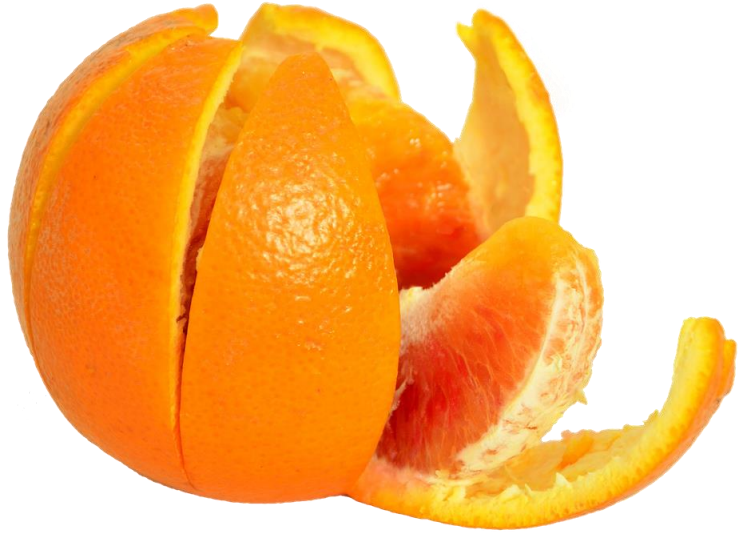
Adding stuff up is hard



+

Aggregation and the WDI

WORLD BANK GROUP

World Development Indicators (WDI)

- Collection of 1500+ development indicators
- Aim is Convenience and Credibility
- Includes national, regional and global estimates

WORLD BANK GROUP

Aggregation: Why and what?

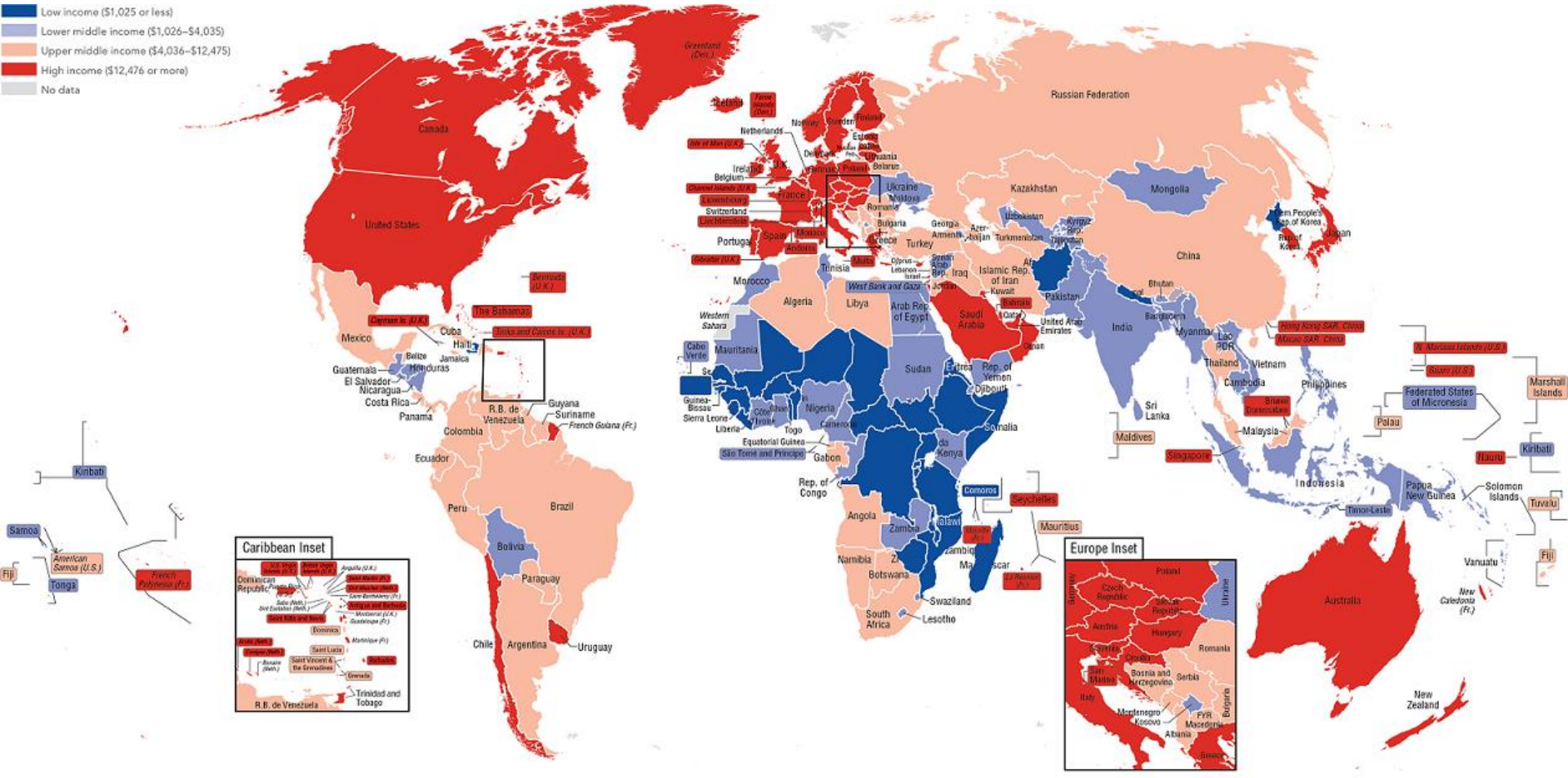
- **Analytical**: Provide benchmarks
  - Regions
  - Income groups

- **Operational**:
  - Lending categories

- **Custom aggregations:**
  - WDI allows for flexible aggregation and definition of groups, regions

WORLD BANK GROUP

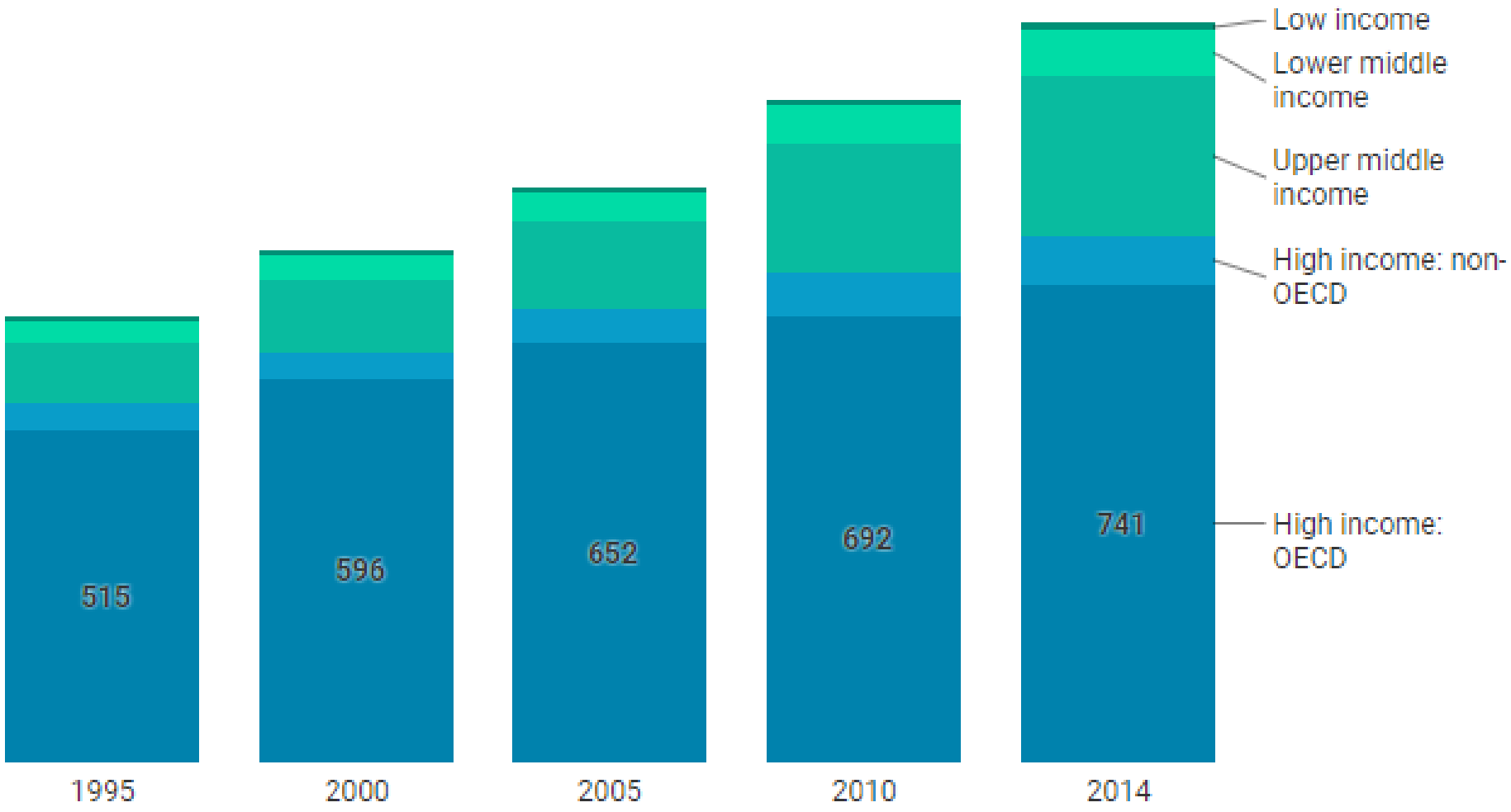
The World by income



WORLD BANK GROUP

Global Wealth Grew 66% between 1995 and 2014

Total wealth (USD Trillions, constant 2014)



Source: World Bank: The Changing Wealth of Nations 2018

WORLD BANK GROUP

Aggregation: Where?

Aggregation is **not uniform** as we have 70+ partners

Aggregations take place:

- **Inside WDI** database
- **Outside WDI** database
  - World Bank experts
  - Non World bank experts

WORLD BANK GROUP

Aggregation: How?

Key decisions:

- How to **weight** individual data points?

  - Population, countries, size of the economy

- How to handle **missing values**?

  - Exclude, threshold, impute

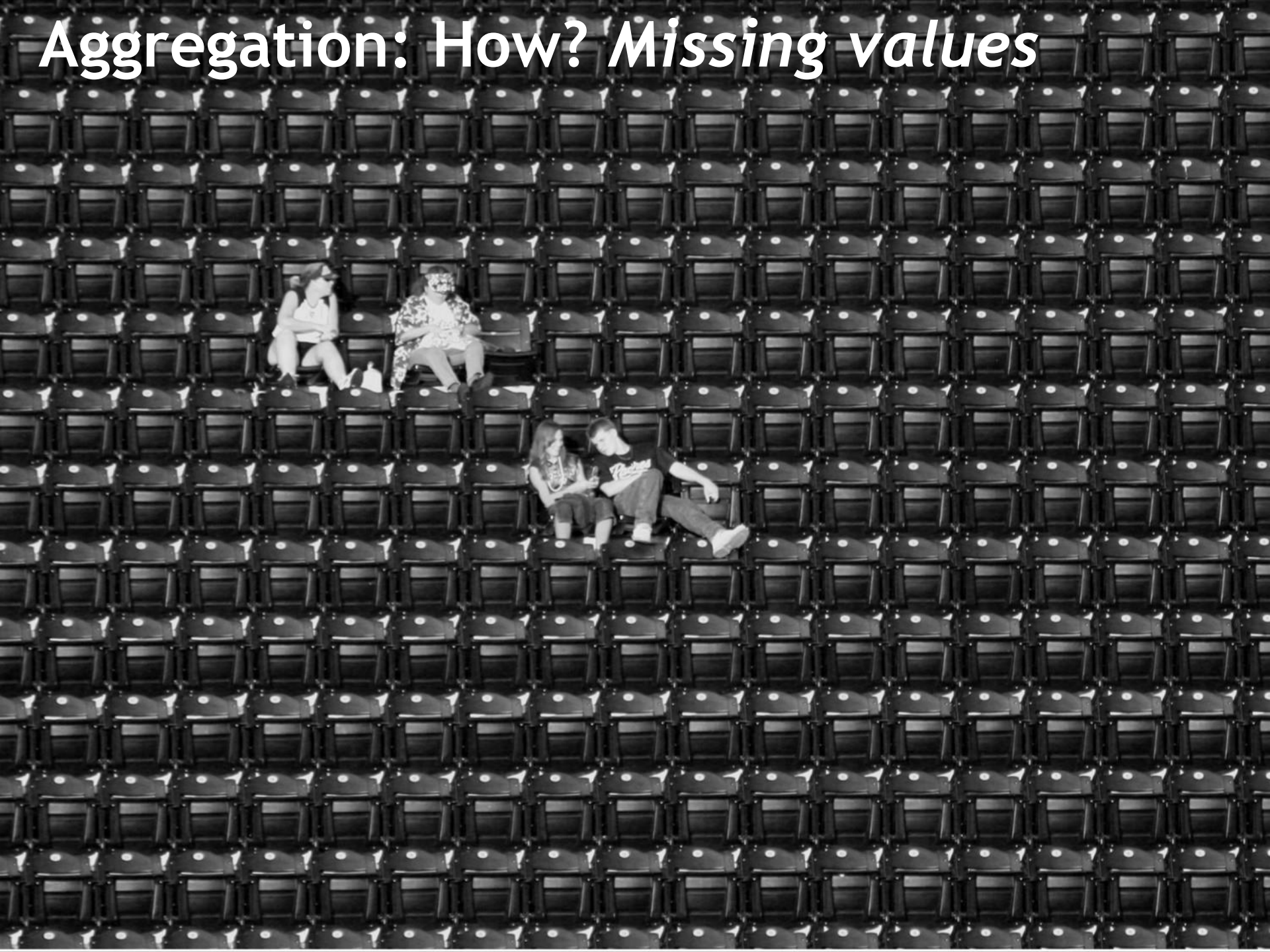
WORLD BANK GROUP

Aggregation: How? *Weighting*

Typical **weights** used in WDI

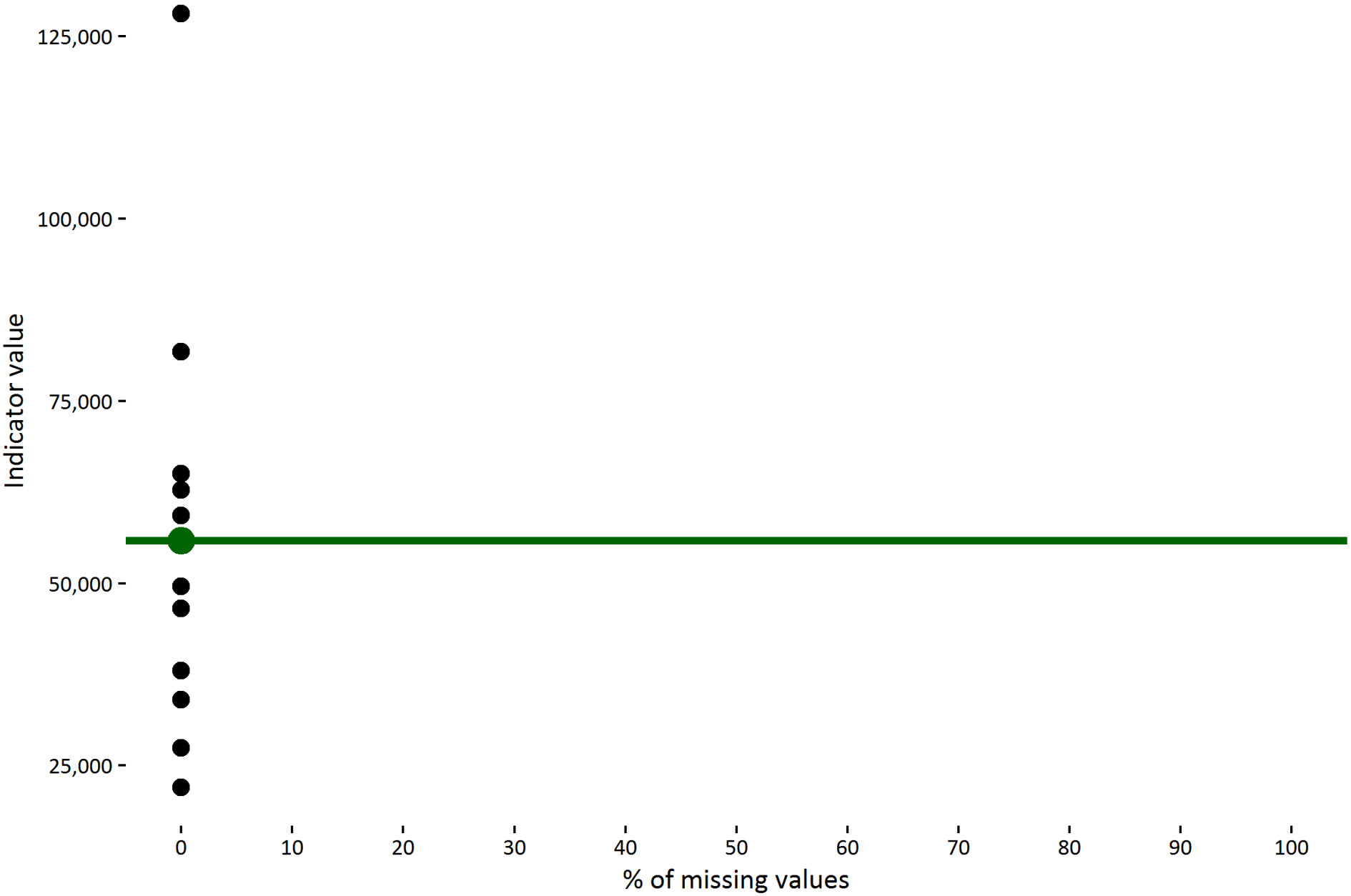
- **Population**
  - Prevalence of undernourishment (% of population)
  - Fertility rate (weight: women between 15 and 49)
- **Economy size**
  - Gross domestic savings (% of GDP)
- **Country (simple average)**
  - Equity of public resource use (country level index)

Aggregation: How? *Missing values*



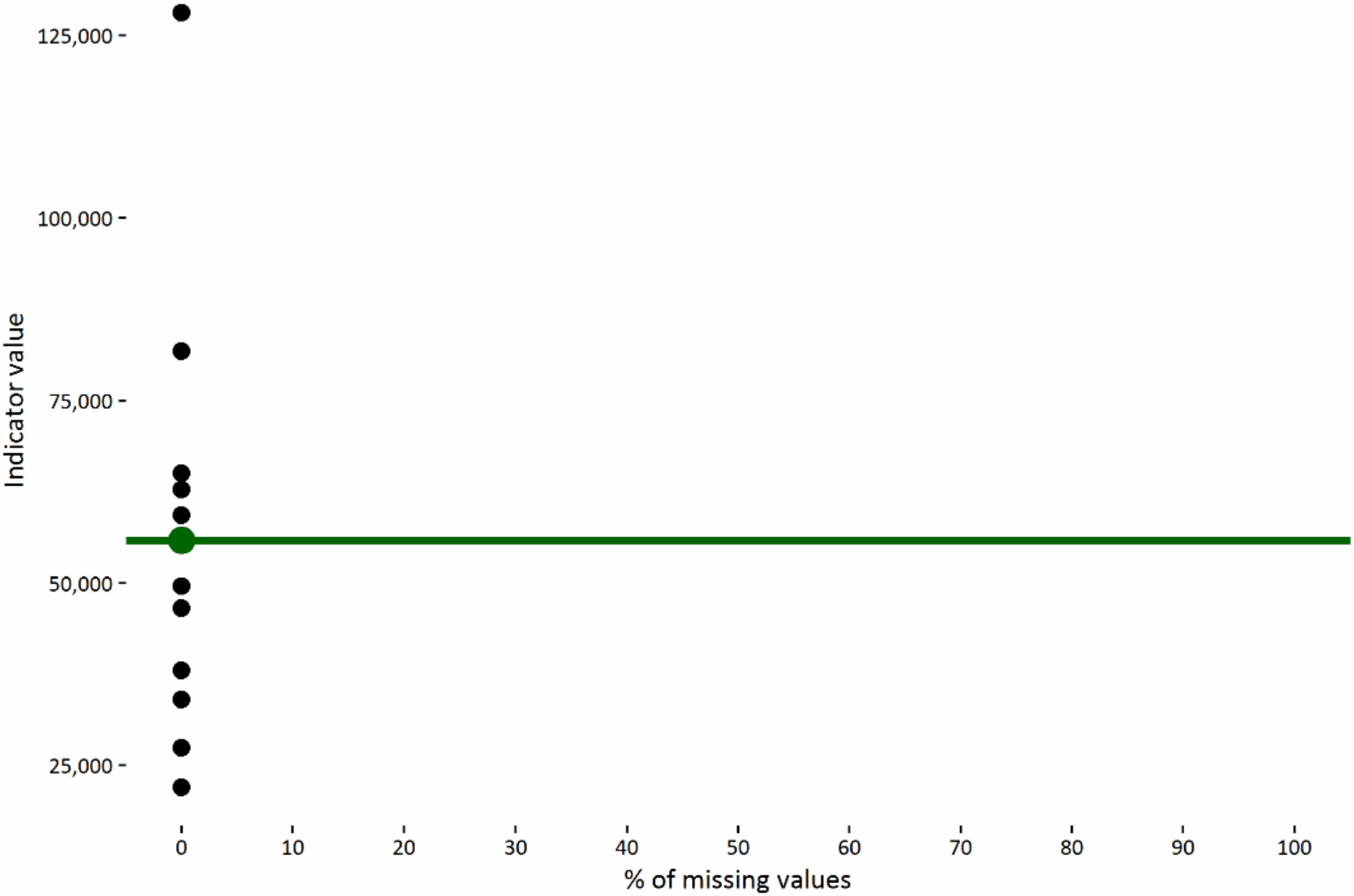
Aggregation with missing data

What happens when missing are removed?



Aggregation with missing data

What happens when missing are removed?



Aggregation thresholds

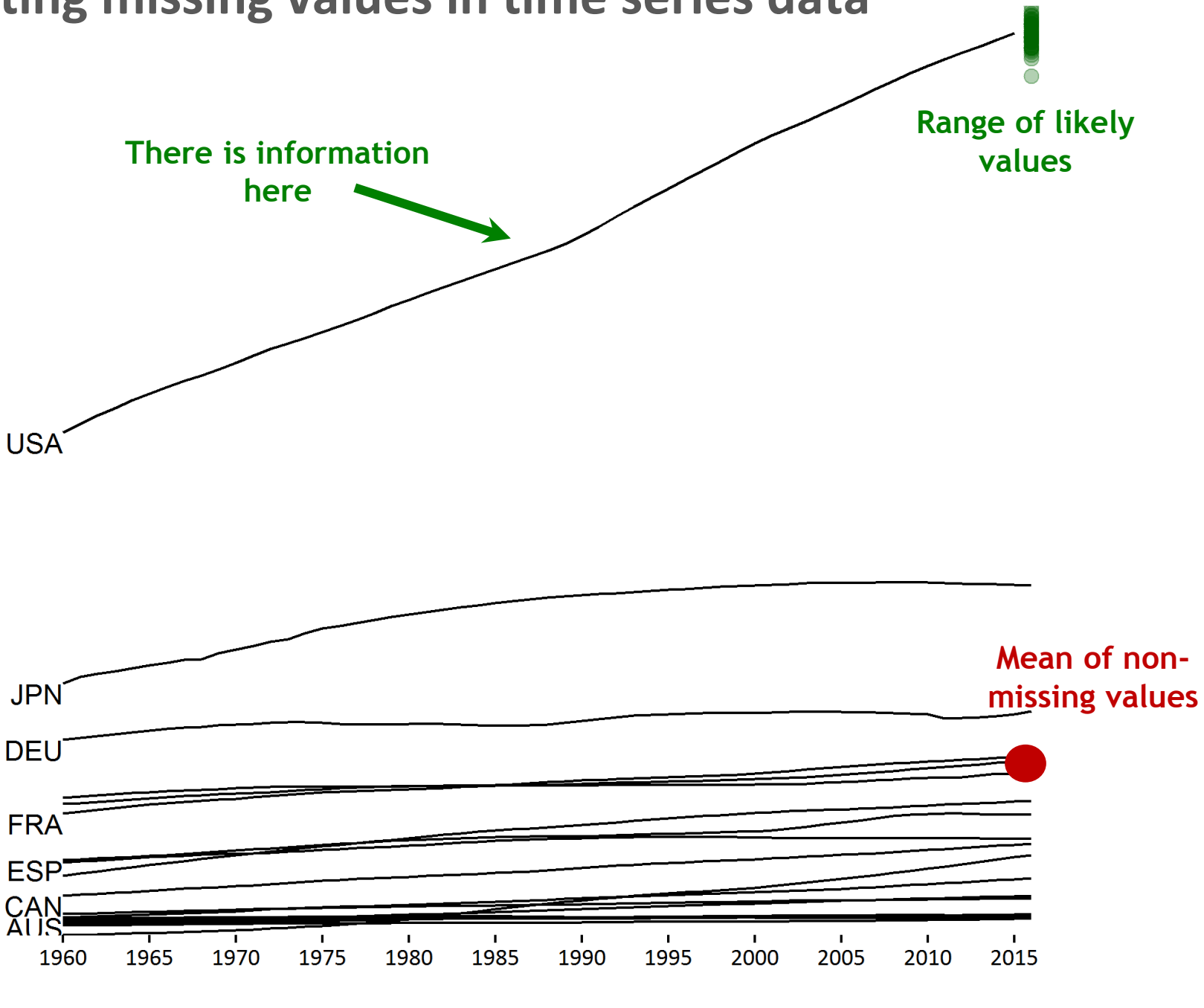
Typical **WDI** thresholds:

- Sum and averages: 67%
- Median: 50%

Non-WDI thresholds:

- Poverty: 40% (over period of +/- 2 years from reference year)
- Safely managed water and sanitation: 30%
- Basic drinking water, sanitation, and hygiene services: 50%

Imputing missing values in time series data



Aggregation: Summing it up!

- Development of **better standards** and methods (urban / rural)
- Be **more transparent**: Better labeling and documentation
- **Global** approach
- Systematic **imputation** of missing values

  Leverage all available information to estimate to provide best estimates

WORLD BANK GROUP

THANK YOU!

Umar Serajuddin	Sr. Economist	userajuddin@worldbank.org
Malarvizhi Veerappan	Sr. Data Scientist	mveerappan@worldbank.org
Tony Fujs	Data Scientist	tfujs@worldbank.org

WORLD BANK GROUP

Aggregation thresholds: Poverty

Aggregate if at least 40% of group total population covered by survey +/- 2 years from reference year.

Alignment of country-level estimates in time

Aggregation and population coverage

Inclusion of high income countries

Previously, global poverty estimates assumed that the "high income economies" not in the PovcalNet regions had zero people living in extreme poverty.

At the recommendation of Commission on Global Poverty (World Bank, 2017), we now adopt a truly global approach to poverty measurement and no longer impose this assumption.

From October 2017, the global poverty estimate includes poverty estimates from several high income countries and aggregated for the group “Other High Income Countries”

Allows for global aggregation of poverty also for higher poverty lines (than $1.90) for which the zero poverty assumption unreasonable.

WORLD BANK GROUP