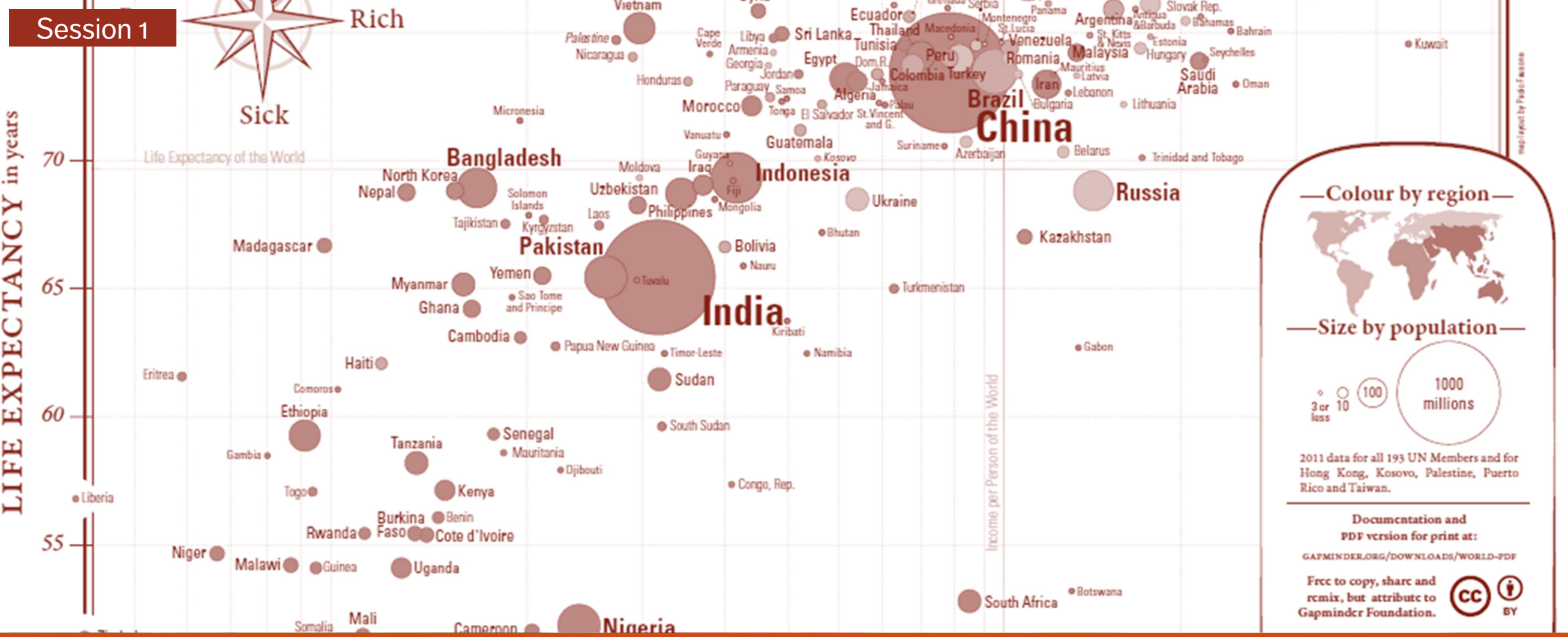
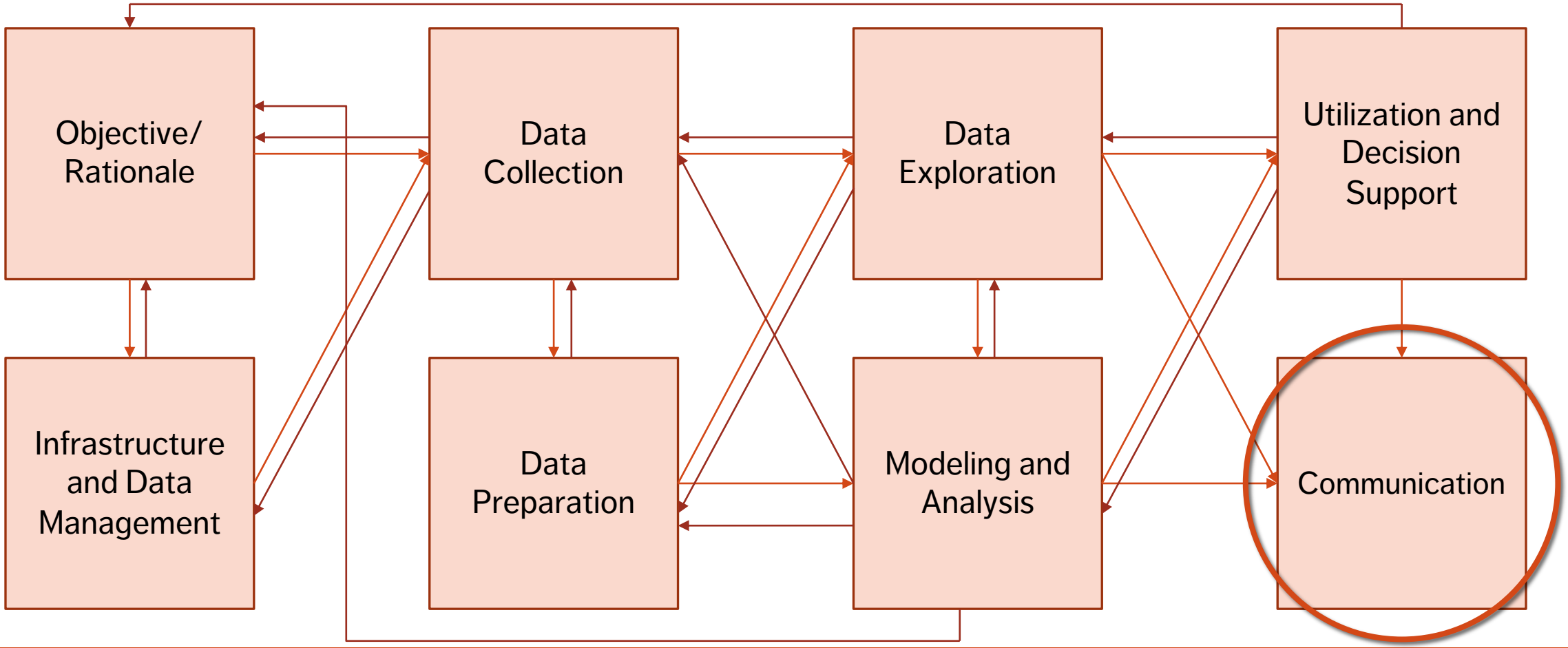


2. Data Visualization for Communication

The (Messy) Analysis Process



[E. Tufte, Beautiful Evidence]

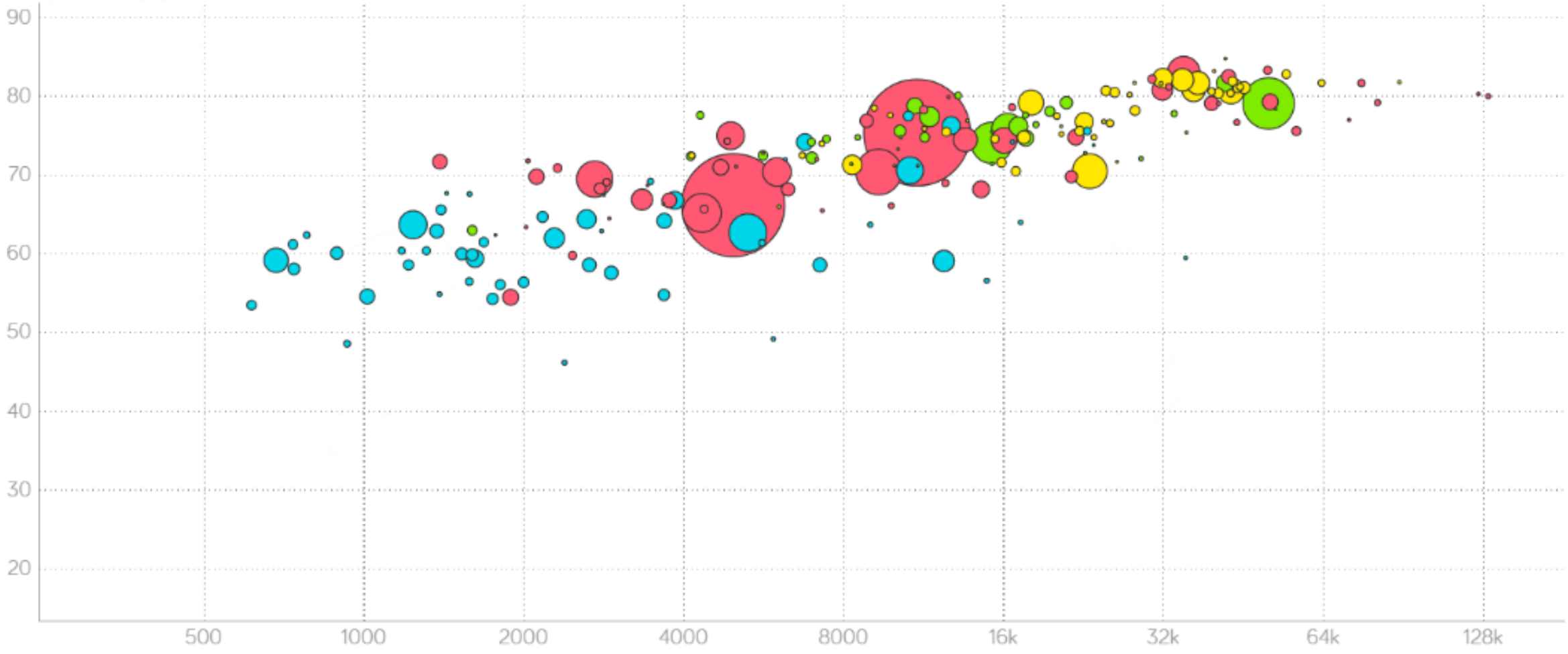
Principles of Analytical Design

Reasoning and **communicating** our thoughts are intertwined with our lives in a causal and dynamic multivariate Universe.

There is a **symmetry** to visual displays of evidence: consumers should be seeking exactly what producers should be providing, namely

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

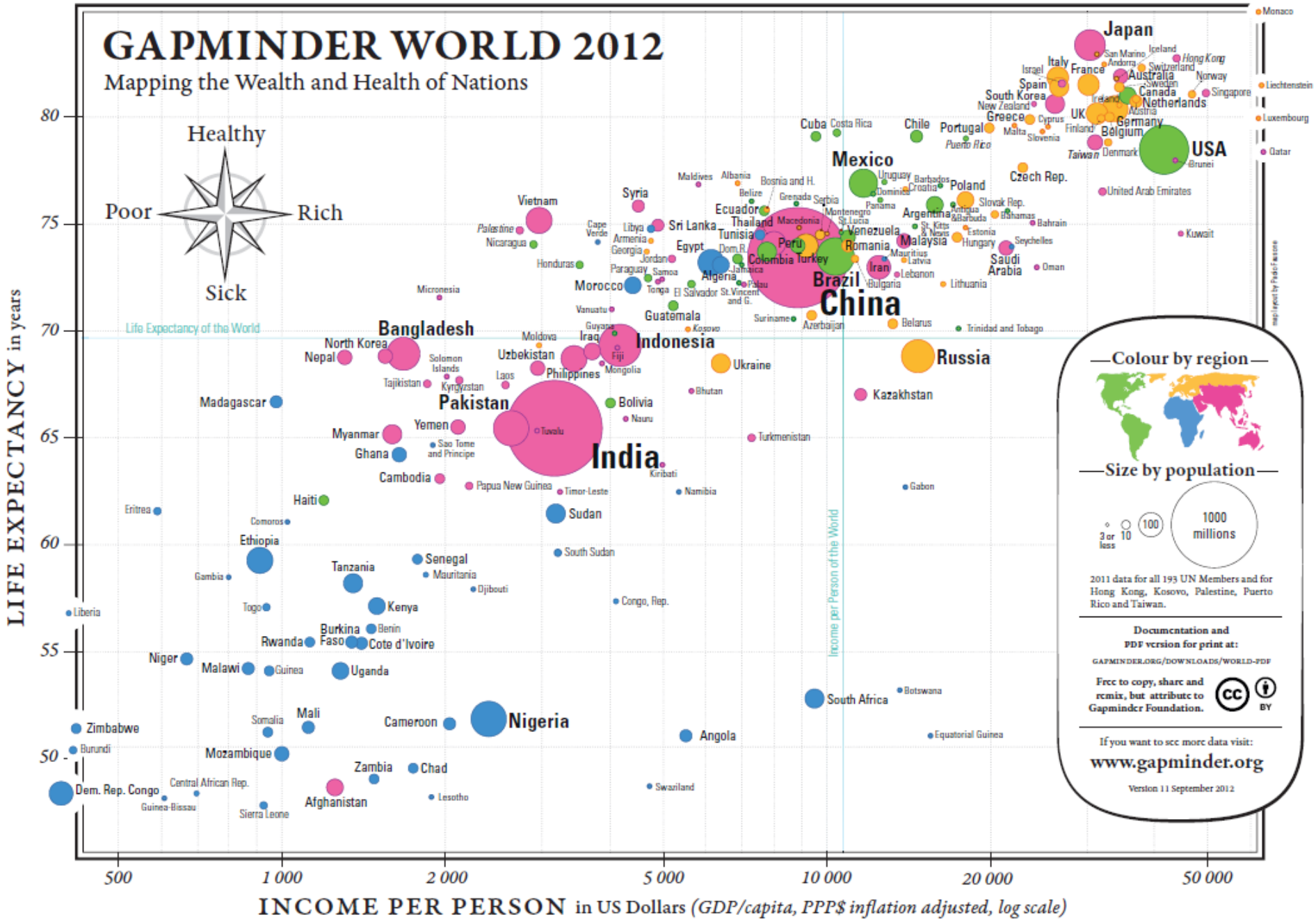
[Gapminder Foundation]



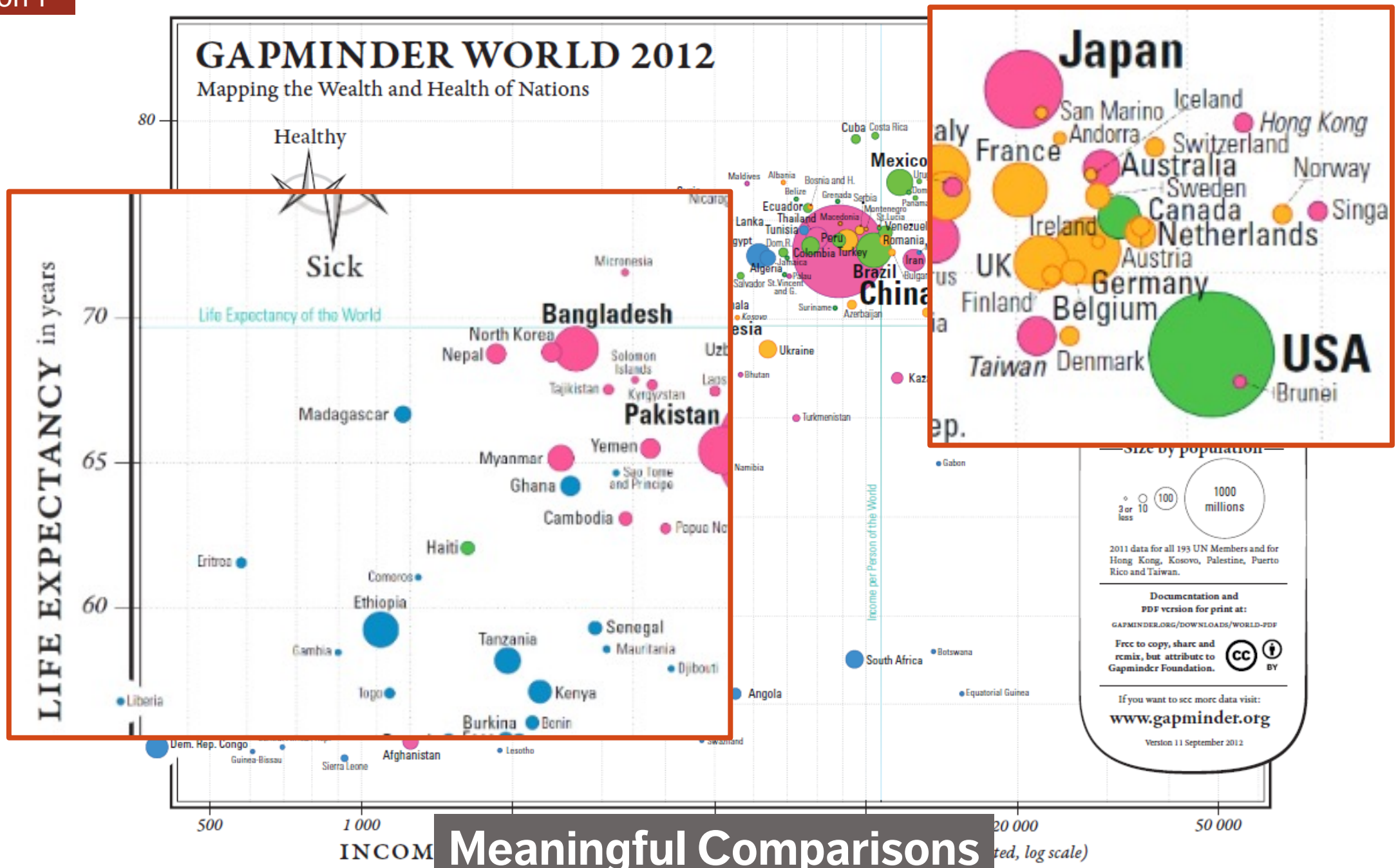
Non-Integrated Data

Session 1

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

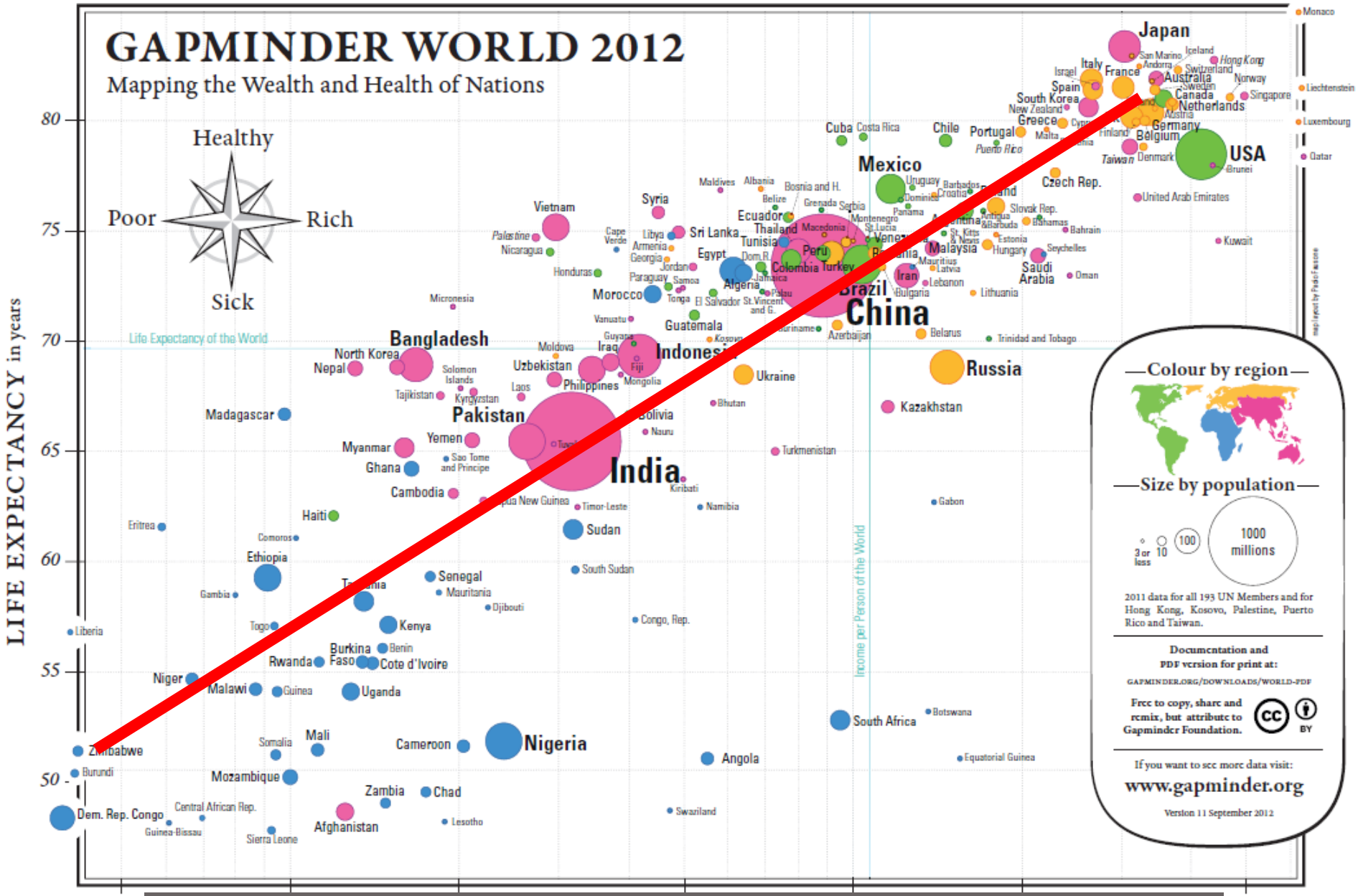


[Gapminder Foundation, https://www.ted.com/talks/hans_rosling_shows_the_best_stats_you_ve_ever_seen]



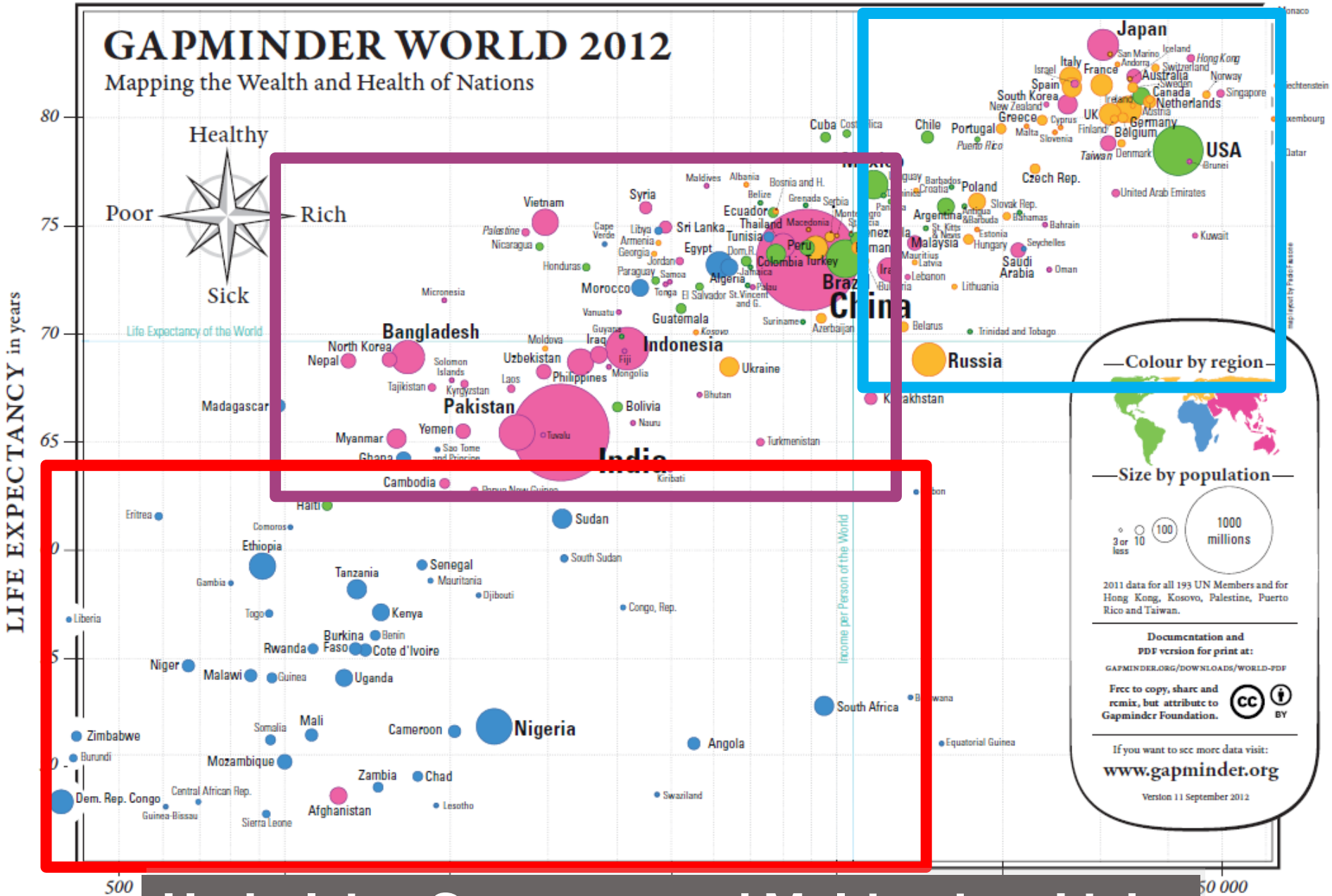
Meaningful Comparisons

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



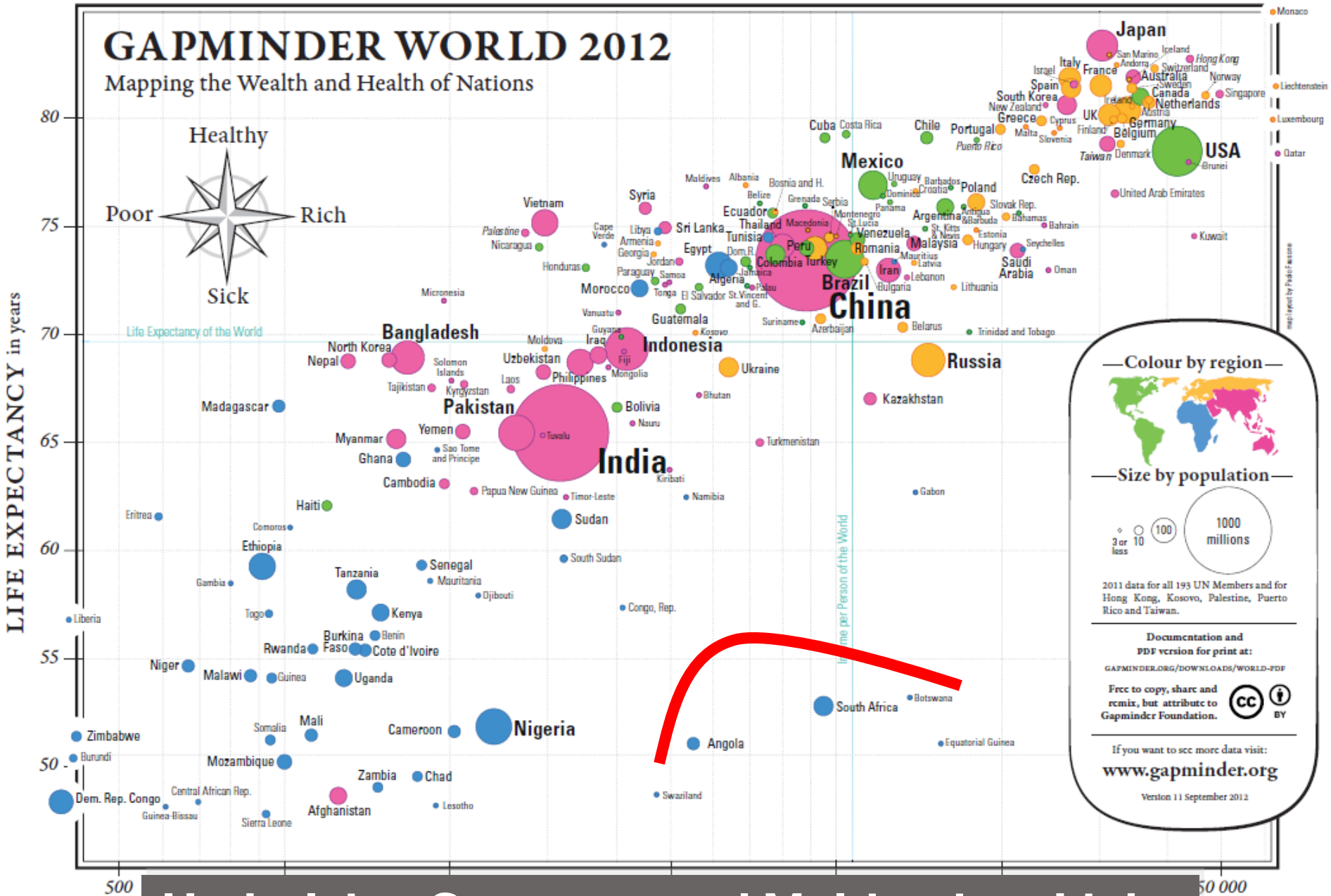
Underlying Structure and Multivariate Links

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



Underlying Structure and Multivariate Links

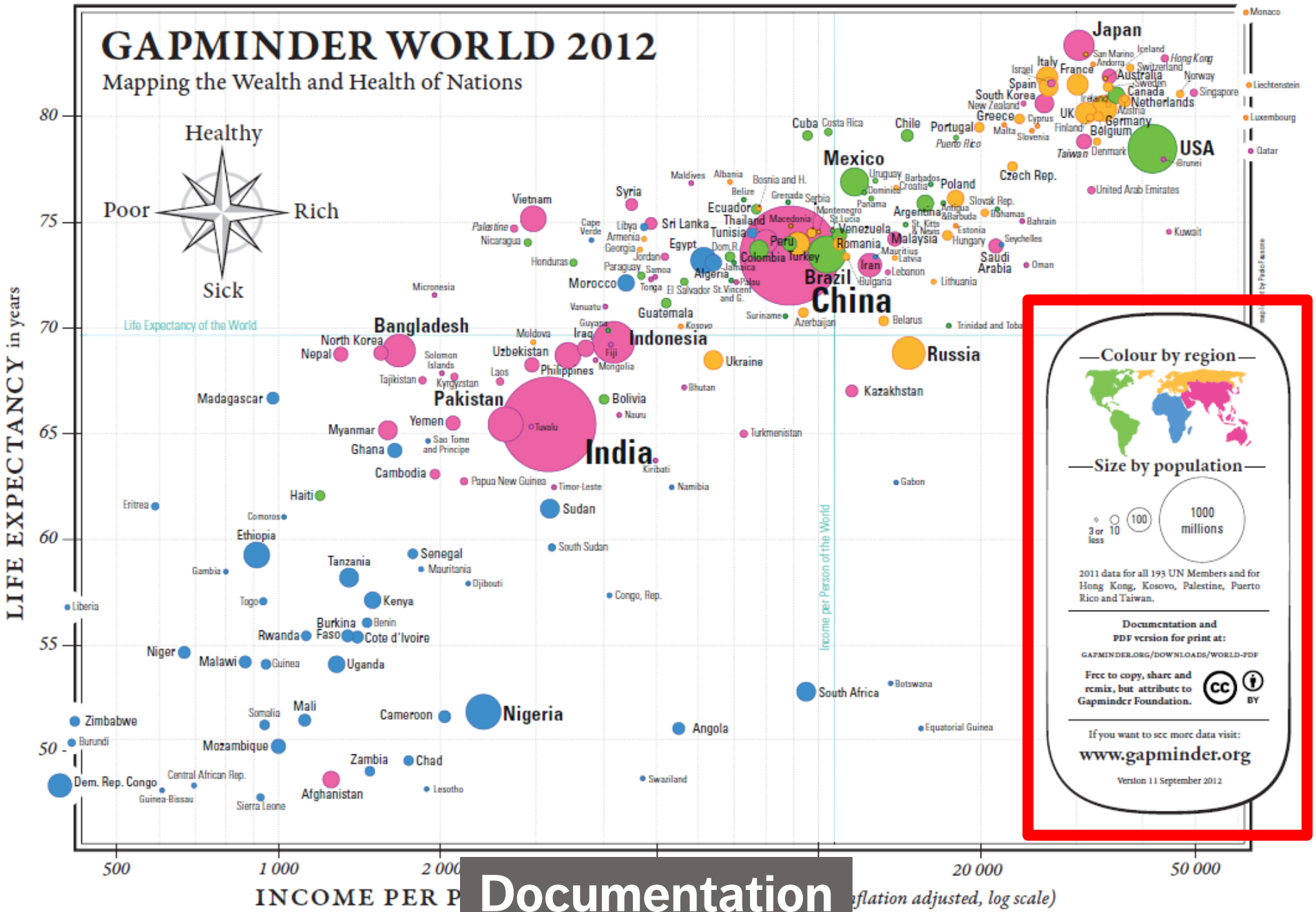
[Gapminder Foundation, https://www.ted.com/talks/hans_rosling_shows_the_best_stats_you_ve_ever_seen]



Underlying Structure and Multivariate Links

Session 1

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



Documentation

[http://flowingdata.com/tag/pitfalls]

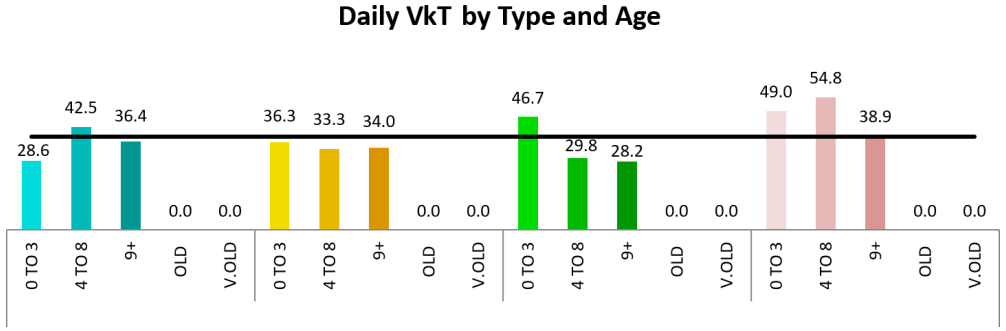
Basic Rules

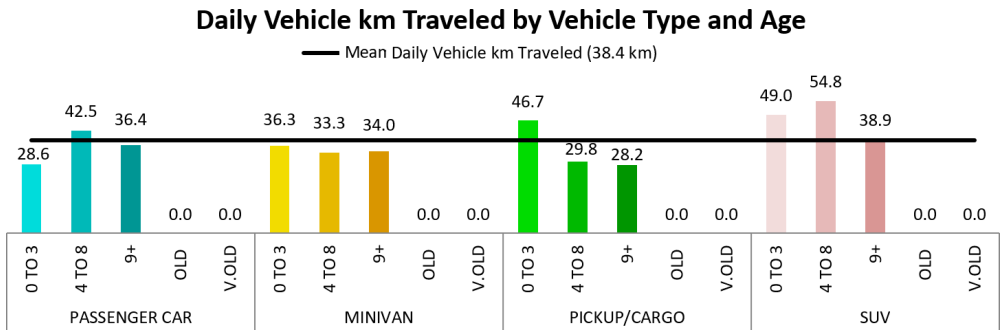
1. Check the data

outliers, spikes, anomalies

2. Explain encoding

don't assume the reader knows what everything means





3. Label axes

knowing the scale is important

[http://flowingdata.com/tag/pitfalls]

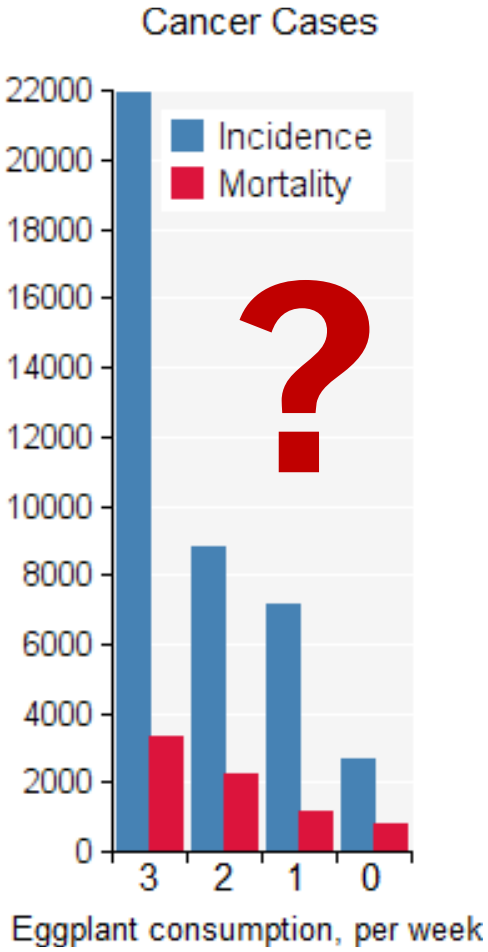
Basic Rules

4. Include units
eliminate the need for guesswork

5. Keep your geometry in check
circles and 2D shape are sized by area, bar charts by length

6. Include your sources
protect yourself, and let those who want to dig deeper do so

7. Consider your audience
a poster can be wordy, a presentation should be minimalist

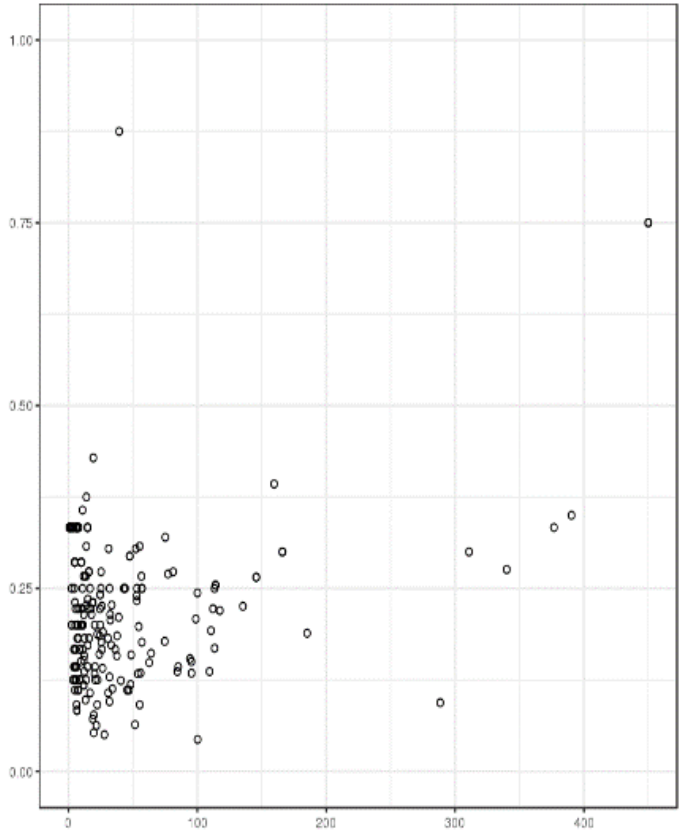


[http://promise.site.uottawa.ca/SERepository/datasets/cm1.arff]

Representing Multivariate Data

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

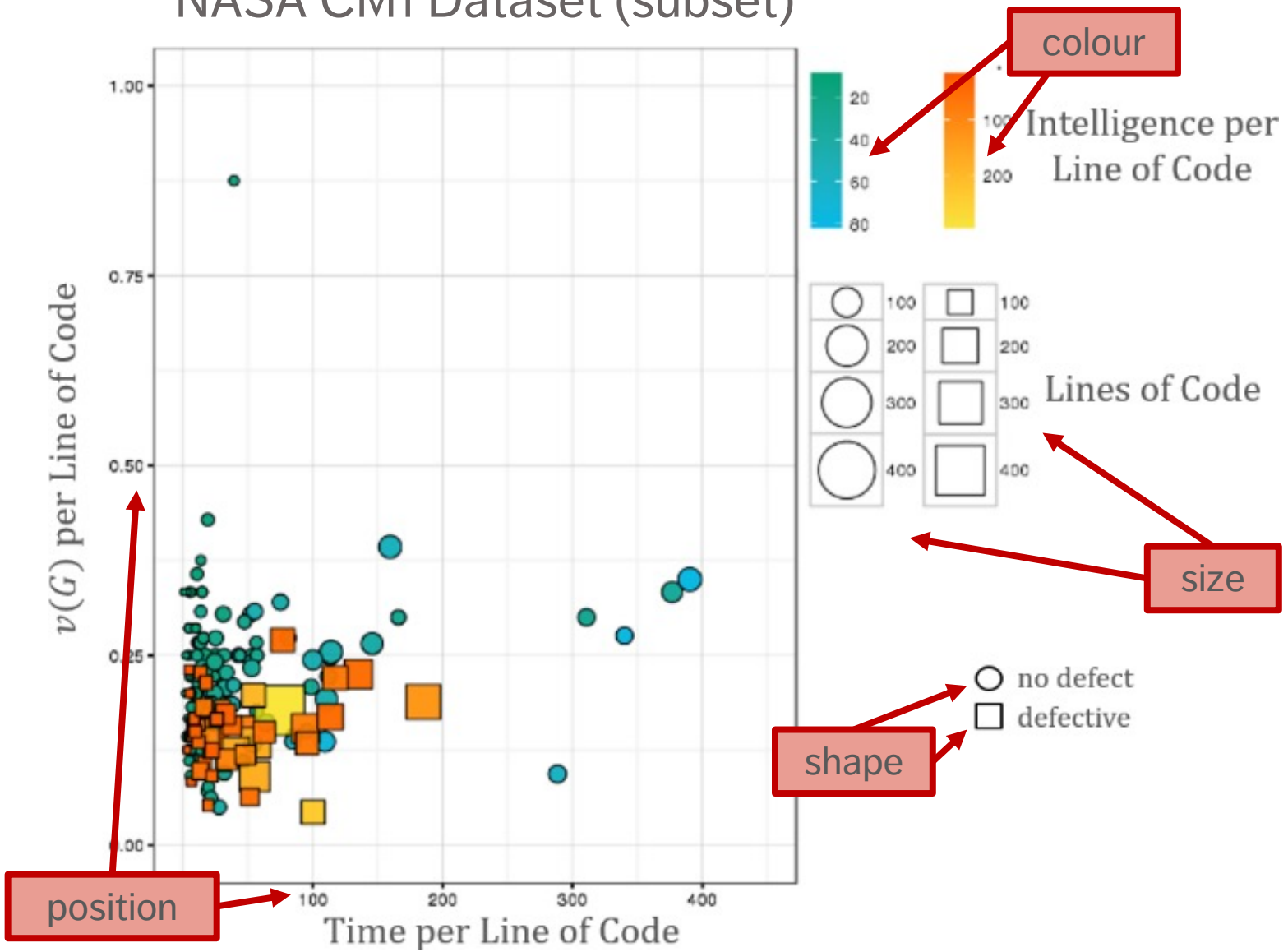
- size
- color
- value
- texture
- line orientation
- shape
- (motion?)



NASA CM1 Dataset (subset)

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

NASA CM1 Dataset (subset)

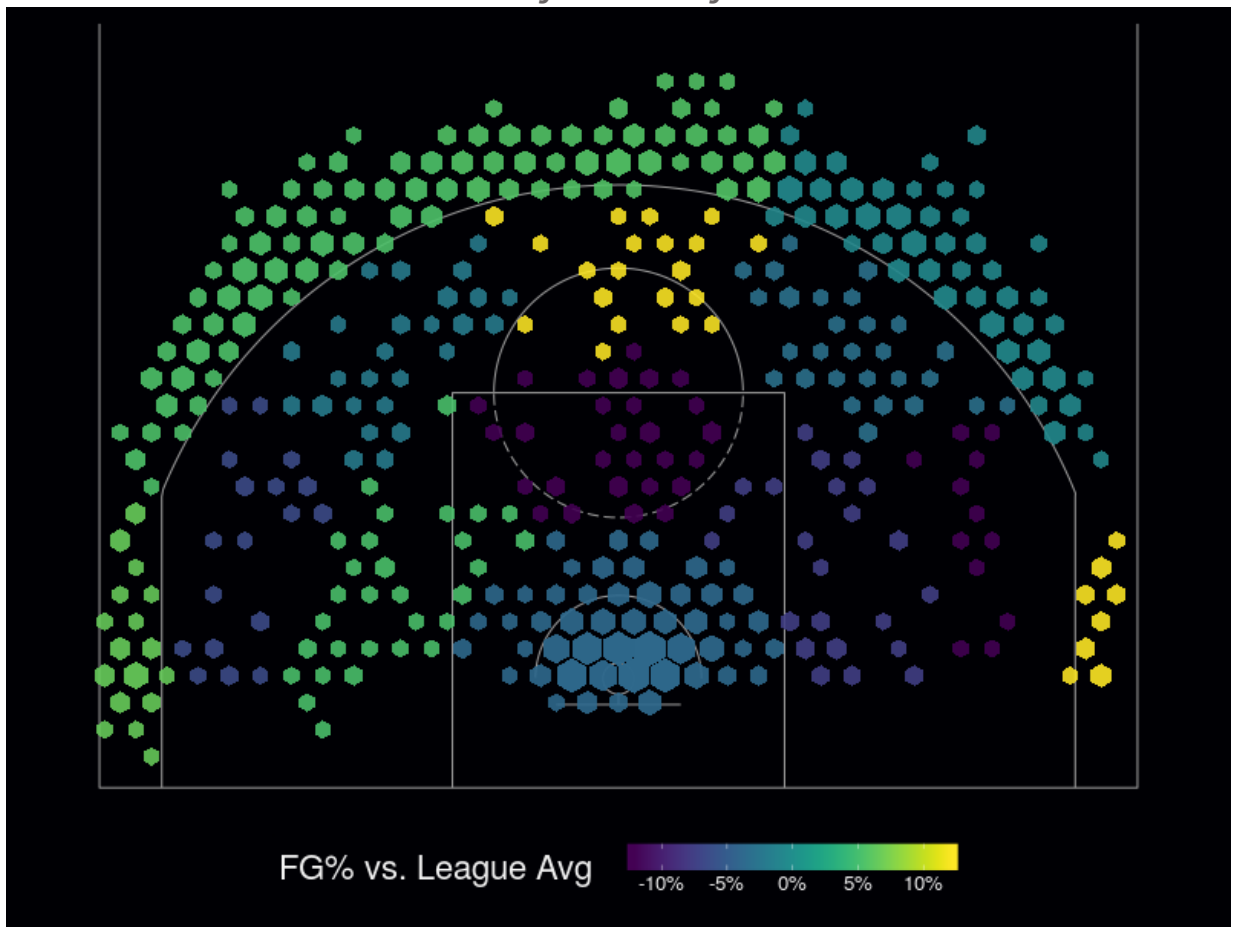


Session 1

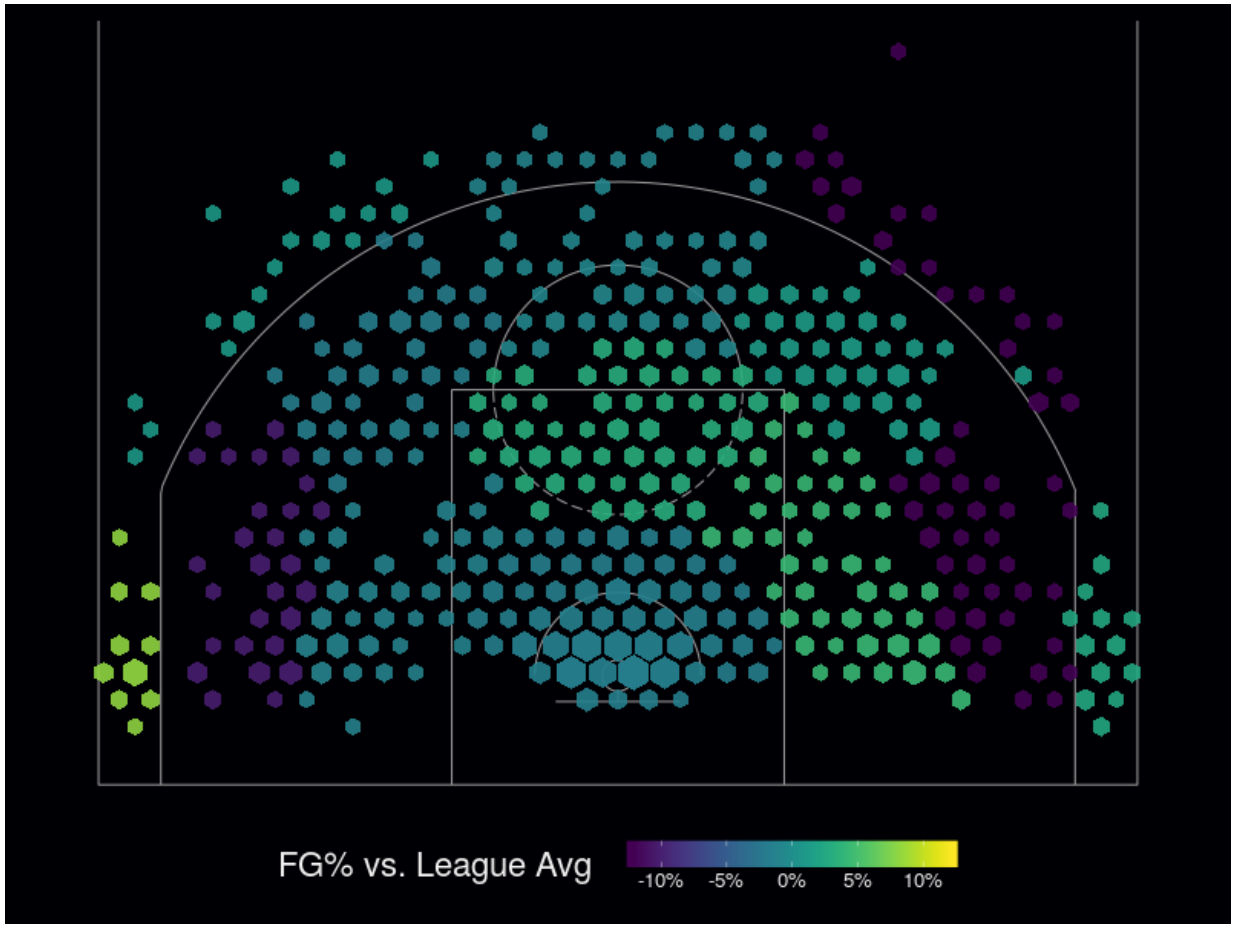
[T.W. Schneider]

NBA FG% Against League Average (‘15–’16)

Kyle Lowry

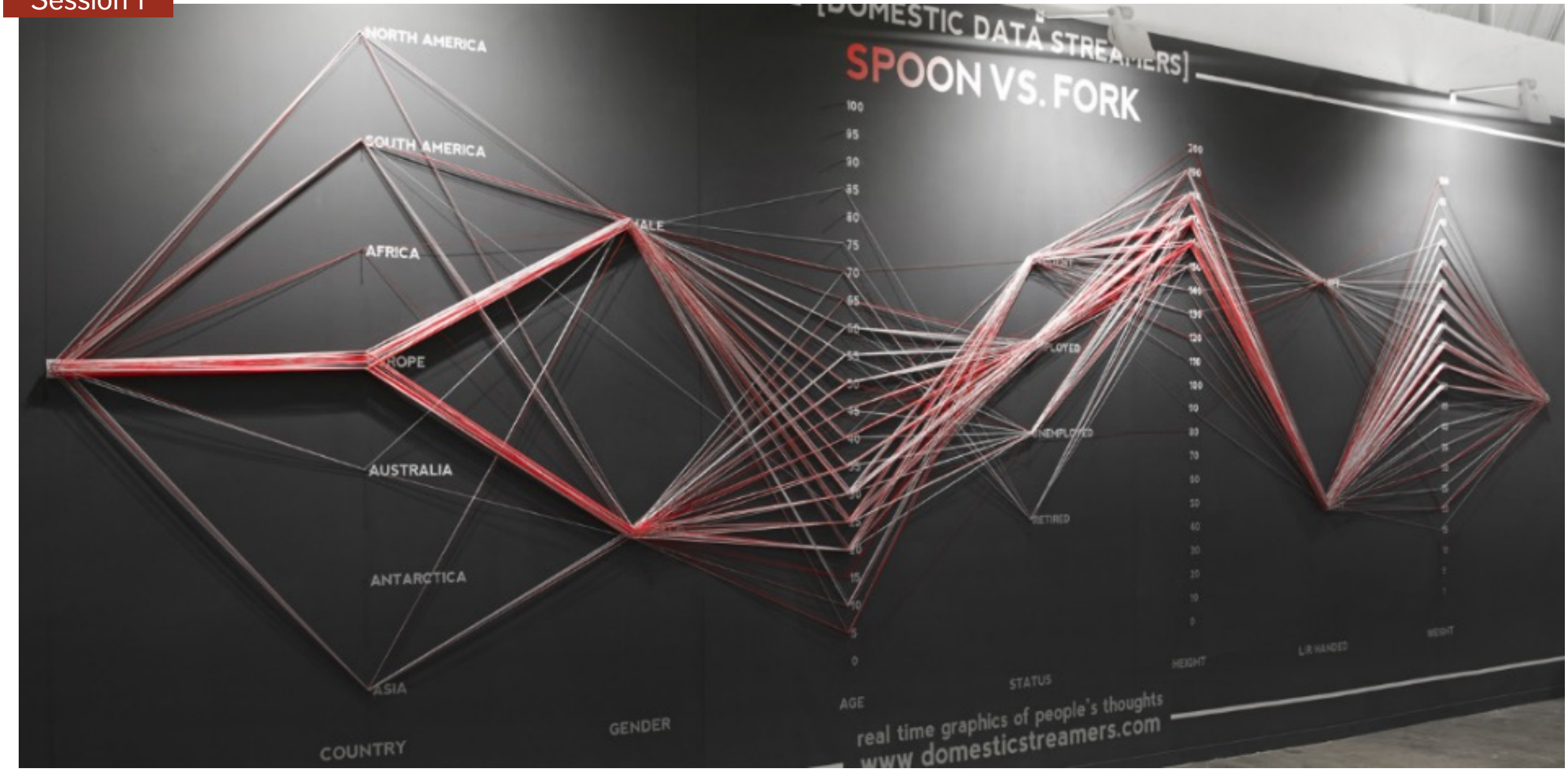


DeMar DeRozan



What comparisons can you make? Do you understand the encoding? The context?

[domesticstreamers.com]

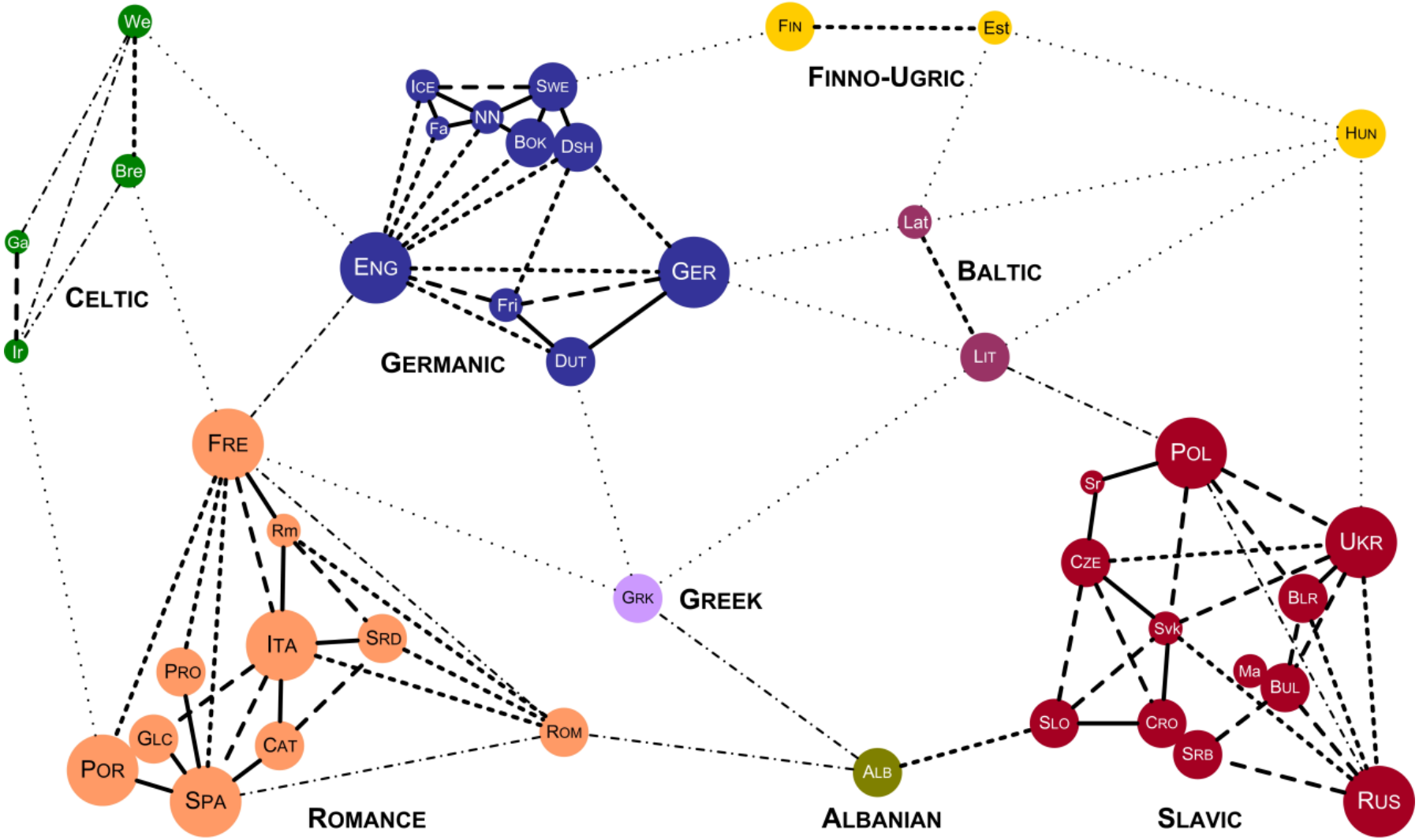


Are there any issues with data collection? Where do you think this event took place?
Is the spoon/fork question a red herring?

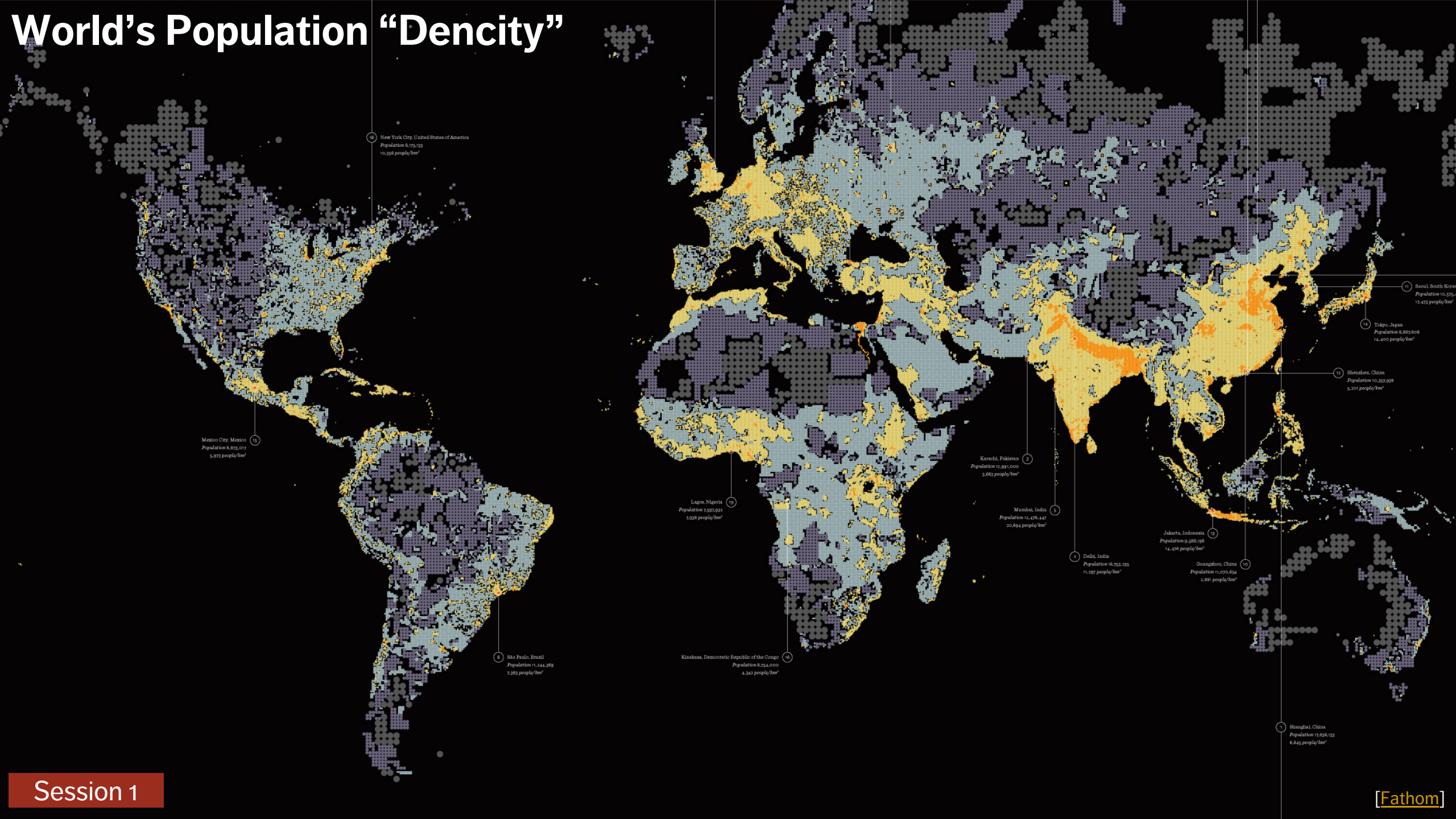
Session 1

Lexical Distance Among the Languages of Europe

[T.Elms, Etymologikon]



World's Population "Dencity"

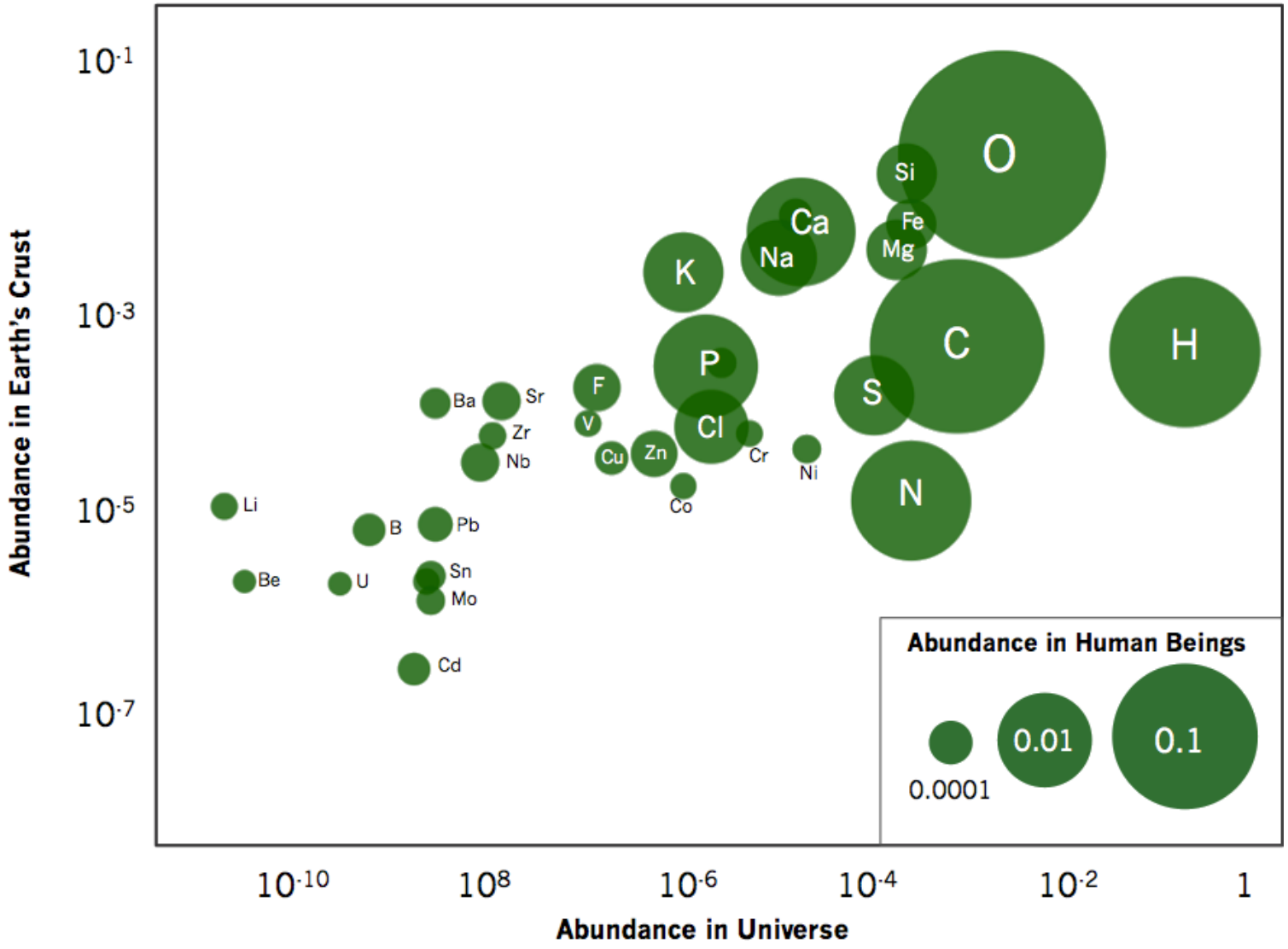


Session 1

[Fathom]

[M. Twardos, The Information Diet]

Abundance of Chemical Elements



Suggested Reading

Data Visualization for Communication

Data Understanding, Data Analysis, Data Science
Data Visualization and Data Exploration

Data and Charts

- Presenting Results
- Multivariate Elements in Charts

Fundamental Principles of Analytical Design

The Practice of Data Visualization
Basics of Data Visualization

Representing Multivariate Observations

Communicating Analysis Results

Hall-of-Fame / Hall-of-Shame

Exercises

Data Visualization for Communication

1. Study the charts presented in this section (or any other chart of your choosing) using the principles of analytical design. Are the basic rules followed?

2. What chart/design elements could be added to enhance the presentations?