FUNDAMENTAL PRINCIPLES OF DATA VIZ

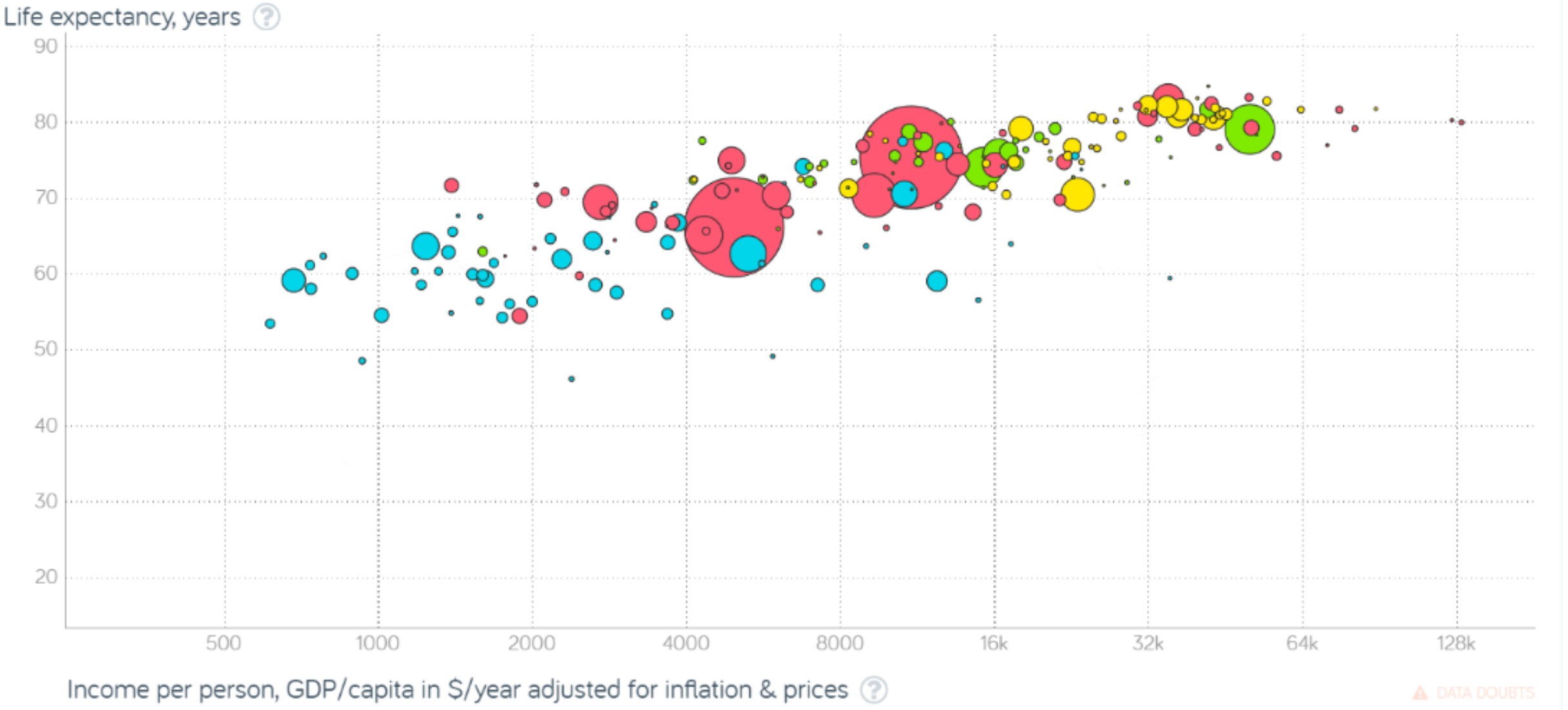
[E. Tufte, Beautiful Evidence]

FUNDAMENTAL PRINCIPLES OF ANALYTICAL DESIGN

Symmetry to visual displays of evidence: consumers should be seeking exactly what producers should be providing, namely

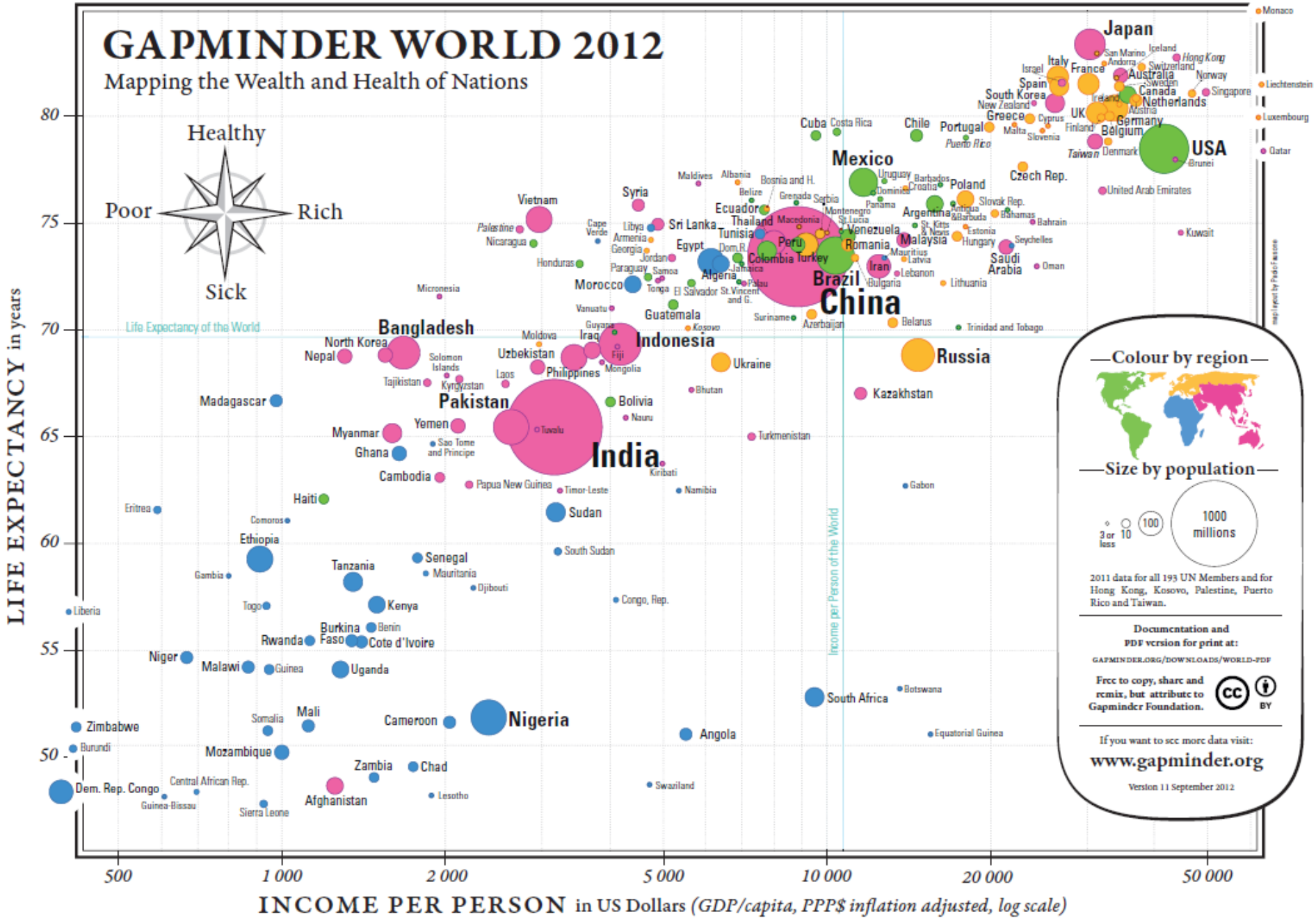
- meaningful comparisons
- causal networks and underlying structure
- multivariate links
- integrated and relevant data
- honest documentation
- primary focus on content

[Gapminder Foundation]

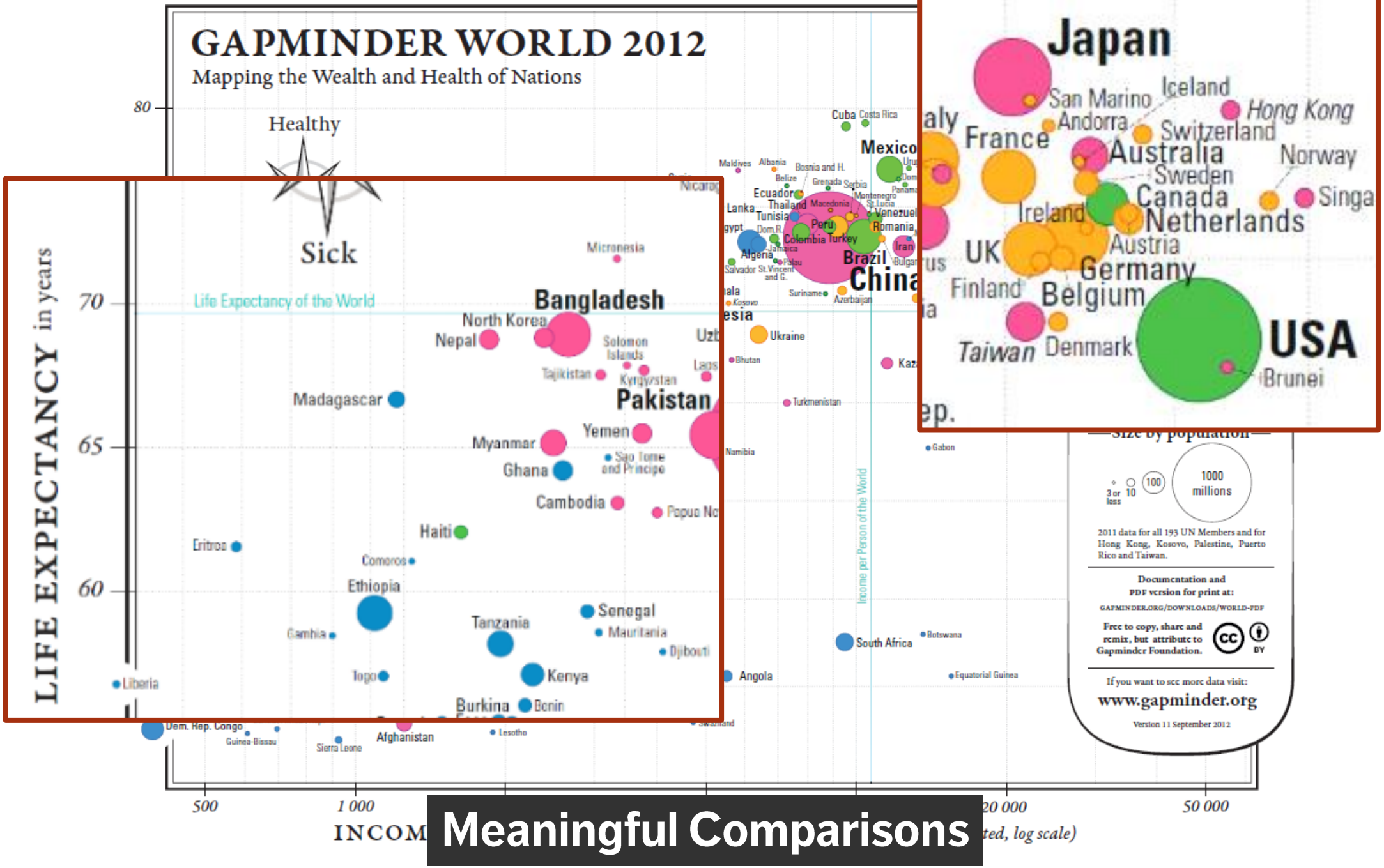


Non-Integrated Data

[Gapminder Foundation, https://www.ted.com/talks/hans_rosling_shows_the_best_stats_you_ve_ever_seen]

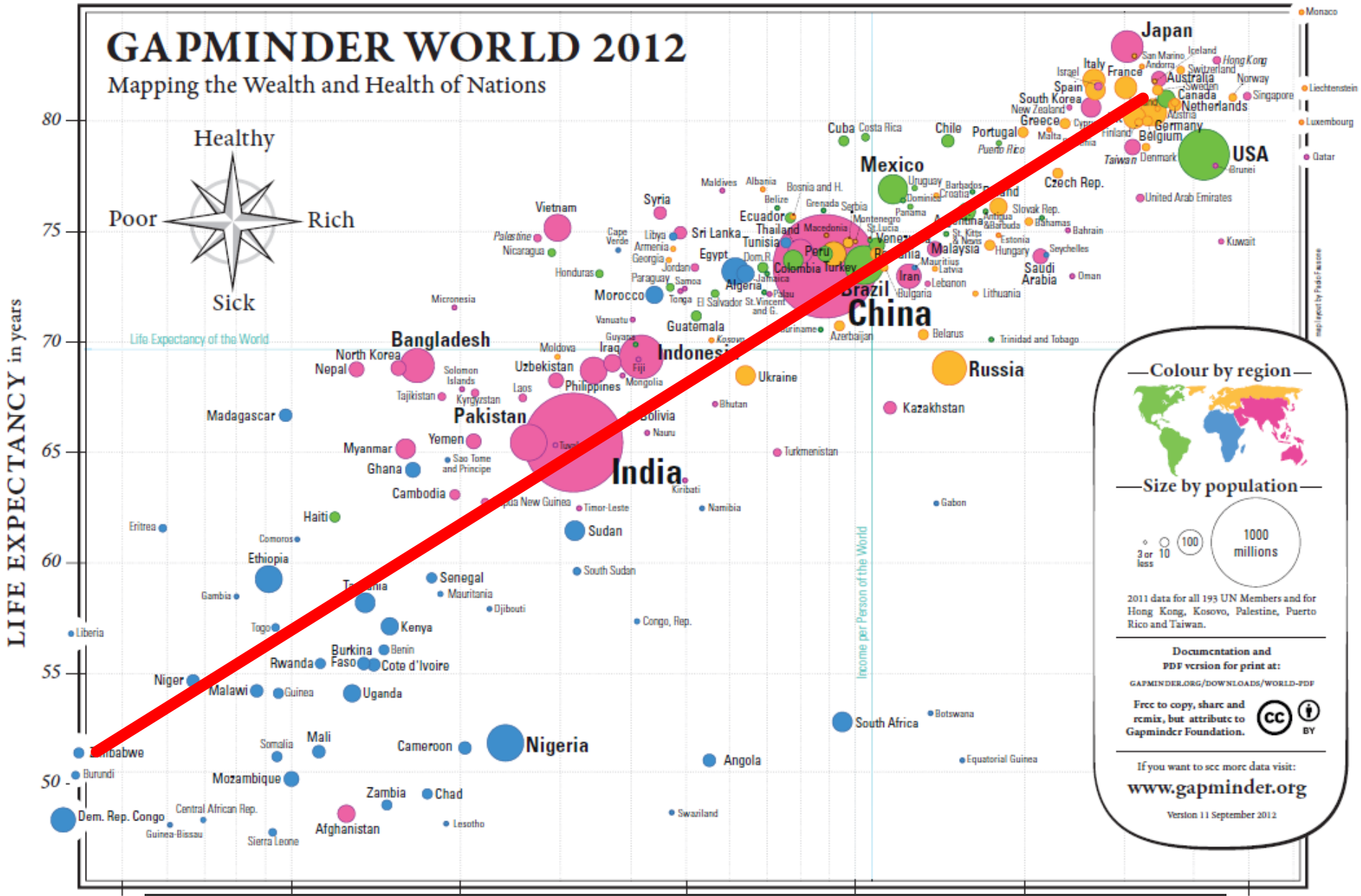


[Gapminder Foundation, https://www.ted.com/talks/hans_rosling_shows_the_best_stats_you_ve_ever_seen]



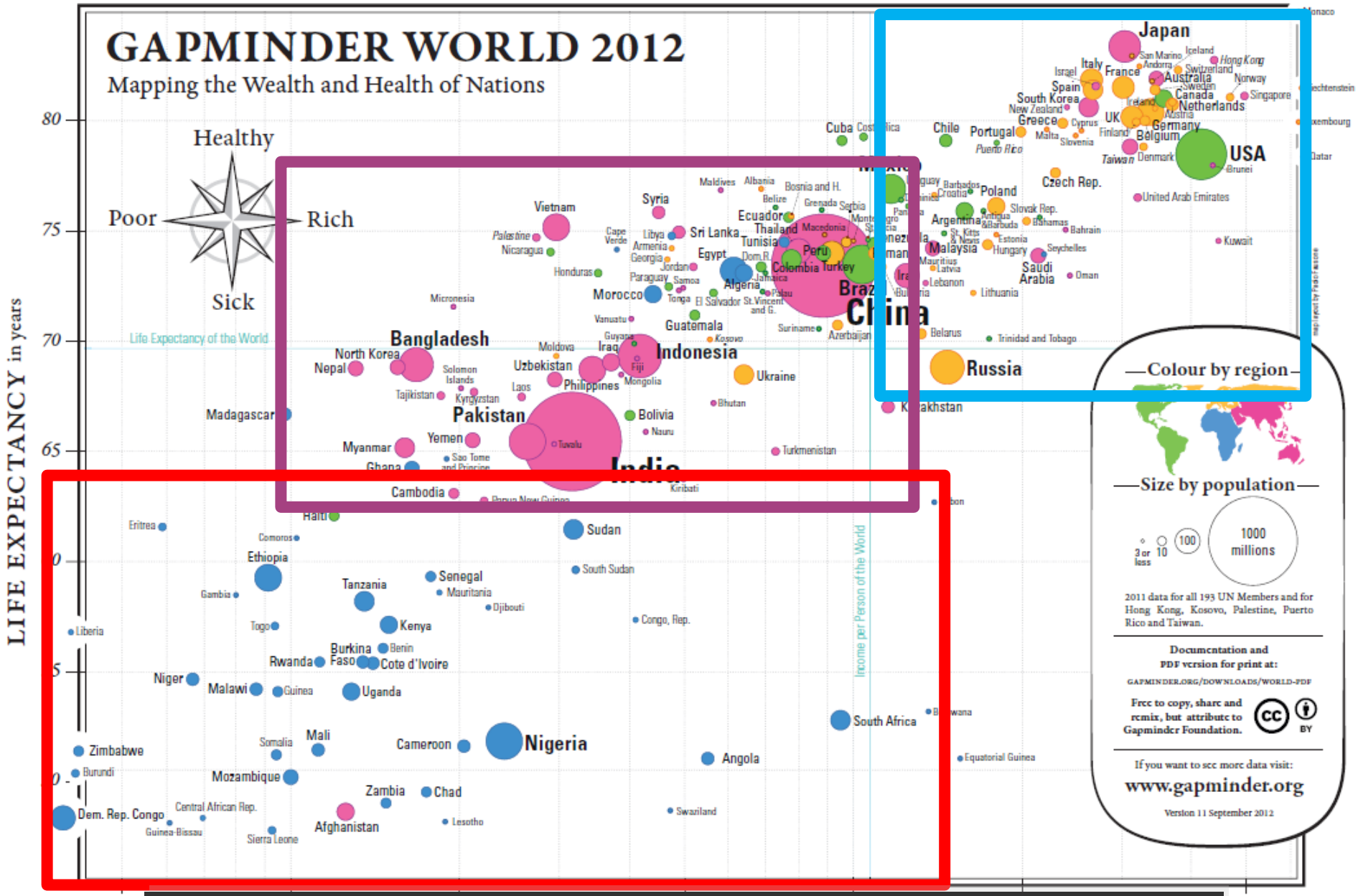
Meaningful Comparisons

[Gapminder Foundation, https://www.ted.com/talks/hans_rosling_shows_the_best_stats_you_ve_ever_seen]



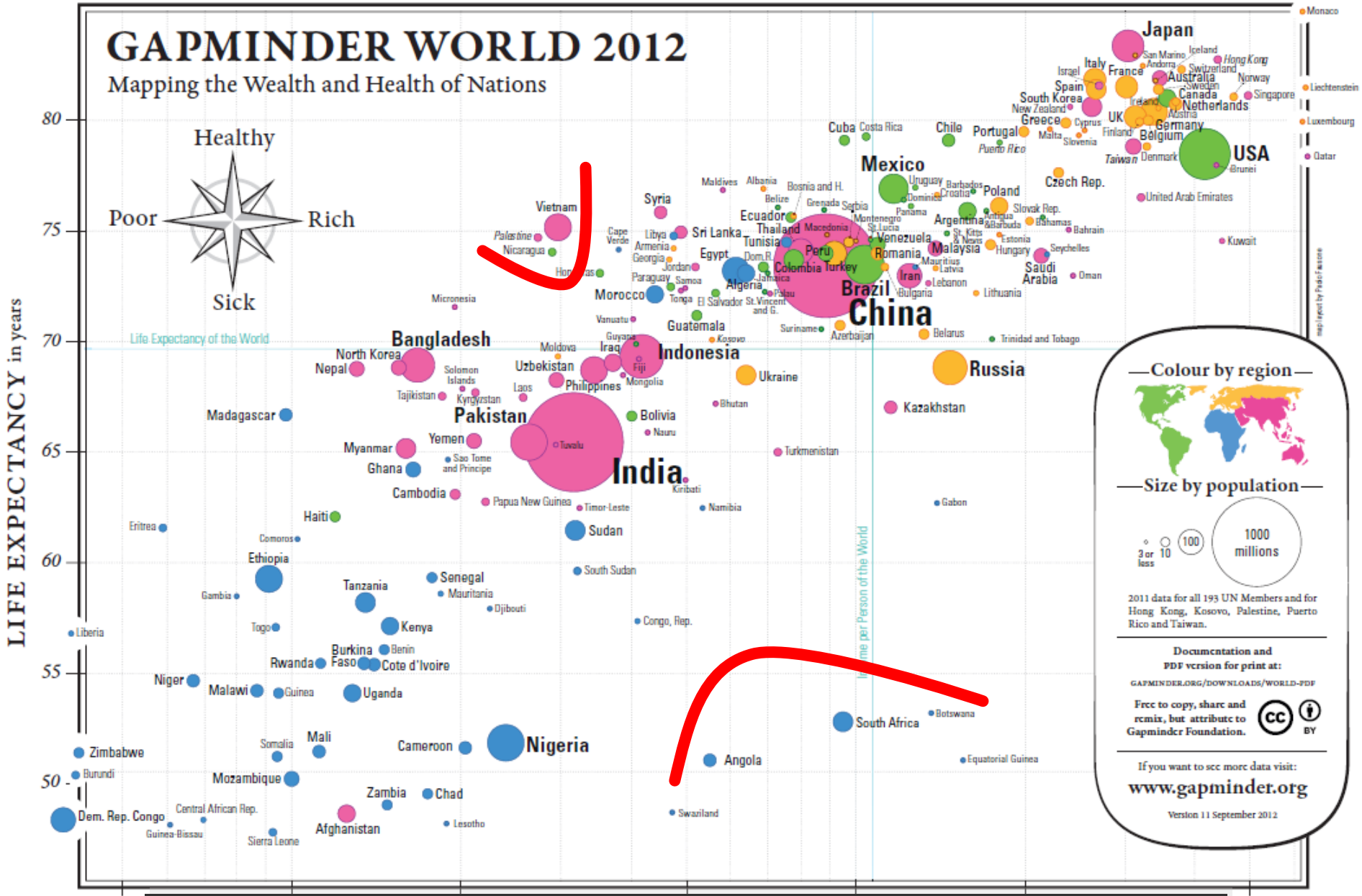
Underlying Structure and Multivariate Links

[Gapminder Foundation, https://www.ted.com/talks/hans_rosling_shows_the_best_stats_you_ve_ever_seen]



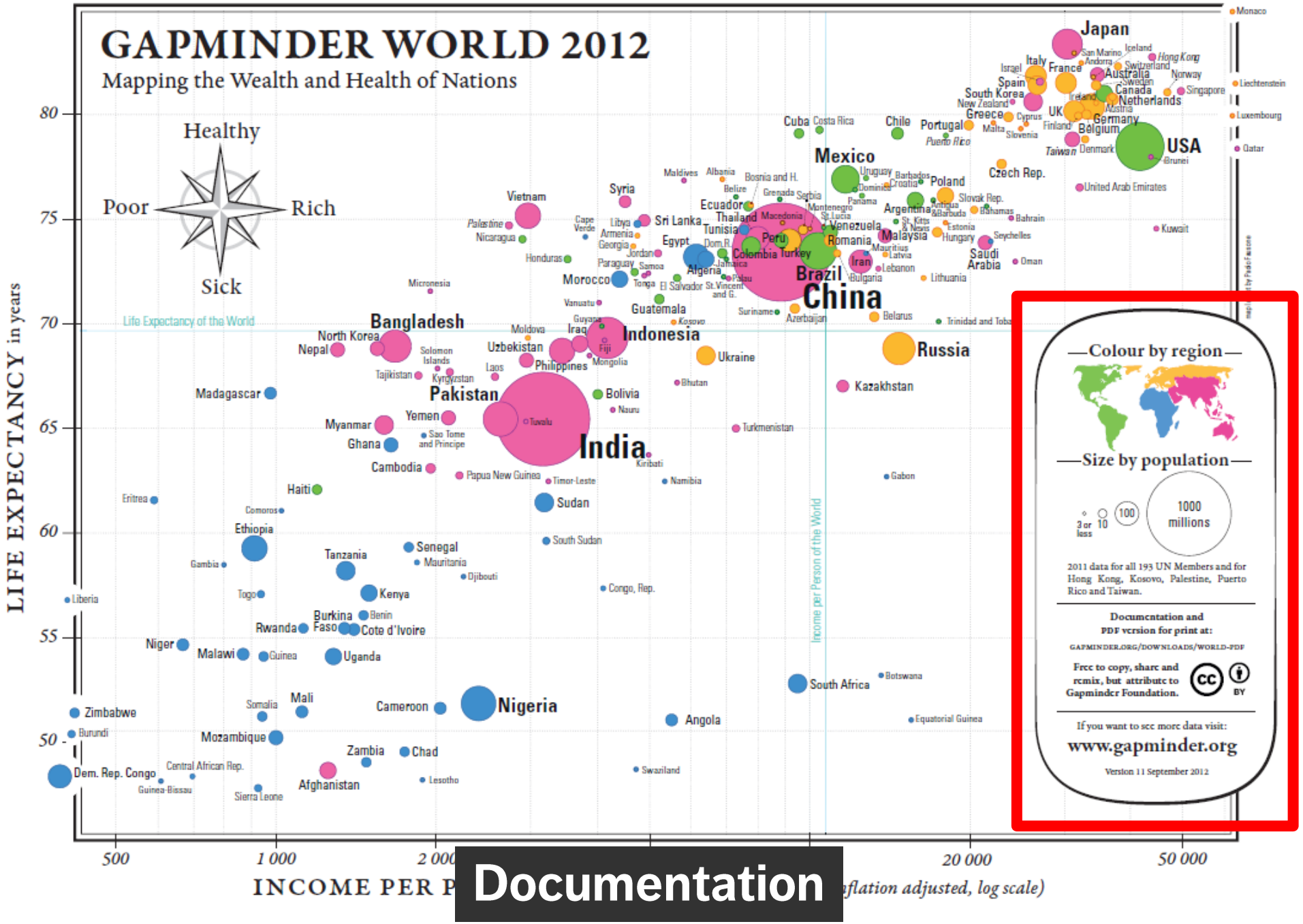
Underlying Structure and Multivariate Links

[Gapminder Foundation, https://www.ted.com/talks/hans_rosling_shows_the_best_stats_you_ve_ever_seen]



Underlying Structure and Multivariate Links

[Gapminder Foundation, https://www.ted.com/talks/hans_rosling_shows_the_best_stats_you_ve_ever_seen]



Documentation

PRESENTING ANALYSIS RESULTS

Graphics should be clear and engaging.

Not every pretty picture tells a story, but if a story can't be told with pretty pictures, perhaps it's time to re-think the story...

Graphical representation techniques appear regularly – it's too early to tell which ones will stand the test of time.

Don't be afraid to try something new if it helps **convey the message**.

DESIGN ELEMENTS

Is the point getting across? Integrated data helps convey the message.

Not all **retinal variables** are equally effective when it comes to convey or represent information. Experiment as needed to find the optimal choice for the given context.

Adding design elements can enhance our understanding of the data.

How we spot patterns affect what we get out of data presentations.

Data displays are not just about picking a random visualization method. The result varies depending on the structure of the data and the (combinations of) questions.

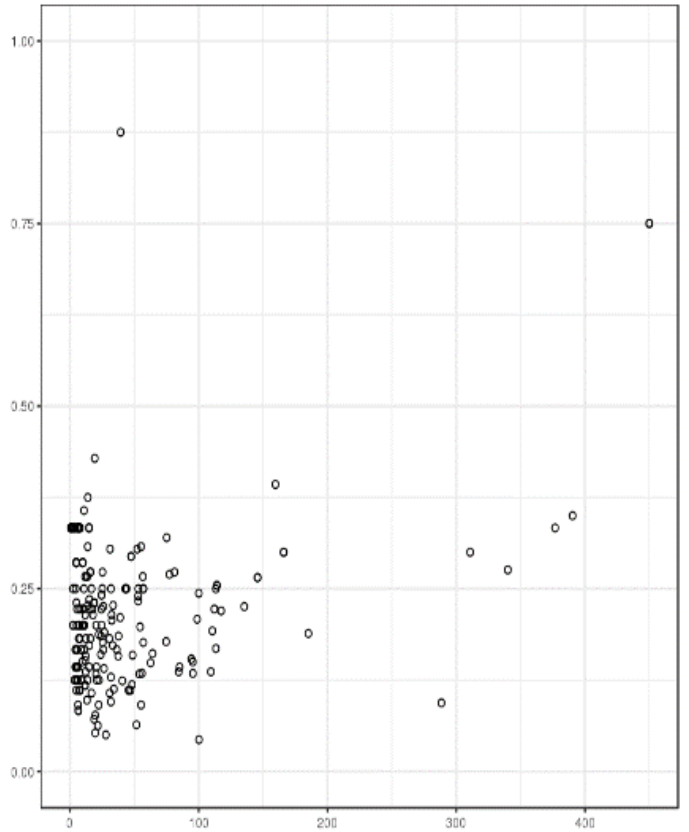
[http://promise.site.uottawa.ca/SERepository/datasets/cm1.arff]

REPRESENTING MULTIVARIATE OBSERVATIONS

2 variables can be represented by position in the plane. Additional factors can be depicted with:

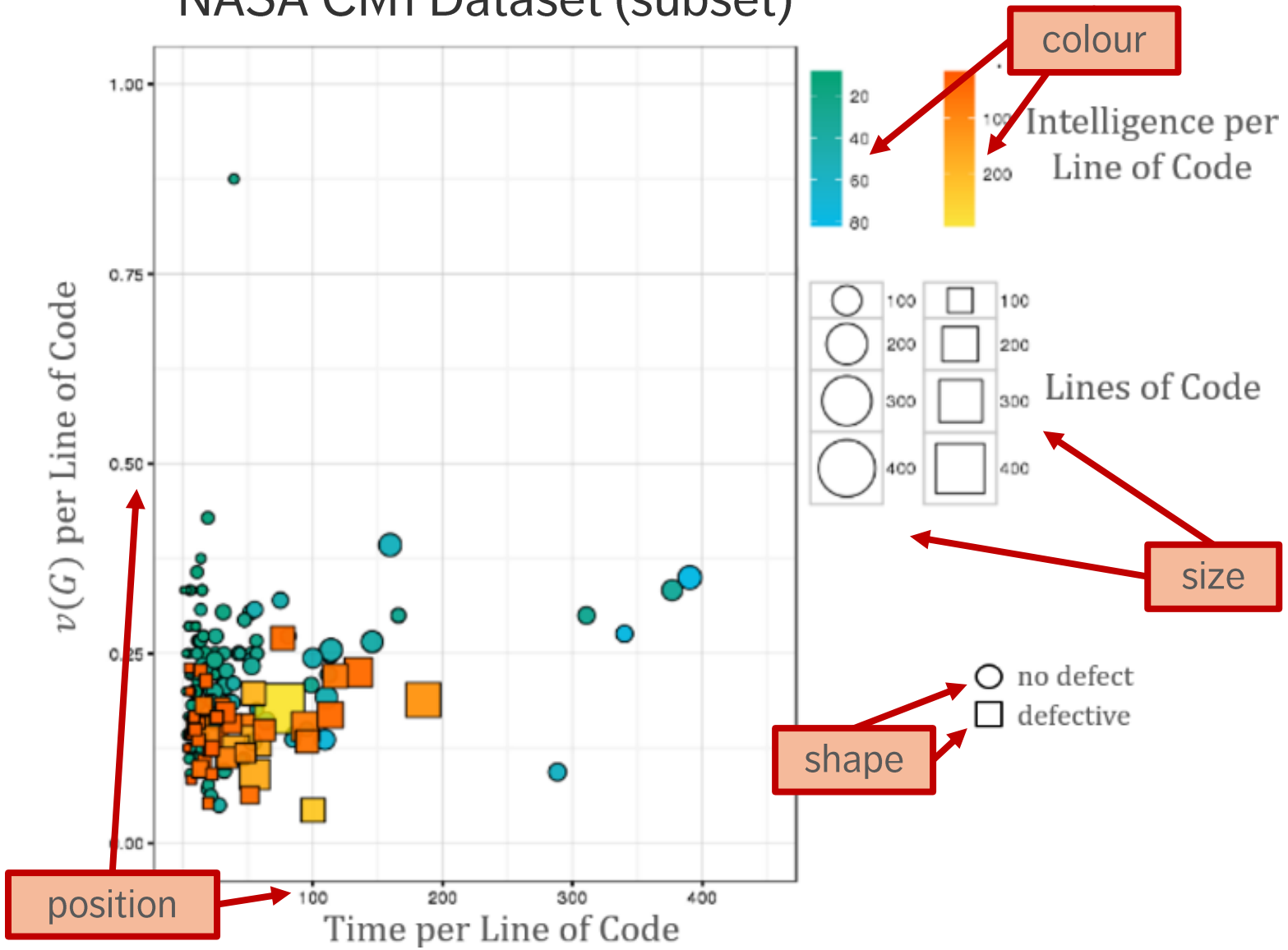
- size
- color
- value
- texture
- line orientation
- shape
- (motion?)

NASA CM1 Dataset (subset)



[http://promise.site.uottawa.ca/SERepository/datasets/cm1.arff]

NASA CM1 Dataset (subset)



DISCUSSION

What are some data visualization rules you have picked up over the years?

Are any of them hard rules?

Which of the rules/principles shown here seem more fundamental?