THE UNIVERSITY OF AUCKLAND
NEW ZEALAND
Te Whare Wānanga o Tāmaki Makaurau

WEEK 4

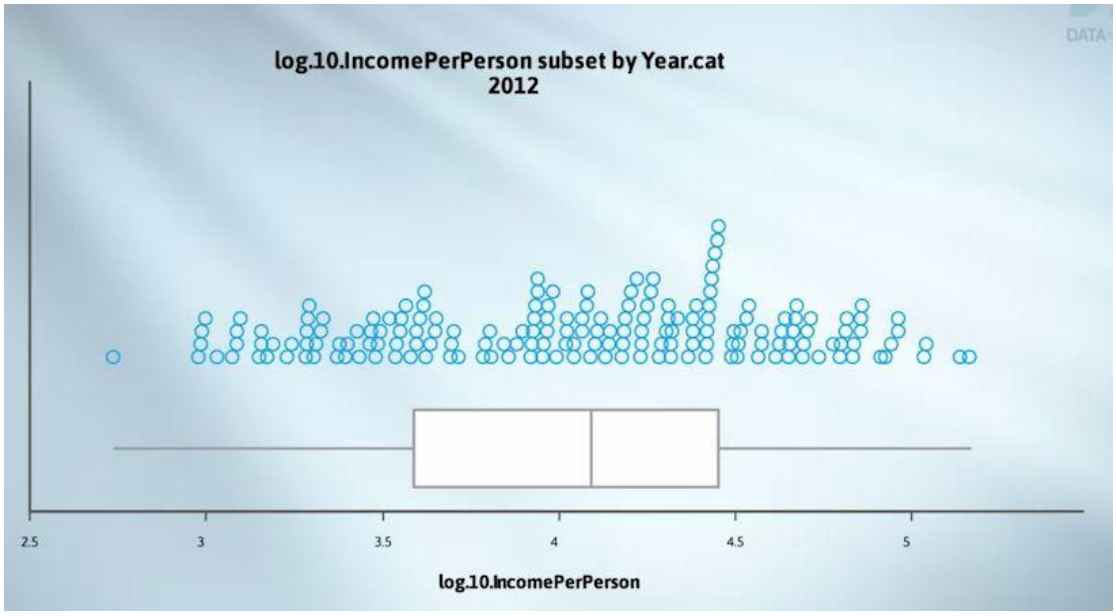
4.10 OUR CHANGING HEALTH AND WEALTH by Chris Wild

Hi. In this video, we're going to explore how the health and wealth of nations has changed over time. It'll show tools such as colour, size, and subsetting being used profitably together.

Our data is country-level data from Gapminder. Hans Rosling has explored these issues in some of his videos. We'll give the links later.

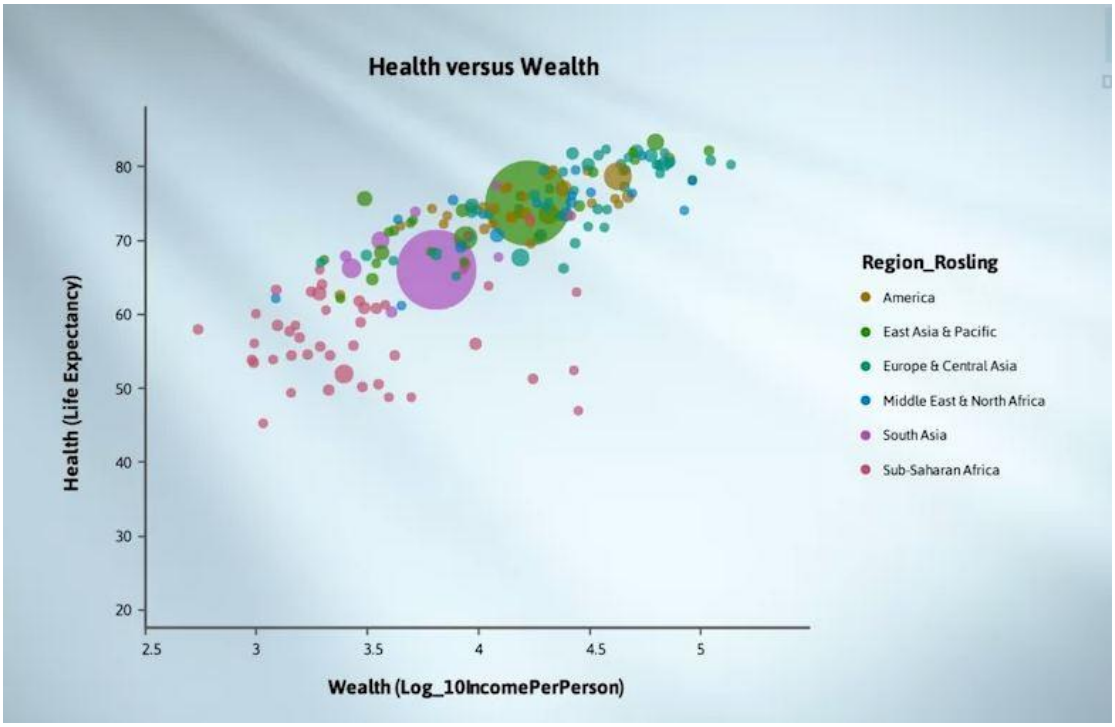
The wealth measure he uses is a little more complicated than income per person. IncomePerPerson is very skewed, with everyone piled up against zero in a way that makes little distinction between income levels at the low end.

Rosling plots IncomePerPerson on a log scale, which is equivalent to this.



For those who know nothing about logs, it might be helpful to know that moving to the right by a full unit on the scale of this graph is equivalent to a ten-fold increase

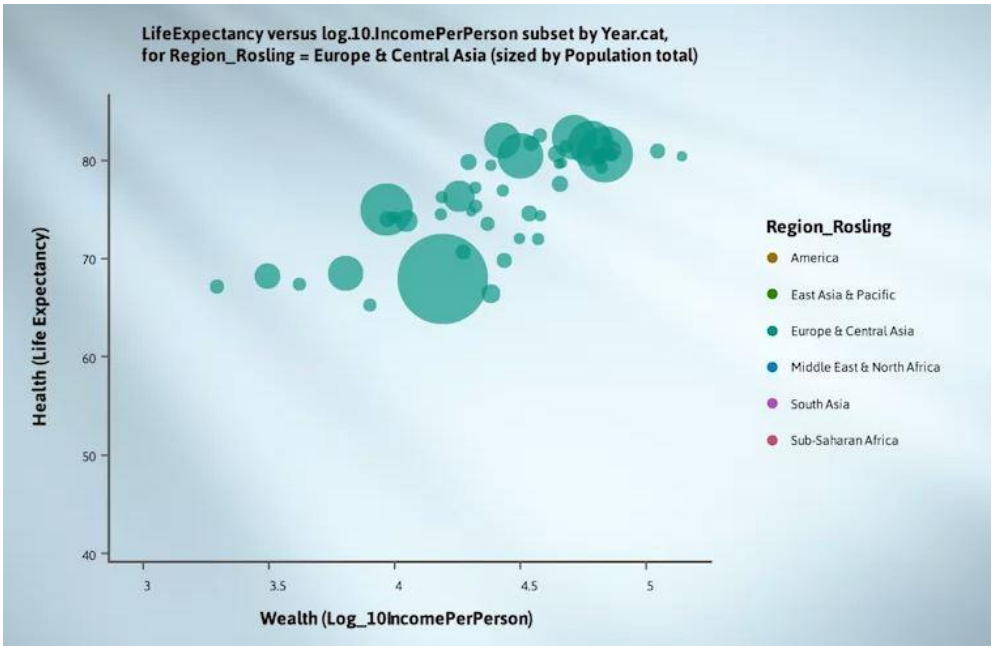
in income. A half-unit move to the right corresponds to a little over a three-fold increase in income. But we will not be interpreting the details of any of this, just spotting some general patterns.



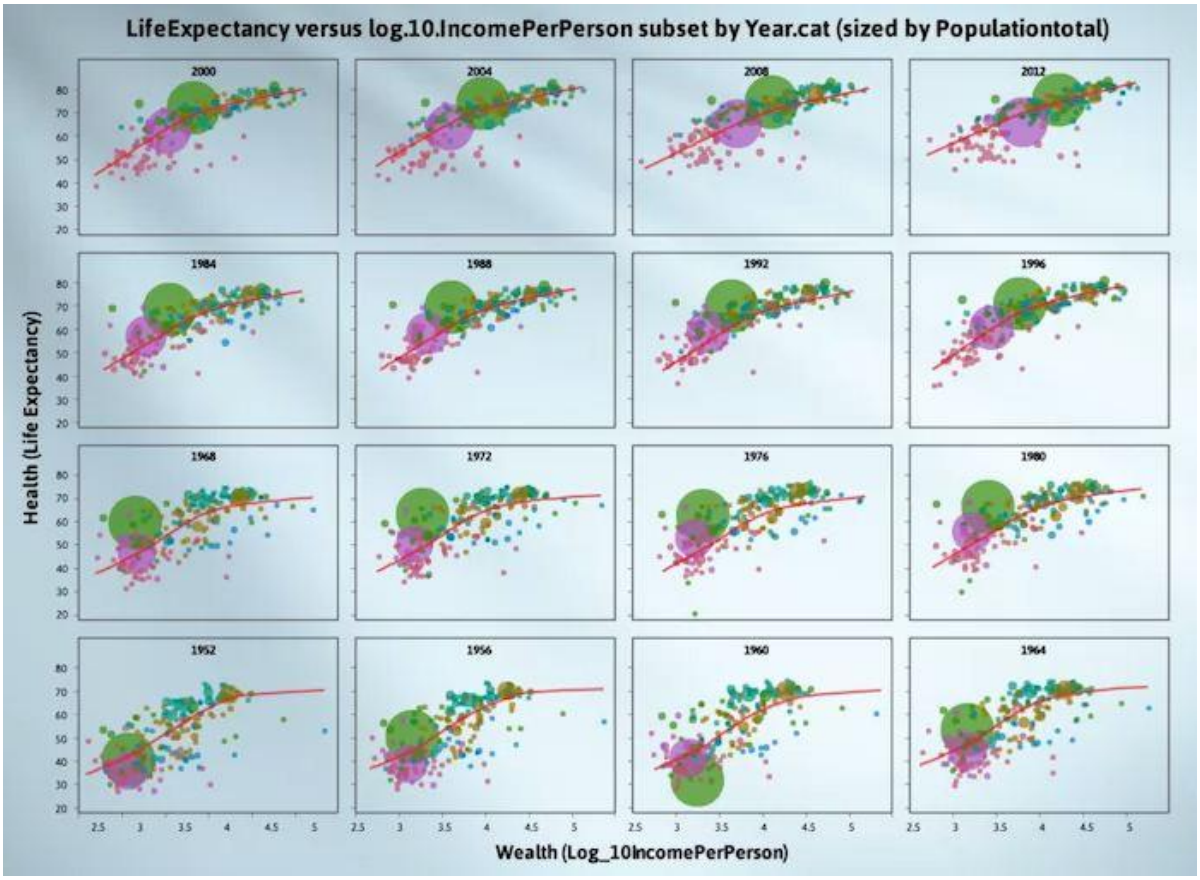
Here's our health versus wealth relationship for 2012. The points here represent countries. I've coloured the points by the region in the world they come from and sized them by how large their population is.

This is doing two things. First, it's conveying information on the relative population sizes. Second, it makes the large countries, which represent the experiences of many more people, much more prominent than the tiny countries. China (in green) and India (in a purplish shade) stand out by being by far the biggest.
The third biggest ball (in brown) belongs to the United States. The one region that stands out from all others is sub-Saharan Africa, whose points tend to fall lower (less healthy) and further to the left (poorer) than all the other regions.

But that's not the most interesting story in the data. The most interesting is what's been happening over time. With this data, we're going to subset on a categorical version of a year and on region. This enables me to see any particular region in any particular year.



Here we are looking at Europe and Central Asia in 2012. The largest discs for this grouping are, in decreasing order, the Russian Federation, Germany, and Turkey.

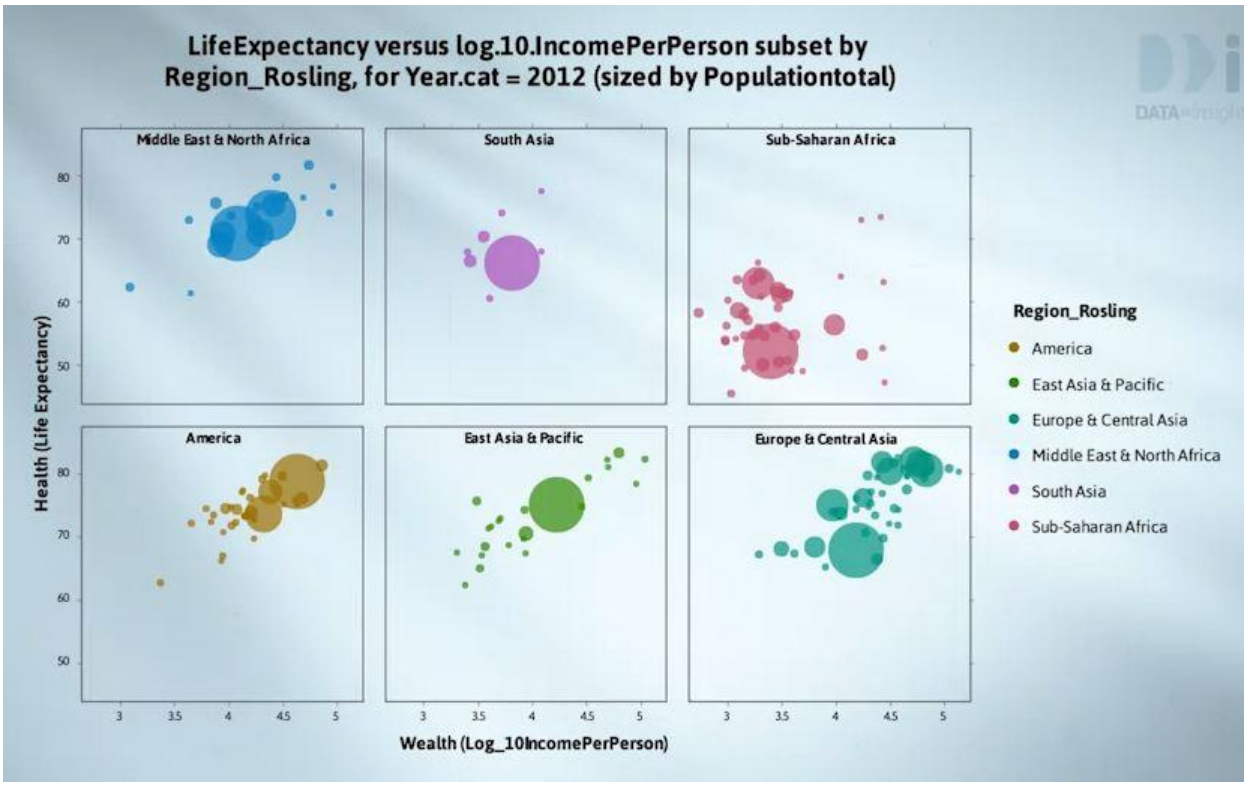


Or I can look at all regions together over time. You'll see that I've added a smoother as a unifying feature to make it easier to see what's happening to the whole point cloud over time.

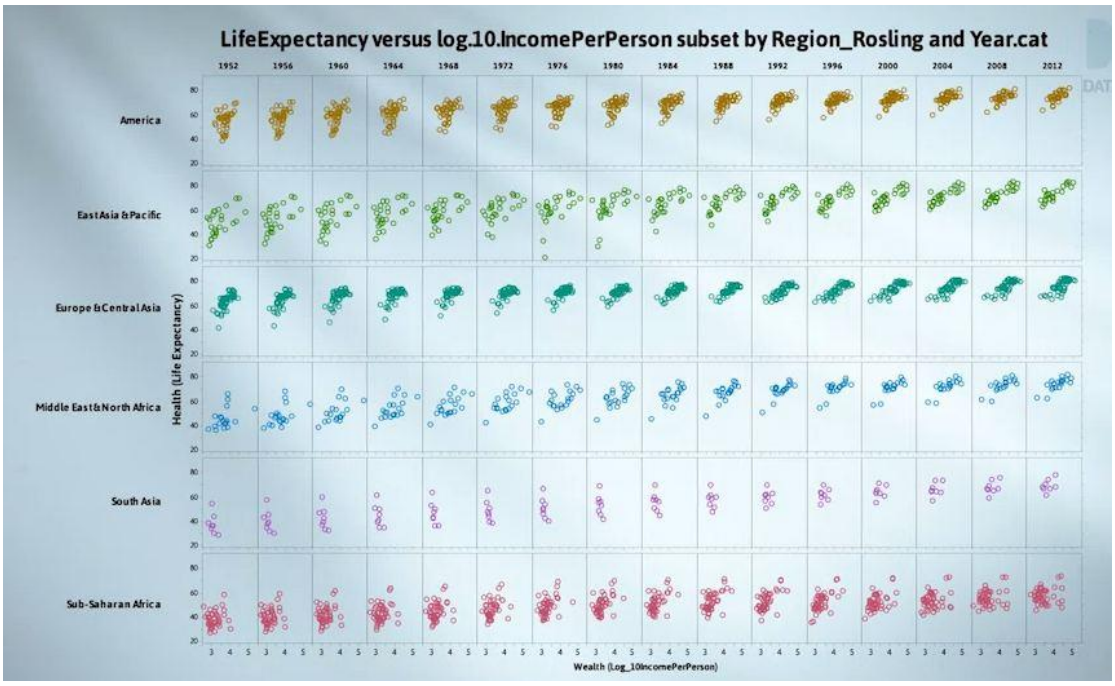
I find what's happening here easier to see if I play it through time rather than looking at a grid of plots. One thing we can see, the biggest story, is a general movement of the whole point cloud over time from the bottom left(low wealth and health) towards the upper right(increasing wealth and health).

Let's look at that again. Looking closely, we can see the same thing in the grid of plots. To go forward in time, you should read by row starting from 1952 in the bottom left. A secondary story we can quite easily see is the progress of China and India, the two biggest discs, from near the worst-performing countries in 1952 to the middle of the pack for India by 2012 (top right) and considerably higher for China. A third story is the improvement in relative position of the Middle East and North Africa over time.

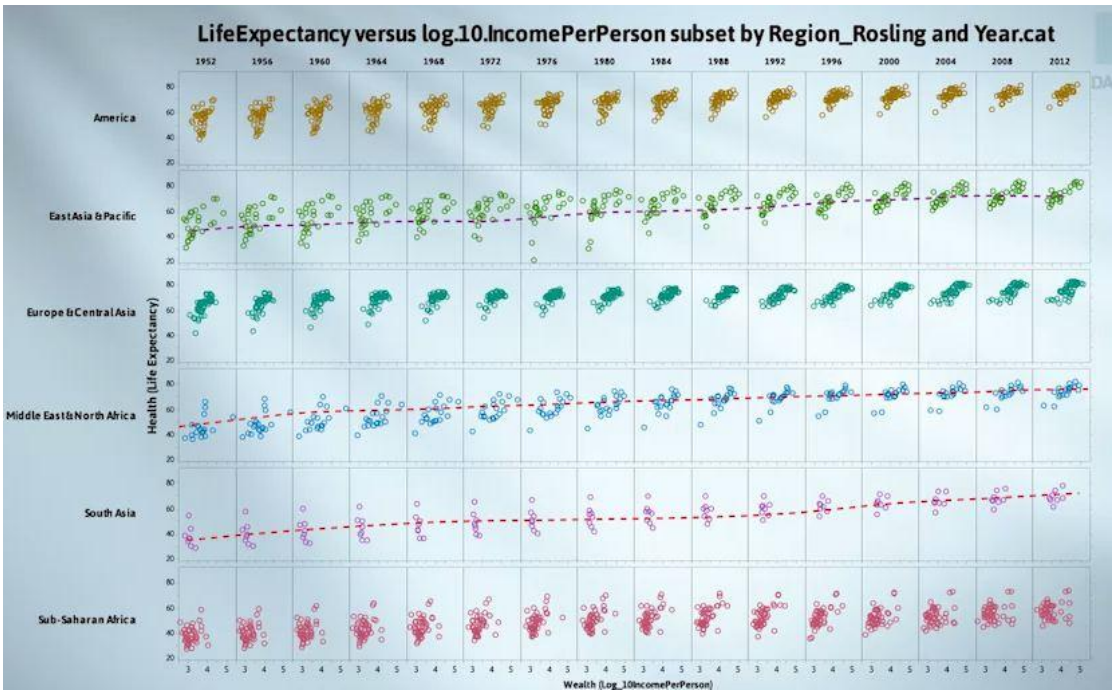
Using subsetting, we can also see all regions separately for a particular year (again emphasising the considerable worst position of sub-Saharan Africa in 2012) …



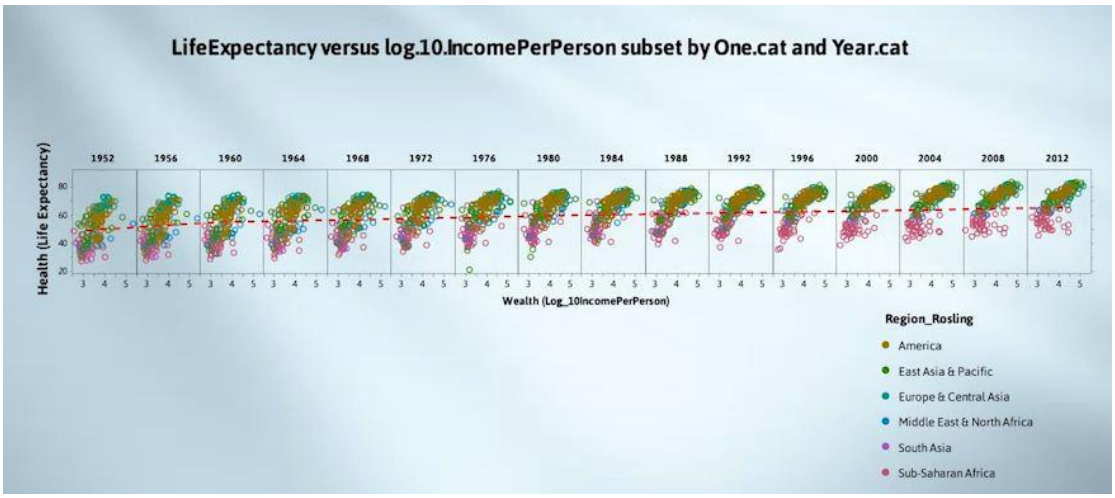
… or all regions (as the rows) by all years (as the columns).



Scanning across a row shows us what's been happening in a given region over time. This is pushing our ability to see things on all but large high-resolution screens. But it's still possible to see the progression over time, with point clouds moving slowly upwards and towards the right.



This is particularly evident for the regions highlighted here.



Here are all the regions together over time as a single row.

There are a lot of questions behind what we're seeing here that would worry someone seriously interested in issues addressed by these graphs — worries about how reliable the data is, worries about how well indexing of the currencies and use of exchange rates so it affects real in-country experiences of income levels, and so on.

You can find a lot of material about issues like this on the Gapminder site. But the intricacies of health development are not our mission. The mission of this course is simply to convey how various tools can be used to discover and display patterns in data.