

# CourseKata Video Transcript

## Video Details

**Video Title:** How to Tell if One Variable "Explains Variation" in Another Variable

**Video Link:** <https://player.vimeo.com/video/379319375>

## Video Transcript

**Student (off screen)**

Dr. Ji, what about these examples? What do they tell us about explaining variation?

**Dr. Ji**

Good question!

So let's put up our data of thumb length by year in school.

[SPEAKER GESTURES TO FACETED HISTOGRAM OF THUMB BY YEAR IN SCHOOL]

So this is just freshman thumb length, sophomore thumb length, juniors, and so on. And so let's apply that definition of explaining variation to year in school.

[GESTURES TO DEFINITION OF EXPLAINING VARIATION THAT APPEARS ON SCREEN. THE DEFINITION IS AS FOLLOWS: "EXPLAINING VARIATION: IF KNOWING SOMEONE'S \_\_\_\_ HELPS US MAKE A BETTER PREDICTION OF THEIR THUMB."]

So that would be if knowing someone's year in school helps us make a better prediction of their thumb length. Is that true? Let's look at our data. Well, here's our freshmen. I'm gonna represent it with this hand.

[HOLDS ONE PALM OVER THE DISTRIBUTION OF FRESHMEN]

And here's our sophomores, right?

[HOLDS OTHER PALM OVER THE DISTRIBUTION OF SOPHOMORES]

And so they kind of look like this, right?

[SHOWS WITH HANDS HOW THE DISTRIBUTIONS ARE ALIGNED WITH ONE ANOTHER]

And then the juniors are just right under it, right?

[USES HAND TO SHOW THAT THE DISTRIBUTION OF JUNIORS OVERLAPS SIMILARLY TO THE OTHER GROUPS]

Contrast that to sex. And here, this is thumb length by sex.

[GESTURES TO FACETED HISTOGRAM OF THUMB BY SEX]

And notice here, we saw that shift in the distributions.

[USES HANDS TO SHOW THE SHIFT IN THE OVERLAP OF THE TWO DISTRIBUTIONS]

It wasn't that every single female had a smaller thumb. But in general, the distribution is kind of shifted downward, and the male distribution is shifted upward, right? So there's this, like, kind of separation. Whereas here, there's not that shift.

[GESTURES BACK TO DISTRIBUTION OF THUMB BY YEAR IN SCHOOL AND USES HANDS TO SHOW THAT THE DISTRIBUTION OF EACH GROUP IS ROUGHLY ALIGNED]

We don't see any shifts. They are instead stacked right on top of each other, right? So here

[REFERENCES THE DISTRIBUTION OF THUMB BY YEAR IN SCHOOL]

even if we knew that they were a sophomore, we knew that they were a junior, we wouldn't change our prediction about their thumb length very much. Whereas here we would change our prediction of their thumb length by a little bit. And so when we're looking for this idea of explaining variation in these different faceted histograms, what you want to look for is that shift.

[USES HANDS TO SHOW HORIZONTAL SHIFT OF OVERLAPPING DISTRIBUTIONS]

When there is a shift, we might suspect, maybe this variable helps us make a better prediction of their thumb length. But when you don't see a shift, and they're all basically stacked on top of each other,

[USES HANDS TO SHOW TWO DISTRIBUTIONS THAT ARE ALIGNED]

we might say, "even if I knew that a little bit about year in school, that wouldn't help me make a better prediction of their thumb length".