On The Average II:
Weighted Averages and The Unit of Analysis

or:

If

A Chicken and a Half
Lays an Egg and a Half
In a Day and a Half,

Then

How Many Eggs Does a Chicken
Lay in a Day?

That title was once a common riddle among kids, designed to sound silly, confuse the mind, and otherwise pass the time. It seems to have fallen out of the common "kids culture", which is no great loss, but its solution actually makes a point: If you can get through things like that with your mind still intact, then maybe you can get through real-world problems, deal with data in their naturally bizarre state of arrangement, and still keep your mind intact.

I actually intend to answer the question stated by the title, but I'm going to have to work up to it. Once again the moral of the story is going to be — be careful of the units and, as far as possible, build the units into your equation.

So far, for the mean, I've been using the simple equation

$$\bar{x} = \frac{x_1 + x_2 + \quad + x_n}{n}$$

abbreviated

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

where “x-bar” is the mean, and where x_1, x_2, and all the little x_i’s up to x_n are the n values whose average is being computed.

Properly amended, building-in the unit, that becomes

$$\bar{x}\,\text{units} = \frac{x_1\,\text{units} + x_2\,\text{units} + \quad + x_n\,\text{units}}{n}$$

or

$$\bar{x}\,\text{units} = \frac{1}{n}\sum_{i=1}^{n} x_i\,\text{units}$$

And now the equation preserves the idea that whatever units are used for “x”, the thing being averaged, the average value of x is measured in the same units. If the units are people, the mean is in people. If the unit is dollars, then the mean is in dollars.

$$\bar{x}\,\text{people} = \frac{1}{n}\sum_{i=1}^{n} x_i\,\text{people}$$

$$\$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} \$x_i$$

Now I need something called the “weighted average”. The weighted introduces another set of numbers, the weights, and associates each of the x’s with a weight. Then, for the weighted average you computed the weighted sum and divide by the sum of the weights.

$$\overline{x} = \frac{x_1 w_1 + x_2 w_2 + \ldots + x_n w_n}{w_1 + w_2 + \ldots + w_n}$$

One use of weights is to keep count: If I throw a pair of dice, and they come up 3, 4, 10, 4, and 3, in successive throws of the dice, then I can compute the average giving each toss of the dice a weight of one:

$$\overline{x} = \frac{3*1+4*1+10*1+4*1+3*1}{5}$$

which, is, of course, identical to adding up the successive values of the dice and dividing by 5. or I can use weights to keep track of the number of 3's and 4's and 10's that came up, giving 3 and 4 a weight of two as compared to a weight of one for the 10 (because 3 and 4 each came up twice):

$$\overline{x} = \frac{3*2+4*2+10*1}{2+2+1}$$

It adds a little bit to the formula for the average, but ultimately it is more compact: I can throw the dice 1,000 times, but the average for dice is always going to be 2 times a weight, plus 3 times a weight, plus 4 times a weight, and so forth, divided by the sum of the weights:

$$\overline{x} = \frac{\sum_{i=2}^{12} i\, w_i}{\sum_{i=2}^{12} w_i}$$

That's the weighted average — as arithmetic. But I want you to use the weighted average as a kind of debater's trick — I'm not trying to debate with you or trick you: The debate is the one I carry on in my head, debating with myself to see if I know what I'm doing — tricking myself with devices that get me to see a problem from unfamiliar angles.

Now back to physicians per capita. I think / hope you understand it. Now let me try to use close attention to the units, and use the weighted average to straighten the whole thing out. But one warning: Even if this looks automatic, a "method", a procedure — use it right and it will pay you back with the right answer — even if it looks that way don't believe it. I'm using this as a debater's trick, Levine v/s Levine: The formulas do not know the answer.

So now, quickly, back to doctors and the world

Unit: Country	Attribute: Number of Doctors, 1975	Attribute: Total Population, 1975	Attribute: Doctors Per Person	Attribute: Doctors Per 1,000 People
Afghanistan	656	19,280,000	0.000034	0.034
Angola	384	6,394,000	0.000060	0.060
Argentina	48,687	25,384,000	0.001918	1.918
Austria	15,702	7,538,000	0.002083	2.083
Bahamas	161	200,000	0.000805	0.805
Bahrain	177	260,000	0.000681	0.681
Bangladesh	5,088	73,746,000	0.000069	0.069
Barbados	166	245,000	0.000678	0.678
Belgium	18,510	9,846,000	0.001880	1.880
Benin	95	3,074,000	0.000031	0.031

Now, what's the question and what's the answer: If the question is a question about a typical country or nation, conceived as a unit, with its population, with a national medical system, with a national medical education system, a national public health system, hospital system and payments system. If the question is about the results for countries and for the typical county — then you want the country as the unit of analysis and you want to average the data with respect to the country. What have we got in the table? We have a list of units, the countries, and a list of attributes in each row that describe that unit. So, using a "sample" of one country, using the ".000034" (in row one) — what is the unit? It is "Doctors per person *per country*" This datum provides evidence for one country, Afghanistan. That last "per ____", at the bottom of the stack names the unit. And the unit in the weight matches the unit of analysis. When I take the weighted average it gets the weight "1 country". In the second row I have attributes of another unit, another country, Angola. The .000060 in the second row is "Doctors per person per country" (for this second country). Now, using the unit of analysis as the unit of weight,

$$\bar{x} = \frac{\sum_{i=2}^{12} i\, w_i}{\sum_{i=2}^{12} w_i}$$

$$\bar{x} = \frac{\left[0.000034 \frac{doctors}{person} \Big/ country\right] * 1 country + \left[0.000060 \frac{doctors}{person} \Big/ country\right] * 1 country + \left[0.001918 \frac{doctors}{person} \Big/ country\right] * 1 country + \ldots}{138\, countries}$$

I get the average of _____ $\left(\frac{doctors}{person} \Big/ country\right)$ And what does this thing,. this average, describe: Look at the units on the average, and there it is, at the bottom of the stack: the country, the average country.

Now, by contrast, suppose I had done it the other way — adding up the numbers in the column for doctors: At the bottom of the data sheet I show a sum of 2,622,088 *doctors*, two-point-six million doctors. And, this way, what is the unit? What do these doctors belong to? What unit has two-point-six million doctors? The unit is the world itself (or the reporting nations). In the next column the table shows 3,028,196,000 people. And what do these people belong to? The same world. And their ratio describes the world: In the world there is a ratio of .000865891 doctors *per person*.

And now, I have to deal with a "chicken and a half" So, what is my unit? What does an egg and a half in a day and a half describe? An egg and a half in a day in a half is an attribute describing 1.5 chickens. — my unit is the chicken, (although I'm given data about a chicken and a half). So the attribute that describes the chicken and a half is

$$\frac{1.5\ eggs}{1.5\ day\ s}$$

And, putting the rate in the proper relation to the unit it describes, as the rate for a chicken and a half: [1]here is the "formula" in eggs per day per chicken:

$$\frac{1.5\ eggs}{1.5\ days}\Big/ 1.5\ ch\ ickens$$

The units and the unit of analysis are clear: eggs per day *per chicken*. So I just clean up the pesky fractions and get one egg per day for 1.5 chickens:

$$1\frac{egg}{day}\Big/ 1.5\ ch\ ickens$$

or two-thirds of an egg per day per chicken

$$\frac{2}{3}\left(\frac{eggs}{day}\Big/ ch\ icken\right)$$

That is: One chicken, the unit, lays two-thirds of an egg per day. Watch the unit of analysis, watch the chicken, not the egg (or the day).

1 Exercise for the reader: How many days does it take for a chicken to lay an egg? (Solve for x, whose unit will turn out to be a number of days, in the equation $\frac{2}{3}\left(\frac{eggs}{day}\Big/ ch\ icken\right) * x = 1\ \frac{egg}{chicken}$