one first.

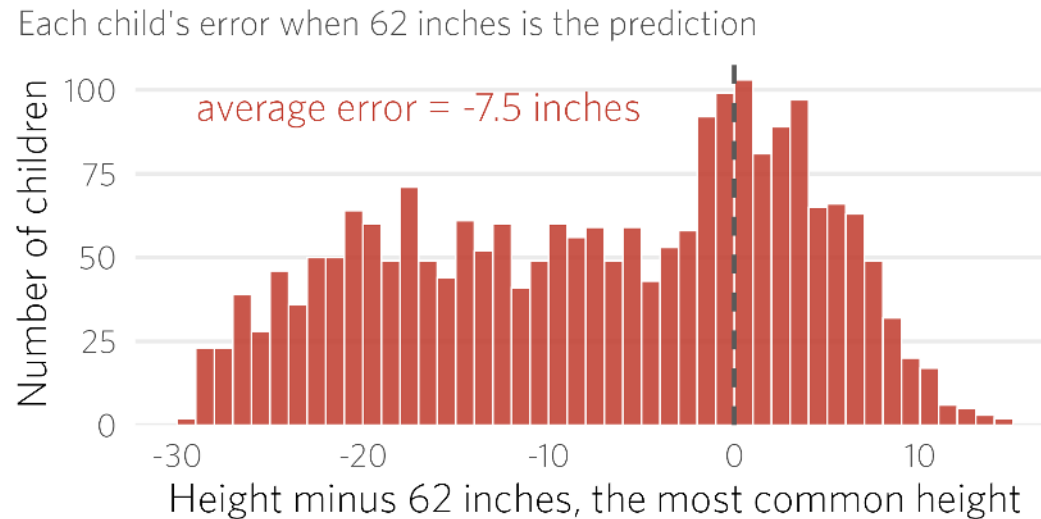
Predict a student's happiness

Which happiness level would you predict for a random student?



“Happy” is right for 105 of the 168 students, and no other single answer is right as often.

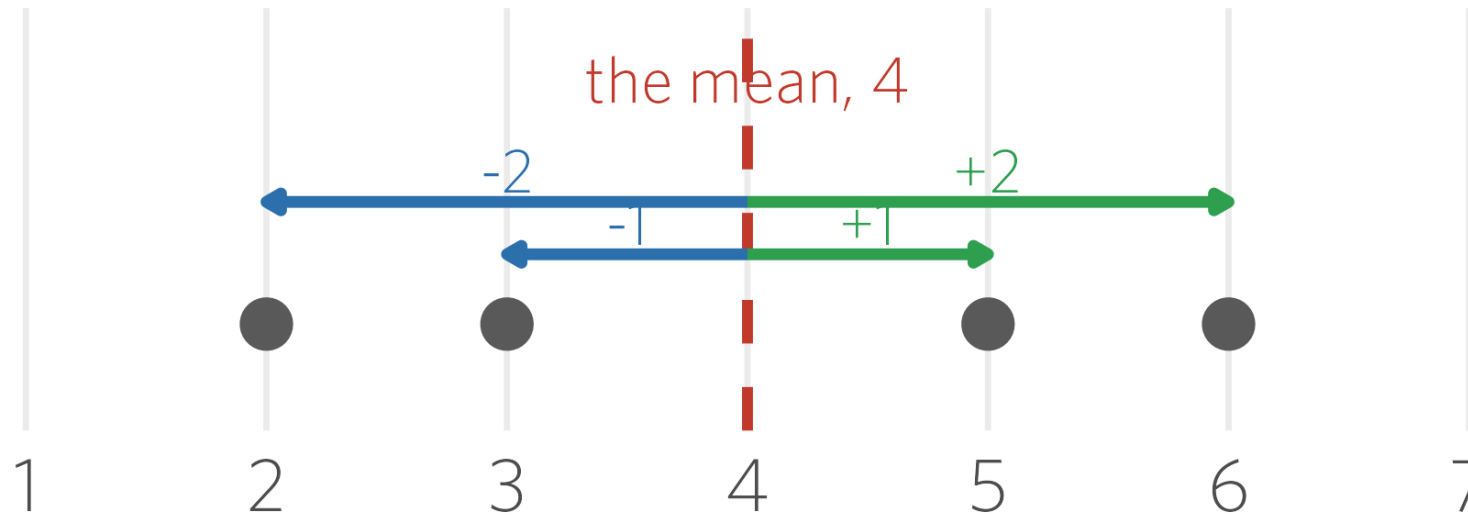
Using the most common height as the model



Suppose we predict 62 inches, the most common height, for every child. Each child's error is their height minus 62. The errors average -7.5 inches, so 62 is too high for most children. We want a number whose errors average out to zero. That number is the mean, which here is 54.5 inches.

The mean is the balancing point

Four values: 2, 3, 5, 6



Here is why the errors from the mean add up to zero, with four values: 2, 3, 5, and 6. The mean is 4. The two values below the mean are 2 and 1 below it, and the two values above it are 1 and 2 above it. The two sides cancel out, so the sum of the deviations is zero.

Does the average teacher earn £54,000?

Both the mean and the median get called “the average”. This summer in England, a dispute about teachers’ pay came down to exactly that difference.

- A government minister said that the average teacher earns about £54,000 a year.
- Teachers wrote in to say that nobody they knew earned anything like that.
- The minister’s number was a real number, computed from real pay records. **Can both sides be right?**

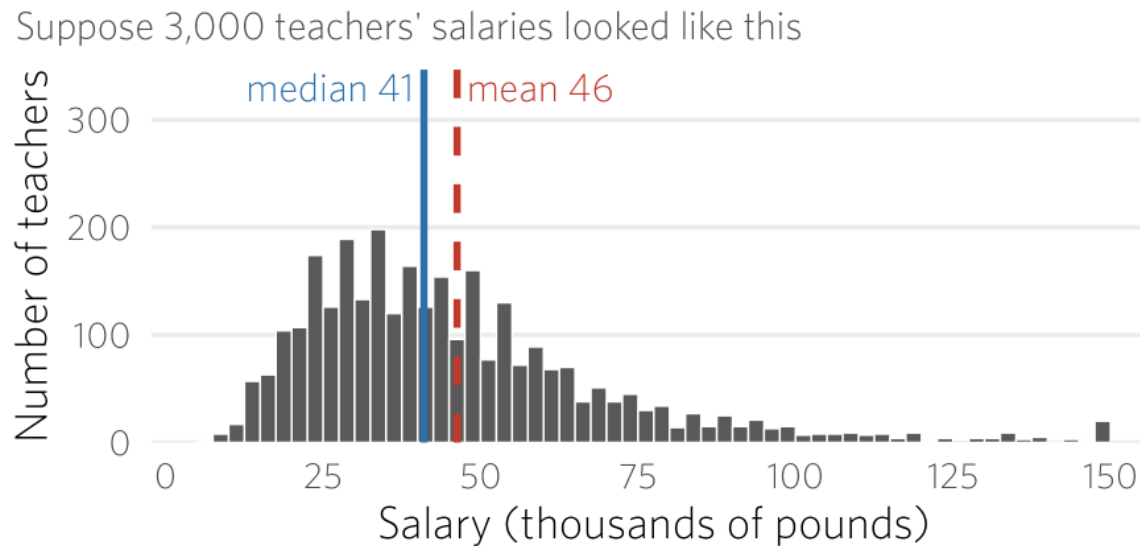
The story is from the BBC radio programme *More or Less*, 8 July 2026

With your neighbor

What shape would the pay distribution have to have for the minister and the teachers both to be right?



Most teachers earn less than the mean



Both sides were right. The £54,000 is the mean over everyone on the teachers' pay scales, including headteachers, whose large salaries raise the mean well above what a classroom teacher earns. The median is the salary in the middle of the list, and the size of the top salaries has no effect on it. Both numbers are "the average", and they answer different questions.

Worksheet - Part 1

Here are the salaries of the seven teachers at one school, in thousands of dollars: **32, 35, 38, 41, 44, 52, 94.**

- a.** Find the median and the mean. (The mean is a whole number.)
- b.** Suppose the teachers' union and the school's budget office each want to report one number as the typical salary at this school. Which number would each of them pick, and why?
- c.** For each teacher, compute the deviation from the mean (the salary minus the mean), keeping the sign. Then add up all seven deviations.

Worksheet - Part 1, continued

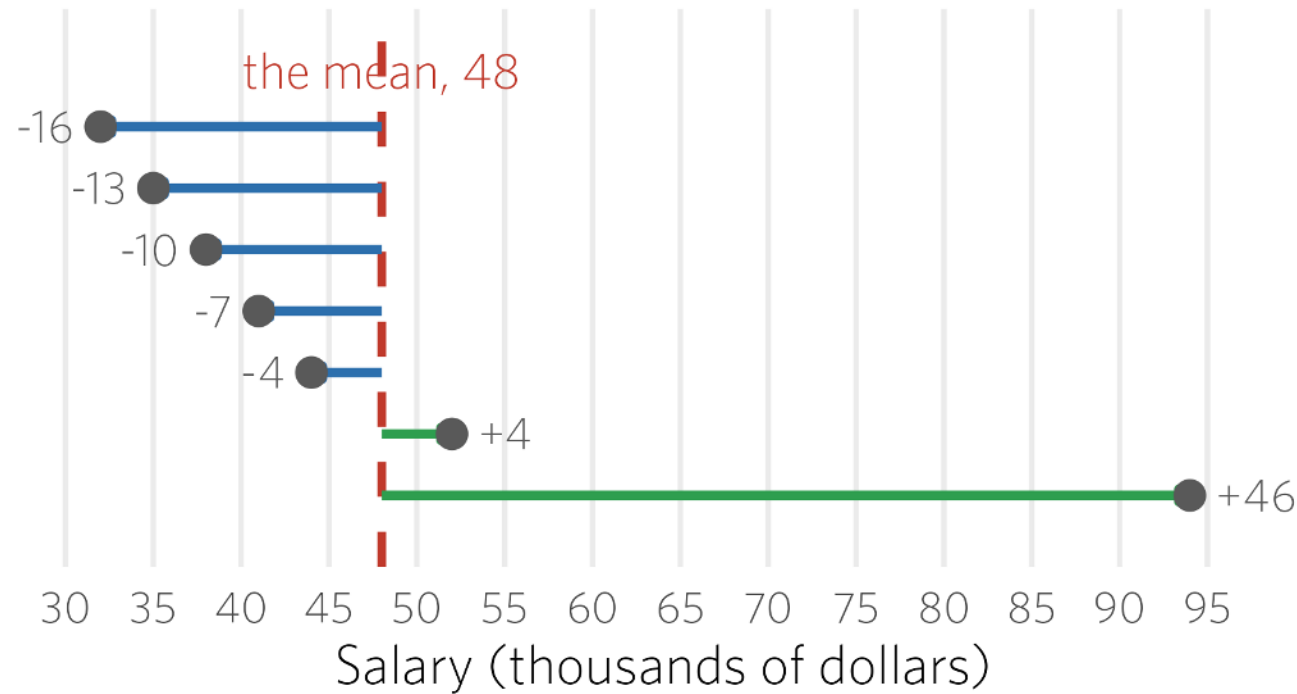
d. Now suppose someone proposed 45 as the model instead of the mean. Compute the seven deviations from 45 and add them up. How is this sum related to the distance between 45 and the mean? Use that relationship to predict what the sum of the deviations from 50 would be, without computing them.

e. (If you have time.) The teacher who earns 94 leaves, and a new teacher who earns 31 is hired. Write the seven new salaries in order, then find the new median and the new mean. Which one changed more, and why?

Median 41 or mean 48?

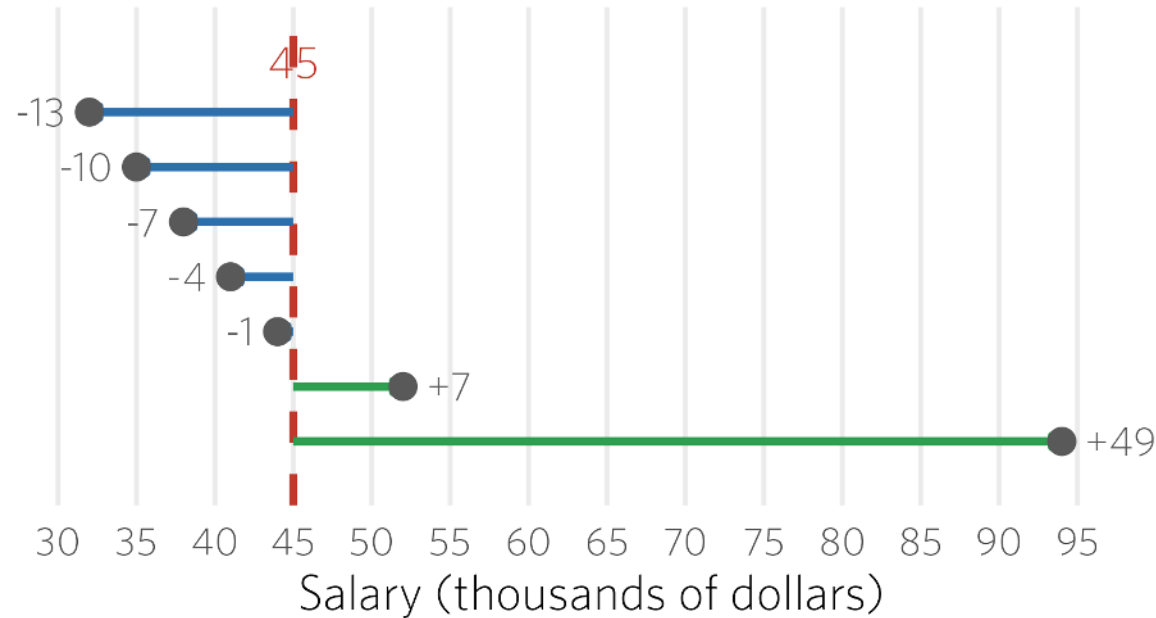
- Both numbers are correct. They summarize the same seven salaries in different ways.
- The median, 41, is the salary of the teacher in the middle of the list. That is the number a teachers' union would quote.
- The mean, 48, is the total pay bill divided by seven, which is what the school spends per teacher. That is the number a budget office would quote.
- Both get called "the average". Whenever you report one, name which one.

The deviations from the mean add to zero



The median is 41 and the mean is 48. The deviations from the mean are -16, -13, -10, -7, -4, +4, and +46. The five below the mean add up to -50, the two above add up to +50, and all seven add up to zero.

The deviations from 45 don't add to zero



The deviations from 45 add up to +21. Each of the seven is 3 larger than the deviation from the mean, because 45 is 3 below the mean, and 7 times 3 is 21. The deviations from 50 are each 2 smaller, so they add up to -14. The mean is the only number that makes the sum come out to zero.

When the 94 leaves and a 31 arrives

- The new salaries are 31, 32, 35, 38, 41, 44, and 52.
- The median moves from 41 to 38. When you replace one value, the median can only move to the next value over in the sorted list.
- The mean moves from 48 to 39. Each salary contributes one seventh of itself to the mean, so swapping 94 for 31 lowers the mean by $63/7$, which is 9.
- In this course, the mean is our single-number model. Every value contributes to it, and the deviations from it add up to zero.

3 Where does “average” make sense?

So far we have averaged one variable at a time: heights, then salaries. Now let's see what happens when we average several variables at once, and when we average things that aren't numbers.

Are you the average student?

Everyone stand up. Sit down as soon as one of these is not true of you.

- Your height is between 5 ft 5 in and 5 ft 8 in.
- You slept between 6 1/2 and 7 1/2 hours last night.
- You studied between 2 1/2 and 4 hours yesterday.

Designing a cockpit for the average pilot



In 1950 the US Air Force ran the same test. It measured 4,063 pilots on more than a hundred body dimensions, in order to design a cockpit that would fit the average pilot. On each dimension, a pilot counted as average if he fell in the middle 30%. **How many of the 4,063 pilots were average on all ten of the dimensions that matter most for a cockpit?**

Gilbert S. Daniels, "The 'Average Man'?", Wright Air Development Center, 1952. Photo: US Air Force, 1952

None of the 4,063

- Not one of the 4,063 pilots was in the middle 30% on all ten dimensions. Fewer than 3.5% were average on even three of them.
- A cockpit built for the average pilot fit nobody. The Air Force made the seats, pedals, and straps adjustable instead.
- The average arm length was a perfectly good summary of 4,063 arm lengths. But the “average pilot”, the one who is average in arm length and leg length and chest size and seven other things all at once, was a person who did not exist.

Worksheet - Part 2

Each of the statements below uses an average. Next to each one, write **A** if the average is a useful summary, **B** if it is a number you can compute but one that is easy to misread, or **C** if it is not meaningful.

1. ____ The average American household has 2.5 people.
2. ____ My roommate got home at 11 pm, midnight, and 1 am this week, so on average they got home at 8 am.
3. ____ The average zip code of the students in this class is 94305.
4. ____ A restaurant's average rating on Yelp is 4.2 stars.
5. ____ The average adult has fewer than two legs.
6. ____ The center of the US population, the average location of all 330 million residents, is in Missouri.
7. ____ You drove to a friend's place at 30 mph and back along the same road at 60 mph, so your average speed for the round trip was 45 mph.

Worksheet · Part 2, continued

- a.** Pick one of the statements you marked B or C. In one sentence, say what goes wrong and what you would report instead.
- b.** Give one example of your own of an average of something that isn't a number, and say what that average would be good for.

The seven, one by one

- **1, A.** No household has 2.5 people in it, but the number is still useful. It is the total number of people divided by the number of households, so a city that expects 10,000 new households can plan for 25,000 new people.
- **2, B.** The 8 am comes from counting hours from midnight: $(23 + 0 + 1)/3 = 8$. If you count from noon instead, the same three times average out to midnight. Clock times wrap around, so the mean depends on where you start the count.
- **3, C.** Zip codes are labels, not amounts. You can compute their mean, but it doesn't mean anything. Report the most common zip code instead.
- **4, B.** Stars are an ordered scale, not amounts. A 4.2 could be mostly 4s and 5s or a mix of 1s and 5s, and it could come from two reviews or from two thousand.

The seven, continued

- **5, B.** True, but misleading. A few people have fewer than two legs, so the mean is a little under 2, and nearly everyone is above average. The most common value, 2, is the number that describes the group.
- **6, A.** The mean of 330 million locations is itself a location. It is the point where a map of the country would balance if every resident were a small weight on it. The Census Bureau computes it after every census, and it has moved west and south since 1790.
- **7, B.** Average speed is total distance divided by total time. Say the trip is 30 miles each way. Going out at 30 mph takes an hour, coming back at 60 mph takes half an hour, and 60 miles in 1.5 hours is 40 mph. The slow half took twice as long, so it counts twice as much.

Let's hear some

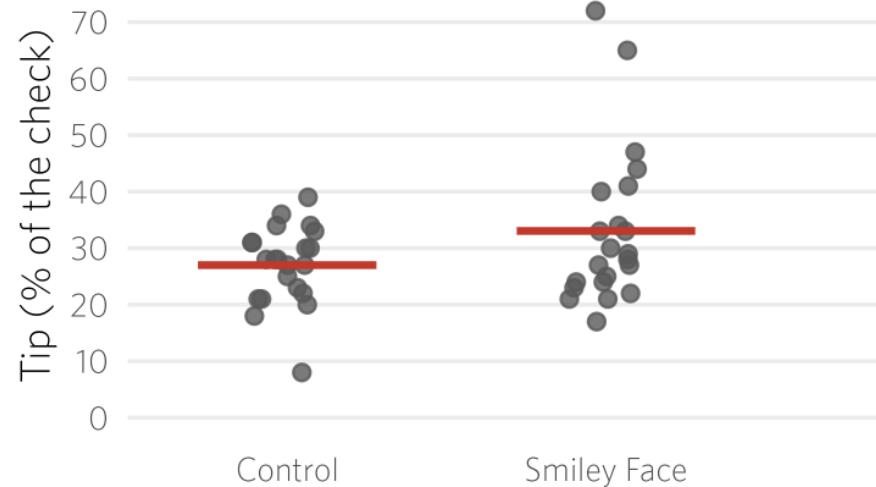
What examples did you come up with of an average of something that isn't a number? What is each one good for?

4 Could it be chance?

Every mean so far has summarized a single group. Now we compute a mean for each of two groups and ask whether the difference between them could have come from chance alone.

The tipping experiment

44 tables, randomly assigned: 22 got a smiley face on the check



Here is a model with two numbers instead of one, a mean for each group. In 1996, a server drew a smiley face on the checks of 22 tables, chosen at random out of 44, and left the other 22 checks plain. The smiley-face tables tipped a little more on average.

Rind and Bordia, *Journal of Applied Social Psychology*, 1996; the 44 tables CourseKata uses

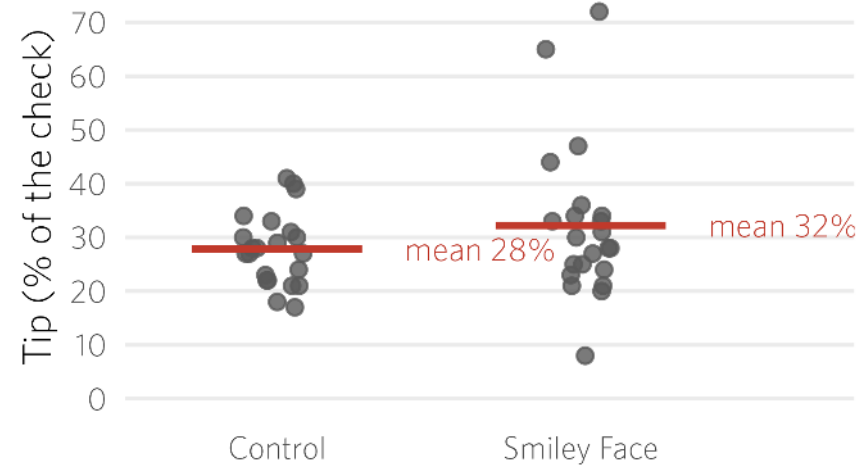
Two reasons the groups can differ

- The smiley face made people leave bigger tips.
- The random assignment. Which 22 tables got a smiley face was decided by the equivalent of a coin flip for each table, not by anything about the tables themselves. Some tables tip generously no matter what is on the check, and the coin flips may have put a few more of those tables in the smiley-face group.

Both of these can be true at the same time, even in a perfectly run experiment.

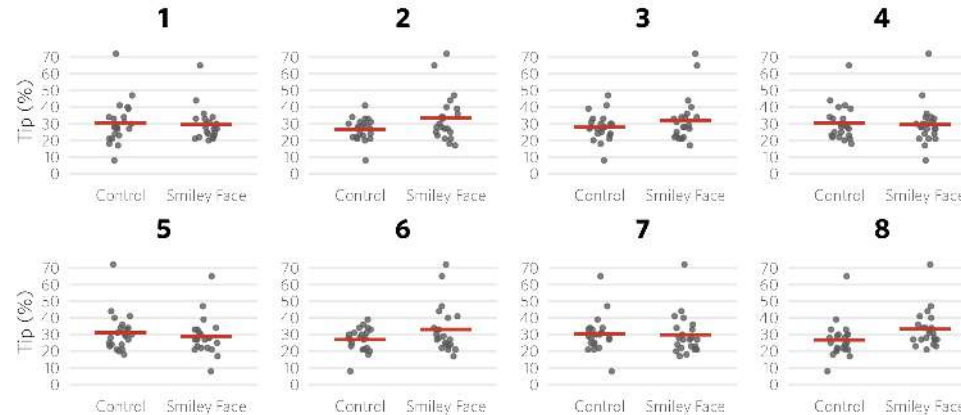
One shuffle of the tips

Shuffle 1: the same 44 tips, dealt out at random to the 44 ta



Keep the 44 tables in their two groups, but reassign the 44 tips to the tables at random. The two groups still differ a little, and this time the random reassignment is the only thing that could have produced the difference. One shuffle shows us one possible outcome. To see the whole range of outcomes, we need many shuffles.

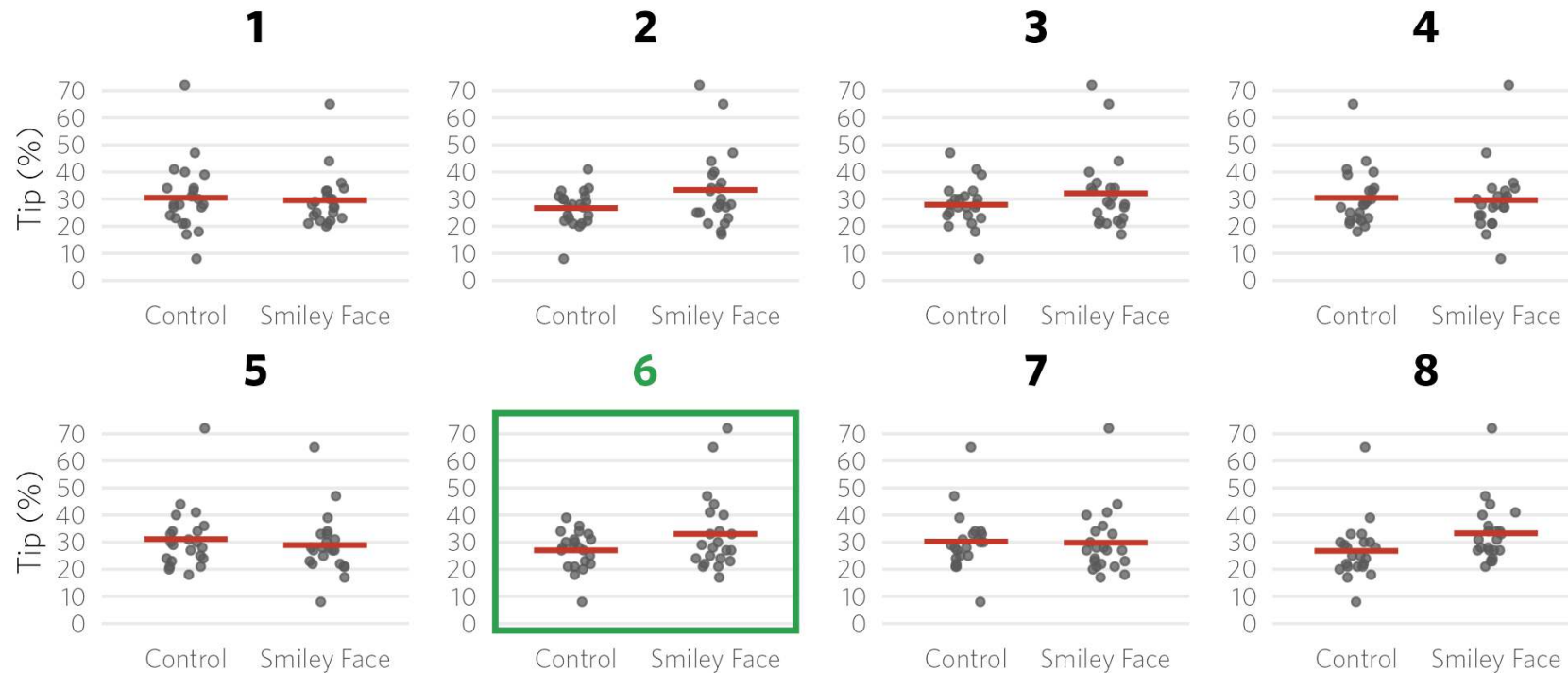
Worksheet - Part 3



One of these eight plots shows the real data from the tipping experiment: 44 restaurant tables, 22 of which got a check with a smiley face drawn on it and 22 of which (labeled Control) got a plain check. The thick bar in each group is the group's mean tip. In the other seven plots, the same 44 tips were reassigned to the two groups at random.

- Circle the number of the plot you think is the real one.
- In one sentence, explain what made you pick that plot.

The real one is panel 6



In panel 6, the real data, the smiley-face tables tipped 33% on average and the control tables tipped 27%, a difference of 6 points. In panels 2 and 8, the shuffled difference is larger than that, and in panel 3 it is nearly as large.

What the lineup shows

- The real difference was 6 points. Two of the seven shuffles produced a larger difference than that, with no smiley face involved at all. So these 44 tables can't tell us whether the smiley face raised tips. A difference this size can come from the random assignment alone.
- If the real groups had differed by 15 points, and no shuffle had come anywhere near that, we would have had evidence that the smiley face did something.
- If we announced that smiley faces raise tips based on these 44 tables, that would be a false alarm, because the coin flips alone produce differences of 6 points. Shuffling is how we check before we announce anything.

Before Friday

- Finish CourseKata **Chapters 4 and 5**.
- Friday: **The Vault of Fallen Labels**. Check your discussion section's Slack channel for your time slot.
- Milestone 2 is due Sunday.

Hand in your worksheet on the way out

On the last line, write down one thing you learned today.

Thanks!