

The European Commission's science and knowledge service

Joint Research Centre



The Essentials of Composite Indicators and Scoreboards

Welcome

William Becker

February 17th 2020

Competence Centre on Composite Indicators and Scoreboards

JRC Ispra - Italy



Support to EU Services
in any policy area



Handbook
on Constructing
**Methodology and
guidelines**

METHODOLOGY
AND USER GUIDE

Launched 02/2016
✓ **17 years of expertise**



Support to international
organisations



Training and
Community of Practice

Competence Centre on Composite Indicators and Scoreboards

In-house development

Measuring Gender Equality in EU regions



ASEM Sustainable Connectivity Portal



The Cultural and Creative Cities Monitor



Social scoreboard

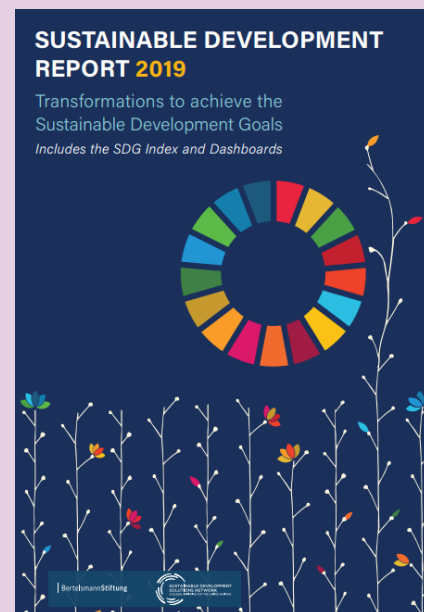
Supporting the European Pillar of Social Rights



Review of indicator frameworks

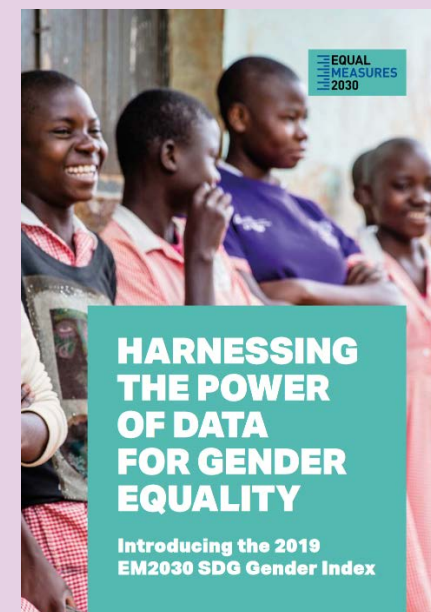
SUSTAINABLE DEVELOPMENT REPORT 2019

Transformations to achieve the
Sustainable Development Goals
Includes the SDG Index and Dashboards



HARNESSING THE POWER OF DATA FOR GENDER EQUALITY

Introducing the 2019
EM2030 SDG Gender Index



100+

Competence Centre on Composite Indicators and Scoreboards

Methods & Tools

COIN Data Hub

One-stop-shop gathering data on 400+ indices

COIN Tool

DIY tool to guide and build composite indicators

Handbook

Step-by-step guidance to build composite indicators

Events & Training

COIN Open Days

One-day workshops with developers

JRC Week on Composite Indicators and Scoreboards

Training and Community of Practice



19-23 Oct (TBC)
Ispra, Italy

Outline of the day

09:00	Welcome: Who we are, outline of the day and a round table
09:20	Introduction to composite indicators and scoreboards: what are (composite) indicators and scoreboards, why use them, policy monitoring, basic steps in building a composite indicator
10:00	Refresher on statistics: mean, median, variance, probability distributions, outliers, correlation, etc.
10:20	Data processing: data treatment (missing data and outliers), normalization, weighting and aggregation.
11:15	Coffee break
11:30	Analysing and summarizing: statistical coherence, robustness and sensitivity, story-telling and visualization.
12:30	Lunch
13:30	COIN Tool practical session



THANK YOU

william.becker@ec.europa.eu
jrc-coin@ec.europa.eu

COIN in the EU Science Hub
<https://ec.europa.eu/jrc/en/coin>

COIN tools are available at:
<https://composite-indicators.jrc.ec.europa.eu/>

The European Commission's
Competence Centre on Composite
Indicators and Scoreboards



The European Commission's science and knowledge service

Joint Research Centre



European
Commission

The European Commission's Competence Centre on Composite Indicators and Scoreboards

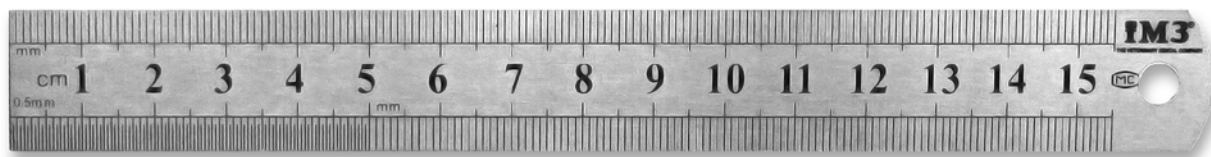
An introduction to composite indicators and scoreboards

William Becker

February 17th 2019

What are indicators?

Length



Weight



Cost



Population



What are indicators?

Indicators are **indirect** measurements of a concept that we can't measure directly (otherwise we would just measure it).

Measured quantity + target concept = indicator

Concepts are typically complex, multi-dimensional, often socio-economic. This means we often need multiple indicators to do the job.

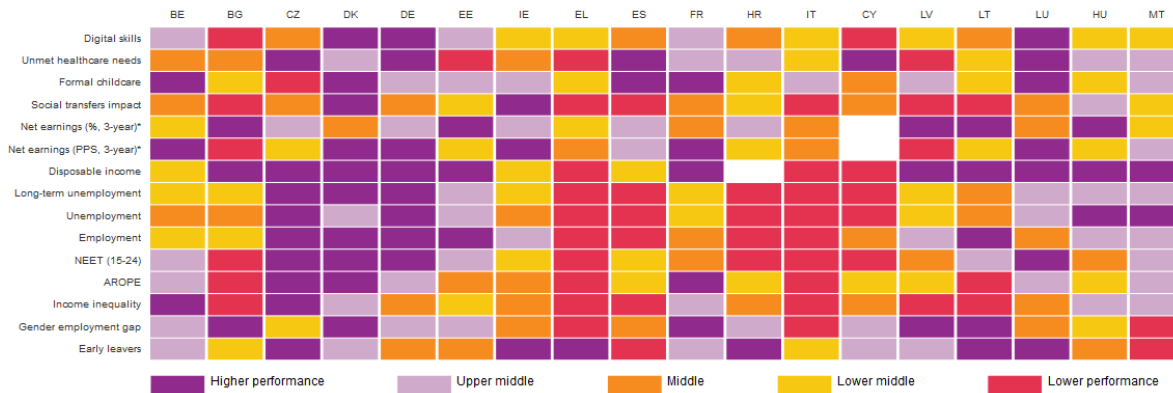
This also gives the **why**: indicators are usually used when there is no direct way to measure.

Consequence: indicators are not a precise science.

What are composite indicators and scoreboards?

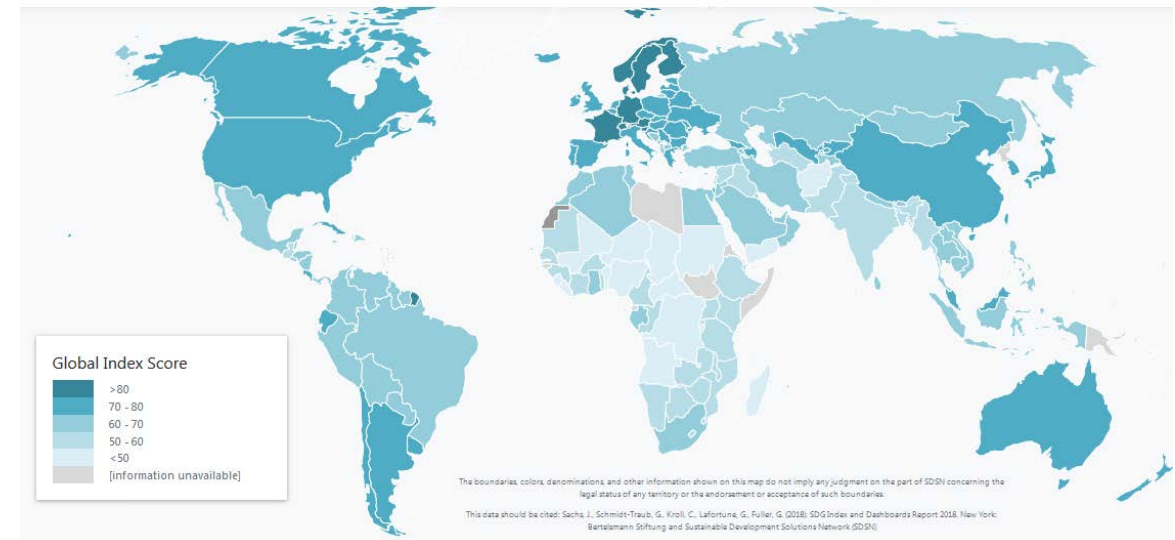
Scoreboards are collections of individual indicators that are related to a common concept.

Composite Indicators are aggregations of individual indicators compiled into a single score.



SDG Index and Dashboards Report 2018

Global Responsibilities: Implementing the goals



A composite indicator is effectively a summary of a (structured) scoreboard

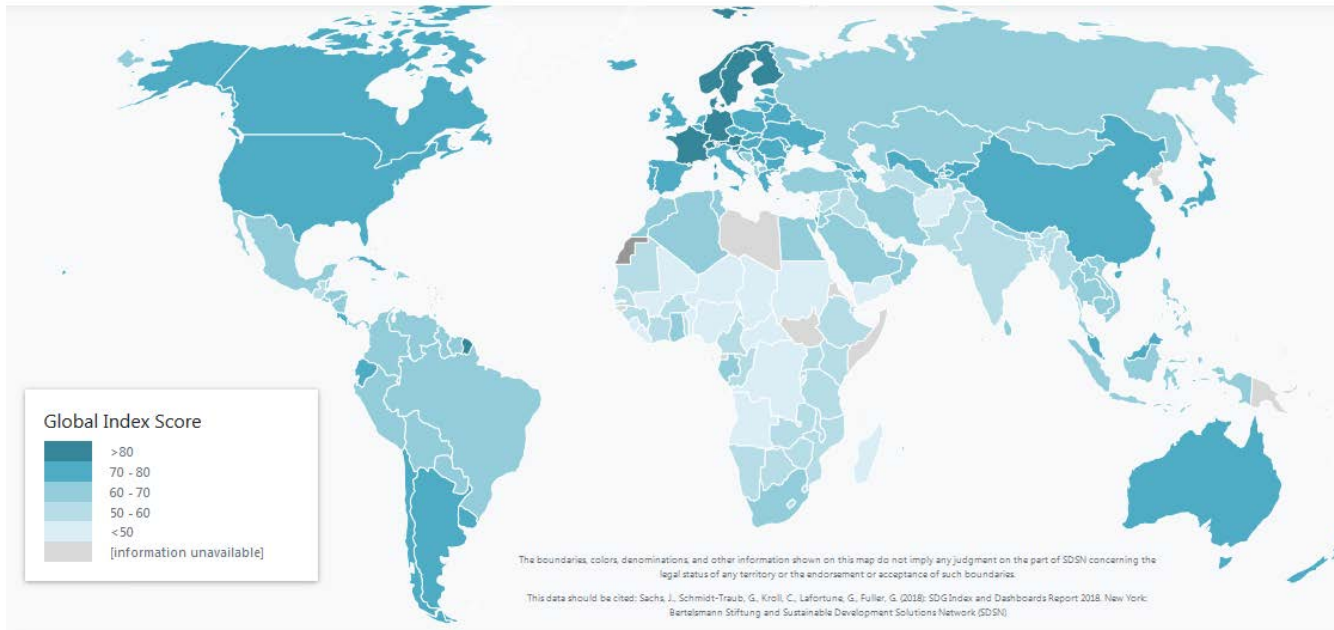
Why aggregate?

SDG Index and Dashboards Report 2018

Global Responsibilities: Implementing the goals



BertelsmannStiftung



You don't have to aggregate (sometimes it is not even recommended), but advantages include:

- o A simplified summary measure which is easy to communicate to stakeholders
- o Ranking of entities according to a complex concept: analysis of the “big picture”

Why aggregate?

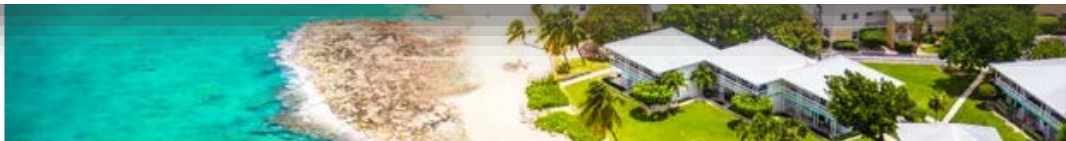
In wake of Brexit, EU to put Cayman Islands on tax haven blacklist

Decision on British overseas territory comes less than two weeks after UK left bloc



Last year the UK and its “corporate tax haven network” was judged to be by far the world’s greatest enabler of corporate tax avoidance by the [Tax Justice Network](#).

British territories and dependencies made up four of the 10 places said to have done the most to “proliferate corporate tax avoidance” on the corporate tax haven index.



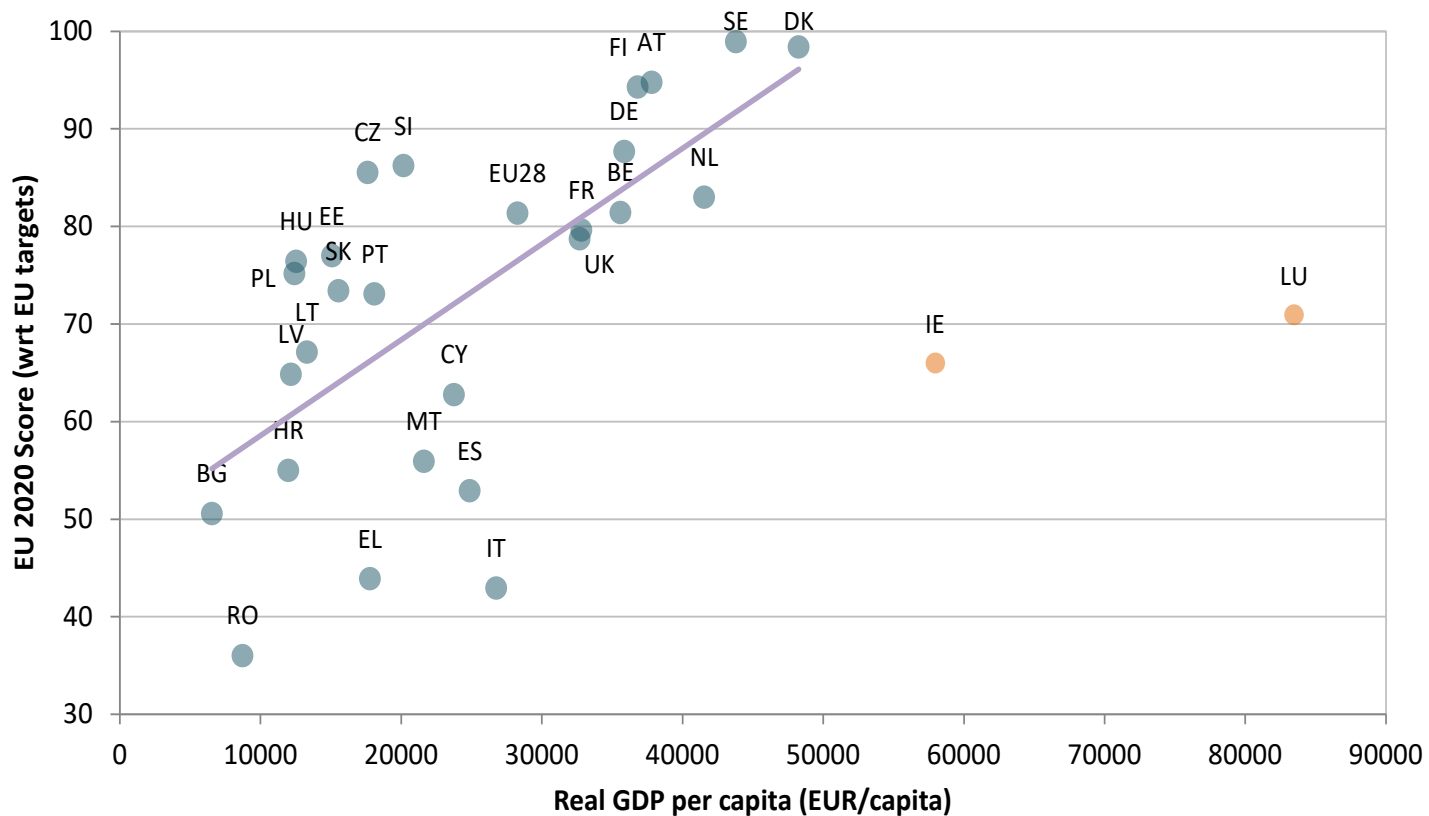
▲ Part of the Cayman Islands coastline. The British overseas territory is to be put on an EU blacklist of tax havens.
Photograph: Jodi Jacobson/Getty Images

The [Cayman Islands](#), a British overseas territory, is to be put on an EU blacklist of tax havens, less than two weeks after the UK’s withdrawal from the bloc.

You don’t have to aggregate (sometimes it is not even recommended), but advantages include:

- o A simplified summary measure which is easy to communicate to stakeholders
- o Ranking of entities according to a complex concept: analysis of the “big picture”
- o Powerful messages for media, advocacy

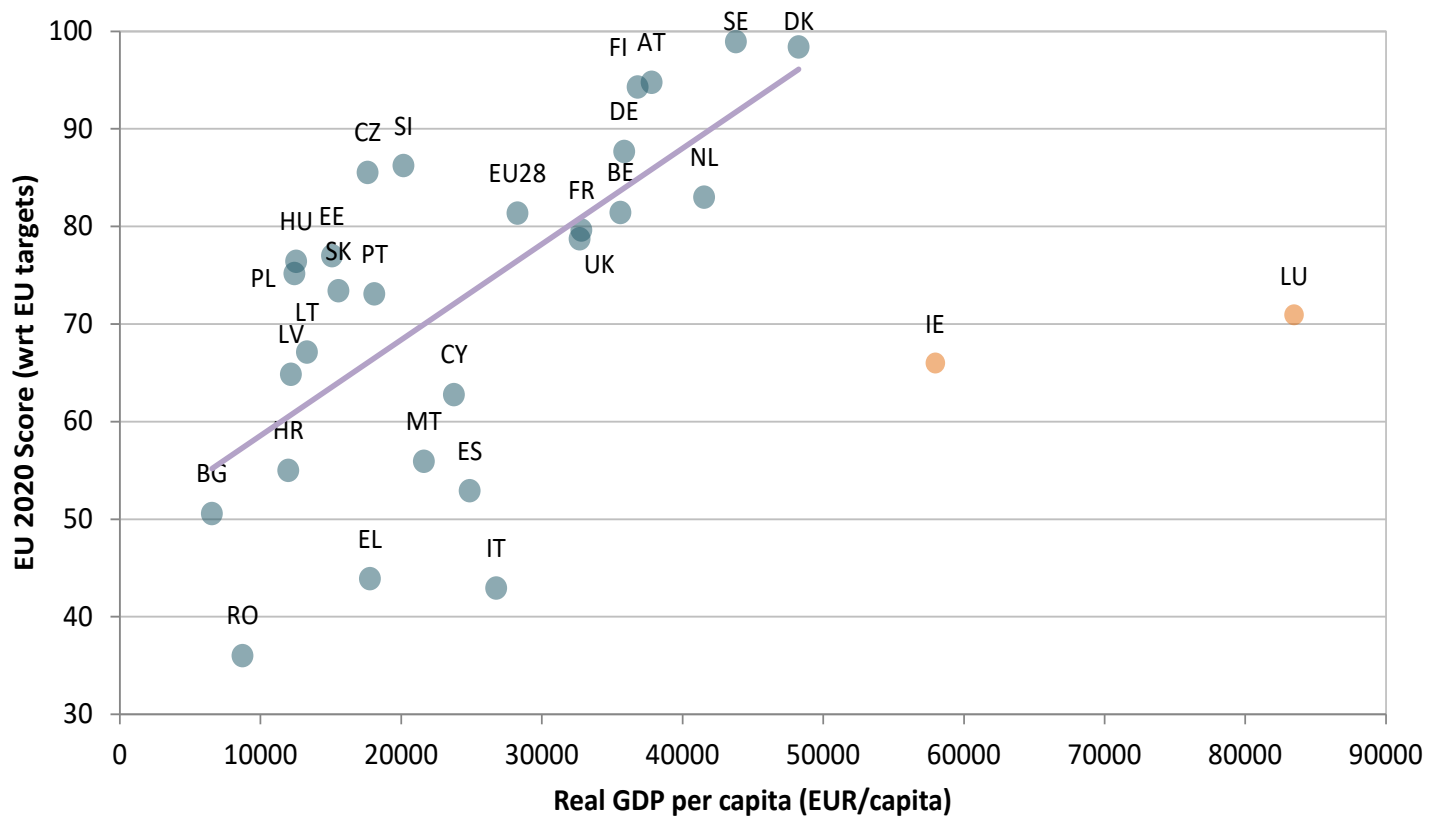
Why aggregate?



You don't have to aggregate (sometimes it is not even recommended), but advantages include:

- o A simplified summary measure which is easy to communicate to stakeholders
- o Ranking of entities according to a complex concept: analysis of the “big picture”
- o Powerful messages for media, advocacy
- o Allows overall comparisons with other measures
- o An access point for a complex data set

Why aggregate?



You don't have to aggregate (sometimes it is not even recommended), but advantages include:

- o A simplified summary measure which is easy to communicate to stakeholders
- o Ranking of entities according to a complex concept: analysis of the “big picture”
- o Powerful messages for media, advocacy
- o Allows overall comparisons with other measures
- o An access point for a complex data set

The Social Scoreboard

- JRC project on request of President Juncker's CAB and SecGen
- Develop an online visualisation tool for the Scoreboard accompanying the European Pillar of Social Rights
- A compass for better working and living conditions

Social Dimensions



Equal opportunities and access to the labour market

Areas

- 01 Education, skills and lifelong learning
- 02 Gender equality in the labour market
- 03 Inequality and upward mobility
- 04 Living conditions and poverty
- 05 Youth



Dynamic labour markets and fair working conditions

Areas

- 06 Labour force structure
- 07 Labour market dynamics
- 08 Income, including employment-related



Public support / Social protection and inclusion

Areas

- 09 Impact of public policies on reducing poverty
- 10 Early childhood care
- 11 Healthcare
- 12 Digital access





map



Country Analysis



Year Heatmap

European Union (28 countries)		
2010	2016	
13.9	10.7	↓
9.3	10.8	↑
***	***	***
33.8	39.1	↑

Country comparison



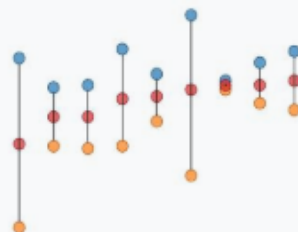
Country heatmap



Time line



Regional map



Regional Analysis



Regional Heatmap

All data is downloadable

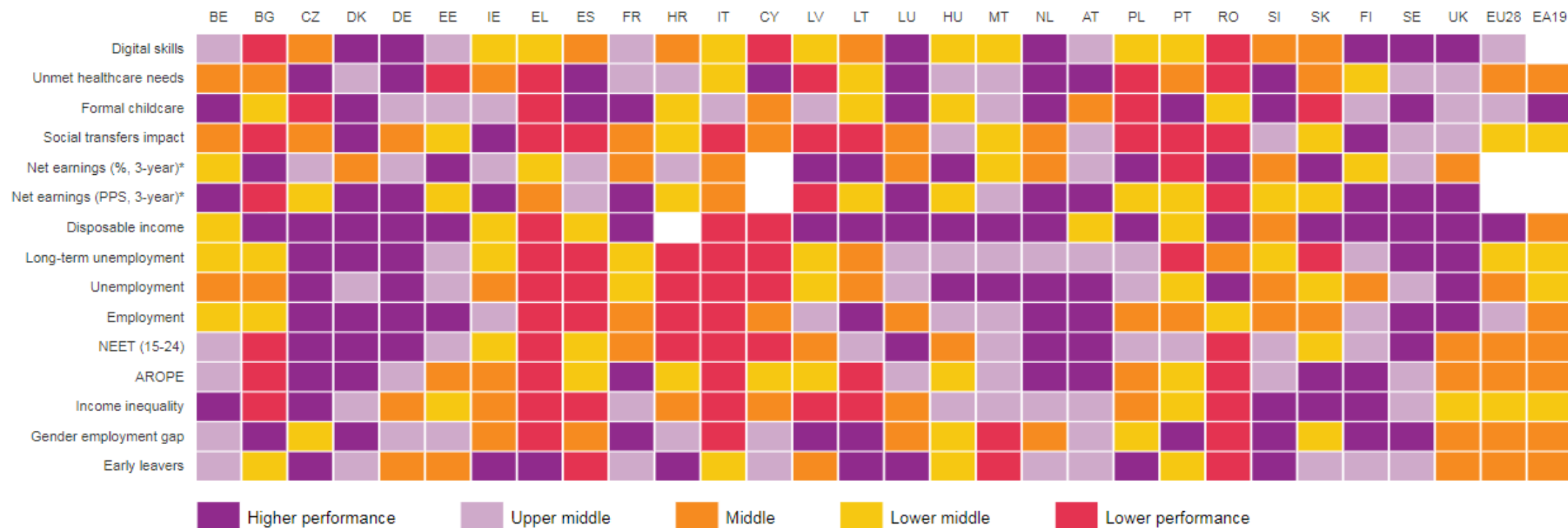
[Home](#)[map](#)[Country analysis](#)[Year heatmap](#)[Comparison table](#)[Country heatmap](#)[Time line](#)[Profile](#)

Select Year



Sort countries by performance

2016: 14 headline indicators



Higher performance

Upper middle

Middle

Lower middle

Lower performance



[Home](#)[map](#)[Country analysis](#)[Year heatmap](#)[Comparison table](#)[Country heatmap](#)[Time line](#)[Profile](#)

Select Year



Sort countries by performance

2016: 14 headline indicators



Cultural and Creative Cities Monitor

- JRC project on request of JRC Director General, in agreement with Commissioner Navracsics
- How can we measure culture and creativity in European cities?
- Support for culture-led development

+22 CITIES compared to the 2017 edition

190

Cultural and Creative cities

98

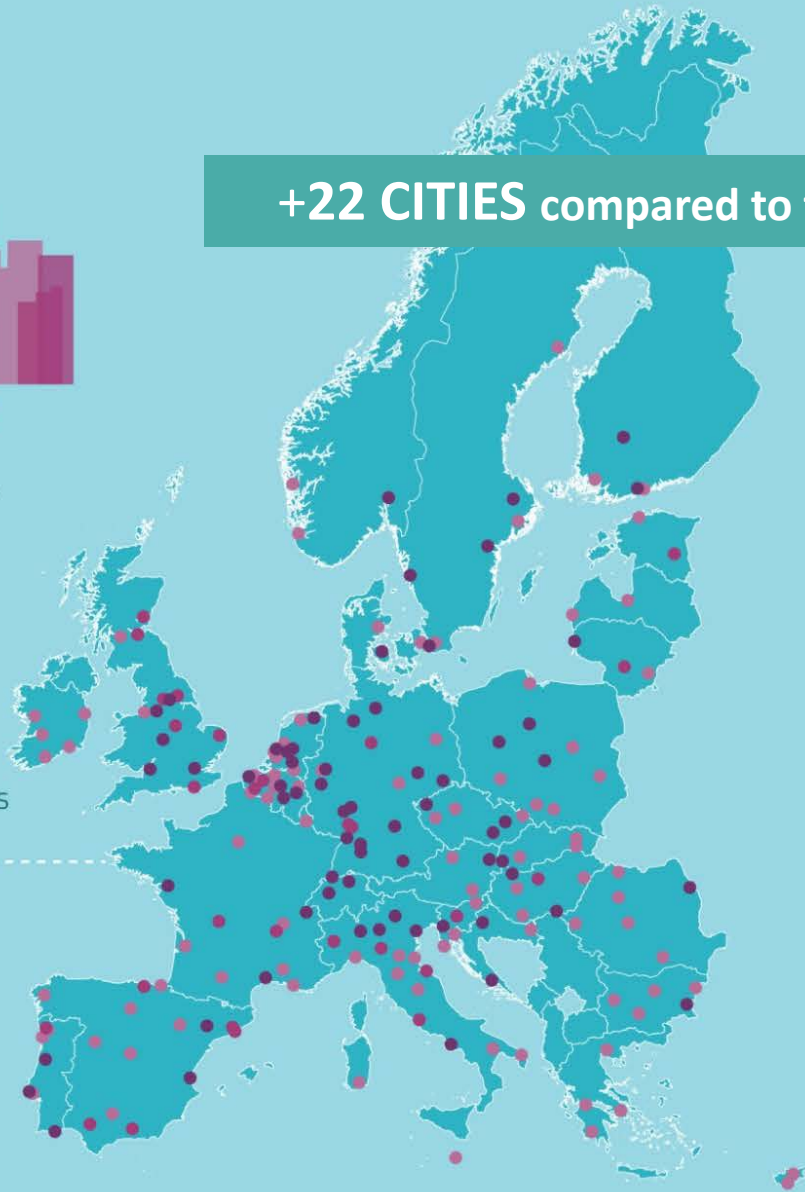
European Capitals of Culture

33

UNESCO Creative Cities

59

Cities hosting at least two international cultural festivals

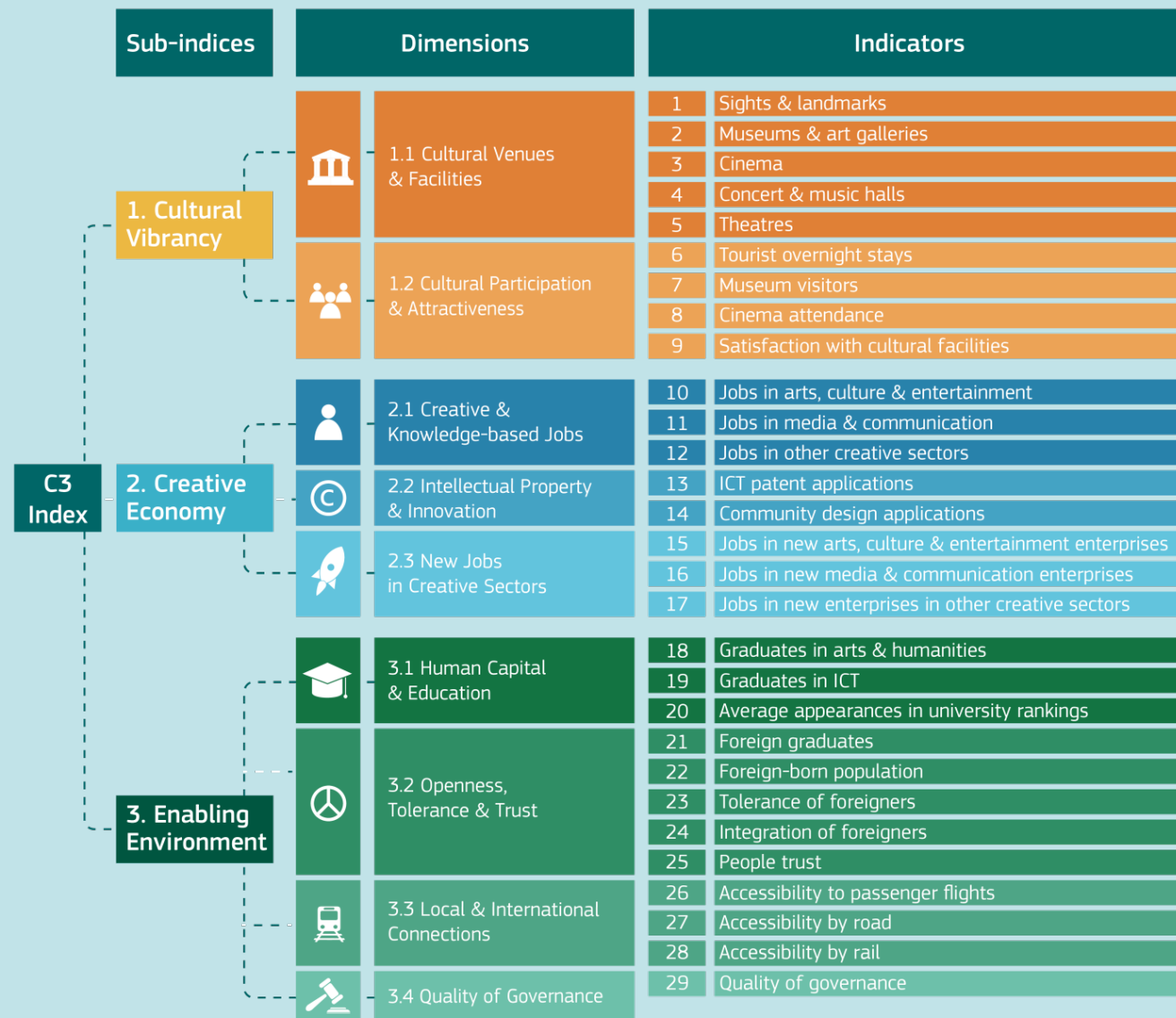


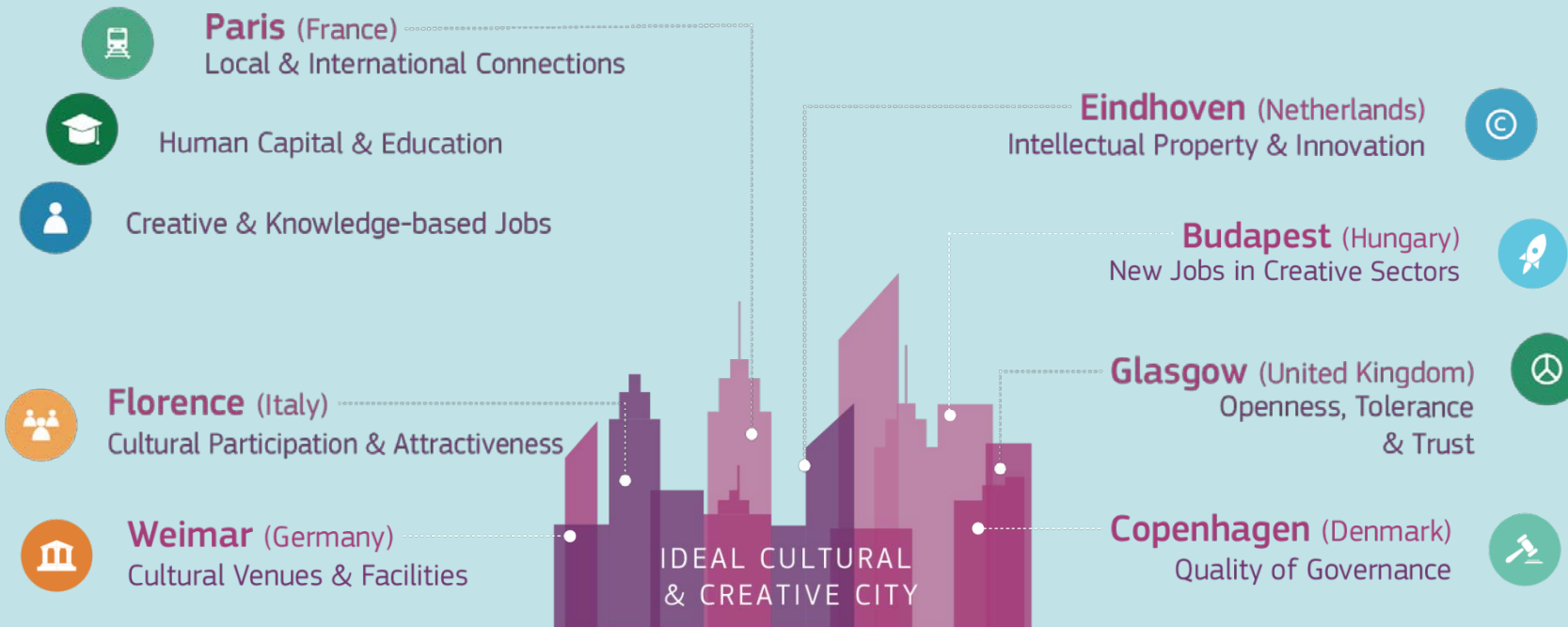
Indicators	
1	Sights & landmarks
2	Museums & art galleries
3	Cinema
4	Concert & music halls
5	Theatres
6	Tourist overnight stays
7	Museum visitors
8	Cinema attendance
9	Satisfaction with cultural facilities
10	Jobs in arts, culture & entertainment
11	Jobs in media & communication
12	Jobs in other creative sectors
13	ICT patent applications
14	Community design applications
15	Jobs in new arts, culture & entertainment enterprises
16	Jobs in new media & communication enterprises
17	Jobs in new enterprises in other creative sectors
18	Graduates in arts & humanities
19	Graduates in ICT
20	Average appearances in university rankings
21	Foreign graduates
22	Foreign-born population
23	Tolerance of foreigners
24	Integration of foreigners
25	People trust
26	Accessibility to passenger flights
27	Accessibility by road
28	Accessibility by rail
29	Quality of governance

Dimensions		Indicators	
	1.1 Cultural Venues & Facilities	1	Sights & landmarks
		2	Museums & art galleries
		3	Cinema
		4	Concert & music halls
		5	Theatres
	1.2 Cultural Participation & Attractiveness	6	Tourist overnight stays
		7	Museum visitors
		8	Cinema attendance
		9	Satisfaction with cultural facilities
	2.1 Creative & Knowledge-based Jobs	10	Jobs in arts, culture & entertainment
		11	Jobs in media & communication
		12	Jobs in other creative sectors
	2.2 Intellectual Property & Innovation	13	ICT patent applications
		14	Community design applications
	2.3 New Jobs in Creative Sectors	15	Jobs in new arts, culture & entertainment enterprises
		16	Jobs in new media & communication enterprises
		17	Jobs in new enterprises in other creative sectors
	3.1 Human Capital & Education	18	Graduates in arts & humanities
		19	Graduates in ICT
		20	Average appearances in university rankings
	3.2 Openness, Tolerance & Trust	21	Foreign graduates
		22	Foreign-born population
		23	Tolerance of foreigners
		24	Integration of foreigners
		25	People trust
	3.3 Local & International Connections	26	Accessibility to passenger flights
		27	Accessibility by road
		28	Accessibility by rail
	3.4 Quality of Governance	29	Quality of governance

Sub-indices	Dimensions		Indicators	
1. Cultural Vibrancy		1.1 Cultural Venues & Facilities	1	Sights & landmarks
			2	Museums & art galleries
			3	Cinema
			4	Concert & music halls
			5	Theatres
		1.2 Cultural Participation & Attractiveness	6	Tourist overnight stays
			7	Museum visitors
			8	Cinema attendance
			9	Satisfaction with cultural facilities
		2.1 Creative & Knowledge-based Jobs	10	Jobs in arts, culture & entertainment
		2.2 Intellectual Property & Innovation	11	Jobs in media & communication
			12	Jobs in other creative sectors
		2.3 New Jobs in Creative Sectors	13	ICT patent applications
			14	Community design applications
			15	Jobs in new arts, culture & entertainment enterprises
			16	Jobs in new media & communication enterprises
			17	Jobs in new enterprises in other creative sectors
		3.1 Human Capital & Education	18	Graduates in arts & humanities
		3.2 Openness, Tolerance & Trust	19	Graduates in ICT
			20	Average appearances in university rankings
			21	Foreign graduates
			22	Foreign-born population
			23	Tolerance of foreigners
		3.3 Local & International Connections	24	Integration of foreigners
			25	People trust
			26	Accessibility to passenger flights
		3.4 Quality of Governance	27	Accessibility by road
			28	Accessibility by rail
			29	Quality of governance

Sub-indices	Dimensions		Indicators	
1. Cultural Vibrancy		1.1 Cultural Venues & Facilities	1	Sights & landmarks
			2	Museums & art galleries
			3	Cinema
			4	Concert & music halls
			5	Theatres
		1.2 Cultural Participation & Attractiveness	6	Tourist overnight stays
			7	Museum visitors
			8	Cinema attendance
			9	Satisfaction with cultural facilities
2. Creative Economy		2.1 Creative & Knowledge-based Jobs	10	Jobs in arts, culture & entertainment
			11	Jobs in media & communication
			12	Jobs in other creative sectors
		2.2 Intellectual Property & Innovation	13	ICT patent applications
			14	Community design applications
		2.3 New Jobs in Creative Sectors	15	Jobs in new arts, culture & entertainment enterprises
			16	Jobs in new media & communication enterprises
			17	Jobs in new enterprises in other creative sectors
		3.1 Human Capital & Education	18	Graduates in arts & humanities
			19	Graduates in ICT
			20	Average appearances in university rankings
		3.2 Openness, Tolerance & Trust	21	Foreign graduates
			22	Foreign-born population
			23	Tolerance of foreigners
			24	Integration of foreigners
			25	People trust
		3.3 Local & International Connections	26	Accessibility to passenger flights
			27	Accessibility by road
			28	Accessibility by rail
		3.4 Quality of Governance	29	Quality of governance







BRATISLAVA

(SLOVAKIA)

GDP

1

Population

3

Employment

1

Groups of peers

Group by:

POP

GDP

EMP

European Capital of culture

UNESCO Creative city

International Cultural festivals

Selection criteria

Rank by:

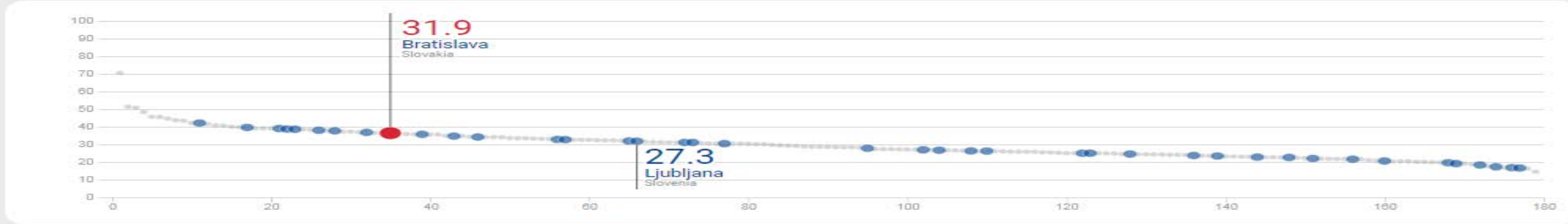
Cultural and Creative Cities Monitor



Index

Sub indices

Dimensions



Performance by Dimension

POP

GDP

EMP

Sub index performance

 Cultural Vibrancy

25.7

 Creative Economy

44.4

 Enabling Environment

19.3

Dimension performance

 Cultural Venues & Facilities

28.9

 Creative & Knowledge Based Jobs

35.3

 Human Capital & Education

24.6

 Cultural Participation & Attractiveness

22.5

 Intellectual Property & Innovation

7.4

 Openness, Tolerance & Trust

12.5

 New Jobs in Creative Sectors

72

 Local & International Connections

20.3

 Quality of Governance

28.6



City profiles



Country profiles



Docs & data



Create your own Monitor



Add your own city



Policy scenario builder



Provide data



"Did you know that...?"



Cultural gems



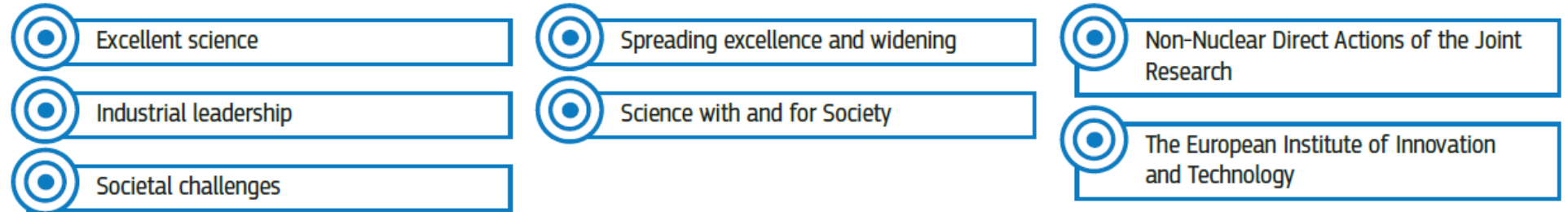
Performance framework



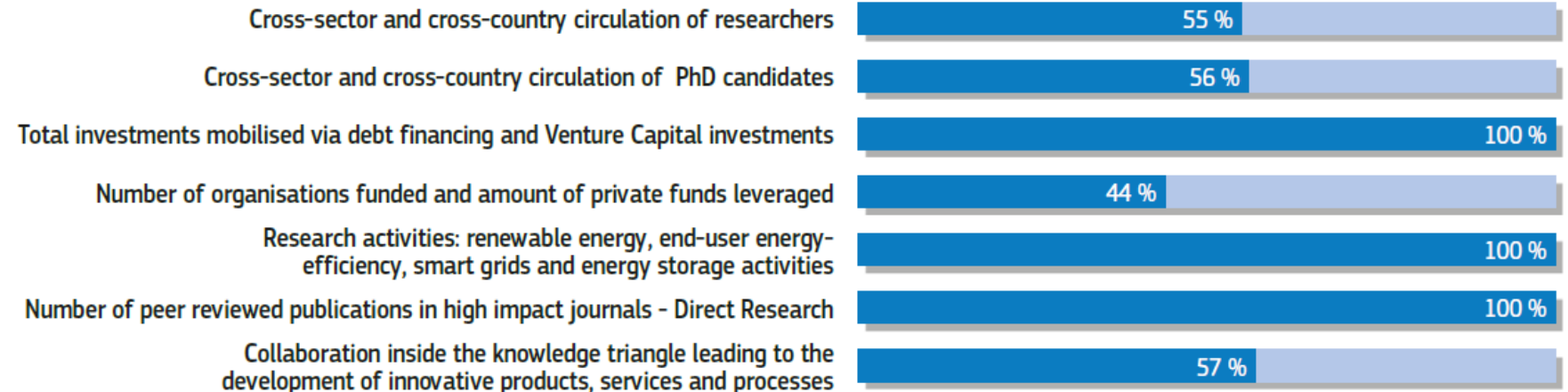
GENERAL OBJECTIVE

- To build a society and an economy based on knowledge and innovation across the whole Union, while contributing to sustainable development

SPECIFIC OBJECTIVES / STRUCTURE OF THE PROGRAMME



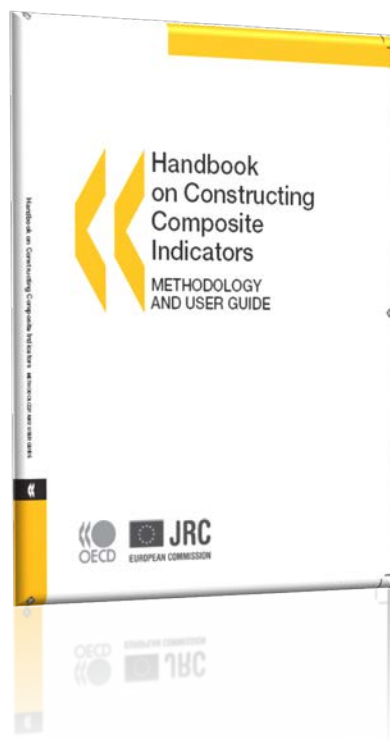
HORIZON 2020 PROGRESS TOWARDS THE TARGET (2020)



■ % of target achieved by 2017

Horizon 2020 Programme Monitoring

Steps to building a composite indicator



<https://composite-indicators.jrc.ec.europa.eu/>

Step 10. Visualisation and communication

Step 9. Make sense of the data

Step 8. Robustness and uncertainty analysis

Step 7. Statistical and conceptual coherence

Step 6. Aggregation

Step 5. Weighting

Step 4. Normalisation – bring indicators onto a common scale

Step 3. Data analysis and treatment

Step 2. Data and indicator selection

Step 1. Conceptual framework

STEP 0: DEFINE OBJECTIVES

Objectives: ask yourself



What exactly are you trying to measure?

Who will use the end product?



What kinds of **questions** should your framework be able to answer?

What level of **detail**: EU-level, National, Regional, Other –
e.g. company, university, etc.



Monitoring what?



State monitoring

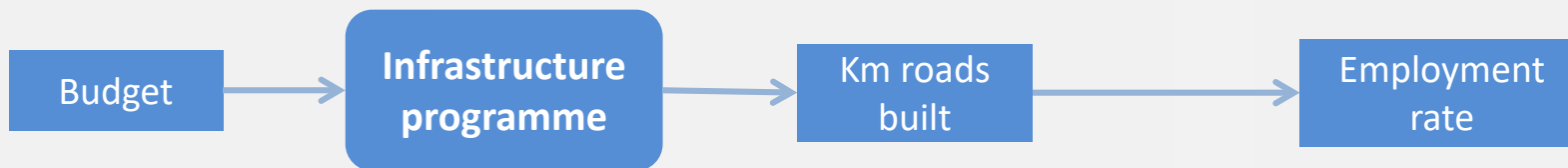
Measure the “state” of a concept in each country/unit. Enables comparisons, trends, etc.

Policy effect monitoring

Measure the impacts/effects of a policy or programme on society



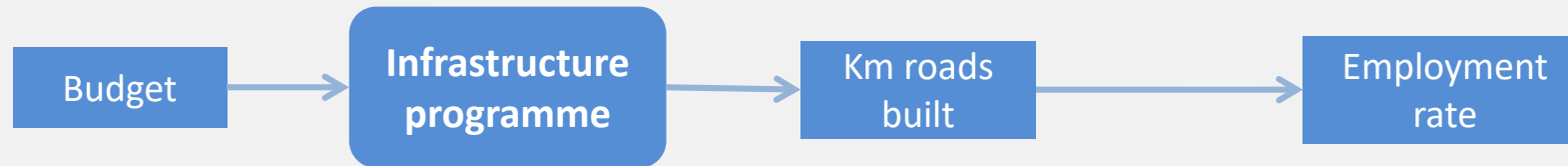
Mix of indicators and **impact evaluation**



Type	Input		Output		Impact
Time-lag	Near-instant		1-2 years		3-5+ years
Causality	Very strong		Strong		Weak
Relevance	Low		Medium		High

Monitoring

Impact evaluation



Type	Input		Output		Impact
Time-lag	Near-instant		1-2 years		3-5+ years
Causality	Very strong		Strong		Weak
Relevance	Low		Medium		High

Monitoring

Impact evaluation

Monitoring

Budget

Km roads
built

- Indicator frameworks and composite indicators
- Input and output indicators, some impact indicators
- *Indicates* performance – a proxy measure for impacts

We can't really monitor impacts due to time lags and external effects

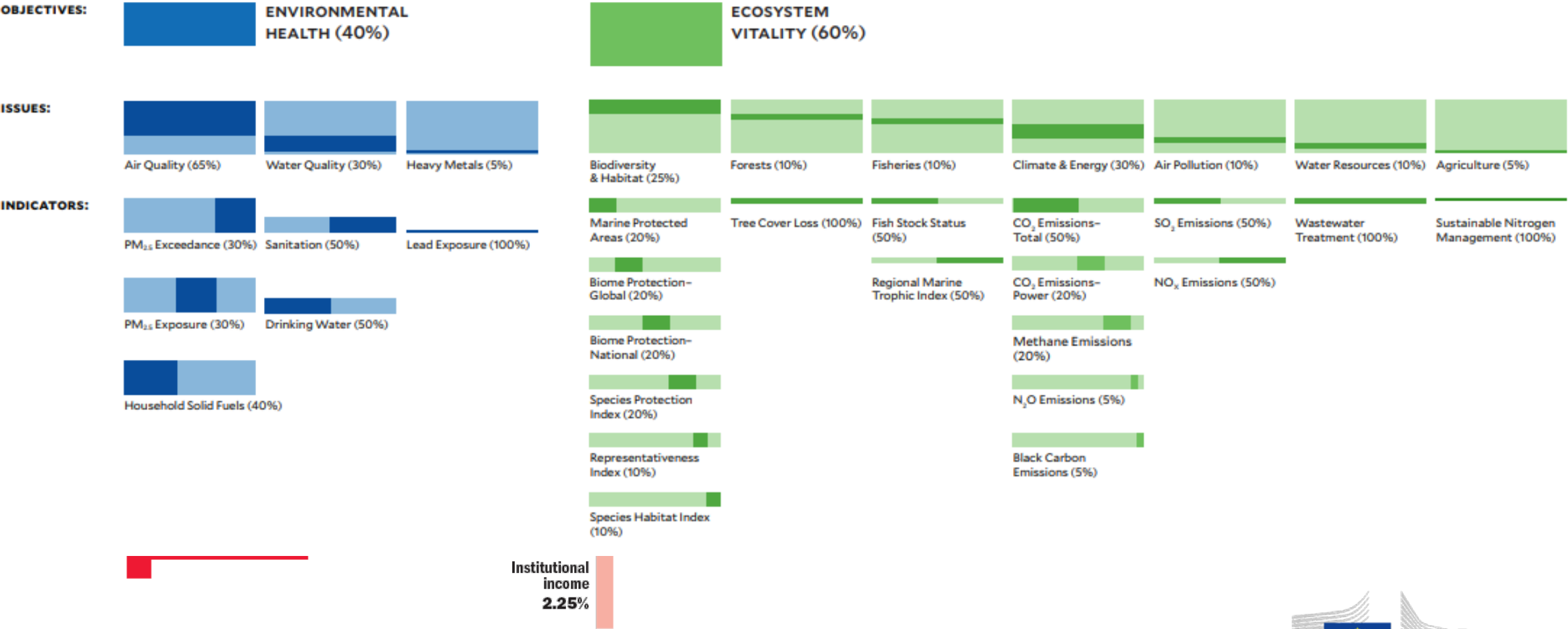
Impact evaluation

Employment
rate

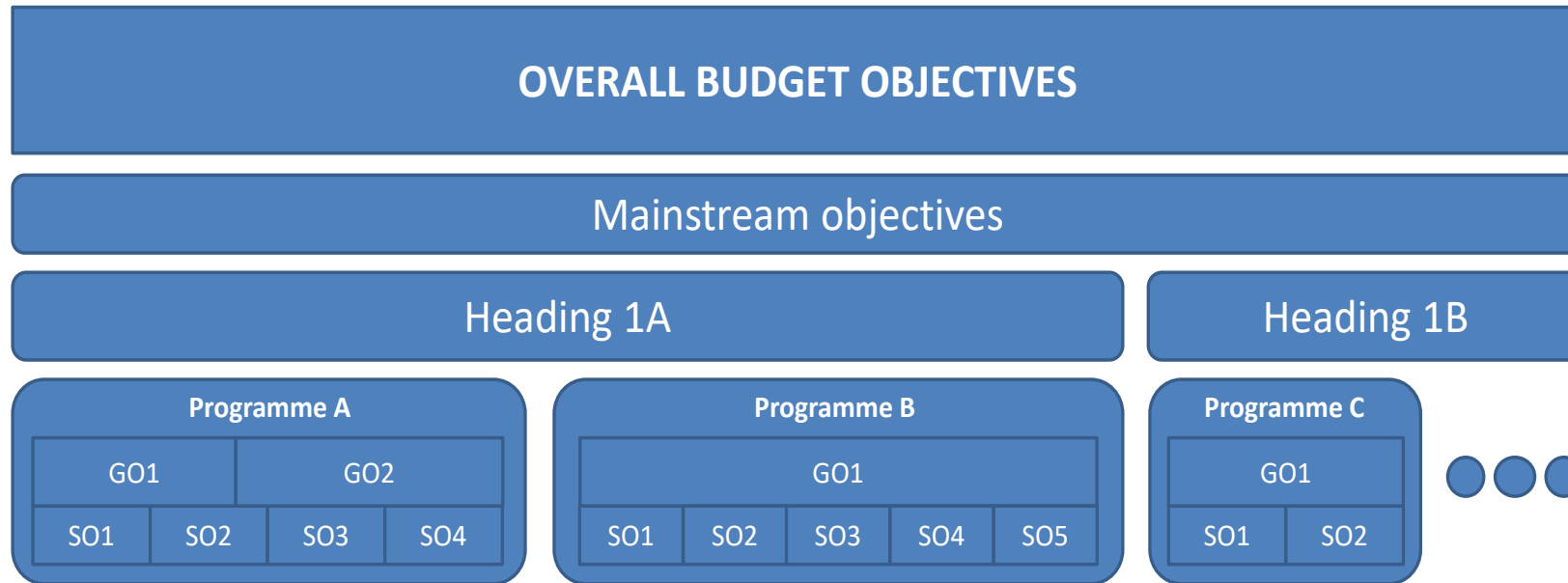
- Focuses on a specific variable which represents programme objectives
- Counterfactual impact evaluation
- Requires micro-data, not EU or national level
- Aims to *measure* performance, not *indicate*
- Retrospective, data/time/skill-intensive

Separates effects/impacts from programme from external effects using statistical techniques

State monitoring frameworks



Performance monitoring framework



Objectives



Indicators

To monitor performance you need:

- A framework of objectives
- A matching framework of indicators
- (ideally) Baselines and targets for each indicator

Step 1: Conceptual framework

Concept mapping

Box 1: Definition of ASEM connectivity adopted by ASEM

Connectivity is about bringing countries, people and societies closer together. It facilitates access and is a means to foster deeper economic and people-to-people ties. It encompasses the hard and soft aspects, including the physical and institutional social-cultural linkages that are the fundamental supportive means to enhance the economic, political-security, and socio-cultural ties between Asia and Europe which also contribute to the narrowing of the varying levels of development and capacities.

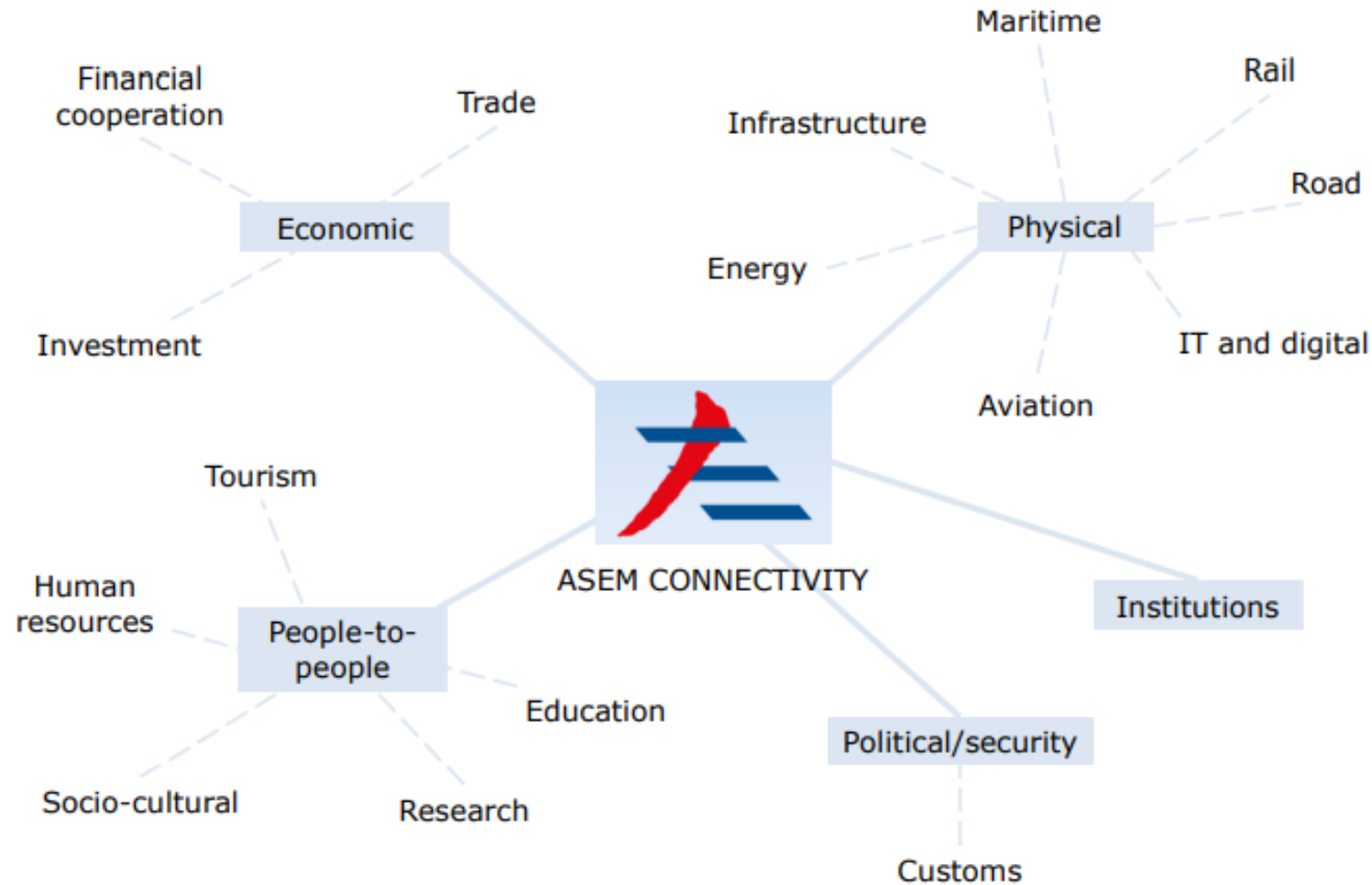
Bearing in mind the Asia-Europe Cooperation Framework (AECF) 2000, ASEM connectivity aims to establish the sense of building ASEM partnership of shared interests. It upholds the spirit of peace, development, cooperation and mutual benefit. It will also adhere to and effectively implement relevant international norms and standards as mutually agreed by ASEM partners.

ASEM Connectivity covers all modes of transport (aviation, maritime, rail and road) and also includes, among others, institutions, infrastructure, financial cooperation, IT, digital links, energy, education and research, human resources development, tourism, cultural exchanges as well as customs, trade and investment facilitation.

ASEM connectivity covers all the three pillars of ASEM - economic, political and socio-cultural. It should be result-oriented, and in support of the following key principles: level playing field, free and open trade, market principles, multi-dimensionality, inclusiveness, fairness, openness, transparency, financial viability, cost-effectiveness and mutual benefits. It should also contribute to the materialisation of the principles, goals and targets of The 2030 Agenda for Sustainable Development. Sustainability is one of the important quality benchmarks for the connectivity initiatives in the ASEM context.

Step 1: Conceptual framework

Concept mapping



Step 1: Conceptual framework

Literature review + Index/scoreboard review

Index	Unit of analysis	Scope	Dimension											
			Economic	Political	People-to-people						Environmental	Sectoral/Physical		
					Communication	Culture	Education	Research	Tourism	Migration		Transport	Energy	ICT
KOF Index of Globalisation	Country	Globalisation	✓	✓	✓	✓	✗	✗	✗	✗	✗	✗	✗	✓
A.T. Kearney Index of Global Cities	City	Globalisation	✓	✓	✓	✓	✓	✗	✓	✓	✗	✓	✗	✓
McKinsey Global Institute Connectedness Index	Country	Globalisation	✓	✗	✓	✗	✗	✓	✓	✓	✗	✓	✓	✓
DHL Global Connectedness 2016	Country	Globalisation	✓	✗	✓	✗	✗	✗	✓	✓	✗	✗	✗	✓
Huawei Global Connectivity Index	Country	ICT Globalisation	✓	✗	✓	✗	✗	✓	✗	✗	✗	✗	✗	✓
The Economist ICT Globalisation Index	Country	ICT Globalisation	✓	✗	✓	✗	✗	✓	✗	✗	✗	✗	✗	✓
The Connectivity Scorecard	Country	ICT Globalisation	✗	✗	✓	✗	✓	✓	✗	✗	✗	✗	✗	✓
Maastricht Globalisation Index	Country	Globalisation	✓	✓	✓	✗	✗	✗	✓	✓	✓	✗	✗	✓
New Globalisation Index	Country	Globalisation	✓	✓	✓	✗	✓	✓	✓	✓	✓	✗	✗	✓
A.T. Kearney/Foreign Policy Globalisation Index	Country	Globalisation	✓	✓	✓	✗	✗	✗	✓	✗	✗	✗	✗	✓
CSGR Globalisation Index	Country	Globalisation	✓	✓	✓	✓	✗	✗	✓	✓	✗	✗	✗	✓
Coverage by dimension			91%	55%	100%	27%	27%	45%	64%	55%	18%	18%	9%	100%

Step 1: Conceptual framework

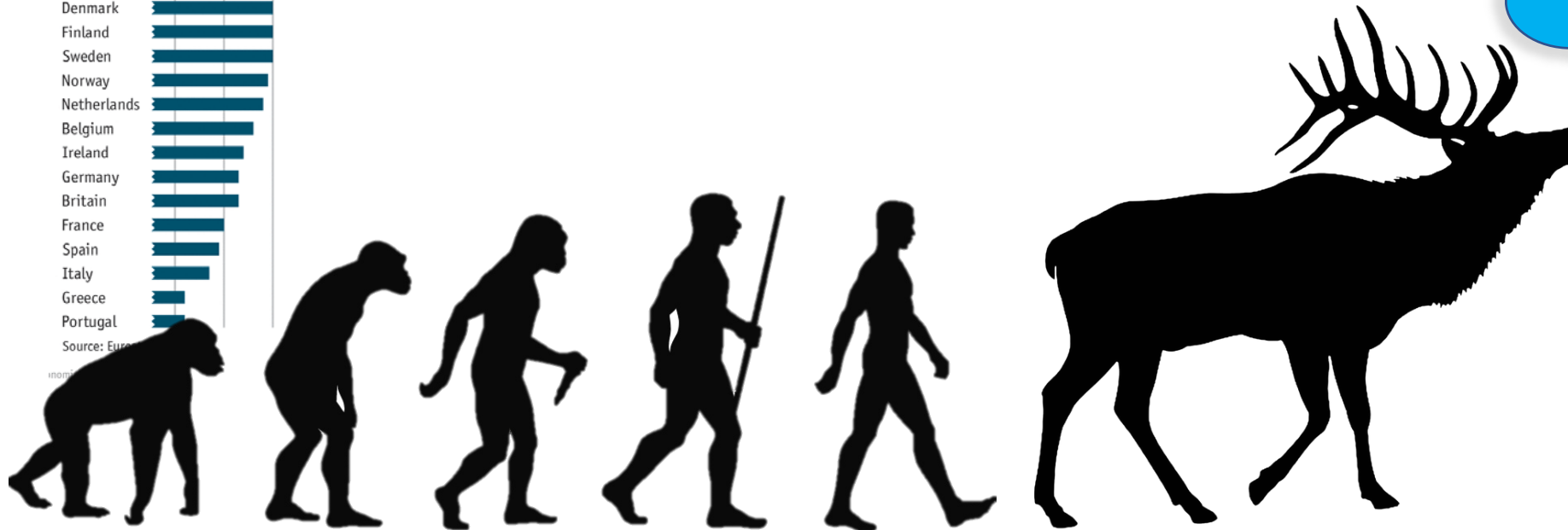
Added value

Happiness in Europe

Self-reported, 2013, 10=most happy



Source: Eurostat



**= How
Scandinavian
are you?**

Step 1: Conceptual framework

Back to the objectives



What exactly are you trying to measure?

Who will use the end product?



What kinds of **questions** should your framework be able to answer?

What level of **detail**: EU-level, National, Regional, Other –
e.g. company, university, etc.



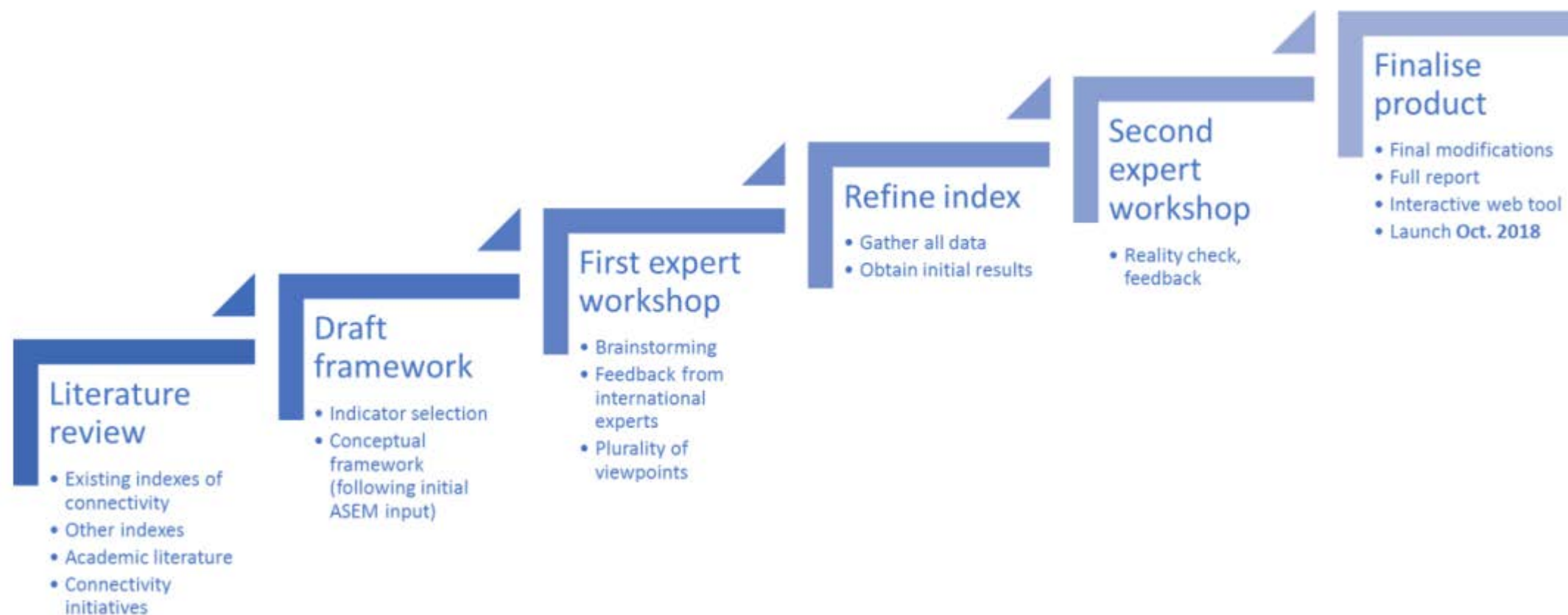
Step 1: Conceptual framework

Draft framework



Step 1: Conceptual framework

Expert workshops and iterations



Step 2: Indicator selection

Better Regulation guidelines on monitoring

- **Comprehensive:** *The monitoring system put in place must cover the objectives of the intervention. In other words, indicators used to monitor programmes should be related to the programme objective(s).*
- **Proportionate:** *the system put in place needs to reflect the importance placed on different aspects of the intervention.*
- **Minimise overlap:** *avoid duplication and the creation of unnecessary data collection burdens by concentrating only on the gaps which need to be filled.*
- **Timeliness:** *Consideration also needs to be given to when the evidence will be used – different legal requirements may play a role not just in deciding what is needed, but also the when it is collected.*
- **Accessibility:** *in principle, all evidence gathered should be made available to the general public..*

Step 2: Indicator selection

RACER criteria on indicators

- **Relevant**, i.e. closely linked to the objectives to be reached. They should not be overambitious and should measure the right thing (e.g. a target indicator for health care could be to reduce waiting times but without jeopardising the quality of care provided).
- **Accepted** (e.g. by staff, stakeholders). The role and responsibilities for the indicator need to be well defined (e.g. if the indicator is the handling time for a grant application and the administrative process is partly controlled by Member States and partly by the EU then both sides would assume only partial responsibility).
- **Credible** for non-experts, unambiguous and easy to interpret. Indicators should be simple and robust as possible. If necessary, composite indicators might need to be used instead – such as country ratings, well-being indicators, but also ratings of financial institutions and instruments. These often consist of aggregated data using predetermined fixed weight values. As they may be difficult to interpret, they should be used to assess broad context only.
- **Easy** to monitor (e.g. data collection should be possible at low cost).
- **Robust** against manipulation (e.g. administrative burden: If the target is to reduce administrative burdens to businesses, the burdens might not be reduced, but just shifted from businesses to public administration).

Indicator criteria

State monitoring

- ✓ Relevance
- ✓ Data Availability
- ✓ Data timeliness
- ✓ Reliability
- ✓ Value added
- ✓ Context and metadata
- ✓ Intensive/extensive
- ✓ Granularity



Indicator criteria

State monitoring

- ✓ Relevance
- ✓ Data Availability
- ✓ Data timeliness
- ✓ Reliability
- ✓ Value added
- ✓ Context and metadata
- ✓ Intensive/extensive
- ✓ Granularity

Performance monitoring (additional)

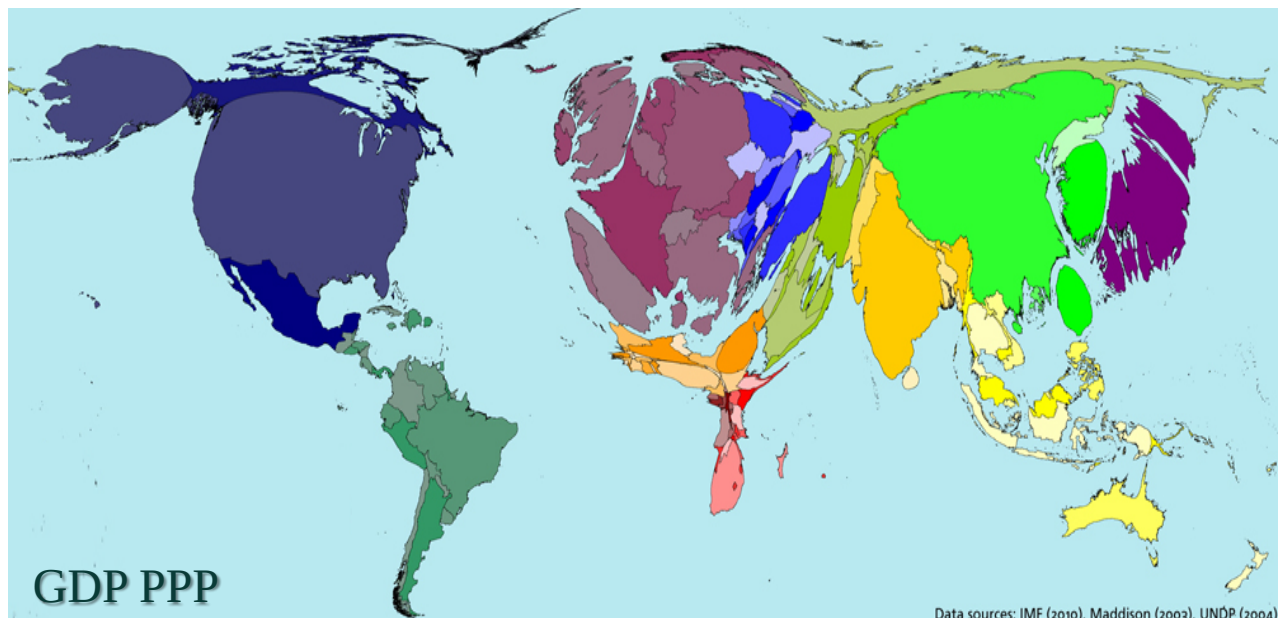
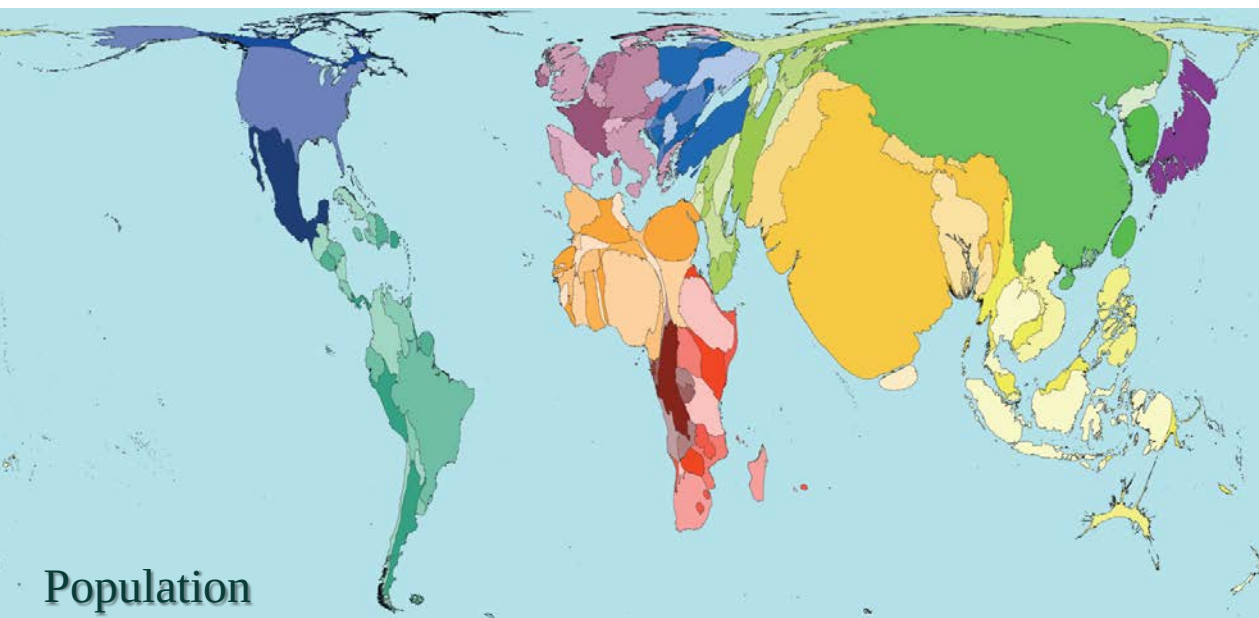
- ✓ Matched objectives, targets, baselines
- ✓ Attributable (causal link)
- ✓ Focused
- ✓ Robust
- ✓ Reactive
- ✓ Legal issues (related to availability)

At the framework level:

- ✓ Fewer indicators is usually better
- ✓ Balanced number of indicators

Statistical steps

- Denomination (intensive vs. extensive)



Data sources: IMF (2010), Maddison (2003), UNDP (2004)

Statistical steps

- Denomination (intensive vs. extensive)
- Imputation (estimate missing data points)
- Data treatment (transform data to improve scaling)
- Normalisation (put indicators onto a common scale)
- Weighting and aggregation (a variety of methods available)
- Correlation analysis, PCA (check statistical properties of index)
- Uncertainty analysis (robustness of your results to changes in methodology)

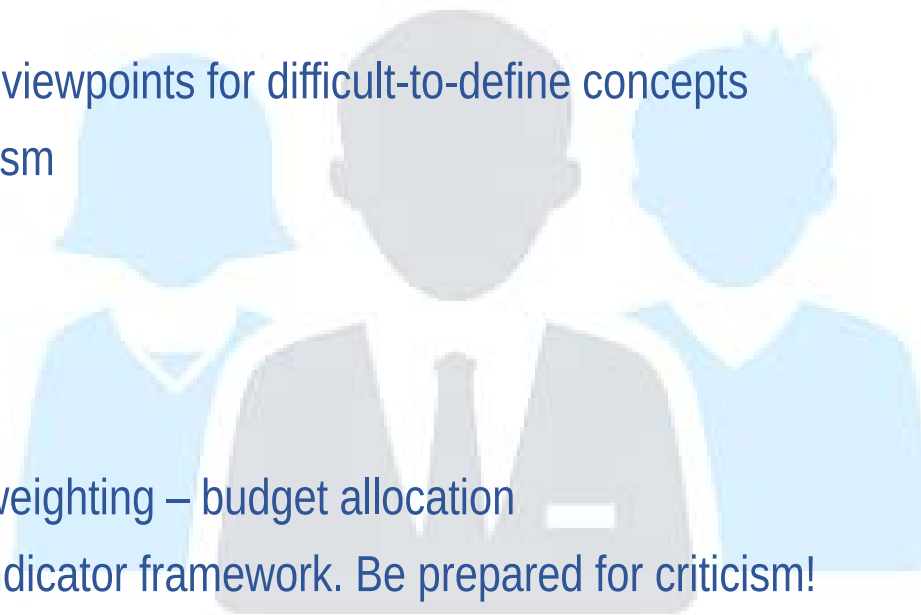
Participation

Stakeholder and expert engagement is extremely important at all stages, if possible.

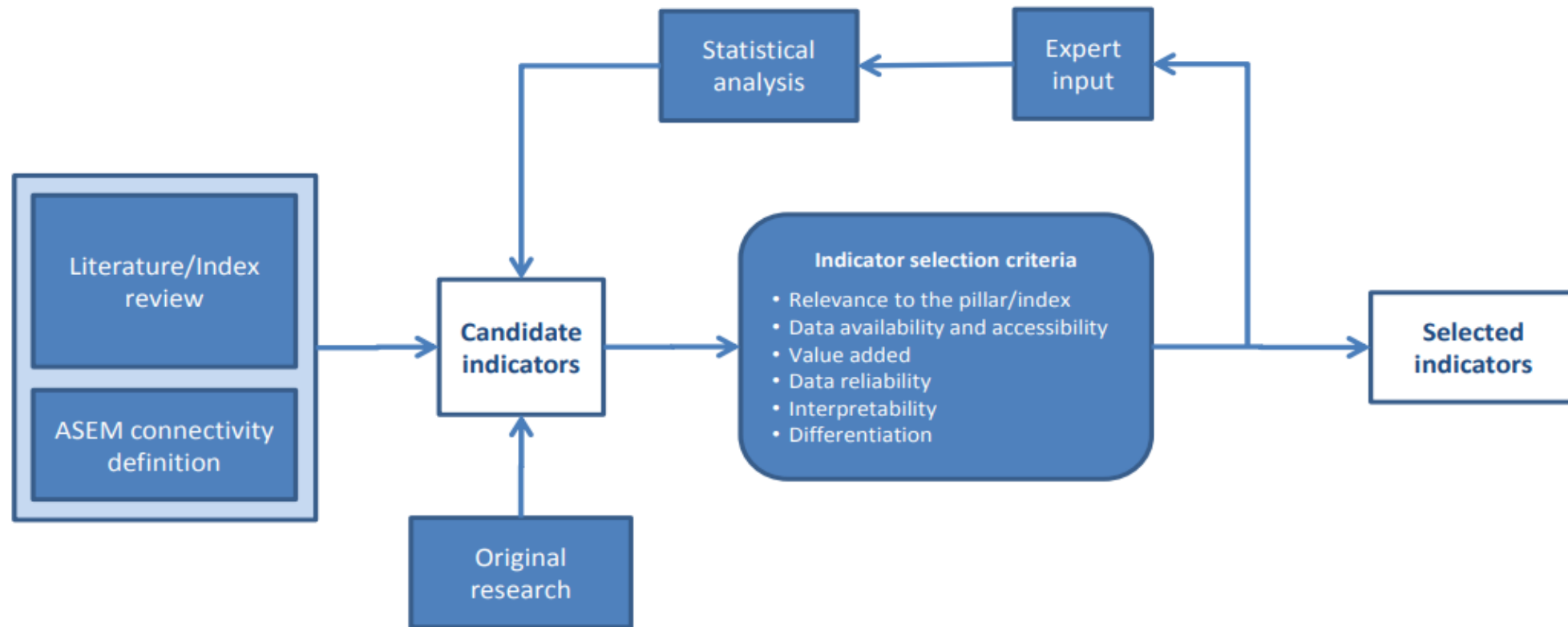
- Improves quality
- Adds legitimacy – plurality of viewpoints for difficult-to-define concepts
- Helps deflect/anticipate criticism

How to do this

- Expert workshops
- Participatory approaches to weighting – budget allocation
- Present early prototypes of indicator framework. Be prepared for criticism!



Step 1: Conceptual framework



Summary: Stumbling blocks

Time – it takes longer than you think, to do a good job

Uncertainty and subjective decisions:

How do we define the concept?

Which indicators to include? Do the indicators equally cover each dimension of the concept?

How to deal with missing data?

How do we weight and aggregate?

Causality issues if you are dealing with policy/programme monitoring

Gaming – think carefully about the incentives created by the indicators/index

Expectations – match them with reality.



COIN Data Hub

Human Capital Index (World Bank) World Happiness Ranking (UN SDSN)

Education for all Development Index (UNESCO)

Global Slavery Index (Walk Free Foundation)

Human Development Index (UNDP)

World Press Freedom Index (Reporters Without Borders)

Better Life Index (OECD)

Living Planet Index (WWF)

Global Innovation Index (WIPO, INSEAD, Cornell)

Best Countries for Business (Forbes)

World Health Report (WHO)





Visualisation and analysis

Granularity

Trend analysis

Priorities for intervention

...



Culture / Creativity

The 2-way (environment $\times$ sex) main effect was significant for both variables ($p < 0.05$). Males expressed more activity in the open field than females, and this was more pronounced in the enriched environment.

[illegible]

☐ A. *Staphylococcus aureus* (Gram-positive cocci in clusters)

PHILLIPS

© 2000 Blackwell Science Ltd

ENVIRONMENTAL

(C) 1980-1981

TYPE OF LINE TO POLARIZE

☐ **APPROPRIATE CONTROLS** (1)

2010

Adrian Davies Development Index (ADDI)

Copyright © 2004 John Wiley & Sons, Inc.

100

1000000

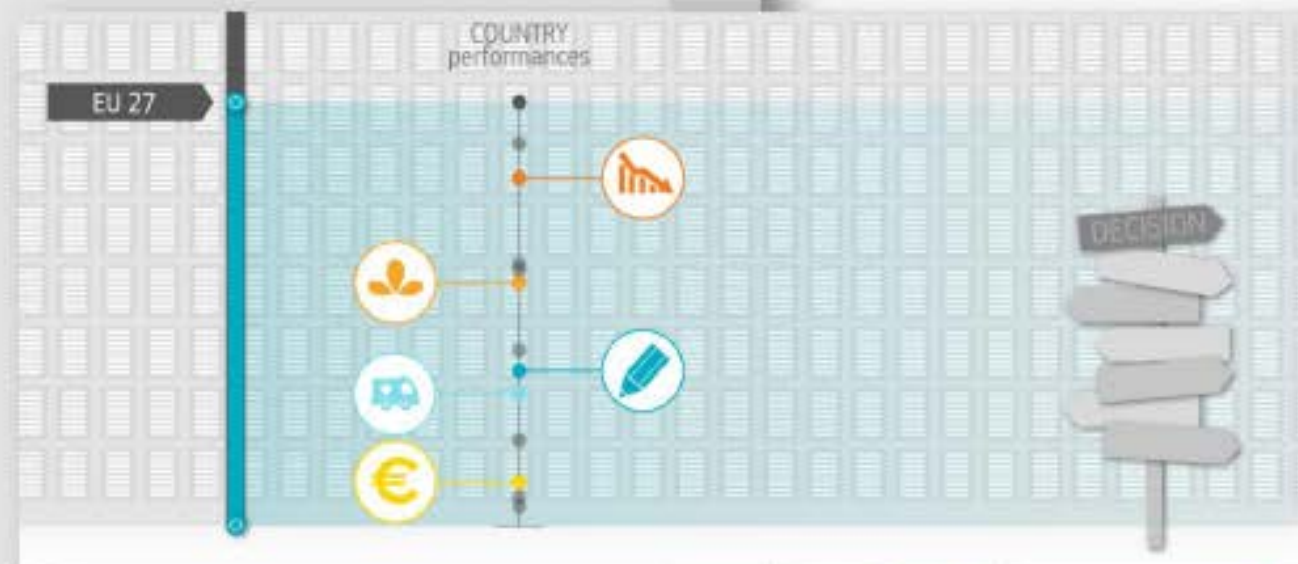
2003

Agencies Participating in the Study

Fuller, Twiss (1972); Tuck, Rees (1974) (see also references); (1) the
affinity (2) = 4.3 (continued, 172)

1999

2000



Austria

Select the Country ▼

EU28

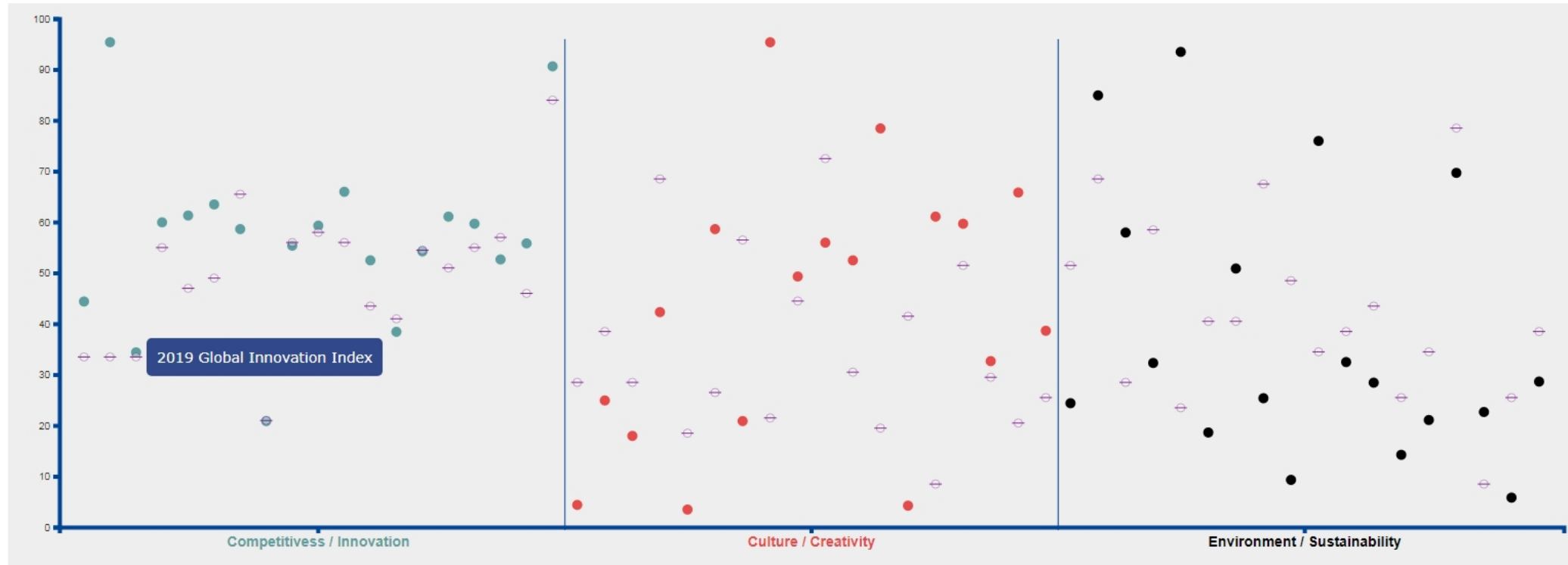
International

EC Priorities Categories

- Competitiveness & Innovation
- Corruption & Governance
- Culture & Creativity
- Environment & Sustainability
- Fairness / Inequality
- Health / Nutrition
- Information / Digital transformation
- Education / Skills
- Markets / Consumers
- Poverty / Social Rights
- Resilience / Risk management
- Ageing / Migration

Active filters

- ☒ Competitiveness & Innovation
- ☒ Culture & Creativity
- ☒ Environment & Sustainability



COIN DATA HUB

COUNTRY PROFILE

Austria

Select the Country ▾

EU28

International

EC Priorities

Categories

A European Green Deal

An economy that works for people

A Europe fit for the digital age

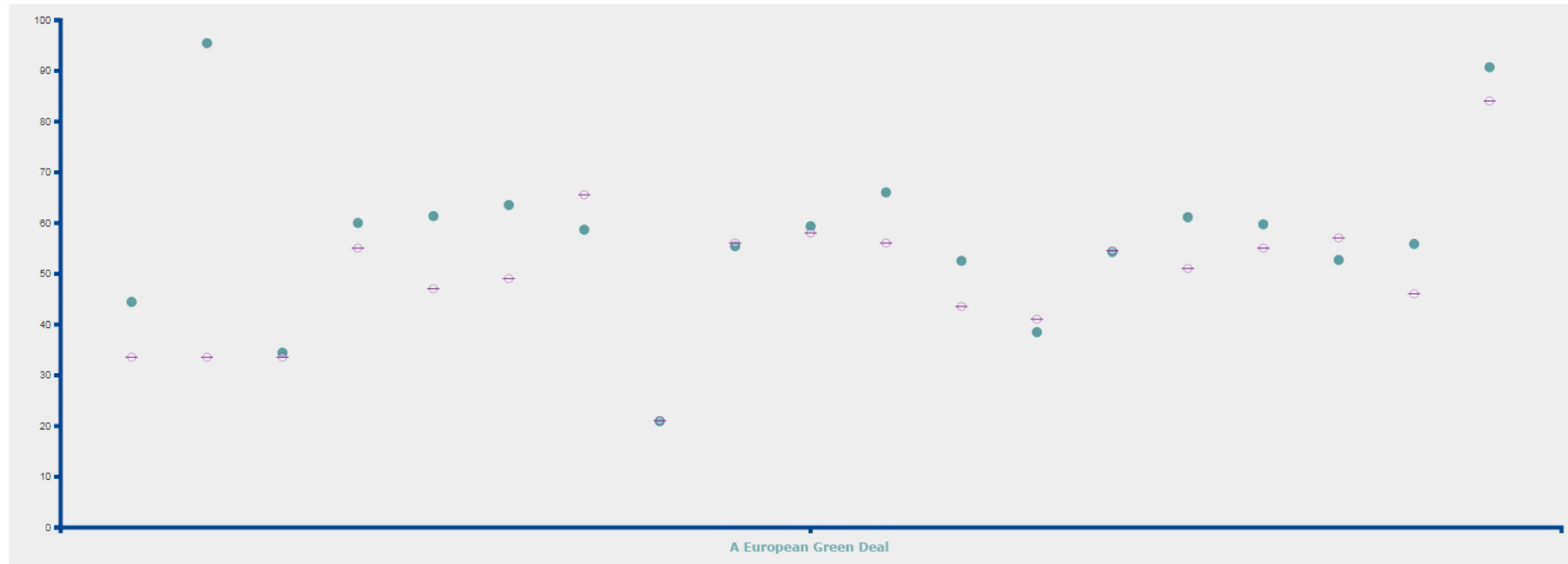
Protecting our European way of life

A stronger Europe in the world

A new push for European
democracy

Active filters

A European Green Deal





THANK YOU

Welcome to email us at: jrc-coin@ec.europa.eu

COIN in the EU Science Hub

<https://ec.europa.eu/jrc/en/coin>

COIN tools are available at:

<https://composite-indicators.jrc.ec.europa.eu/>

The European Commission's
Competence Centre on Composite
Indicators and Scoreboards



The European Commission's science and knowledge service

Joint Research Centre



Statistics Refresher

William Becker

COIN 2019 - 16th JRC Annual Training on Composite Indicators & Scoreboards
04-06/11/2019, Ispra (IT)

Notation

An indicator:

x

e.g. GDP

j th value of an indicator:

x_j

e.g. GDP for country j

i th indicator (in a group):

x_i

e.g. group is GDP, life expectancy,
median income, ...

j th value, i th indicator:

$x_{i,j}$

e.g. GDP for country j

Number of countries

N

Some data

x_i
9.6
10
9.3
7.1
5
8
7.5
8.2
6.4
1

COIN Training 2018 satisfaction scores



Rank

x_i	Rank
9.6	2
10	1
9.3	3
7.1	7
5	9
8	5
7.5	6
8.2	4
6.4	8
1	10

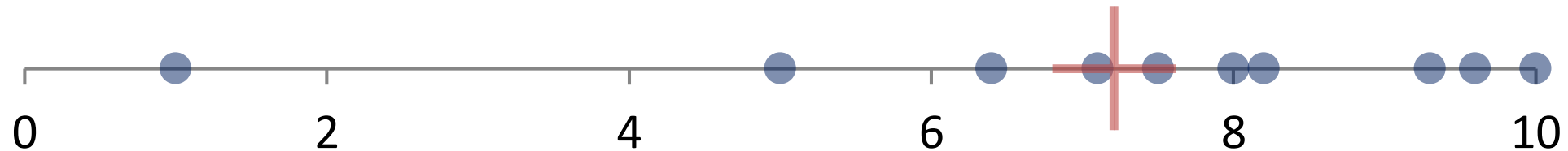
COIN Training 2018 satisfaction scores



Mean

$$\mu_i = E(x_i) = \frac{1}{N} (x_{i,1} + x_{i,1} + x_{i,1} + \cdots + x_{i,N}) = \sum_{j=1}^N x_{i,j}$$

$$\mu_i = \frac{1}{10} (9.6 + 10 + 9.3 + 7.1 + 5 + 8 + 7.5 + 8.2 + 6.4 + 1)$$



Median

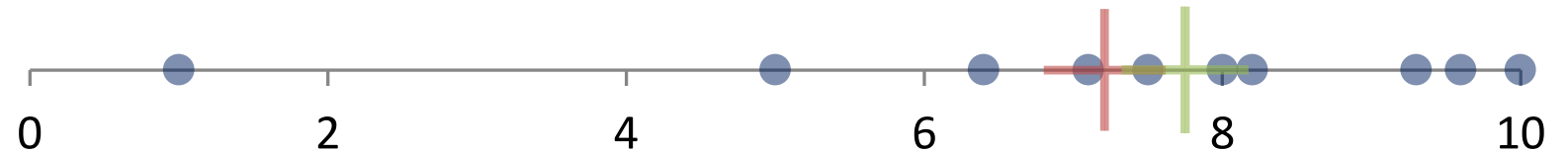
x_i
1
5
6.4
7.1
7.5
8
8.2
9.3
9.6
10



The **middle value of the list of numbers**. Half of the values should be higher, half should be lower.

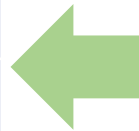
For an even number of values, it is the midway point between the two middle values.

The median is not equal to the mean unless it is a symmetric distribution.



Percentiles

x_i
1
5
6.4
7.1
7.5
8
8.2
9.3
9.6
10



20th
percentile



MEDIAN
(50th
percentile)

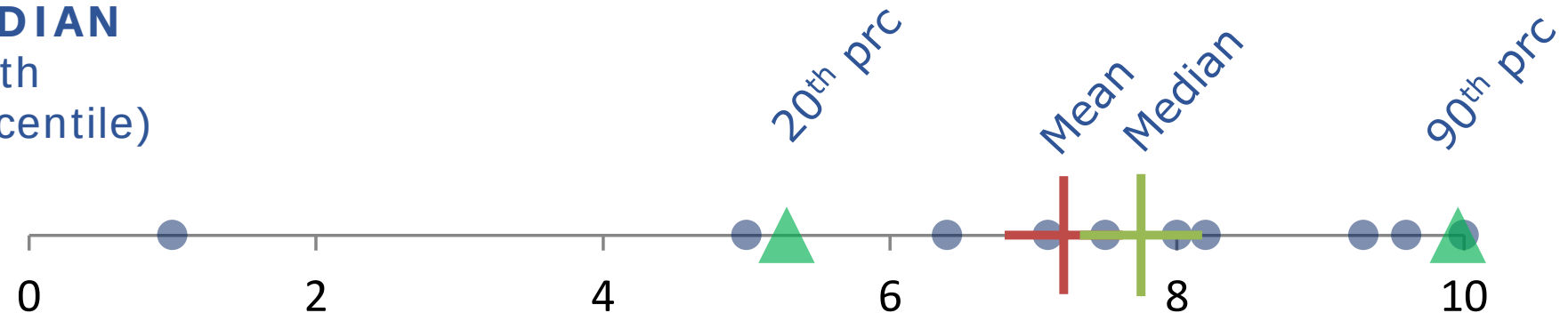


90th
percentile

Percentiles are a generalisation of the median.

Median: value such that 50% of values are below it

Xth percentile: value such that X% of values are below it

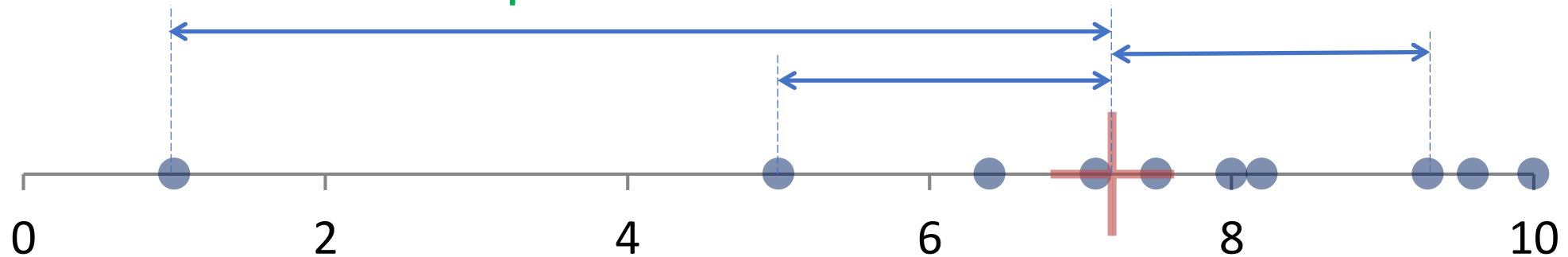


Variance

A measure of dispersion: how different are values from one another?

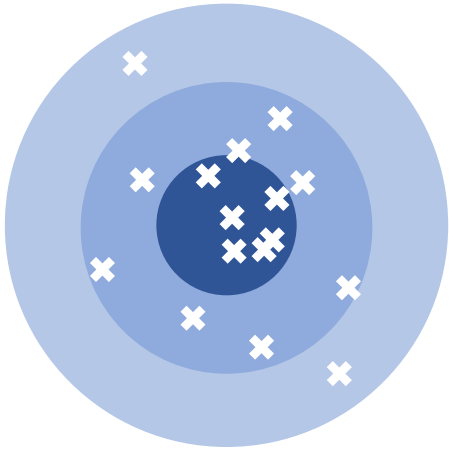
$$\text{var}(x_i) = \sigma_i^2 = \frac{1}{N} \sum_{j=1}^N \underbrace{(\mu_i - x_{i,j})^2}$$

σ_i is the standard deviation

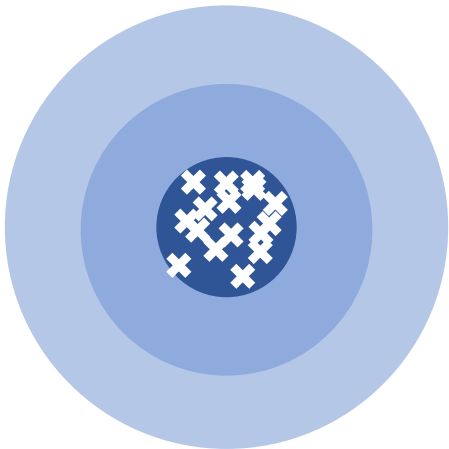


The average squared distance of data points from the mean

Variance and standard deviation



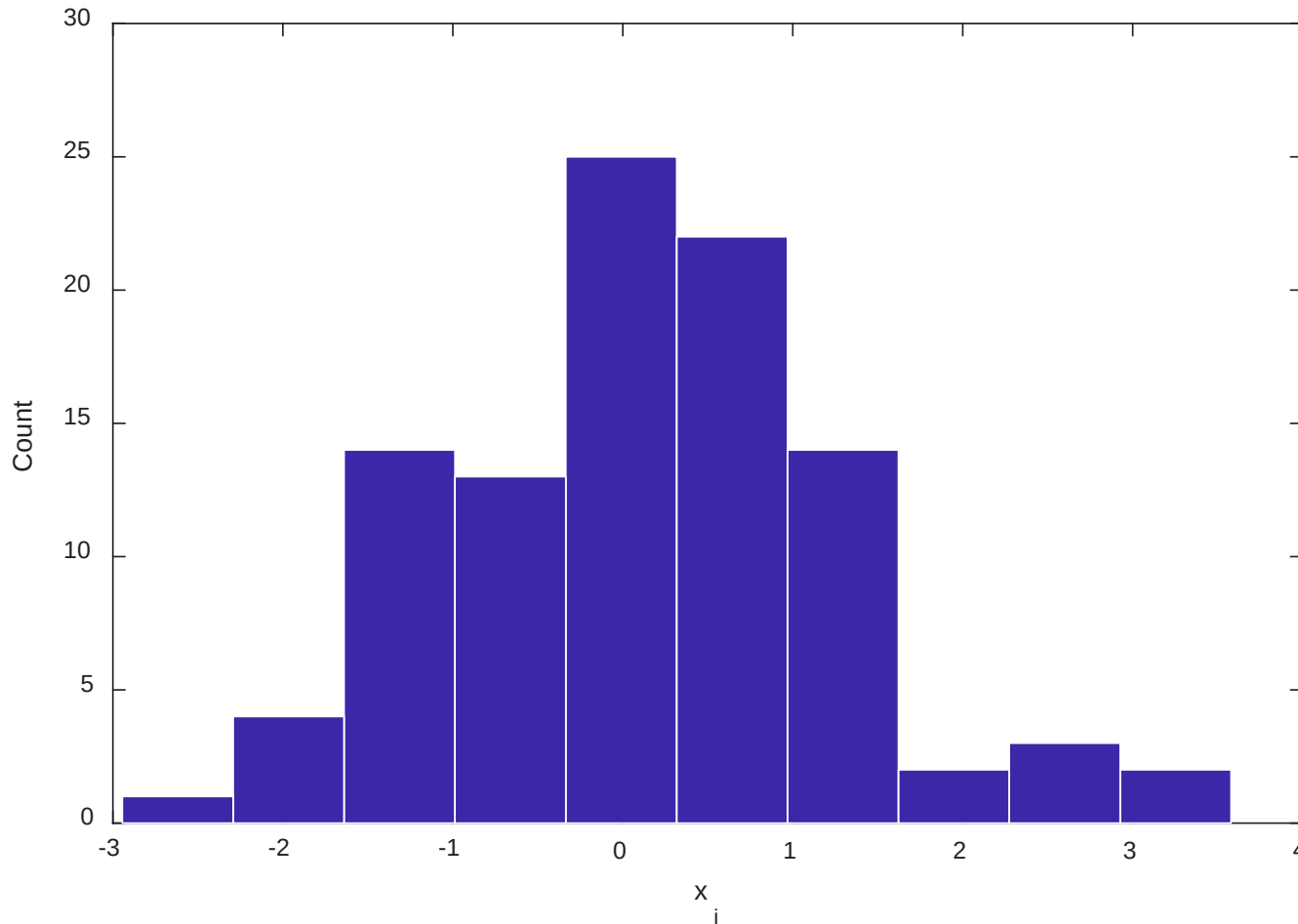
High variance



Low variance

Standard deviation is the square root of the variance.

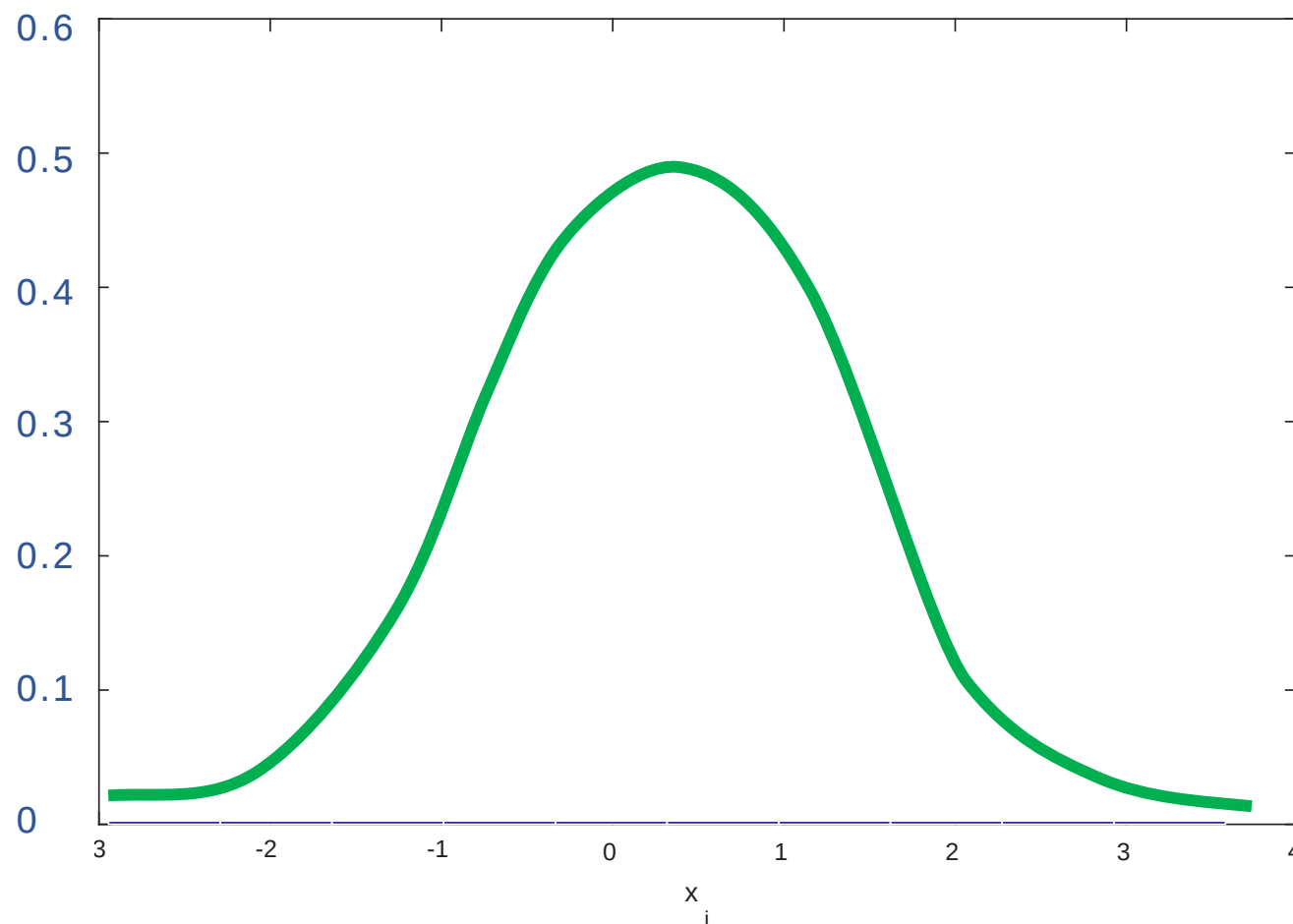
Probability distributions



A **histogram** shows how data is distributed into equally-spaced “bins”.

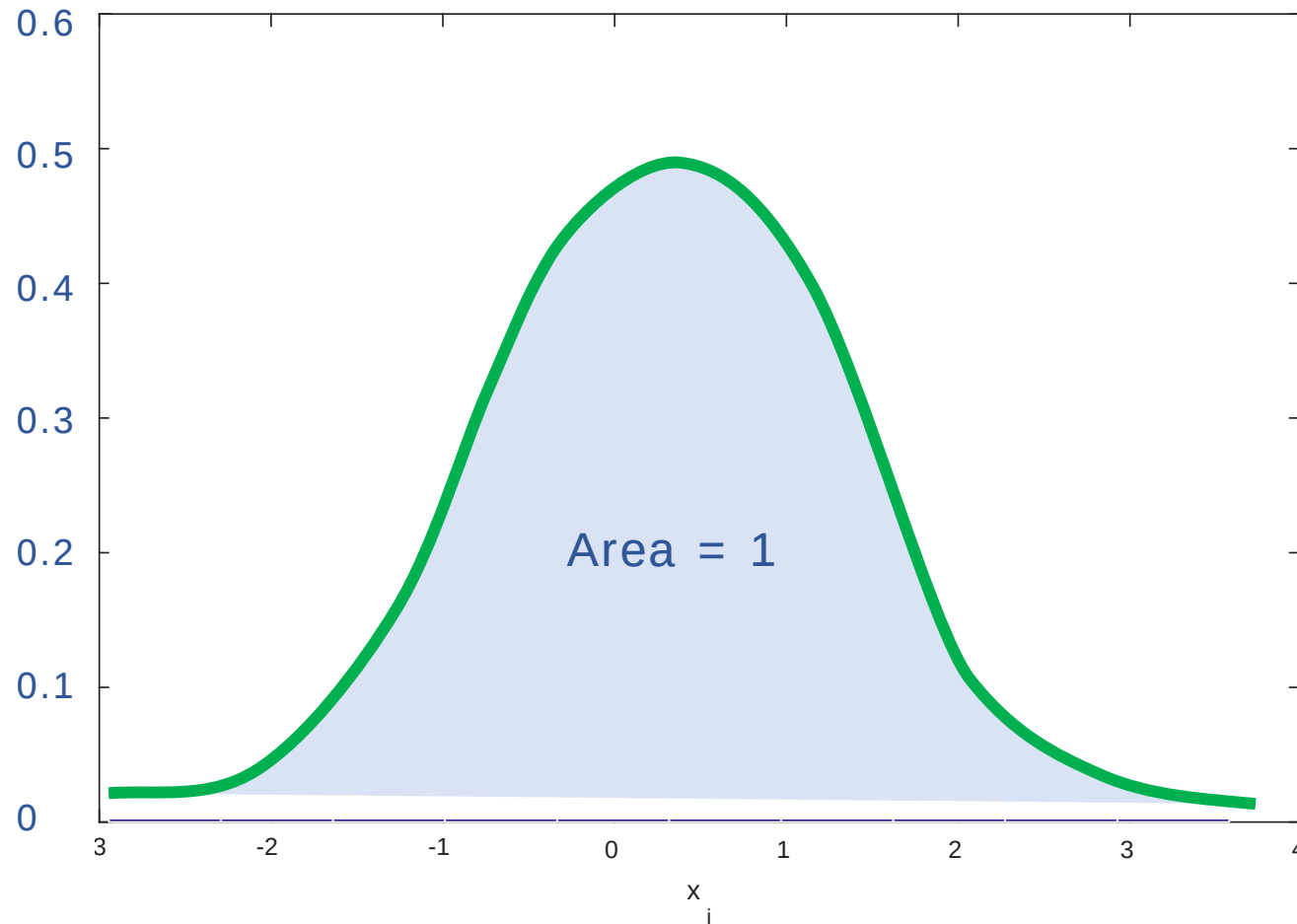
Each bar gives an indication of how frequently (how probable) it is that a value will fall inside that range.

Probability distributions



A probability density function (pdf) shows the relative probability of different values occurring.

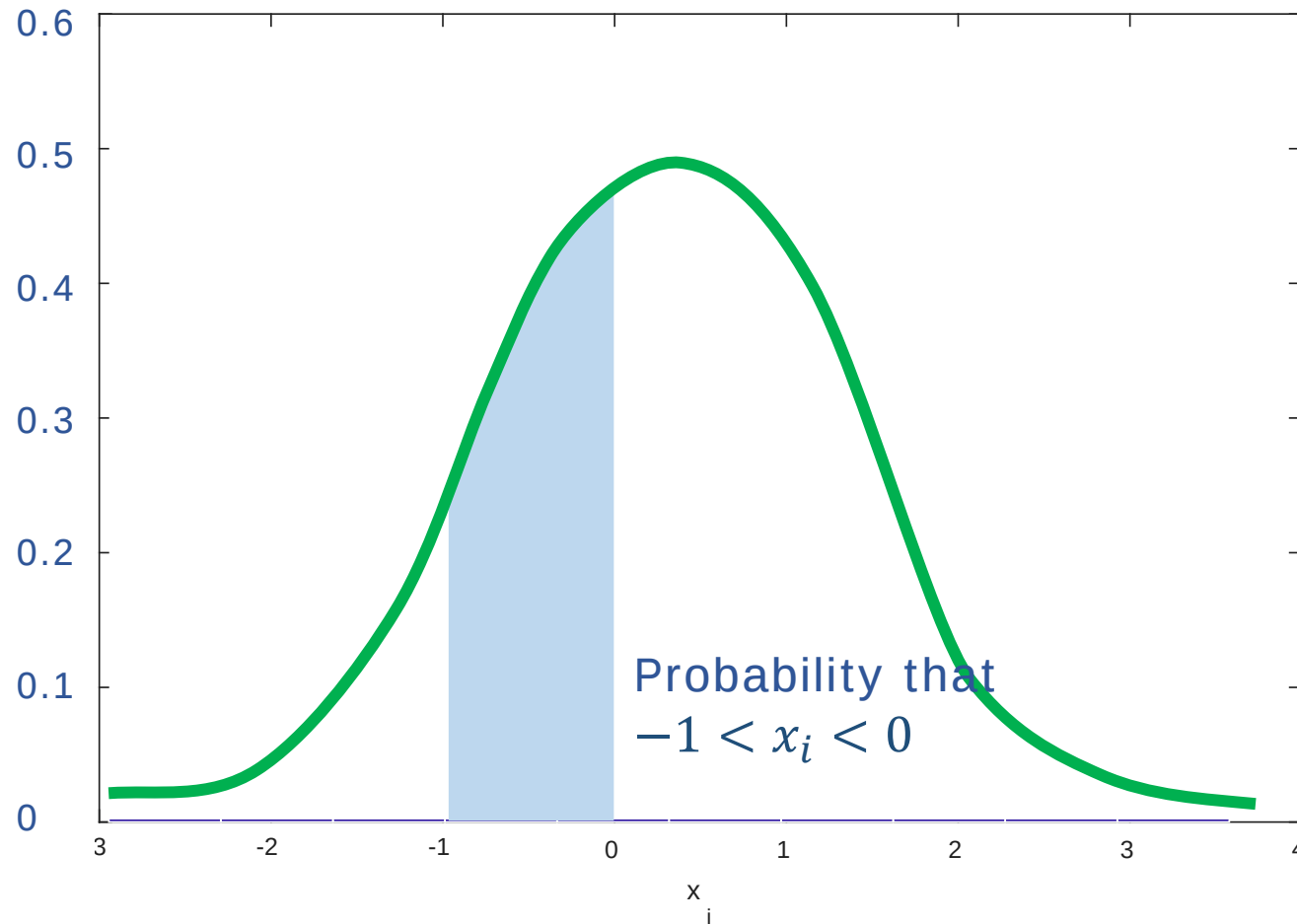
Probability distributions



A probability density function (pdf) shows the relative probability of different values occurring.

The total area under the pdf (integral) is equal to 1 by definition.

Probability distributions

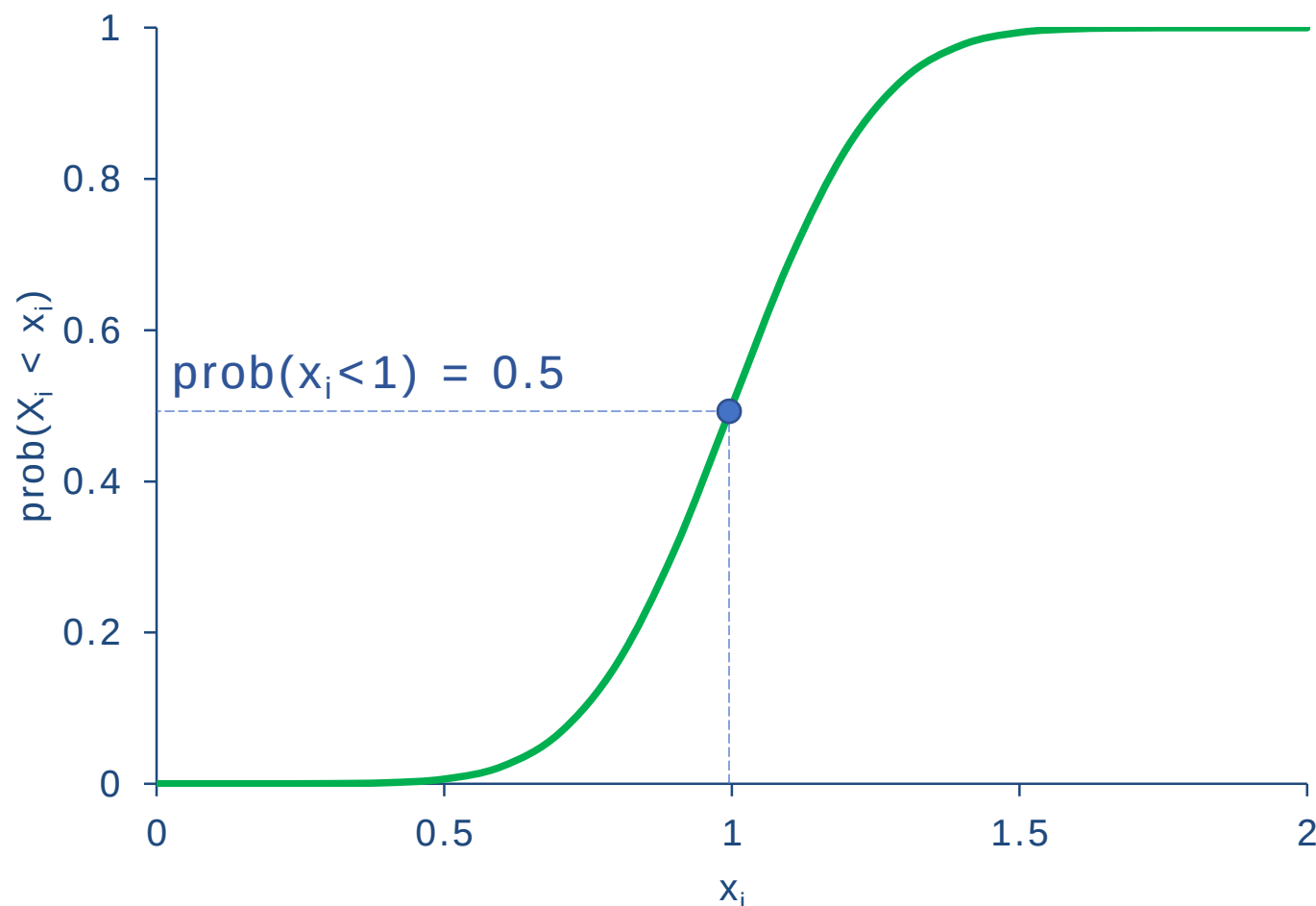


A probability density function (pdf) shows the relative probability of different values occurring.

The total area under the pdf (integral) is equal to 1 by definition.

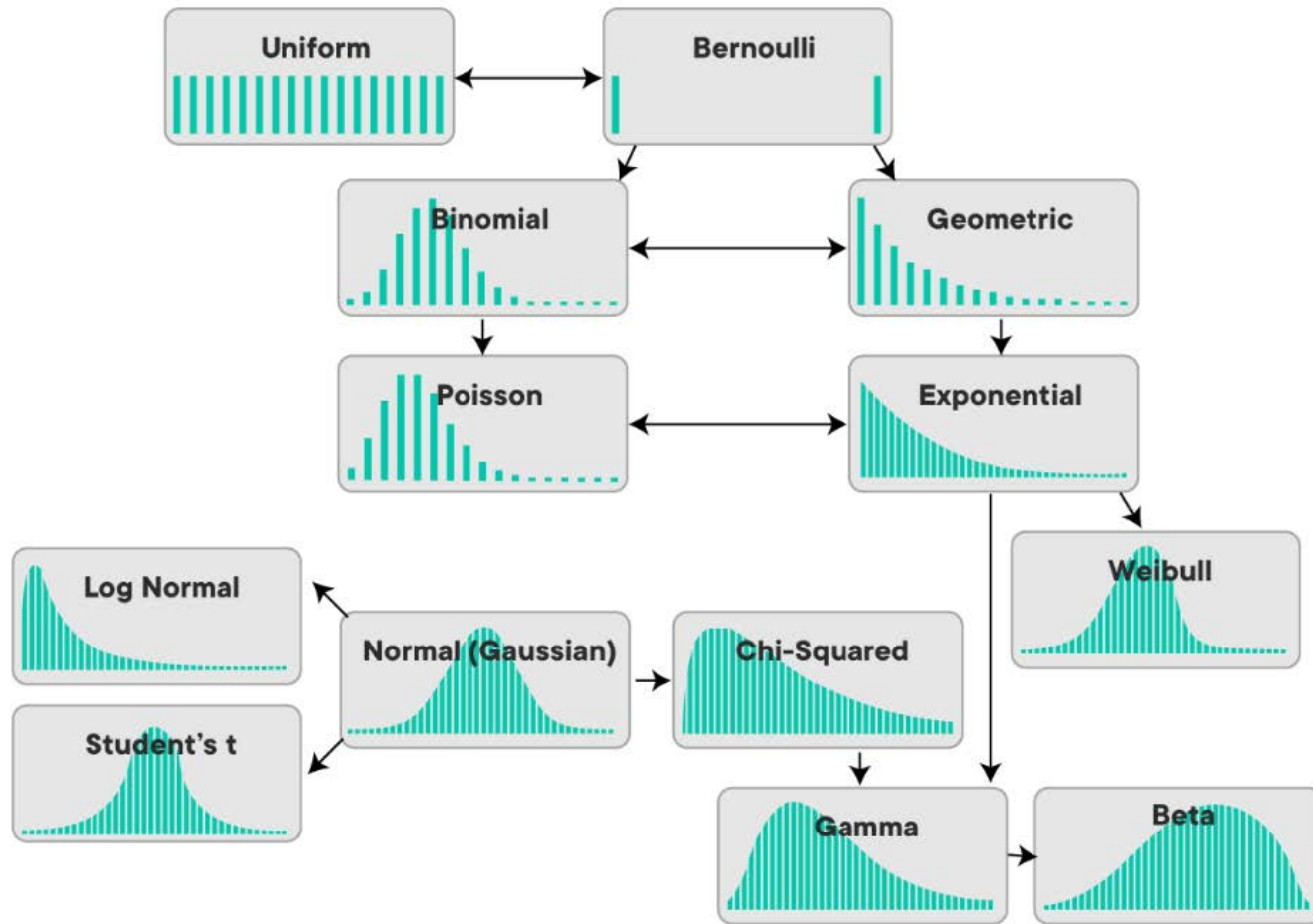
The area between two points shows the probability of a value falling in that range.

Probability distributions



A cumulative density function (cdf) shows the probability that the variable is less than a given value.

Probability distributions



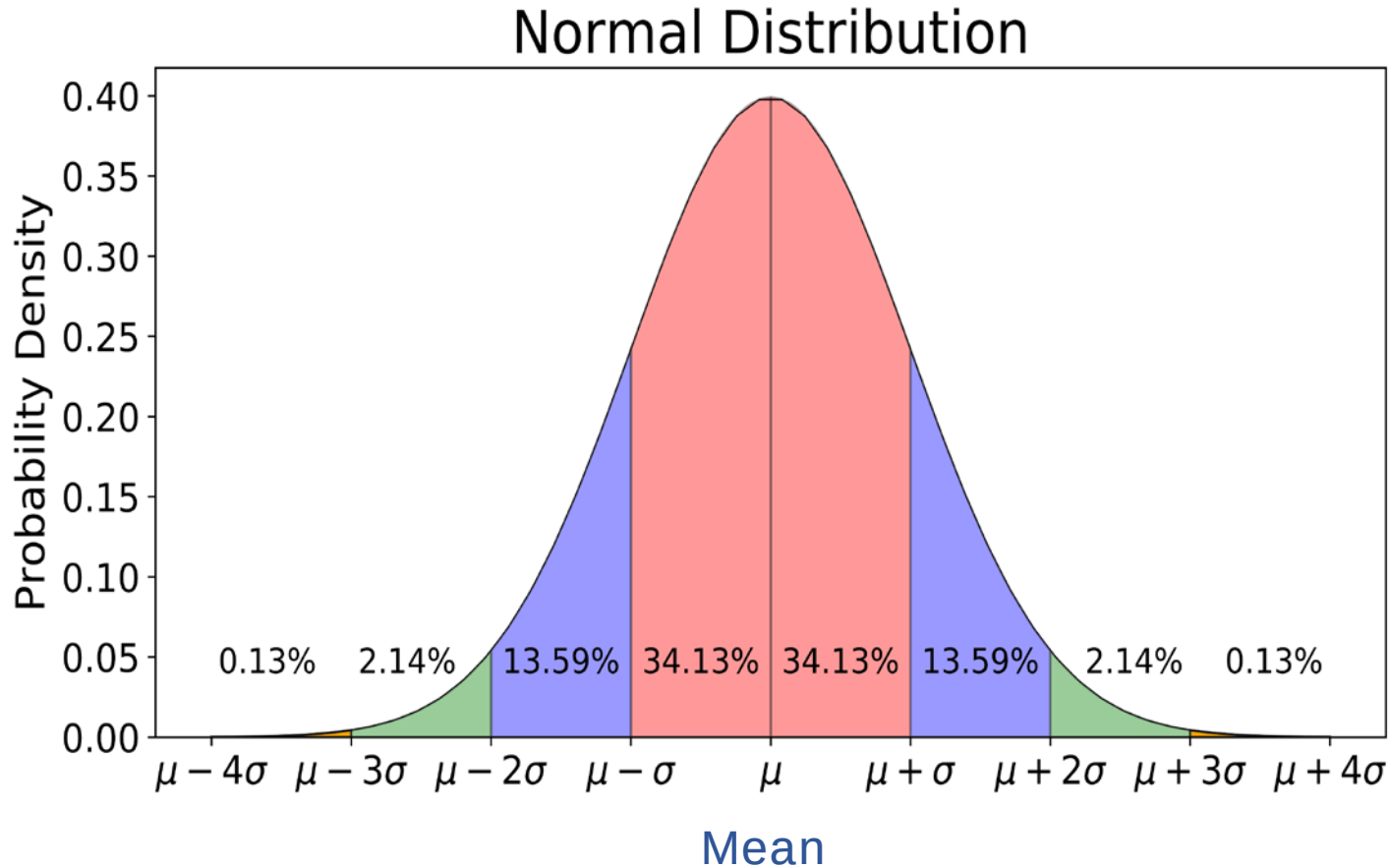
There are many theoretical types of probability distribution.

They aim to mimic real distributions observed in nature, economics, physics, etc.

Discrete distributions can only take a finite amount of values.

Continuous distributions can take an infinite number of values (in a finite or infinite range).

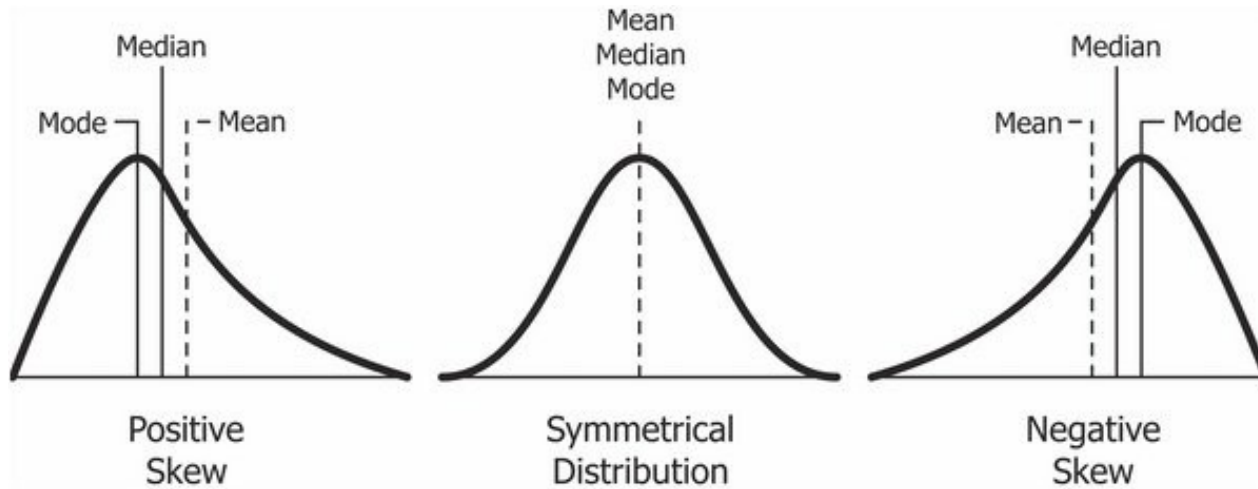
Normal distribution



The Normal (Gaussian) distribution is the best distribution!

- Very common due to **central limit theorem**: sum of lots of random variables converges to normal
- **Makes things easy**:
 - Linear
 - Analytical
 - No skew/kurtosis
 - etc.

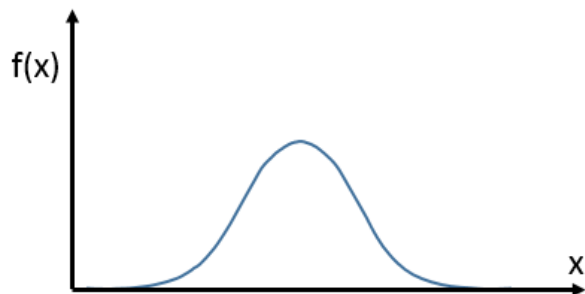
Skew and kurtosis



Skew shows the extent to which values are clustered towards one end of the scale. Measure of lack of symmetry.

Kurtosis measures how heavy the tails of the distribution are, compared to a normal distribution.

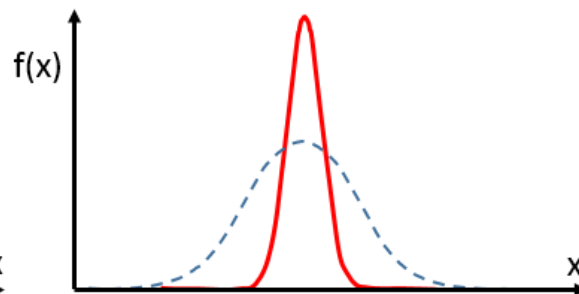
Be careful: “kurtosis” may mean *true kurtosis*, or *excess kurtosis* ($= \text{true kurtosis} - 3$)



Normal

Kurtosis = 3

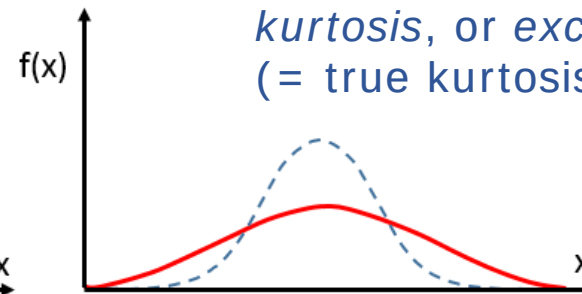
Excess kurtosis = 0



Leptokurtic

Kurtosis > 3

Excess kurtosis > 0



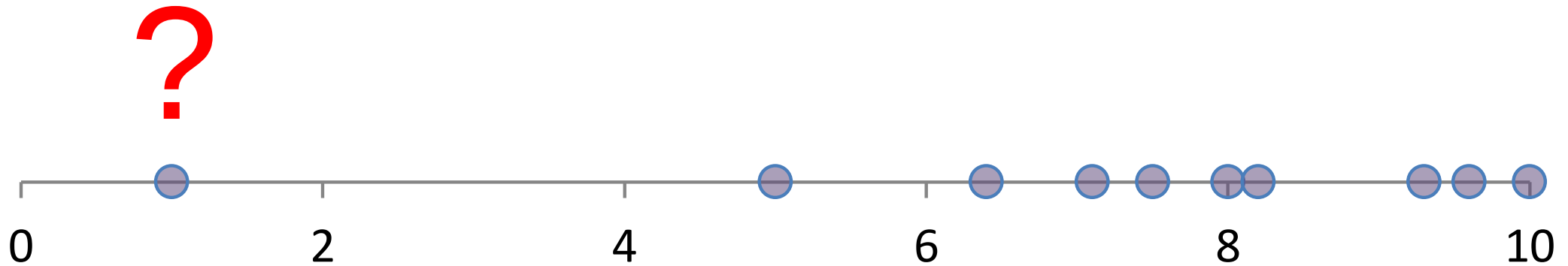
Platykurtic

Kurtosis < 3

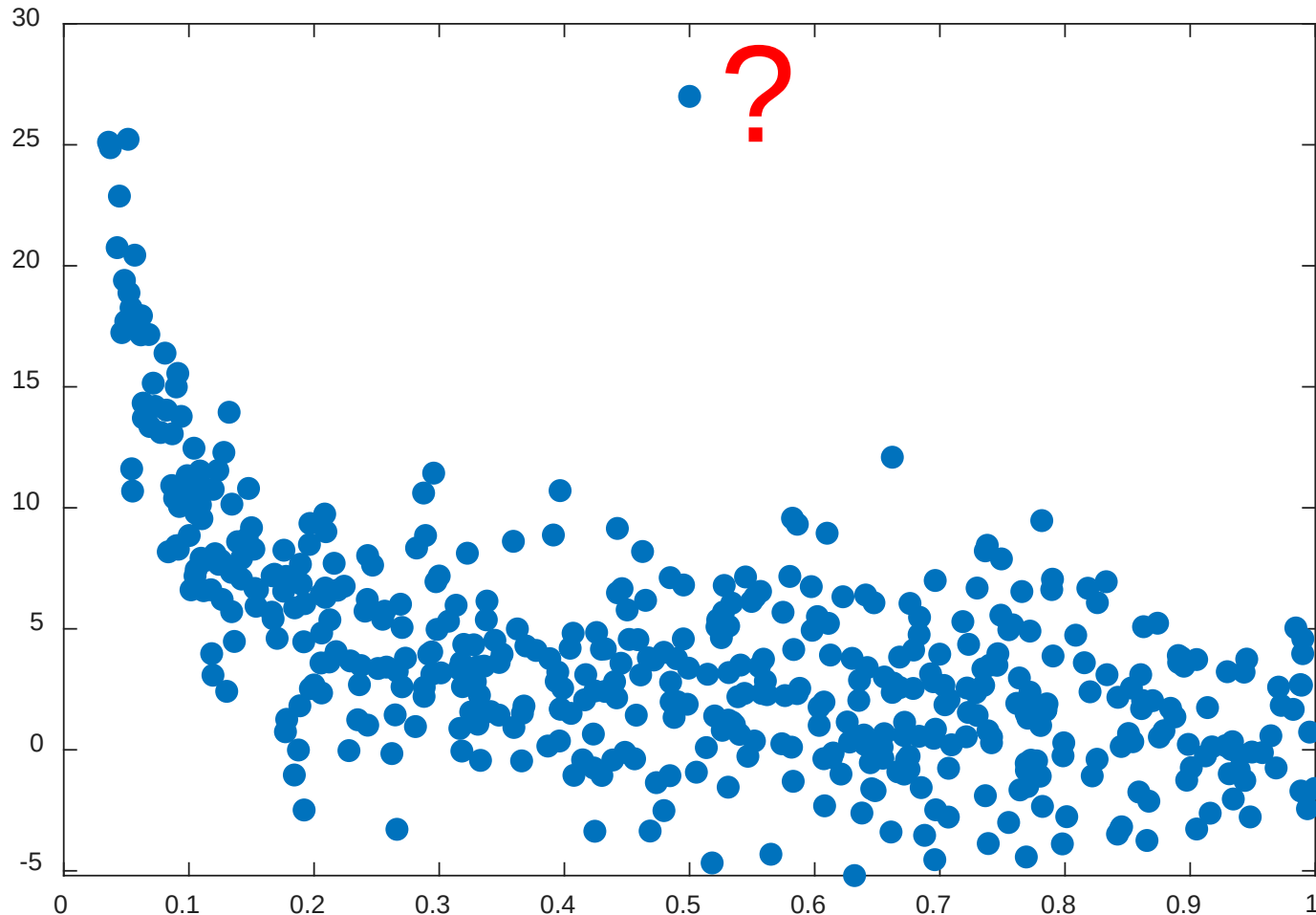
Excess kurtosis < 0

Outliers

An outlier is an observation that differs significantly from other observations.



Outliers



An outlier is an observation that differs significantly from other observations.

Outlier identification can be performed via:

- Inter-quartile range
- Skew/kurtosis measures
- Multivariate analysis
- Etc.

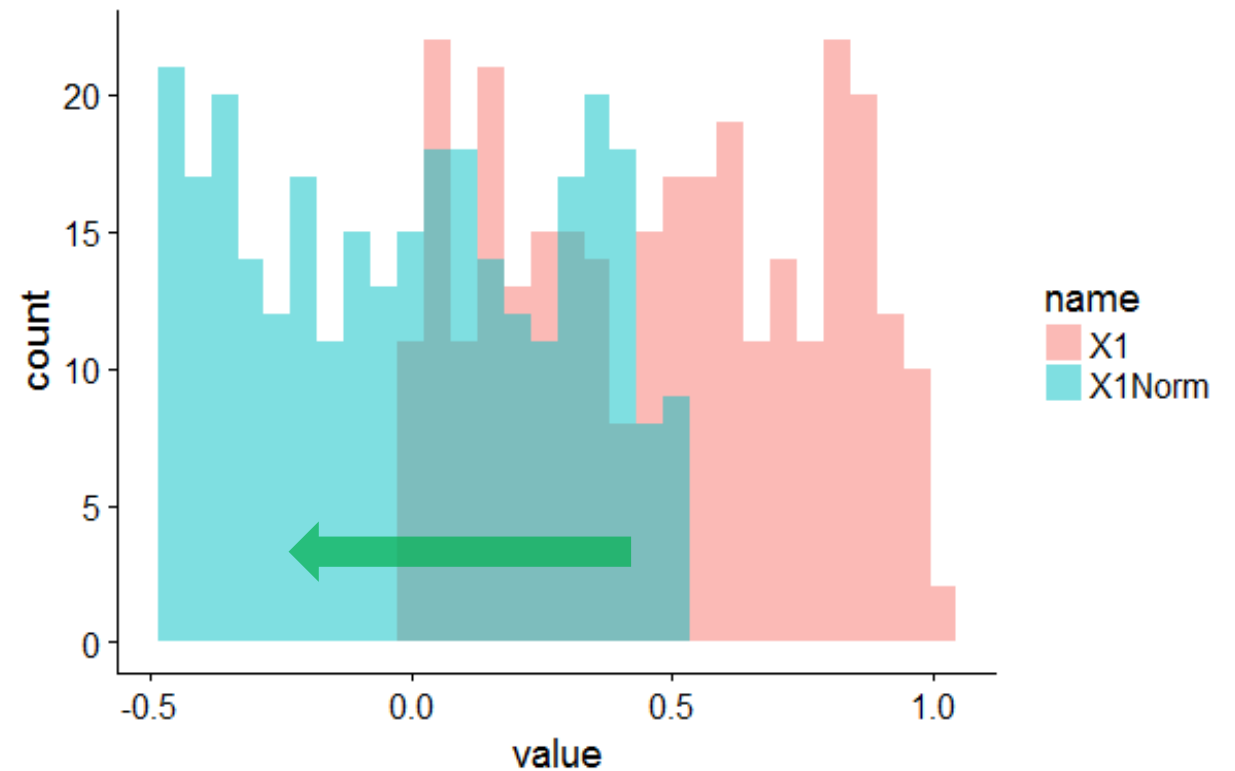
Measures of dependence

Fundamentally, we are interested in how the random variable x_1 depends on the random variable x_2 .

Covariance

$$\text{cov}(x_1, x_2) = E[(x_1 - \mu_1)(x_2 - \mu_2)]$$

$(x_2 - \mu_2)$ is shifting x_2 so that it has a mean of zero



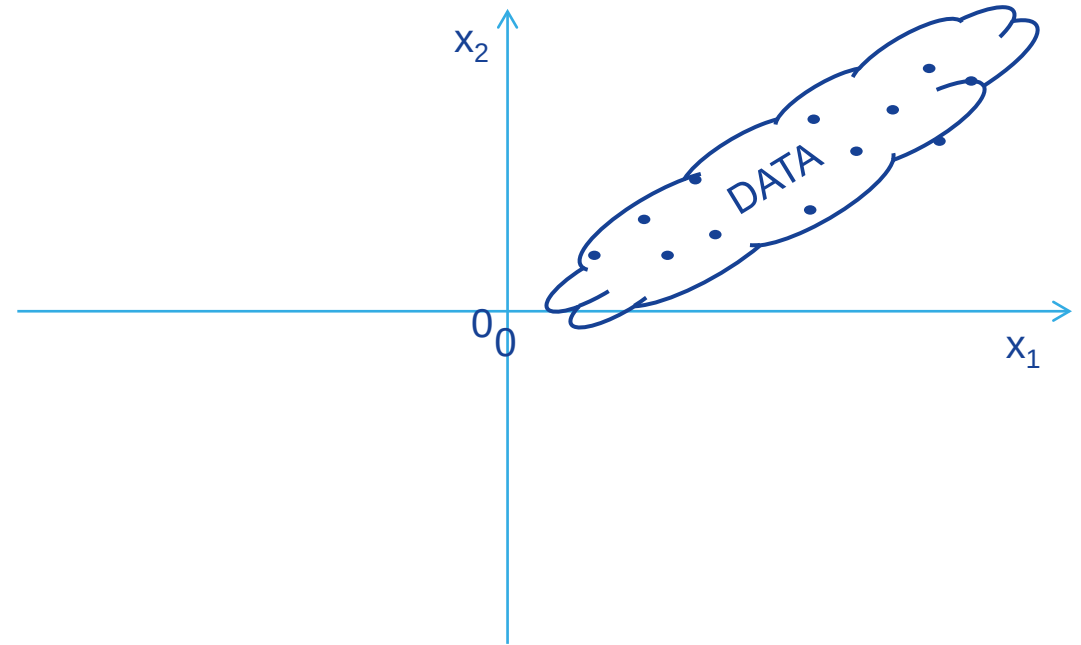
Measures of dependence

Fundamentally, we are interested in how the random variable x_1 depends on the random variable x_2 .

Covariance

$$\text{cov}(x_1, x_2) = E[(x_1 - \mu_1)(x_2 - \mu_2)]$$

$(x_2 - \mu_2)$ is shifting x_2 so that it has a mean of zero



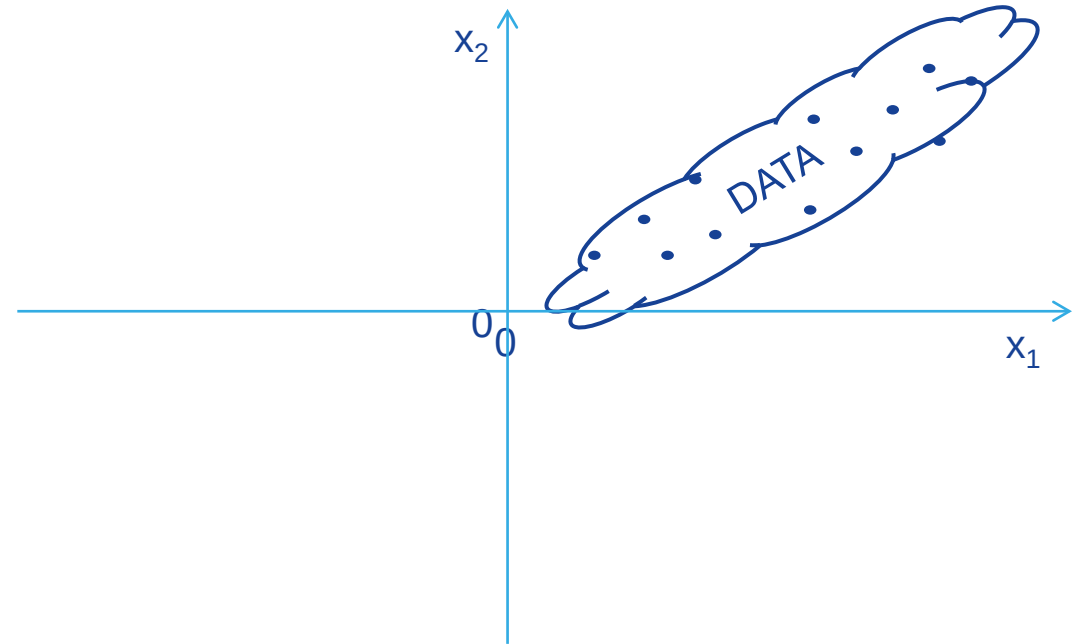
Measures of dependence

Fundamentally, we are interested in how the random variable x_1 depends on the random variable x_2 .

Covariance

$$\text{cov}(x_1, x_2) = E[(x_1 - \mu_1)(x_2 - \mu_2)]$$

$(x_2 - \mu_2)$ is shifting x_2 so that it has a mean of zero



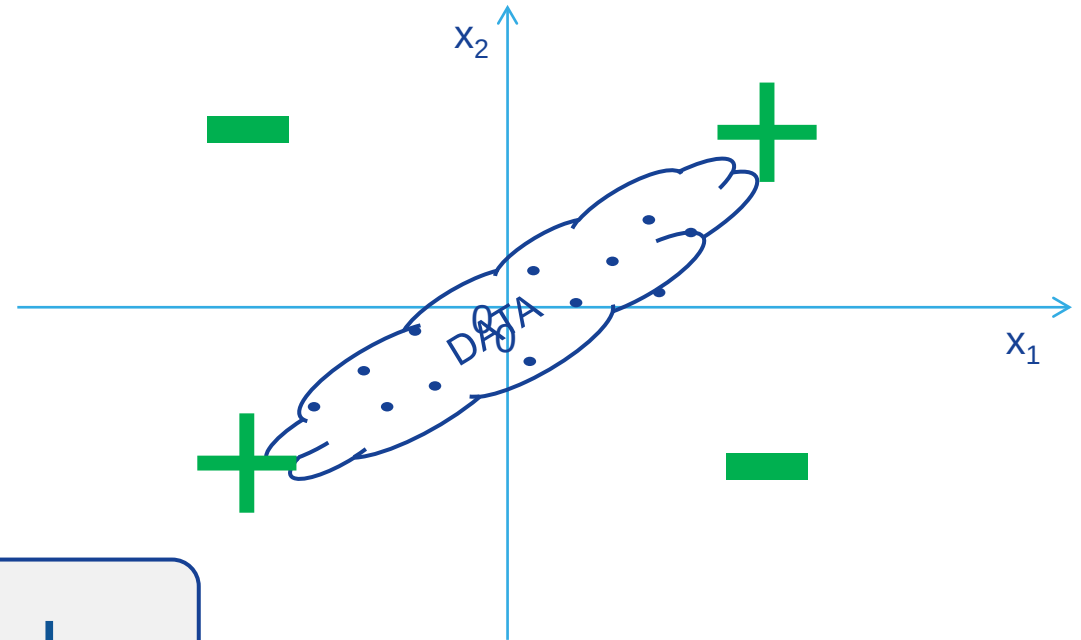
Measures of dependence

Fundamentally, we are interested in how the random variable x_1 depends on the random variable x_2 .

Covariance

$$\text{cov}(x_1, x_2) = E[(x_1 - \mu_1)(x_2 - \mu_2)]$$

$(x_2 - \mu_2)$ is shifting x_2 so that it has a mean of zero



2+

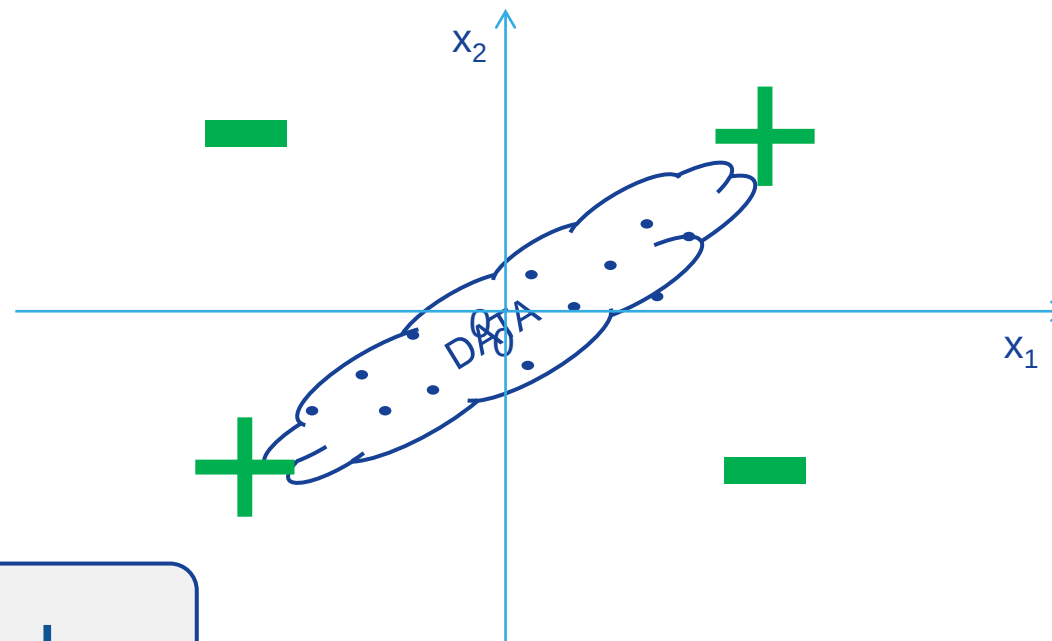
Measures of dependence

Fundamentally, we are interested in how the random variable x_1 depends on the random variable x_2 .

Covariance

$$\text{cov}(x_1, x_2) = E[(x_1 - \mu_1)(x_2 - \mu_2)]$$

$(x_2 - \mu_2)$ is shifting x_2 so that it has a mean of zero



2+

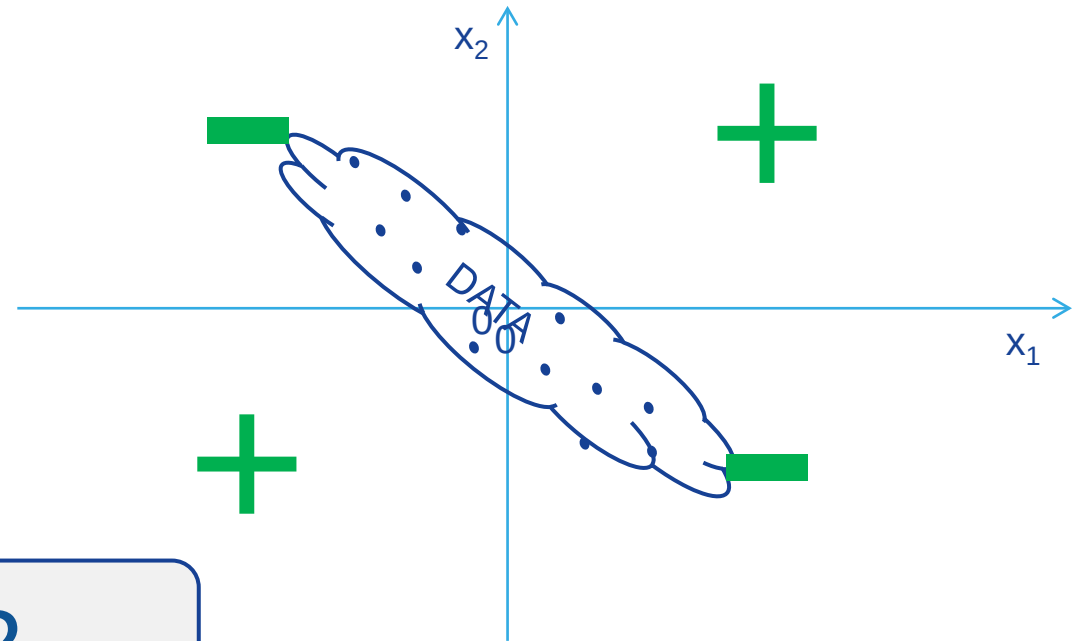
Measures of dependence

Fundamentally, we are interested in how the random variable x_1 depends on the random variable x_2 .

Covariance

$$\text{cov}(x_1, x_2) = E[(x_1 - \mu_1)(x_2 - \mu_2)]$$

$(x_2 - \mu_2)$ is shifting x_2 so that it has a mean of zero



2-

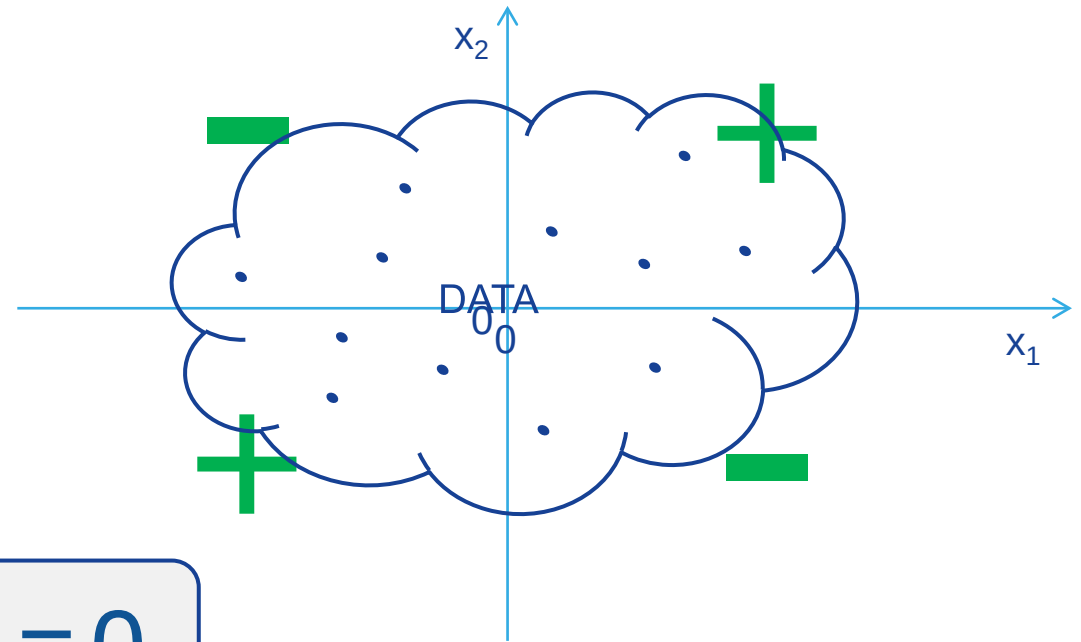
Measures of dependence

Fundamentally, we are interested in how the random variable x_1 depends on the random variable x_2 .

Covariance

$$\text{cov}(x_1, x_2) = E[(x_1 - \mu_1)(x_2 - \mu_2)]$$

$(x_2 - \mu_2)$ is shifting x_2 so that it has a mean of zero



$$(2+) + (2-) = 0$$

Measures of dependence

Fundamentally, we are interested in how the random variable x_1 depends on the random variable x_2 .

Covariance

$$\text{cov}(x_1, x_2) = E[(x_1 - \mu_1)(x_2 - \mu_2)]$$

Correlation

$$R(x_1, x_2) = \text{corr}(x_1, x_2) = \frac{\text{cov}(x_1, x_2)}{\sigma_1 \sigma_2}$$

Standardises covariance so that $R \in [-1, 1]$: 1 or -1 is perfect correlation, 0 is no correlation. Allows comparability.

Measures of dependence

Correlation

$$R(x_1, x_2) = \text{corr}(x_1, x_2) = \frac{\text{cov}(x_1, x_2)}{\sigma_1 \sigma_2}$$

Coefficient of determination: R^2

$$R^2(x_1, x_2) = [\text{corr}(x_1, x_2)]^2 = \left[\frac{\text{cov}(x_1, x_2)}{\sigma_1 \sigma_2} \right]^2$$

More generally, R_i^2
can be defined as:

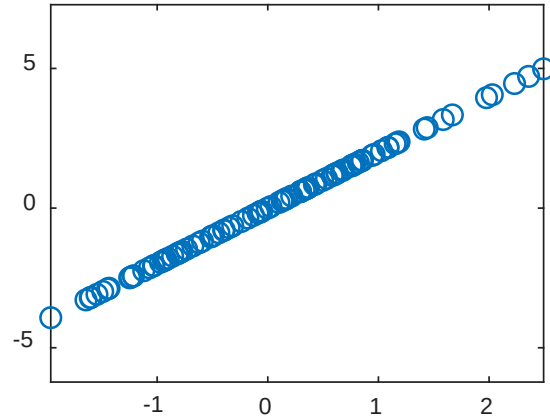
$$\frac{\text{variance explained by regression}}{\text{total variance}}$$

R^2 is a measure of **linear dependence**.

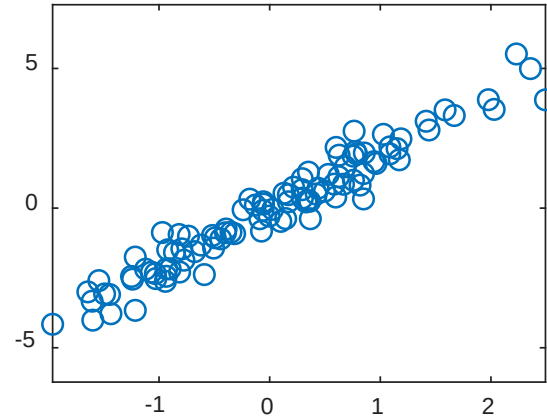
$R^2 \in [0,1]$: higher values indicate stronger dependence.

Measures of dependence

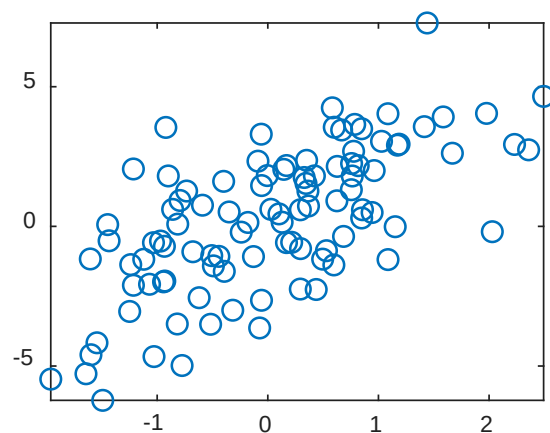
Perfect
positive
correlation
 $R = 1$



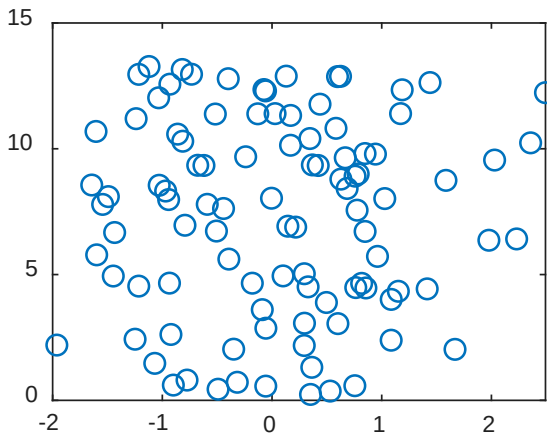
Very strong
positive
correlation
 $R = 0.97$



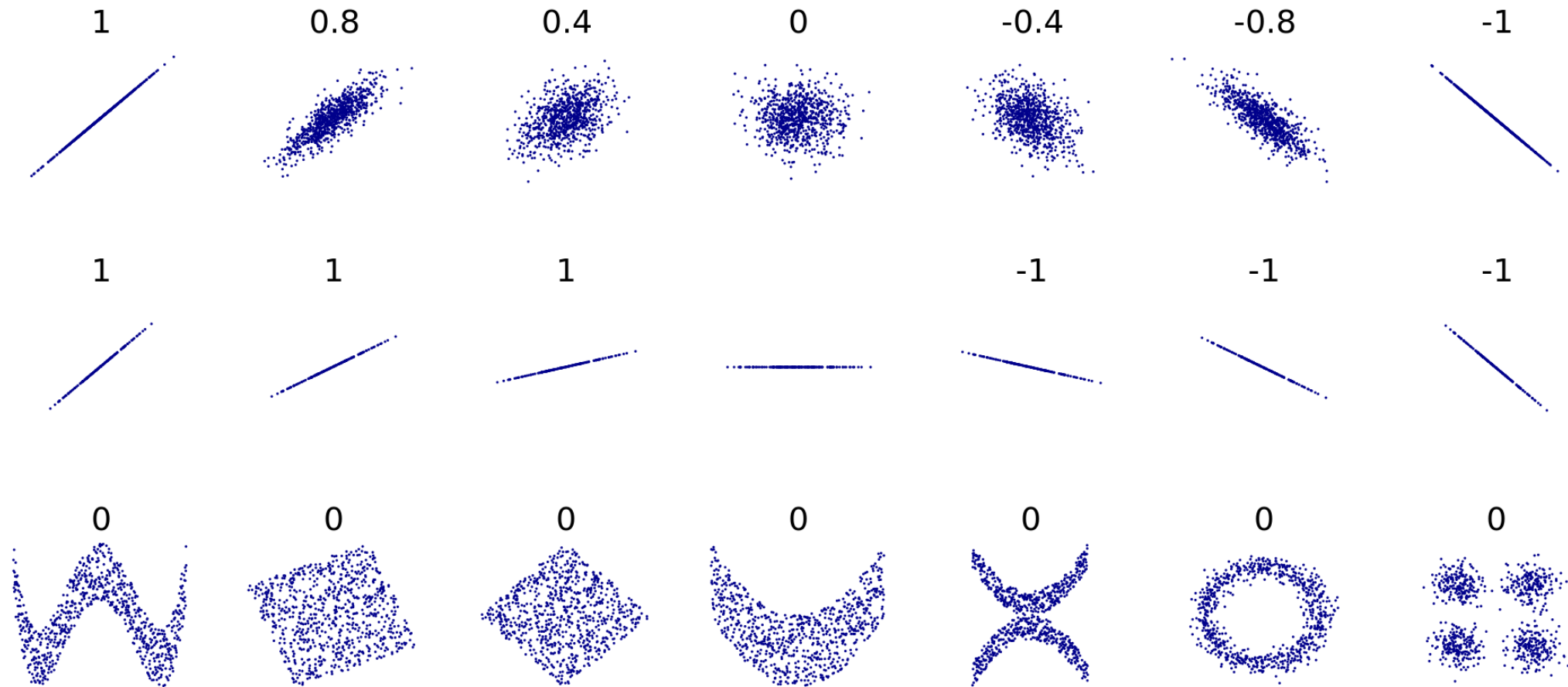
Moderate
positive
correlation
 $R = 0.66$



No
positive
correlation
 $R = 0.02$



Measures of dependence



Measures of dependence

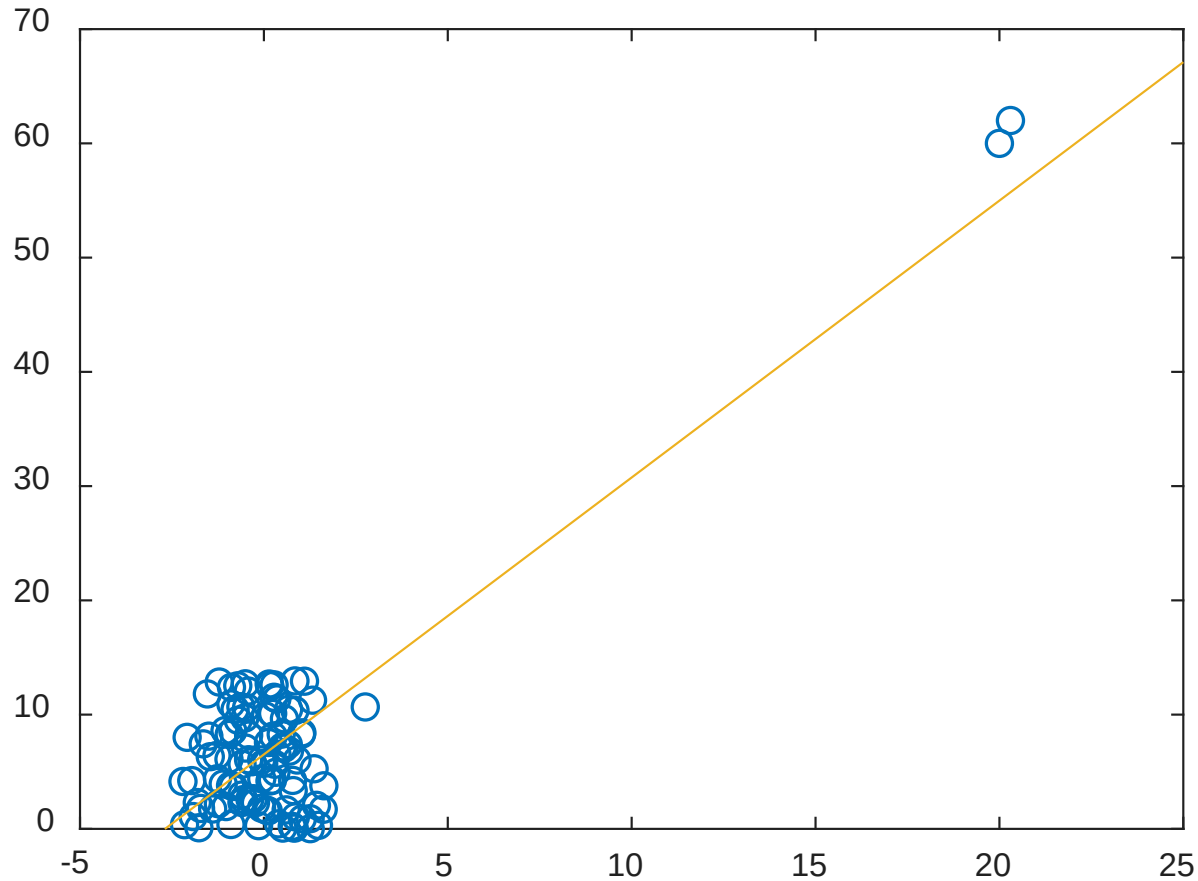
Size of correlation	Interpretation
0.9 to 1 (-0.9 to -1)	Very high positive (negative) correlation
0.7 to 0.9 (-0.7 to -0.9)	High positive (negative) correlation
0.5 to 0.7 (-0.5 to -0.7)	Moderate positive (negative) correlation
0.3 to 0.5 (-0.3 to -0.5)	Low positive (negative) correlation
0 to 0.3 (0 to -0.3)	Negligible correlation

BUT: check for statistical significance!

“Given the sample size (number of points, countries, ...), is the correlation **significantly different** from zero?”

A quick test for significance is $|R| \geq \frac{2}{\sqrt{N}}$

Measures of dependence



What happens to correlation when there are outliers?

correlation:

0.02 ➡ 0.84

Be careful with outliers.

Always plot your data!



THANK YOU

Welcome to email us at: jrc-coin@ec.europa.eu

COIN in the EU Science Hub

<https://ec.europa.eu/jrc/en/coin>

COIN tools are available at:

<https://composite-indicators.jrc.ec.europa.eu/>

The European Commission's
Competence Centre on Composite
Indicators and Scoreboards





The Essentials of Composite Indicators and Scoreboards

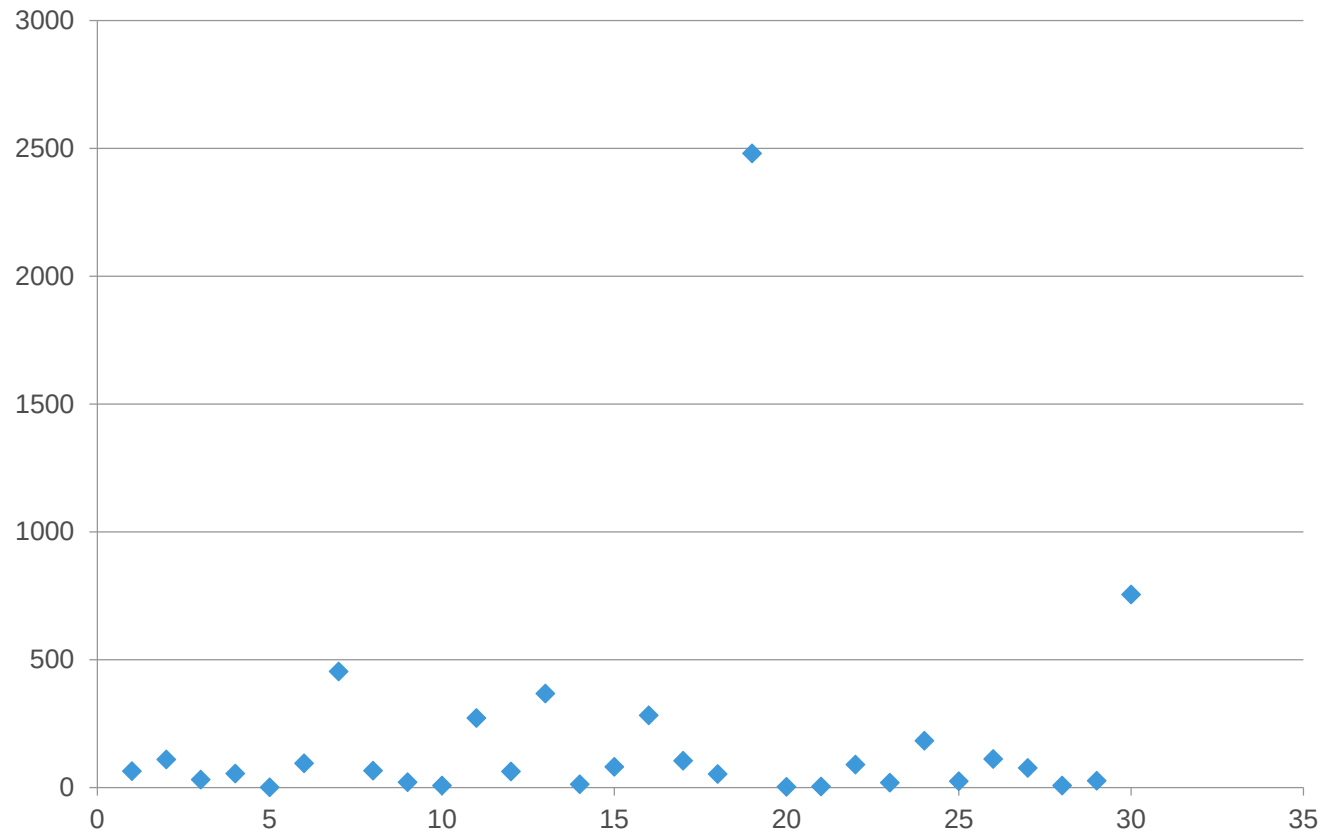
Data Processing

Giulio Caperna
Bruxelles, 17/02/2020

The next steps



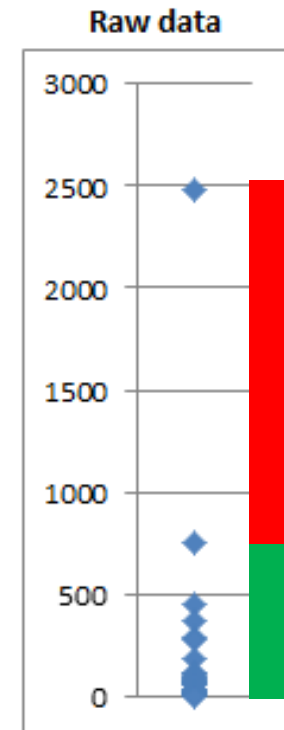
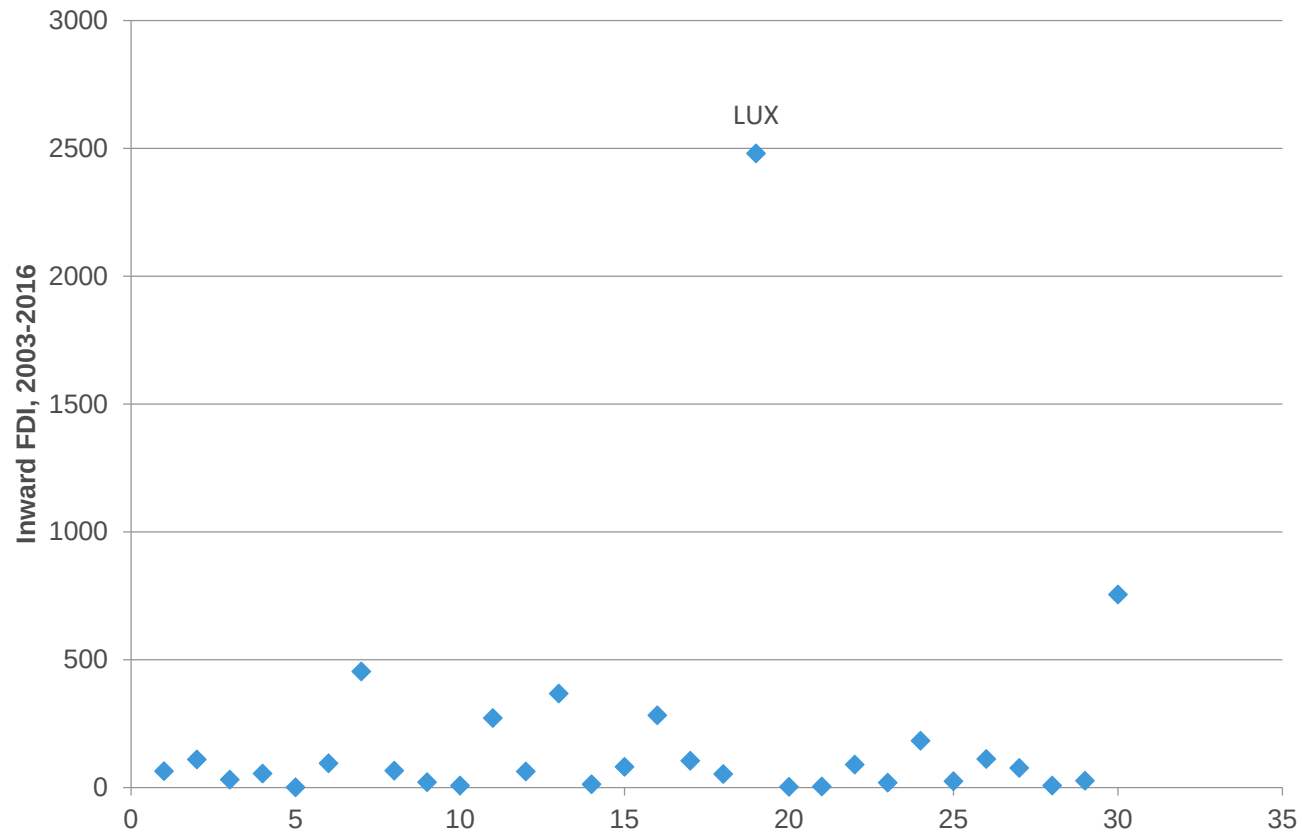
What is an outlier?



“An outlier is an observed value that is so extreme (either large or small) that it seems to stand apart from the rest of the distribution”

(Knoke & Mee, 2002)

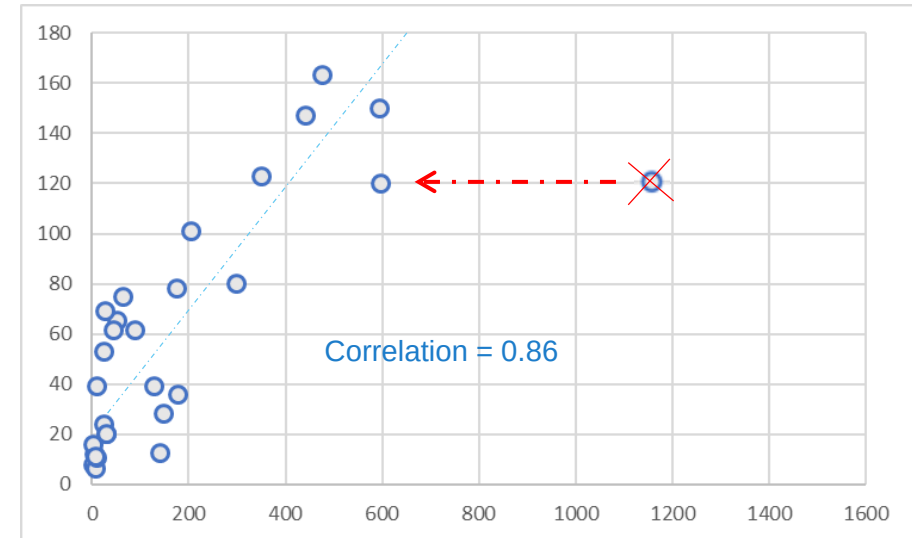
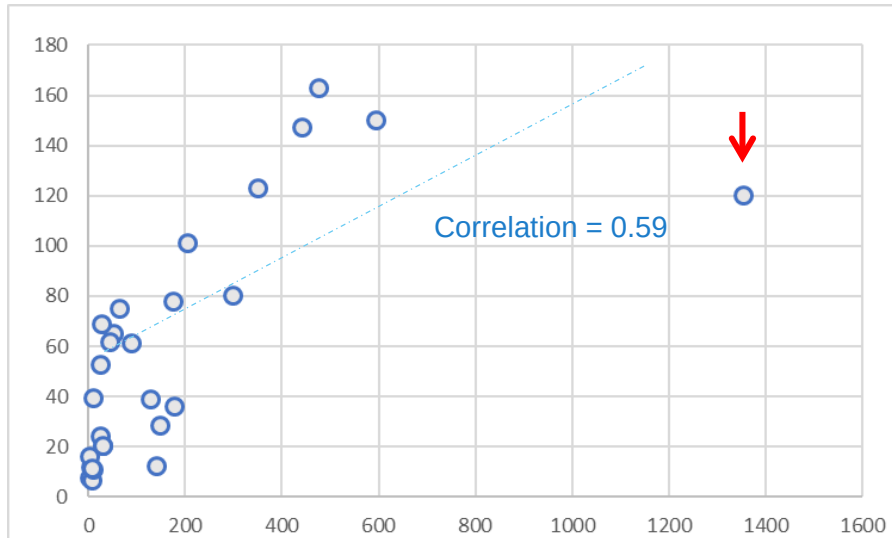
What is wrong with outliers (1)?



The empty range should be small

What is wrong with outliers (2)?

Outliers may spoil descriptive statistics (means, standard deviations, correlations ...) and cause misinterpretations



How to handle them

➤ Outlier identification

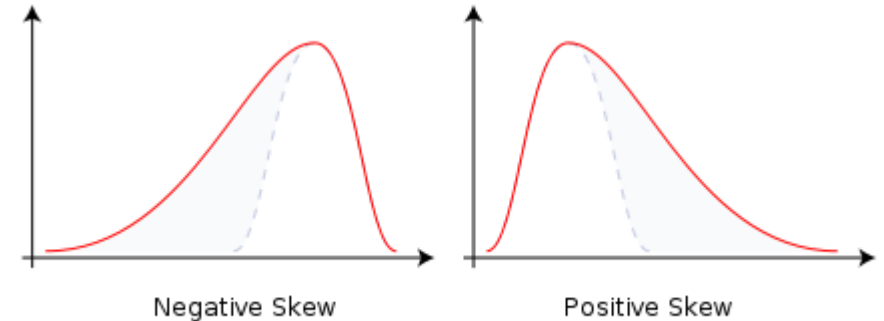
- 1) z-scores
- 2) Interquartile range
- 3) Skewness and Kurtosis

➤ Outlier treatment techniques

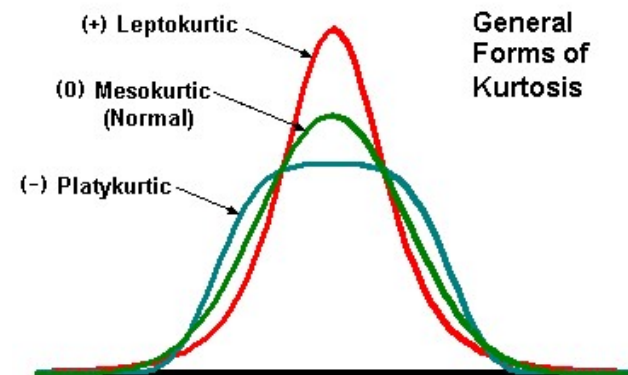
- 1) Winsorisation
- 2) Trimming
- 3) Box-Cox Transformation

Identification by Skewness and Kurtosis (1)

Skewness: measure of the asymmetry of a distribution
(= 0 in the Normal distribution)



Kurtosis: measure of the thickness of the tails of a distribution
(= 3 in the Normal distribution)



(+) higher peak
around the mean
and fatter tails

(-) fatter around
the mean and
thinner tails

Identification by Skewness and Kurtosis (2)

A distribution contains one or more outliers if:

$$|\text{skewness}| > 2$$

and

$$\text{kurtosis} > 3.5$$

Groeneveld and Meeden (1984)

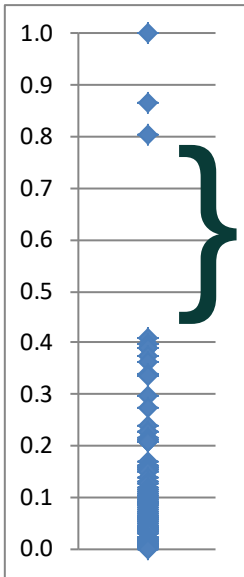
variable	min	p10	p25	mean	p50	p75	p90	max	sd	cv	skewness	kurtosis	N
Var_1	2,12	2,34	2,61	3,26	2,99	3,66	4,76	5,89	0,92	0,28	1,17	3,63	133
Var_2	1,91	2,79	3,16	3,90	3,68	4,43	5,40	6,19	0,97	0,25	0,52	2,54	133
Var_3	2,09	2,47	2,65	3,28	3,01	3,62	4,67	6,02	0,90	0,27	1,28	4,07	133
Var_4	2,20	2,57	3,04	3,62	3,41	4,06	4,94	5,90	0,86	0,24	0,71	2,84	133
Var_5	2,29	2,84	3,20	3,64	3,57	4,05	4,39	5,50	0,61	0,17	0,25	2,80	133
Var_6	2,70	3,10	3,53	4,14	4,16	4,68	5,18	6,01	0,77	0,19	0,17	2,34	133
Var_7	0,00	0,00	0,00	18,55	0,40	3,24	71,09	200,00	44,35	2,39	2,74	9,89	133
Var_8	1,70	2,46	2,81	3,76	3,54	4,61	5,66	6,21	1,17	0,31	0,53	2,21	133

Yet, even if skewness low ...

Treatment of Outliers: Winsorisation

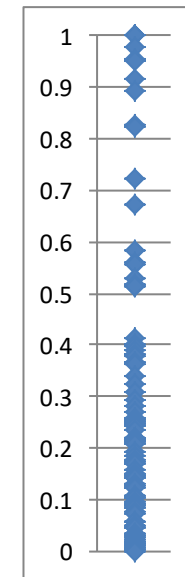
Winsorisation: outliers' values are assigned the next highest(/lowest) value (up to when the indicator's distribution satisfies the desired rule-of-thumb for identification)

No outlier treatment
(minmax normalized data)



Country	Raw data	Winsorized	Trimmed
CHE	17.1	17.1	17.1
SWE	5.9	5.9	5.9
NLD	7.2	7.2	7.2
USA	4.6	4.6	4.6
GBR	18.2	18.2	18.2
DNK	9.9	9.9	9.9
SGP	19.2	19.2	19.2
FIN	7.7	7.7	7.7
DEU	7.7	7.7	7.7
IRL	7.0	7.0	7.0
KOR	1.7	1.7	1.7
ISL	6.5	6.5	6.5
LUX	40.6	19.2	
JPN	3.4	3.4	3.4
FRA	0.0	0.0	0.0

Winsorized
(minmax normalized data)

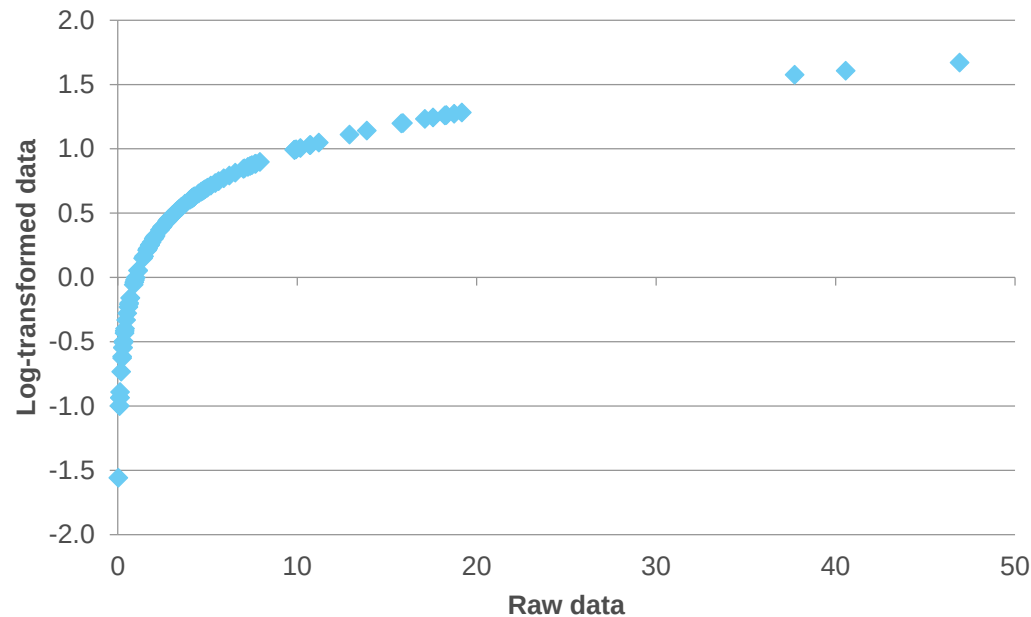
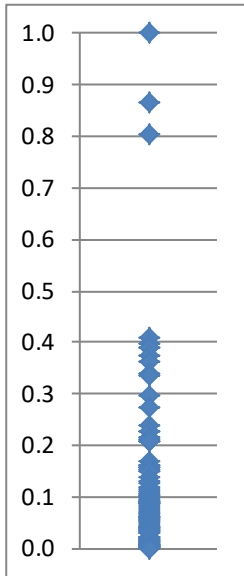


data-points spread homogeneously across the scale

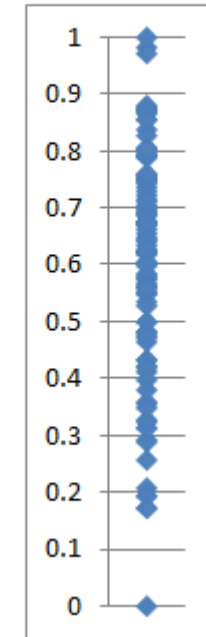
Treatment of Outliers: Transformation

The logarithm of all the observations is computed

No outlier treatment
(minmax normalized data)



Winsorized
(minmax normalized data)



data-points
spread
homogeneously
across the scale

The suggestion from JRC - COIN

➤ JRC suggested outlier identification method

○ Skewness & Kurtosis

-> identifies less outliers, change original data as less as possible

➤ JRC suggested outlier treatment methods

○ Winsorisation, with less than 5 outliers

-> change outliers only

○ Log-transformation, with 5 or more outliers

-> change all observations

What is a missing data?

“Missing data (or missing values) is defined as the data value that is not stored for a variable in the observation of interest”

- *Kang, 2013. The prevention and handling of the missing data.*

Missing variables, missing units

Rows: **units**, cases, observations or subjects

	Var1	Var2	Var3	Var4	Var5	Var6	Var7	Var8	Var9	Var10	Missing per country
C1		x			x						2
C2	x		x					x	x	x	5
C3	x	x				x					3
C4											0
C5											0
C6	x		x		x	x		x		x	6
C7											0
C8											0
C9	x			x		x	x		x	x	6
C10											0
C11						x					1
C12											0
C13											0
C14			x		x						2
C15						x					1
C16			x								1
C17											0
C18		x				x					2
C19											0
C20						x					1
Missing per indicator	4	3	4	1	3	7	1	2	2	3	10

Columns: **variables** measured for each unit

Try to ask yourself **why** the information isn't available...

...and how you can deal with it

Handling Missing Data – Deletion Methods

Listwise deletion - discard

Country	Pupil-teacher ratio, secondary	Tertiary enrolment	Graduates in science and engineering
DNK	11.3	81.5	20.4
SGP	14.9	69.8	N/A
FIN	12.8	87.3	27.9
DEU	12.1	68.3	N/A
IRL	N/A	77.6	23.8
KOR	15.6	95.3	31.9
ISL	N/A	81.3	15.6

Pros: simple; same number of cases for all indicators

Cons: loss of information; loss of precision, and potential bias

vs.

Pairwise deletion - ignore

Country	Pupil-teacher ratio, secondary	Tertiary enrolment	Graduates in science and engineering
DNK	11.3	81.5	20.4
SGP	14.9	69.8	N/A
FIN	12.8	87.3	27.9
DEU	12.1	68.3	N/A
IRL	N/A	77.6	23.8
KOR	15.6	95.3	31.9
ISL	N/A	81.3	15.6

Pros: simple; retains more data compared to listwise

Cons: it is data treatment even if it goes unnoticed (“**shadow imputation**”); might encourage countries not to report bad performances

The real effect of ignoring missing data

“shadow imputation” == assigning mean value of indicators observed for that unit (row mean)

Country	Pupil-teacher ratio, secondary	Tertiary enrolment	Graduates in science and engineering	Mean
DNK	11.3	81.5	20.4	37.7
SGP	14.9	69.8	N/A	42.4
FIN	12.8	87.3	27.9	42.7
DEU	12.1	68.3	N/A	40.2
IRL	N/A	77.6	23.8	50.7
KOR	15.6	95.3	31.9	47.6
ISL	N/A	81.3	15.6	48.5

Country	Pupil-teacher ratio, secondary	Tertiary enrolment	Graduates in science and engineering	Mean
DNK	11.3	81.5	20.4	37.7
SGP	14.9	69.8	42.4	42.4
FIN	12.8	87.3	27.9	42.7
DEU	12.1	68.3	40.2	40.2
IRL	50.7	77.6	23.8	50.7
KOR	15.6	95.3	31.9	47.6
ISL	48.5	81.3	15.6	48.5

Imputation Methods

implicit modelling

- substitution
- cold deck
- hot deck (e.g. k-NN)

explicit modelling

- mean imputation
- regression imputation
- likelihood based (e.g. EM algorithm)

multiple imputation

- creates m -imputed data sets (e.g. Amelia II software)

Hot Deck Imputation: Nearest Neighbour (kNN)

Replaces missing values for a non-respondent (**recipient**) with observed values from a respondent (**donor**) “similar” to the recipient with respect to observed characteristics

Step 1. Compute the distance / similarity between recipient and potential donors

Country	Expenditure on education	Government expenditure on education per pupil, secondary	School life expectancy
SGP	2.9	16.7	12.8
DEU	4.9	23.7	17.3
IRL	5.3	26.0	19.0
KOR	4.6	23.4	16.6
ISL	7.8	18.3	19.6
LUX	4.1	19.4	13.9
JPN	3.8	25.1	15.4
FRA	5.5	26.8	16.3
HKG	3.3	20.4	N/A

Country	Distance	
	Euclidean	Manhattan
SGP	3.68	4.02
DEU	3.74	5.02
IRL	6.00	7.70
KOR	3.31	4.37
ISL	4.97	6.55
LUX	1.30	1.83
JPN	4.79	5.27
FRA	6.85	8.72
HKG	0	0

	closest country
	2nd closest country
	3rd closest country

Step 2. The imputed value for the recipient is the observed value on the most similar unit, or the mean value of the k -closest

Number of neighbours	Distance type	Imputed value	
1NN	Euclidean	13.9	
1NN	Manhattan	13.9	
2NN	Euclidean	15.3	$[= (13.9+16.6)/2]$
2NN	Manhattan	13.4	$[= (13.9+12.8)/2]$

Pros: uses actual values (easy to communicate); does not impose a structure on relationships between variables.

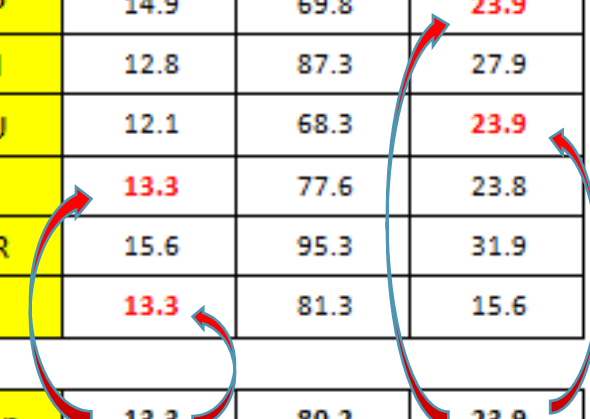
Cons: might be computational-intensive; might reduce variance, but typically less than mean substitution.

Mean (or median or mode) imputation

Unconditional mean imputation (by column)

Country	Pupil-teacher ratio, secondary	Tertiary enrolment	Graduates in science and engineering
DNK	11.3	81.5	20.4
SGP	14.9	69.8	N/A
FIN	12.8	87.3	27.9
DEU	12.1	68.3	N/A
IRL	N/A	77.6	23.8
KOR	15.6	95.3	31.9
ISL	N/A	81.3	15.6

Country	Pupil-teacher ratio, secondary	Tertiary enrolment	Graduates in science and engineering
DNK	11.3	81.5	20.4
SGP	14.9	69.8	23.9
FIN	12.8	87.3	27.9
DEU	12.1	68.3	23.9
IRL	13.3	77.6	23.8
KOR	15.6	95.3	31.9
ISL	13.3	81.3	15.6
Mean	13.3	80.2	23.9



Pros: simple, relies on the observed values from the same variable.

Cons: correlations are affected; variances will be typically under-estimated

Main rules to remember

- *Imputations often unreliable if with more than 40% of missing values*
- **@indicator-wise**: at least 50% of units should have valid data for that indicator – otherwise drop indicator and consider an alternative one
- **@unit-wise**: at least 65-75% of the indicators for the unit should have valid data (at every aggregation level not only overall!)
- Avoid as much as possible using different methods for different indicators

Suggested readings and references (Outliers)

- Aggarwal, C. C. (2015) Outlier analysis. In *Data mining* (pp. 237-263). Springer, Cham.
- Atkinson, A.C., Riani, M. & A. Cerioli (2004) Exploring Multivariate Data with the Forward Search. Springer-Verlag.
- Ghosh, D., & A. Vogt (2012) Outliers: an evaluation of methodologies. American Statistical Association. Section on Survey Research Methods – JSM 2012.
- Groeneveld, R. A., & Meeden, G. (1984). Measuring Skewness and Kurtosis. *The Statistician* 33: 391–99.
- Grubbs, F. E. (1969) Procedures for detecting outlying observations in samples" *Technometrics* 11 (1): 1–21.
- Hawkins, D. (1980) Identification of Outliers. Chapman and Hall.
- Knoke, B. & P. Mee (2002) Statistics for social data analysis. Wadsworth Publishing.
- Tukey, John W (1977). Exploratory Data Analysis. Addison-Wesley.

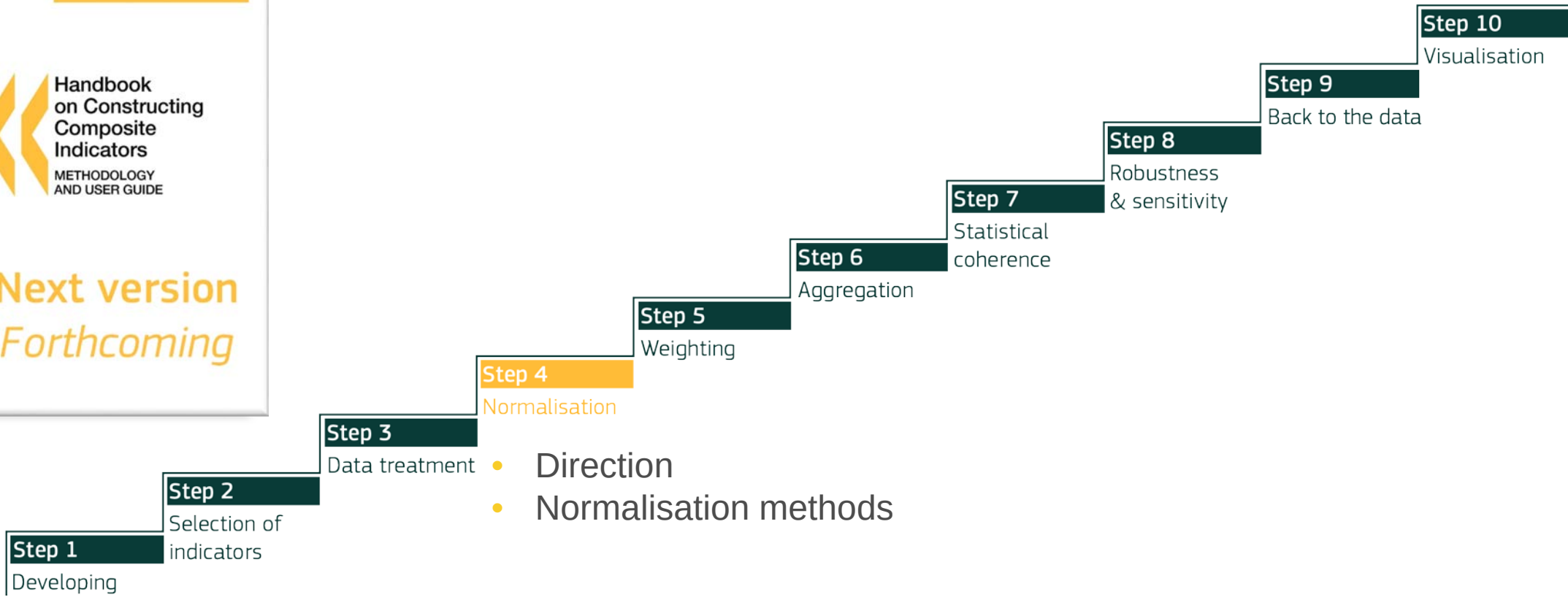
Suggested readings and references (Missing Data)

- Beretta, L. and A. Santaniello (2016). Nearest neighbor imputation algorithms: a critical evaluation. BMC Medical Informatics and Decision Making, 16 (Suppl 3):74.
- Chen, Y. & M.R. Gupta, 2010 EM Demystified: An Expectation-Maximization Tutorial. Department of Electrical Engineering. University of Washington
- Dondersa et al., 2006. Review: A gentle introduction to imputation of missing values. Journal of Clinical Epidemiology. 59: 1087-1091
- Enders, C. K., 2010, Applied Missing Data Analysis. The Guilford Press. Inc: New York, London
- He, Y., 2010, Missing Data Analysis Using Multiple Imputation: Getting to the Heart of the Matter. Circ Cardiovasc Qual Outcomes. 3(1): 98
- Graham, J. W., 2012. Missing data: Analysis and design. New York: Springer.
- Kang, H., 2013, The prevention and handling of the missing data. The Korean Journal of Anesthesiology. 64(5): 402–406.
- Little, R.A & Rubin, D.A., 2002. Statistical Analysis with missing data. Second edition. Hoboken, John Wiley & Sons.
- Schafer, J. L. & Graham, J.W., 2002, Missing Data: Our View of the State of the Art Psychological Methods. 7(2):147–177

Suggested Software for Missing Data

Software Package	link
Freeware	
Amelia	http://gking.harvard.edu/amelia
CAT	http://cat.texifter.com/ (for categorical data)
EMCOV	https://methodology.psu.edu/publications/books/missing
NORM	https://methodology.psu.edu/publications/books/missing
MICE	http://www.stefvanbuuren.nl/mi/index.html
PAN	http://stat.ethz.ch/~maechler/adv_topics_compstat/MissingData_Imputation.html (Free with R, commercial with S-Plus, for clustered data, including longitudinal data).
Commercial Software	
AMOS	https://www.ibm.com/us-en/marketplace/structural-equation-modeling-sem
EQS	http://www.mvsoft.com
HLM	http://www.ssicentral.com/hlm/index.html
LISREL	http://www.ssicentral.com/index.html
Mplus	http://www.statmodel.com
SAS	https://www.sas.com/it_it/home.html
SOLAS	https://www.statcon.de/shop/en/software/statistics/solas
S-Plus	http://www.solutionmetrics.com.au/products/splus/default.html
SPSS	http://www-01.ibm.com/software/analytics/spss/products/statistics/modules/
Stata	http://www.stata.com , <i>mi</i> command; installing <i>ice</i> or <i>mvis</i>
Source: updated from Acock, 2005	

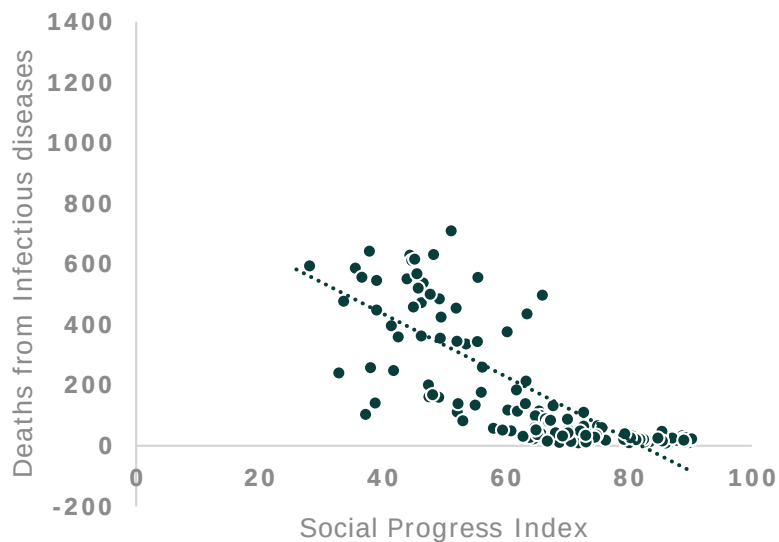
Normalisation of Data



Before Normalisation

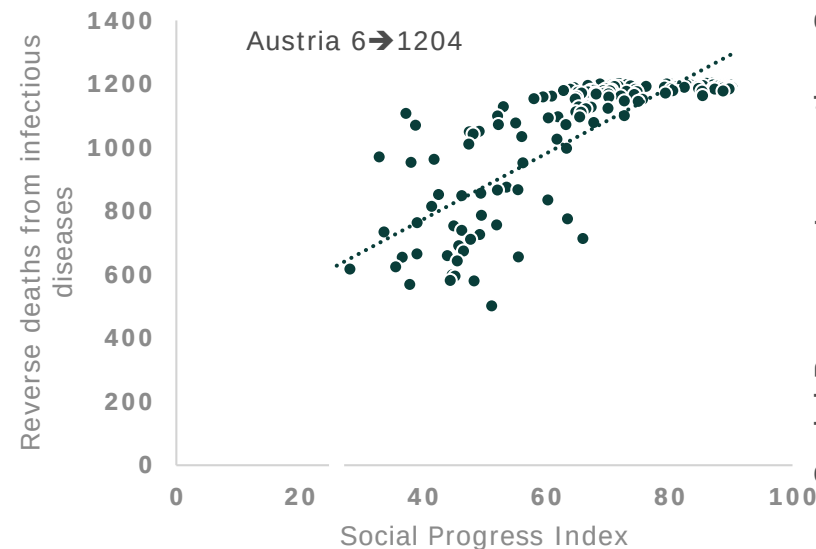
Adjust for direction

**Social Progress Index and
deaths from infectious diseases**



**Social Progress Index and
reverse deaths from infectious diseases**

Reverse I = (max-x)
max=1210



Make sure that higher values in the dataset mean better results, if not, reverse the original direction

What is data normalisation?

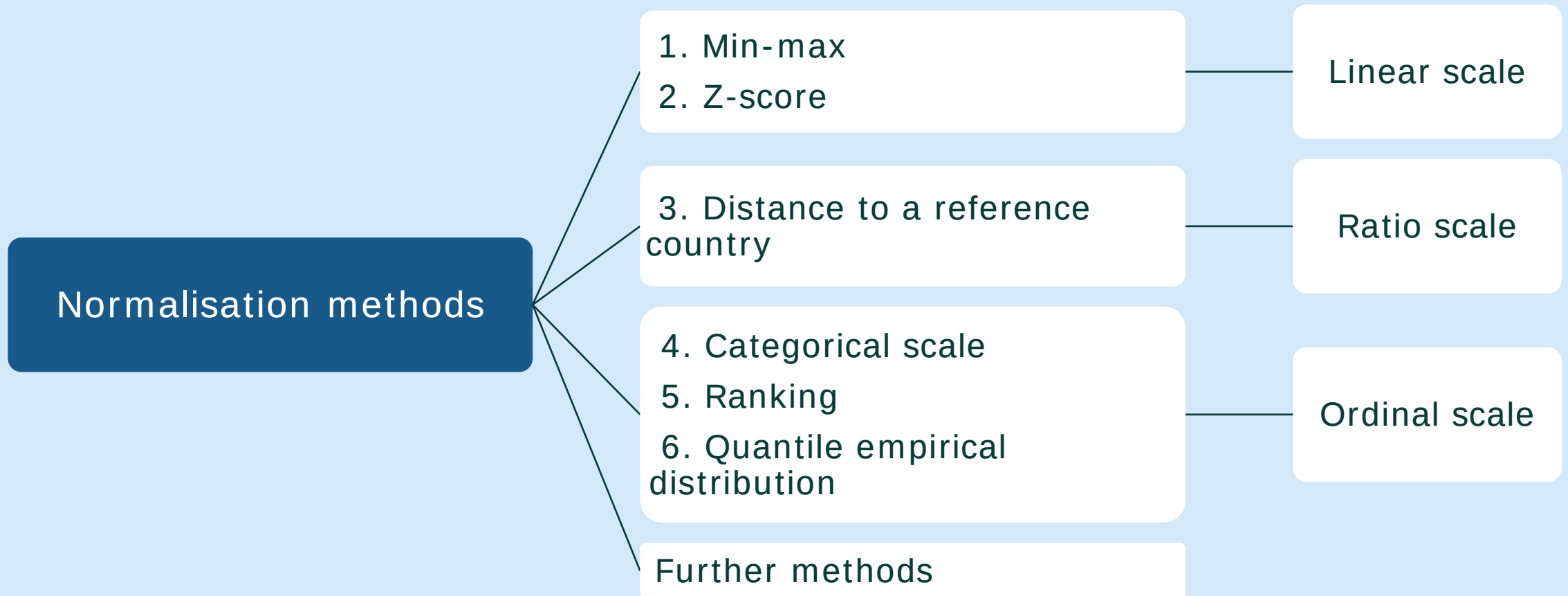
Definition:

... is the adjustment of variables onto a common scale,
prior to any data aggregation.

Aim: comparability across variables by dealing with:

1. different units of measurement
2. different ranges of variation

Overview



1. the normalisation method should respect the **conceptual framework** and **the data properties**
2. different normalisation methods may lead to different rankings

Z- Score normalisation

Indicators before and after z-score normalisation	3. Reading, maths & science scores aged 15 (Pisa score)	4. Recent training	6. High computer skills
Before normalisation			
Mean	486.94	10.85	29.18
Variance	23.44	7.61	7.93
Min	437.49	1.20	7.00
Max	524.29	29.60	46.00
Variation range	[437.49, 524.29]	[1.2, 29.6]	[7, 46]

After normalisation using min-max			
Mean	0.55	0.34	0.57
Variance	0.08	0.07	0.04
Min	0.00	0.00	0.00
Max	1.00	1.00	1.00
Variation range	[0, 1]	[0, 1]	[0, 1]

How are the two indicators different?

- 1. Units of measurement
- 2. Ranges of variation

$$I = \frac{x - \min}{\max(x) - \min(x)}$$

Min-max effects

- Unit of measurement
- μ, σ^2 no adjustments
- I Variation range: [0, 1]
- Extreme values: no adjustments
- Distribution: no adjustments

Z- Score normalisation

Indicators before and after z-score normalisation	3. Reading, maths & science scores aged 15 (Pisa score)	4. Recent training	6. High computer skills
Before normalisation			
Mean	486.94	10.85	29.18
Variance	23.44	7.61	7.93
Min	437.49	1.20	7.00
Max	524.29	29.60	46.00
Variation range	[437.49, 524.29]	[1.2, 29.6]	[7, 46]

After z-score normalisation			
Mean	0	0	0
Variance	1	1	1
Min	-2.11	-1.27	-2.80
Max	1.59	2.46	2.12
Variation range	[-2.11, 1.59]	[-1.27, 2.46]	[-2.8, 2.12]

How are the two indicators different?

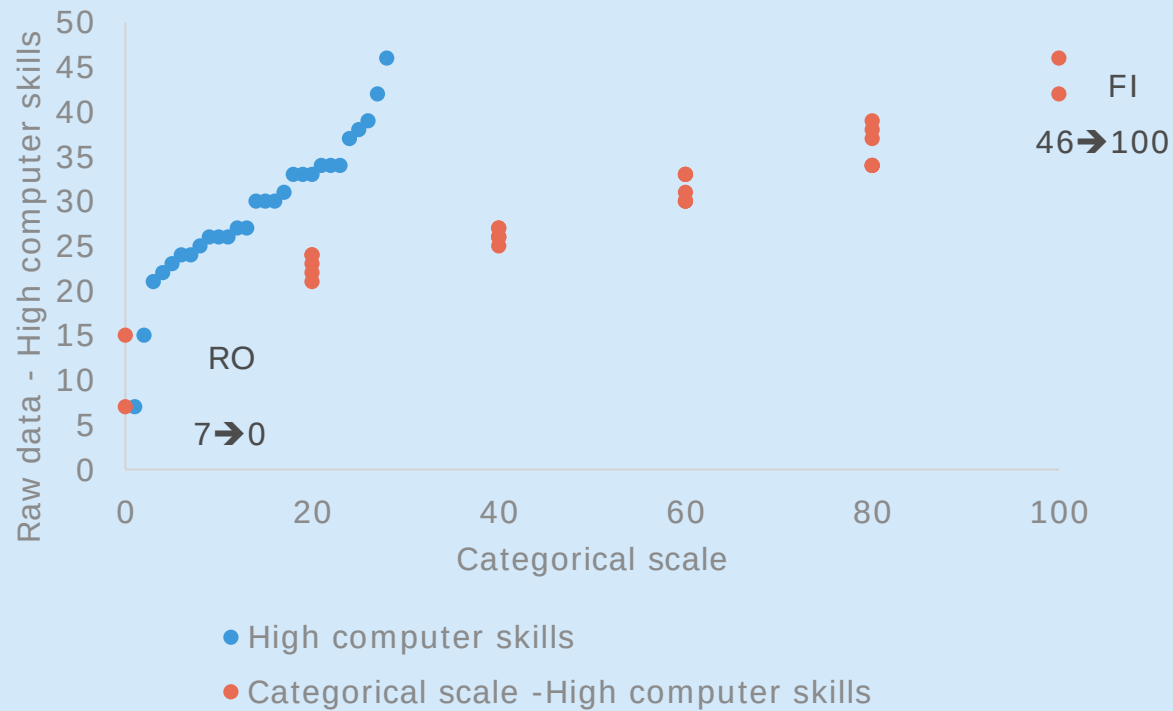
- 1. Units of measurement
- 2. Ranges of variation

$$Z = \frac{x - \mu}{\sigma}$$

Z-score effects

- Unit of measurement
- $I \sim \mu = 0, \sigma^2 = 1$
- Variation range: no adjustments
- Extreme values: no adjustments
- Distribution: no adjustments

Categorical Scale



Indicators after categorical scaling

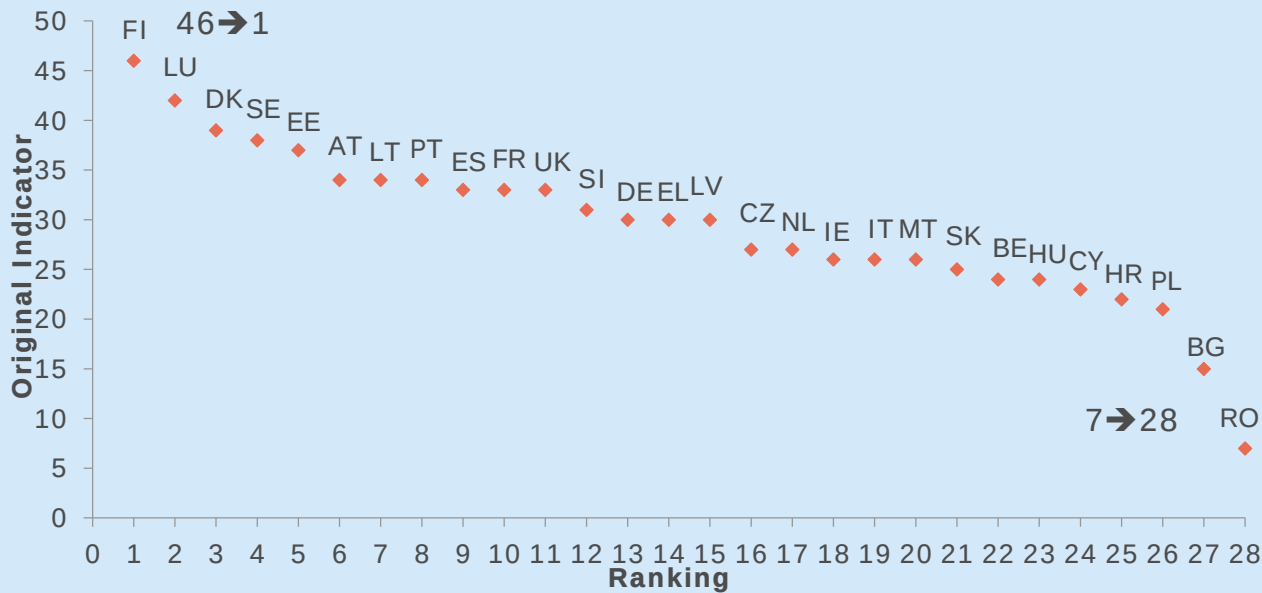
$$I_{q,c}^t = \begin{cases} 0 & \text{if } p^0 \leq x < p^{10} \\ 20 & \text{if } p^{10} \leq x < p^{25} \\ 40 & \text{if } p^{25} \leq x < p^{50} \\ 60 & \text{if } p^{50} \leq x < p^{75} \\ 80 & \text{if } p^{75} \leq x < p^{90} \\ 100 & \text{if } p^{90} \leq x \leq p^{100} \end{cases}$$

Categorical scale effects

- Unit of measurement
- Variation range [0, 100]
- Variance: *depends on the categories*
- Robust to extreme values
- Distribution: No uniform

Ranking

High computer skills Ranking



Indicators after ranking scale

- Scores are replaced by ranks, e.g. the highest score receives rank 1
- Uses only ordinal information, information on levels is not kept

$$I = \text{rank}(x)$$

Ranking scale effect

- Unit of measurement
- Range $[1, n]$, our case $n=28$
- Same variance, our case $\sigma^2 = 65.25$
- Robust to extreme values
- Distribution: Uniform

Source: European Skills Index, 2018

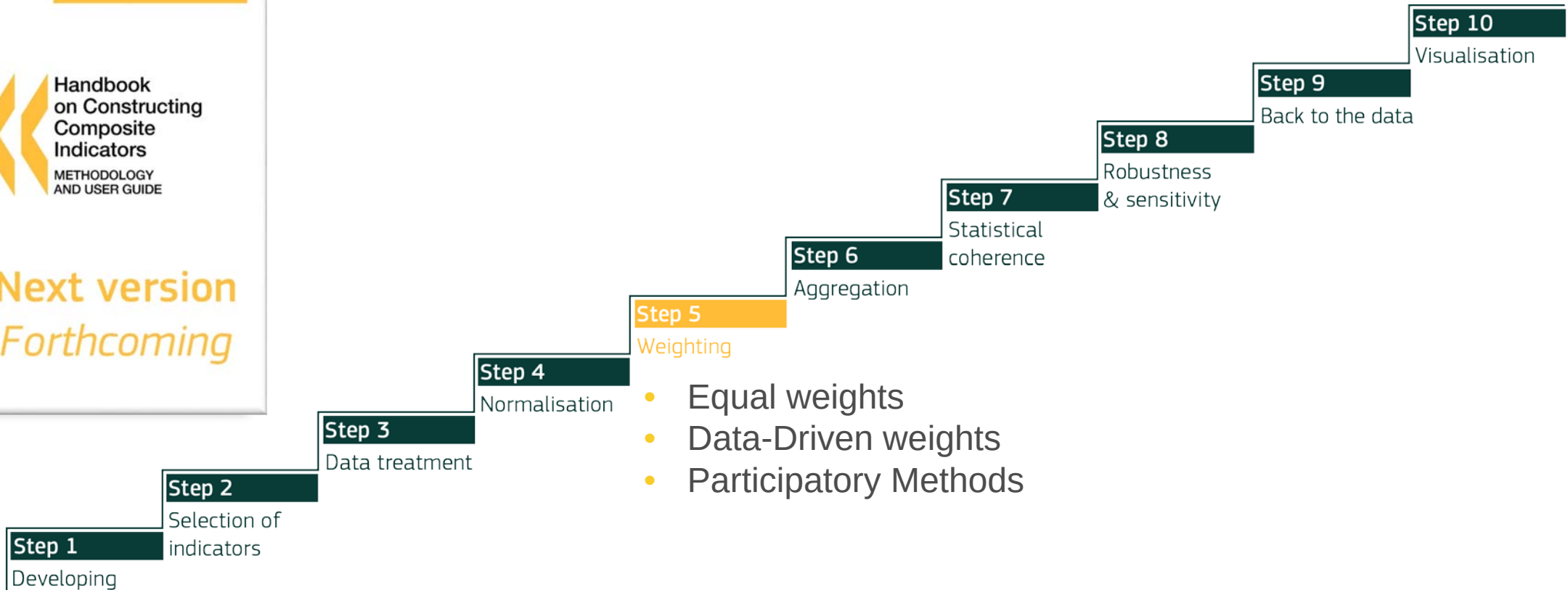
Summary

Adjustments/Method	Quantile empirical distribution	Ranking	Categorical scale	Z-score	Min-max	Distance to a reference country
Unit of measurement	Y	Y	Y	Y	Y	Y
Variance	Y	Y	Y/N	Y	N	N
Range of variation	Y	Y	Y	N	Y	N
Extreme values*	Y	Y	Y	N	N	N
Distribution*	Y	Y	Y/N	N	N	N
Extreme values*	non-sensitive to extreme values					
Distribution*	the distribution will be the same for the normalised indicators					

Suggested readings and references (Normalisation)

- Becker, W., M. Saisana, P. Paruolo, and I. Vandecasteele. 2017. 'Weights and Importance in Composite Indicators: Closing the Gap. Ecological Indicators 80: 12–22.
- Bichindaritz I., Big Data, Genes, and Medicine, the State University of New York (SUNY), <https://www.coursera.org/lecture/data-genes-medicine/data-normalisation-jGN7k>
- Cohen, A., Saisana, M., 2014, Quantifying the Qualitative: Eliciting Expert Input to Develop the Multidimensional Poverty Assessment Tool, Journal of Development Studies 50(1): 35-50.
- European Union, European Centre for the Development of Vocational Training (Cedefop). 2018, European Skills Index technical report. Thessaloniki, Greece.
- OECD/JRC, 2008, Handbook on Constructing Composite Indicators. Methodology and user Guide, OECD Publishing, ISBN 978-92-64-04345-9.
- Paruolo P., Saisana M., Saltelli A., 2013, Ratings and Rankings: voodoo or science?. J Royal Statistical Society A 176(3), 609-634.
- Saisana, M., and Saltelli, A., 2011, Rankings and Ratings: Instructions for use, Hague Journal on the Rule of Law 3(2), 247-268.
- Saisana M., D'Hombres B., Saltelli A., 2011, Rickety Numbers: Volatility of university rankings and policy implications. Research Policy 40, 165–177
- Social Progress Imperative, Social Progress Index, 2018

Definition of Weights



Equal Weights

They are applied when there is no clear reference in the literature about the importance of elements the composite indicator

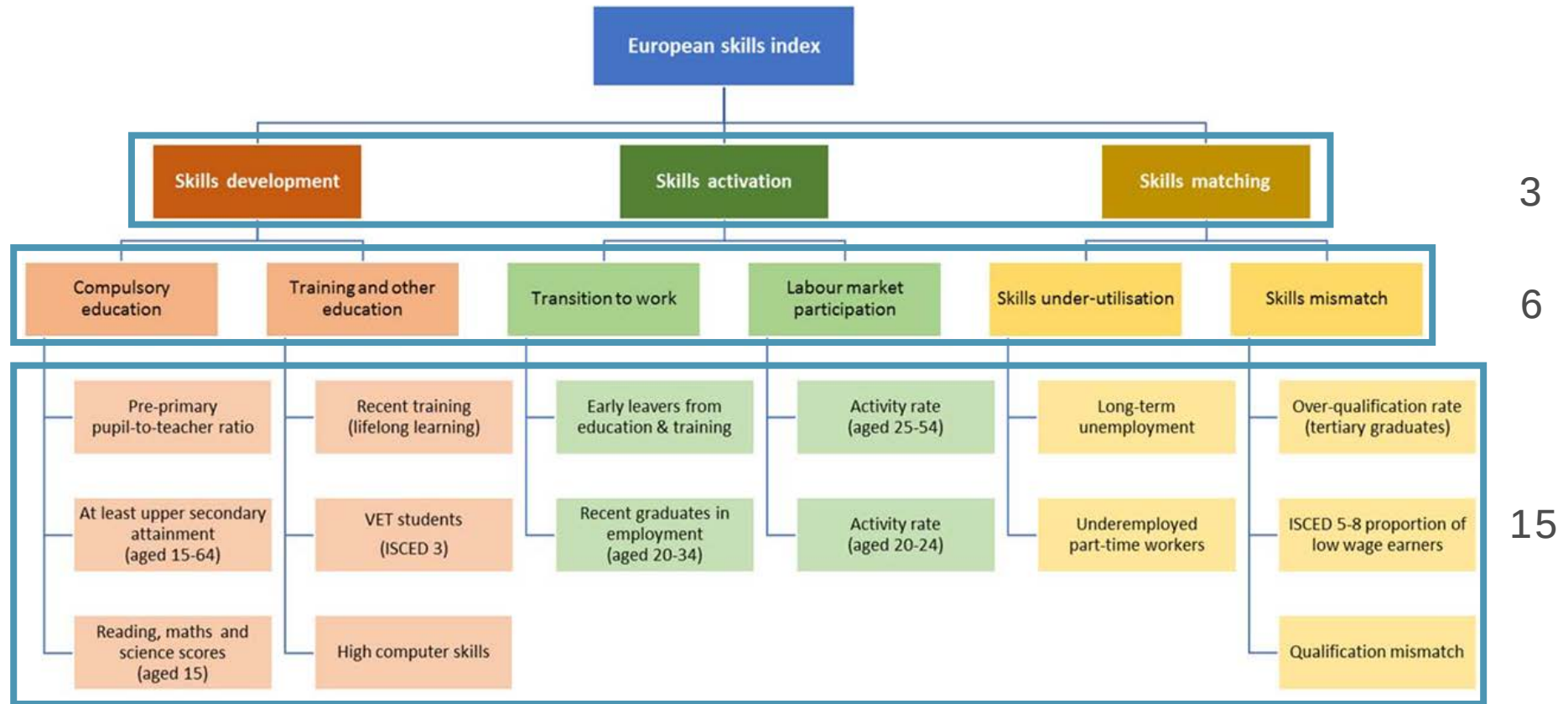
Straightforward method

Easy to communicate

However:

Equal weighing does not guarantee equal importance and equal contribution of the indicators to the composite indicator!

Example: European Skills Index



Source: European Skills Index (2019), Cedefop

Example: European Skills Index



Equal Weights

Pillars	Pearson Correlation Coefficient	R ²
Skills Development	0.80	0.64
Skills Activation	0.81	0.66
Skills Matching	0.64	0.41

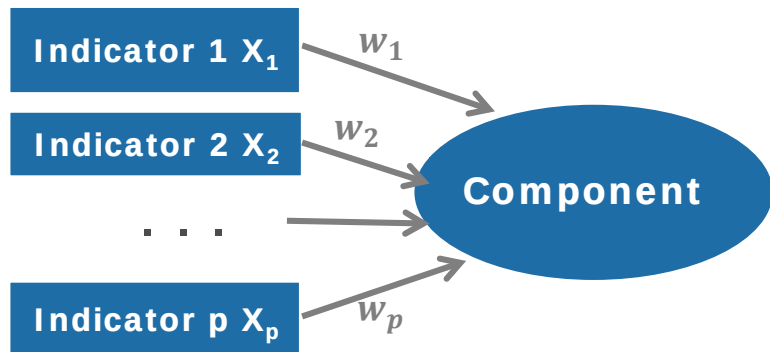
Adjusted Weights

Pillars	Pearson Correlation Coefficient	R ²
Skills Development	0.77	0.59
Skills Activation	0.76	0.58
Skills Matching	0.71	0.50

The unequal weights result in equal importance

Principal Component Analysis (PCA)

Identify a small number of PCs that explain most of the variance observed.



Observed indicators are reduced into components

PCA summarizes information of all indicators and reduces it into a **fewer number of components**

Each principal component PC_i is a new variable computed as **a linear combination** of the original (standardized) variables

$$PC_1 = w_1x_1 + w_2x_2 + \cdots w_px_p$$

Weights derived from PCA



Total variance explained

Component	Eigenvalue	Variance	Cumulative variance
PC1	1.75	58	58
PC2	0.88	29	88
PC3	0.37	12	100

Stopping criterion *Eigenvalue* > 1.0 (or cumulative variance over 0,70)

Pearson correlation coefficient

Pillars	PC1	PC2	PC3
Skills Development	0.88	-0.16	-0.44
Skills Activation	0.84	-0.36	0.41
Skills Matching	0.51	0.85	0.09

PC1 _{norm}
0.40
0.38
0.23

$$ESI = 0.40 P_1 + 0.38 P_2 + 0.23 P_3$$

Weights derived from PCA



Equal Weights

Pillars	Pearson	
	Coefficient	R ²
Skills Development	0.80	0.64
Skills Activation	0.81	0.66
Skills Matching	0.64	0.41

PCA Weights

Pillars	Pearson	
	Coefficient	R ²
Skills Development	0.84	0.70
Skills Activation	0.87	0.75
Skills Matching	0.52	0.27

Stronger difference

Adjusted Weights

Pillars	Pearson	
	Coefficient	R ²
Skills Development	0.77	0.59
Skills Activation	0.76	0.58
Skills Matching	0.71	0.50

All Pillars are equally important!

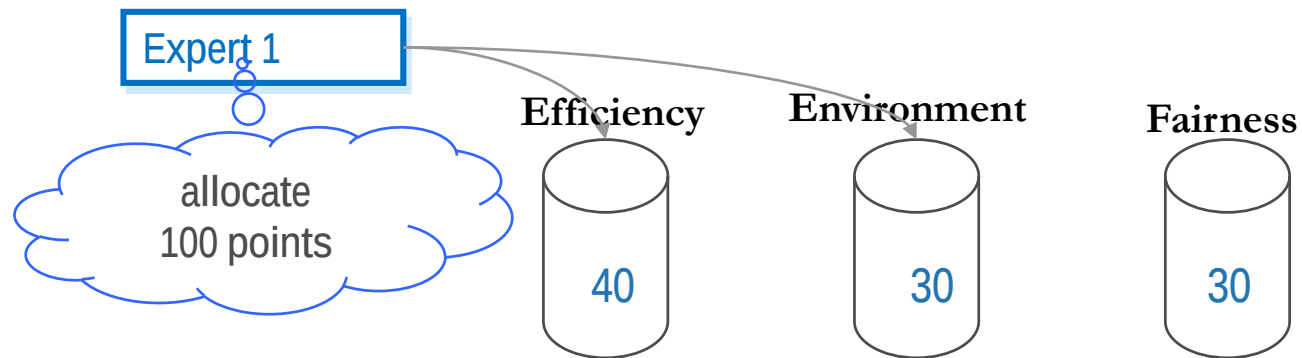
Participatory Methods

- **Weights based on participation**
 - *Budget allocation*
 - *Analytic hierarchy process*
 - *Conjoint analysis*

Budget Allocation - BAL

Phases

1. Selection of **experts/stakeholders** for the evaluation;
 - a. Number
 - b. Background/Expertise
2. **Allocation** of budget to indicators;



Suited for up to **8-10** indicators

3. Calculation of **weights**;
4. Iteration of the budget allocation until convergence is reached (optional)

Analytic Hierarchy Process (AHP)

- **Multi-criteria** decision making method
- Developed by **Thomas Saaty** (1980, 1987)



Recommended for less
than **10** indicators

Phases

1. Selection of **experts/stakeholders** for the evaluation;
2. **Pairwise comparisons** of indicators on a scale 1 to 9 (1: equally important, 9: most important);
3. Calculation of **weights** through the derivation of the priority vector;
4. Estimation of **consistency of the experts' assessment**.

Suggestions for participatory methods

- ❑ When possible- use a **sampling frame** to select the experts & **maximize response rate**
→ **Compensating** experts might increase participation (Chowdury and Squire, 2006)
- ❑ **Experts** with balance of **diverse backgrounds**
- ❑ Collect information on the **characteristics** of the experts (Cooke, 1991)
- ❑ **Randomize** the order of the components, so that some experts evaluate first component A and others component B, and so on.

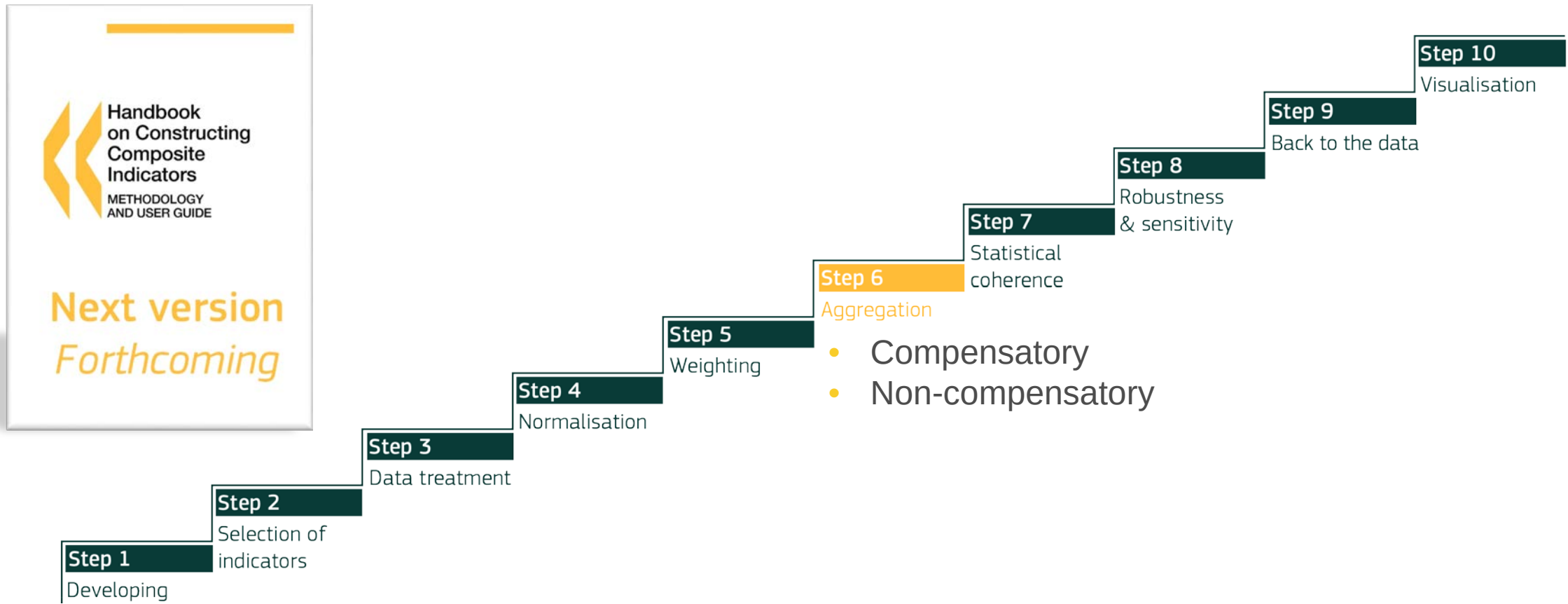
Suggested readings and references (Weighting)

- Becker W., Saisana M., Paruolo P., Vandecasteele I. “Weights and importance in composite indicators: Closing the gap”. Ecological Indicators 80:12-22, 2017
- Härdle W.K., Simar L. “Applied Multivariate Statistical Analysis”, Springer, 2012
- Johnson R., Wichern D., “Applied Multivariate Statistical, 6th edition”, Pearson New International Edition, 2013
- Jolliffe, “Principal Component Analysis, Second Edition”, Springer, 2002
- Paruolo P., Saisana M. and Saltelli A., “Ratings and rankings: voodoo or science?.” Journal of the Royal Statistical Society: Series A (Statistics in Society) 176.3: 609-634, 2013
- Chowdhury S., & Squire L., 2006, Setting weights for aggregate indices: An application to the commitment to development index and human development index. Journal of Development Studies, 42, 761–771.
- Cohen, A., Saisana, M., 2014, Quantifying the Qualitative: Eliciting Expert Input to Develop the Multidimensional Poverty Assessment Tool, Journal of Development Studies 50(1): 35-50.
- Cooke, R. (1991). Experts in uncertainty: Opinion and subjective probability in science. New York: Oxford University Press.
- Green P.E., and Srinivasan V., 1978, Conjoint analysis in consumer research: issues and outlook, Journal of Consumer Research 5: 103-123.
- Hair J.F., Anderson R.E., Tatham R.L., and Black W.C., 1995, Multivariate data analysis with readings, fourth ed. Prentice Hall, Englewood Cliffs, NJ.
- Hair J.F., Black W.C., Babin, Anderson R.E. and R.L., Tatham, 2006, Multivariate data analysis, sixth edition, Pearson Prentice Hall, Upper Saddle River, NJ.

Suggested readings and references (Weighting)

- Kucukusta, D., Denizci Guillet, B., 2014 Measuring spa-goers' preferences: A conjoint analysis approach. *Int. J. Hosp. Manag.* 41, 115–124
- McDaniel C. and Gates R. ,1998, *Contemporary Marketing Research*. West Publishing, Cincinnati, OH.
- Montalto V., Jorge Tacao Moura C., Langedijk S. and Saisana M. 2017. *The Cultural and Creative Cities Monitor*, Scientific and Technical Research Report, JRC105971
- Saaty T. L. ,1980, *The Analytic Hierarchy Process*, New York: McGraw-Hill.
- Saaty R.W.,1987, The analytic hierarchy process: what it is and how it is used, *Mathematical Modelling*, 9: 161-176.
- Saaty, T.L.,1990, The Analytic Hierarchy Process in Conflict Management, *International Journal of Conflict Management*, 1(1):47 – 68
- Saaty, T.L, 2003, Decision-making with the AHP: Why is the principal eigenvector necessary, *European Journal of Operational Research* 1145:85-91
- Saisana M., Saltelli A., 2008, Expert Panel Opinion and Global Sensitivity Analysis for Composite Indicators, *Lecture Notes in Computational Science and Engineering* 62: 251-275.
- Sevigny, E. L., Saisana M., 2013, *Developing the U.S. Drug Consequences Indices, 2000-2009*. Washington, DC: Office of National Drug Control Policy, Executive Office of the President.
- Sydorovych, O., Wossink, A., 2008, The meaning of agricultural sustainability: Evidence from a conjoint choice survey. *Agric. Syst.* 98,

The aggregation method



Aggregation Methods

- Based on Average Scores
 1. Arithmetic Mean
 2. Geometric Mean
- Based on Ranks
 3. Median rank
 4. Majority
 5. Borda's Count
- Based on Pairwise Comparisons
 6. Condorcet
 7. Kemeny
 8. Arrow – Raynaud
 9. Copeland



Arithmetic Mean

The *arithmetic mean* of a list of n real numbers equals:

$$\frac{1}{n} \sum_{i=1}^n x_i = \frac{x_1 + x_2 + \cdots + x_n}{n}$$

This is the simplest, most obvious and most widespread aggregation method

Perfect (and constant) substitutability – *the impact of a unit increase (decrease) in the level of any indicator is the same, i.e. underperformance in one component can be perfectly compensated by equivalent overperformance in another*

Geometric Mean

The ***geometric mean*** of a list of n positive real numbers equals:

$$\sqrt[n]{\prod_{i=1}^n x_i} = \sqrt[n]{x_1 \times x_2 \times \cdots \times x_n}$$

Partial substitutability – compensation of low performance on some indicators by high performance on others is possible only partially, i.e. *unbalanced performance is always penalised* by the aggregation formula when compared to arithmetic aggregation

The Monkey's Dilemma



Safe: 1
Yummy: 10

A.M.: **5.5**

G.M.: 3.2



Arithmetic mean results in more tolerance towards extreme values (high-risk-low-taste and vice versa); therefore, if not willing to trade-off very high (low) risk levels for very good (bad) taste, geometric mean can guide you to a more conservative decision!



Safe: 5
Yummy: 5

A.M.: 5.0

G.M.: **5.0**



Safe: 9
Yummy: 2

A.M.: **5.5**

G.M.: 4.2

Aggregation Methods

- Based on Average Scores
 1. Arithmetic Mean
 2. Geometric Mean
- Based on Ranks
 3. Median rank
 4. Majority
 5. Borda's Count
- Based on Pairwise Comparisons
 6. Condorcet
 7. Kemeny
 8. Arrow – Raynaud
 9. Copeland



Borda's Count

Assigns a score according to the ranking in every indicator

$$C = (0.30 + 0.15) \times 2 + (0.05 + 0.20) \times 1 = 1.15$$

6 indicators/ criteria with (unequal) weights

Ind. 1	Ind. 2	Ind. 3	Ind. 4	Ind. 5	Ind. 6
0.05 weight	0.30 weight	0.15 weight	0.10 weight	0.10 weight	0.20 weight
A	C	C	B	A	B
C	A	B	A	B	C
B	B	A	C	C	A

Borda Points
2
1
0

(weighted) Scores	
A	0.70
B	0.85
C	1.15

Aggregation Methods

- Based on Average Scores
 1. Arithmetic Mean
 2. Geometric Mean
- Based on Ranks
 3. Median rank
 4. Majority
 5. Borda's Count
- Based on Pairwise Comparisons
 6. Condorcet
 7. Kemeny
 8. Arrow – Raynaud
 9. Copeland



What if we don't have numbers?



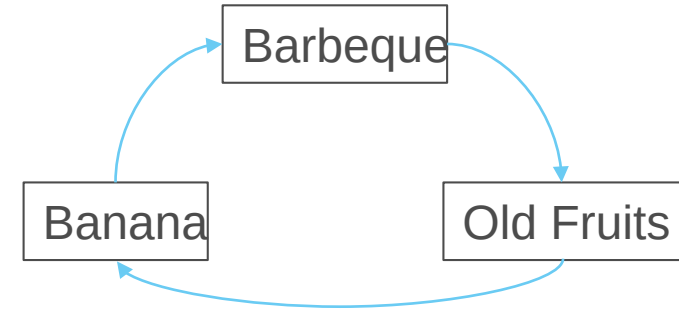
Safety: Danger! Humans
Taste: I love meat!
Distance: Needs little walk



Safety: Safe
Taste: Same old
Distance: Next to me



Safety: Can it be more safe?
Taste: Meh
Distance: Far



Copeland Pairwise Comparison



Safety: Bad
Taste: Very Good
Distance: Average



Safety: Good
Taste: Average
Distance: Very Good



Safety: Very Good
Taste: Bad
Distance: Bad

How does Copeland work? Counts **Victories (+1)** and **Defeats (-1)** (*ties don't count*)

Outranking matrix

Preference	Barbeque	Banana	Old Fruits
Barbeque	-	1	2
Banana	2	-	2
Old Fruits	1	1	-

	Wins	Defeats	Score	Rank
Banana	2	0	2	1
Barbeque	1	1	0	2
Old Fruits	0	2	-2	3

Summary

- Based on Average Scores

1. Arithmetic Mean
2. Geometric Mean

“Absolute” approaches:
each unit’s performance is
almost independent of other
units / output: scores

- Based on Ranks

3. Median rank
4. Majority (Relative Majority)
5. Borda’s Count

“Compensatory”
approaches
(weights interpreted
as trade-offs)

- Based on Pairwise Comparisons

6. Condorcet
7. Kemeny (C-K-Y-L)
8. Arrow – Raynaud
9. Copeland

“Non-Compensatory”
approaches
(weights interpreted as
importance coefficients)

“Relative” approaches:
each unit’s performance
depends on other unit’s
performance:
ranks (*part of numerical
information is lost*)
- no normalisation needed
- not influenced by outliers

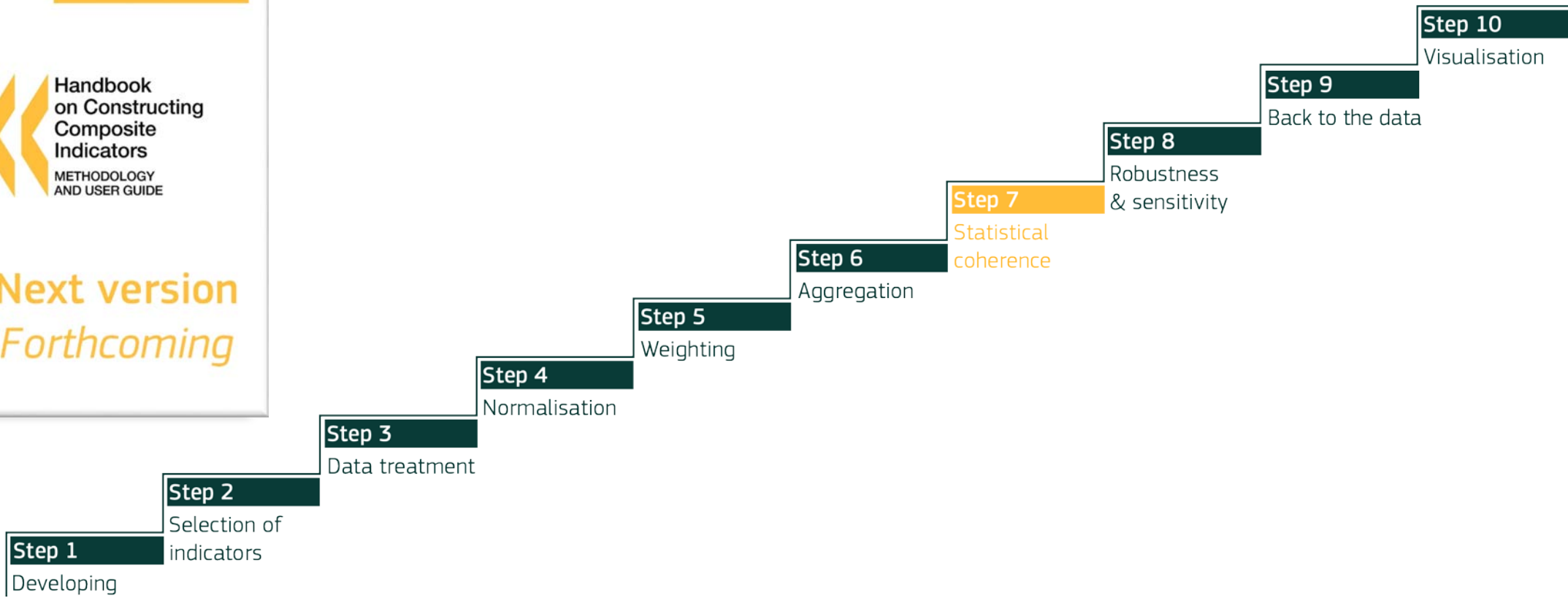
Suggested readings and references (Aggregation)

- Arrow, K. J., and H. Raynaud, H. Social choice and multicriterion decision-making. MIT Press Books, 1.
- Athanasoglou S. (2015) Multidimensional welfare rankings under weight imprecision: a social choice perspective. Social Choice and Welfare 4(4), 719-744.
- Bartholdi III, J., Tovey, C. A., & Trick, M. A. (1989). Voting schemes for which it can be difficult to tell who won the election. Social Choice and Welfare, 6(2), 157-165.
- Kemeny, J. (1959). Mathematics without numbers. Daedalus 88: 577-591.
- Munda, G. 2008. Social Multi-Criteria Evaluation for a Sustainable Economy. Berlin Heidelberg: Springer-Verlag.
- Munda, G., Nardo, M. (2009). Non-compensatory/Non-linear composite indicators for ranking countries: a defensible setting. Applied Economics, 41(12), 1513-1523.
- Tarjan, R. (1972). Depth-first search and linear graph algorithms. SIAM Journal on Computing, 1(2), 146-160.
- Truchon, M. (1998). An extension of the Condorcet criterion and Kemeny orders. Cahier 98-15 du Centre de Recherche en Économie et Finance Appliquées.
- Van Zuylen, A., and D. Williamson (2009). Deterministic Pivoting Algorithms for Constrained Ranking and Clustering Problems. Mathematics of Operations Research, 34, 594-620
- Young, H.P. (1988) Condorcet's Theory of Voting. American Political Science Review 82(2): 1231–1244.
- Young, H.P. (1995) Optimal Voting Rules. Journal of Economic Perspectives, 9, 51-64.
- Young, H.P. and A. Levenglick (1978) A Consistent Extension of Condorcet's election principle. SIAM Journal on Applied Mathematics, 35, 285-300.

Analysing and Summarising

Statistical Coherence, Robustness and Sensitivity, Back to data

Statistical Coherence



Internal coherence of the framework

Indicator level	Pillar level	Sub-Index level	Index
Indicator 1(X)	Pillar 1(Y)	Sub- index 1	Index
Indicator 2			
Indicator 3			
Indicator 4	Pillar 2	Sub- index 1	
Indicator 5			
Indicator 6			
Indicator 7	Pillar 3	Sub- index 2	
Indicator 8			
Indicator 9			
Indicator 10	Pillar 4	Sub- index 2	
Indicator 11			

CORRELATION
 $\rho(x,y) = \text{Cov}(x,y) / (\sigma_x \sigma_y)$

To which extent two variables are **linearly** related

1 perfect positive linear correlation.

-1 perfect negative linear correlation.

0 no linear correlation.

ρ^2 is the fraction of the variance in Y that is explained by X in a simple linear regression.

Statistically **significant** (5%) if $> 2/\sqrt{N}$

Which correlations do we observe?

Indicator level	Pillar level	Sub-Index level	Index
Indicator 1	Pillar 1	Sub- Index 1	Index
Indicator 2			
Indicator 3			
Indicator 4	Pillar 2	Sub- Index 1	
Indicator 5			
Indicator 6			
Indicator 7	Pillar 3	Sub- Index 2	
Indicator 8			
Indicator 9			
Indicator 10	Pillar 4	Sub- Index 2	
Indicator 11			



COIN Tips

7

Check whether indicators:

- Are not related or negatively related to indicators in the same dimension < -0.3
- Are under-represented in their dimension: < 0.3
- Are negatively related to the composite indicator: < -0.3
- Does any indicators dominate the framework: > 0.95

What responses we can find?

- Are indicators allocated in the **right dimension**?
- Are some indicators **over or under represented** in the aggregate **Index**?
- Up to what **level** should we aggregate?

Correlations at Indicator level

Name		Logistics Performance Index	International flights passenger capacity	Liner Shipping Connectivity Index	Border crossings	Trade in electricity	Trade in gas	Average connection speed	Population covered by at least a 4G mobile network
Name	Indicator	ind.01	ind.02	ind.03	ind.04	ind.05	ind.06	ind.07	ind.08
Logistics Performance Index	ind.01		0.99	0.43	0.33	0.13	0.22	0.45	0.58
International flights passenger ca	ind.02	0.99		0.42	0.33	0.14	0.22	0.48	0.58
Liner Shipping Connectivity Index	ind.03	0.43	0.42		-0.17	-0.40	0.07	-0.05	0.06
Border crossings	ind.04	0.33	0.33	-0.17		0.55	0.34	0.17	0.37
Trade in electricity	ind.05	0.13	0.14	-0.40	0.55		-0.08	0.14	0.43
Trade in gas	ind.06	0.22	0.22	0.07	0.34	-0.08		0.26	0.05
Average connection speed	ind.07	0.45	0.48	-0.05	0.17	0.14	0.26		0.50
Population covered by at least a 4	ind.08	0.58	0.58	0.06	0.37	0.43	0.05	0.50	

Example: Social Progress Index

Indicator level	Pillar level	Sub-Index level	Index
Indicators 1-5	Component 1 Nutrition & basic medical health	Dimension 1 Basic human needs	Social Progress Index
Indicators 6-9	Component 2 Water & Sanitation		
Indicators 10-12	Component 3 Shelter		
Indicators 13-16	Component 4 Personal Safety		
Indicators 17-21	Component 5 Access to basic knowledge	Dimension 2 Foundations of wellbeing	
Indicators 22-25	Component 6 Access to Information and Communication		
Indicators 26-29	Component 7 Health and Wellness		
Indicators 30-33	Component 8 Environmental Quality		
Indicators 34-38	Component 9 Personal Rights	Dimension 3 Opportunity	
Indicators 39-42	Component 10 Personal Freedom and Choice		
Indicators 43-47	Component 11 Inclusiveness		
Indicators 48-51	Component 12 Access to advanced education		

Source: Statistical audit of the Social Progress Index(2018), JRC-COIN

Cross - Correlations

-Are indicators allocated in the **right dimension**?

-Are some indicators **over or under represented** in the aggregate score?

Dimension	Component	Indicator		C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	D1	D2	D3	Index
D1: Basic Human Needs	C1: Nutrition and Basic Medical Care	Undernourishment	ind.01	0,85	0,75	0,77	0,54	0,67	0,73	0,69	0,59	0,43	0,64	0,44	0,67	0,80	0,72	0,64	0,76
		Maternal mortality rate	ind.02	0,94	0,88	0,88	0,54	0,88	0,80	0,76	0,77	0,40	0,80	0,39	0,72	0,90	0,86	0,68	0,85
		Child mortality rate	ind.03	0,93	0,87	0,88	0,58	0,90	0,82	0,77	0,78	0,42	0,83	0,43	0,77	0,91	0,88	0,72	0,87
		Child stunting	ind.04	0,88	0,83	0,84	0,65	0,82	0,82	0,82	0,76	0,52	0,81	0,58	0,80	0,88	0,87	0,79	0,88
		Deaths from infectious diseases	ind.05	0,90	0,79	0,82	0,51	0,75	0,72	0,69	0,64	0,32	0,66	0,28	0,63	0,84	0,75	0,56	0,75
	C2: Water and Sanitation	Access to basic drinking water	ind.06	0,88	0,93	0,89	0,55	0,83	0,83	0,75	0,73	0,46	0,76	0,45	0,72	0,91	0,85	0,70	0,86
		Access to piped water	ind.07	0,85	0,89	0,87	0,54	0,77	0,76	0,79	0,75	0,43	0,78	0,49	0,71	0,88	0,82	0,71	0,84
		Access to basic sanitation facilities	ind.08	0,91	0,96	0,92	0,51	0,86	0,82	0,78	0,73	0,33	0,83	0,36	0,78	0,93	0,86	0,68	0,86
		Rural open defecation	ind.09	0,68	0,85	0,69	0,37	0,72	0,67	0,58	0,59	0,25	0,64	0,23	0,62	0,74	0,69	0,51	0,68
	C3: Shelter	Access to electricity	ind.10	0,90	0,89	0,93	0,46	0,82	0,79	0,71	0,70	0,33	0,74	0,34	0,70	0,89	0,81	0,62	0,81
		Quality of electricity supply	ind.11	0,80	0,81	0,91	0,66	0,81	0,81	0,85	0,78	0,41	0,83	0,54	0,80	0,89	0,87	0,76	0,88
		Household air pollution deaths	ind.12	0,89	0,88	0,95	0,49	0,82	0,78	0,83	0,78	0,35	0,80	0,38	0,74	0,90	0,86	0,67	0,85
	C4: Personal Safety	Political killings and torture	ind.13	0,54	0,48	0,48	0,78	0,51	0,64	0,58	0,57	0,85	0,54	0,77	0,46	0,61	0,62	0,77	0,69
		Perceived criminality	ind.14	0,46	0,43	0,46	0,82	0,51	0,49	0,52	0,41	0,43	0,56	0,57	0,47	0,58	0,52	0,58	0,58
		Traffic deaths	ind.15	0,53	0,43	0,47	0,73	0,50	0,51	0,63	0,58	0,61	0,44	0,54	0,52	0,57	0,59	0,62	0,62
		Homicide rate	ind.16	0,40	0,32	0,37	0,74	0,45	0,44	0,47	0,37	0,25	0,41	0,34	0,49	0,48	0,46	0,44	0,48



Pillar, Sub-index and Index

SPI components, dimensions and index		C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	D1	D2	D3	Index
Nutrition and Basic Medical Care	C1	1,00												0,96	0,91	0,75	0,91
Water and Sanitation	C2	0,91	1,00											0,95	0,89	0,72	0,89
Shelter	C3	0,93	0,93	1,00										0,96	0,91	0,73	0,91
Personal Safety	C4	0,63	0,54	0,58	1,00									0,73	0,71	0,78	0,77
Access to Basic Knowledge	C5	0,89	0,87	0,88	0,64	1,00								0,91	0,92	0,78	0,91
Access to Information and Communications	C6	0,86	0,85	0,85	0,68	0,85	1,00							0,90	0,95	0,87	0,94
Health and Wellness	C7	0,83	0,80	0,86	0,71	0,79	0,84	1,00						0,88	0,93	0,81	0,91
Environmental Quality	C8	0,79	0,77	0,81	0,62	0,79	0,81	0,83	1,00					0,83	0,92	0,81	0,89
Personal Rights	C9	0,46	0,41	0,39	0,69	0,48	0,65	0,52	0,58	1,00				0,52	0,61	0,83	0,67
Personal Freedom and Choice	C10	0,83	0,83	0,85	0,64	0,85	0,85	0,81	0,80	0,51	1,00			0,88	0,89	0,87	0,91
Inclusiveness	C11	0,47	0,42	0,45	0,73	0,52	0,61	0,62	0,60	0,81	0,62	1,00		0,55	0,63	0,87	0,71
Access to Advanced Education	C12	0,79	0,78	0,80	0,63	0,81	0,85	0,81	0,78	0,49	0,85	0,58	1,00	0,83	0,87	0,86	0,89
Basic Human Needs	D1													1,00			0,96
Foundations of Wellbeing	D2													0,95	1,00		0,98
Opportunity	D3													0,81	0,88	1,00	0,93



Example: Sustainable Development Goals Index (2019)

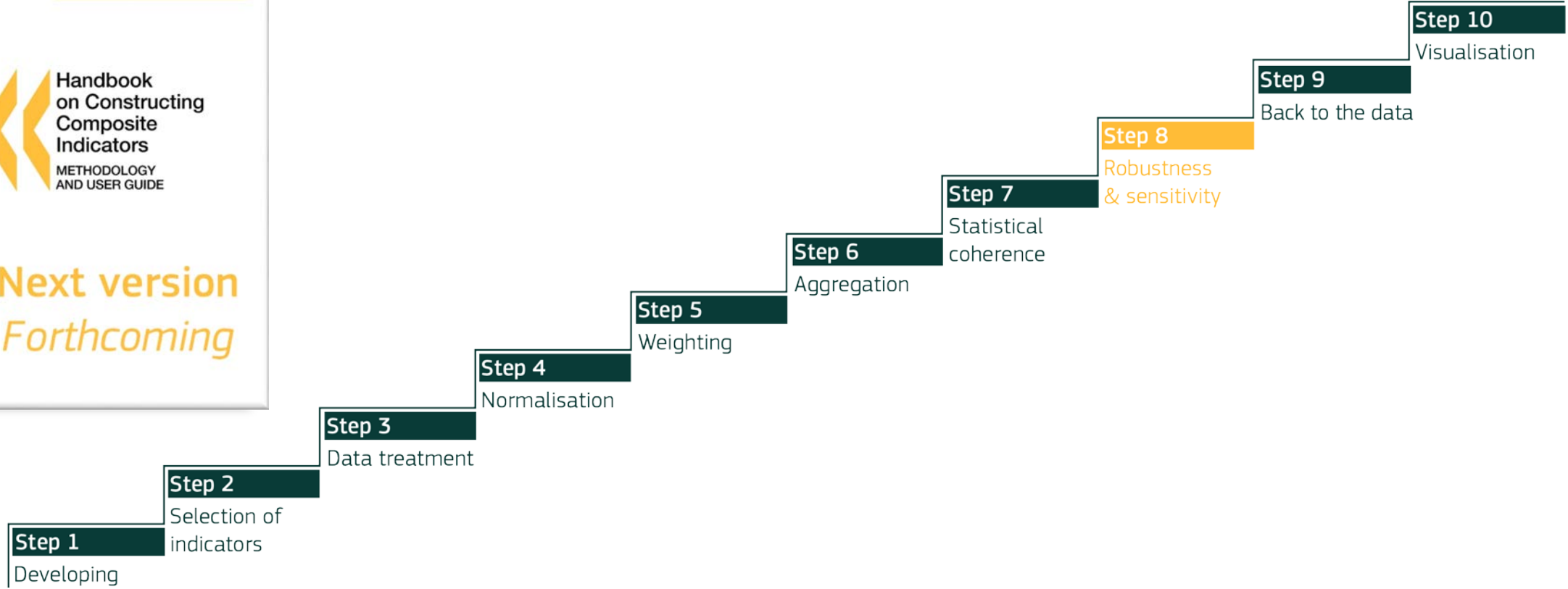
SDG	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	Index
1	1.00																	
2	0.49	1.00																
3	0.84	0.64	1.00															
4	0.77	0.61	0.84	1.00														
5	0.35	0.54	0.59	0.61	1.00													
6	0.69	0.66	0.81	0.73	0.68	1.00												
7	0.88	0.50	0.85	0.81	0.46	0.71	1.00											
8	0.50	0.60	0.68	0.62	0.59	0.63	0.51	1.00										
9	0.66	0.66	0.82	0.67	0.59	0.75	0.68	0.63	1.00									
10	0.36	0.29	0.36	0.17	0.01	0.21	0.19	0.24	0.38	1.00								
11	0.54	0.46	0.68	0.66	0.64	0.69	0.59	0.53	0.55	0.10	1.00							
12	-0.59	-0.53	-0.76	-0.59	-0.52	-0.67	-0.60	-0.46	-0.86	-0.32	-0.50	1.00						
13	-0.28	-0.17	-0.33	-0.32	-0.18	-0.19	-0.27	-0.10	-0.32	-0.05	-0.13	0.50	1.00					
14	-0.17	-0.01	-0.11	-0.10	0.09	-0.04	-0.13	0.06	-0.05	-0.19	-0.04	0.04	0.02	1.00				
15	-0.09	0.15	-0.02	-0.03	0.17	0.12	-0.05	0.14	0.09	0.01	-0.01	-0.02	0.22	0.24	1.00			
16	0.64	0.59	0.81	0.67	0.50	0.65	0.60	0.58	0.78	0.43	0.63	-0.72	-0.33	-0.11	0.04	1.00		
17	0.19	0.07	0.21	0.22	0.14	0.12	0.29	-0.03	0.11	0.01	0.16	-0.12	-0.35	-0.02	-0.08	0.14	1.00	
Index	0.84	0.71	0.93	0.86	0.67	0.86	0.86	0.73	0.83	0.40	0.73	-0.68	-0.20	-0.01	0.14	0.79	0.24	1.00

SDG 12: Responsible Consumption and Production
SDG 13: Climate Action

Take home message

- Analysis of correlation alone is not the final source of every decision
- When an inconsistency is spotted, try ask yourself why
- Is never to late for a revision or an improvement of the theoretical framework
- Acknowledge the inconsistencies that can not be eliminated
- Transparency is the main source of reliability of a Composite Indicator

The next steps



Our tools for a better result



Quality Assurance

Ensuring **statistical coherence**
(previous section)

How do indicators contribute to the composite?



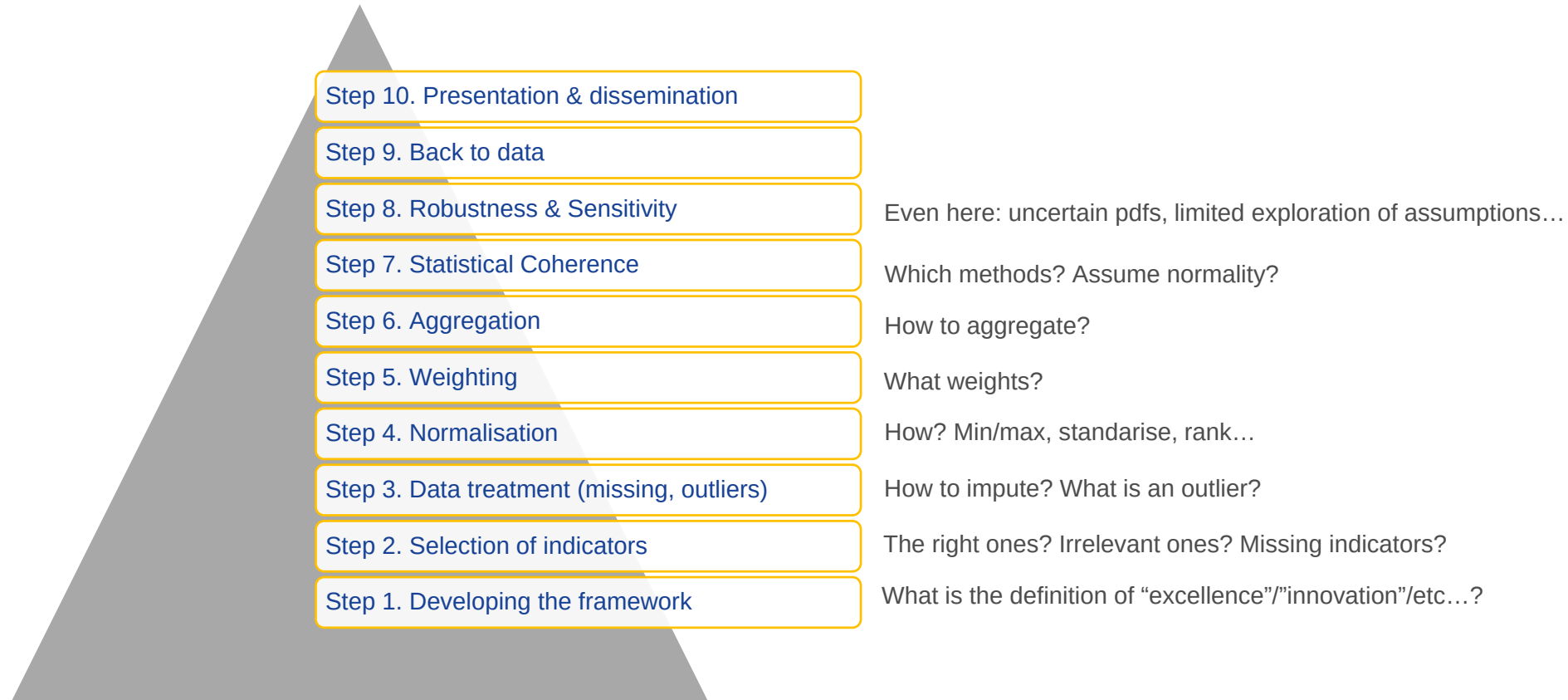
Robustness

How **sensitive** is the composite indicator to its assumptions?

- Transparency
- Exploration of uncertainty
- Anticipate criticism

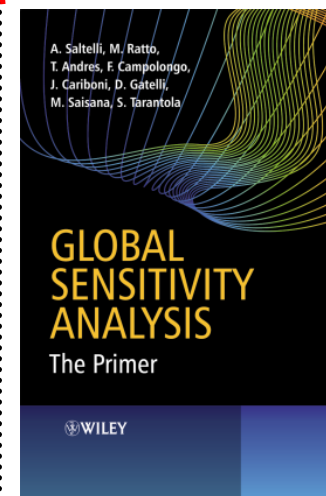
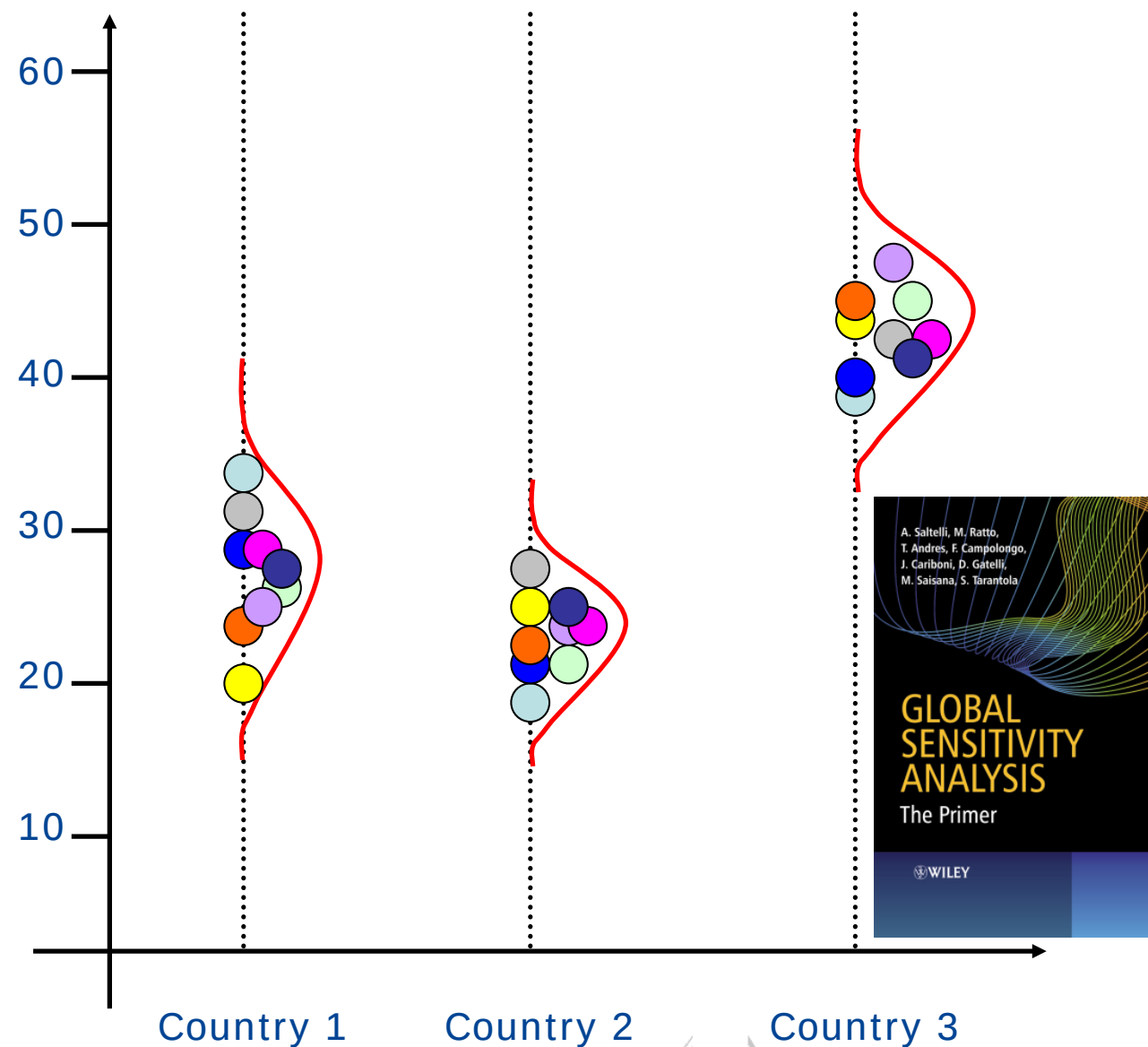
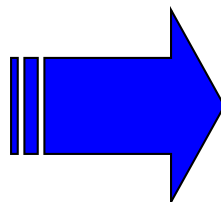
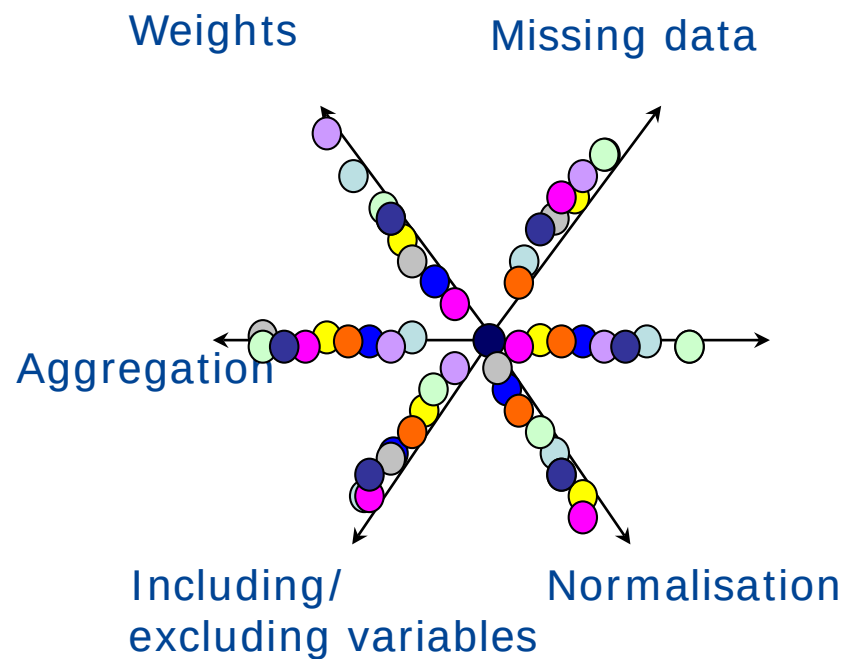
We use **Monte Carlo** methods and **sensitivity analysis**

What is uncertain in a Composite Indicator



But, the same is true for any model.

Space of alternatives



Example: Global Talent Competitiveness Index

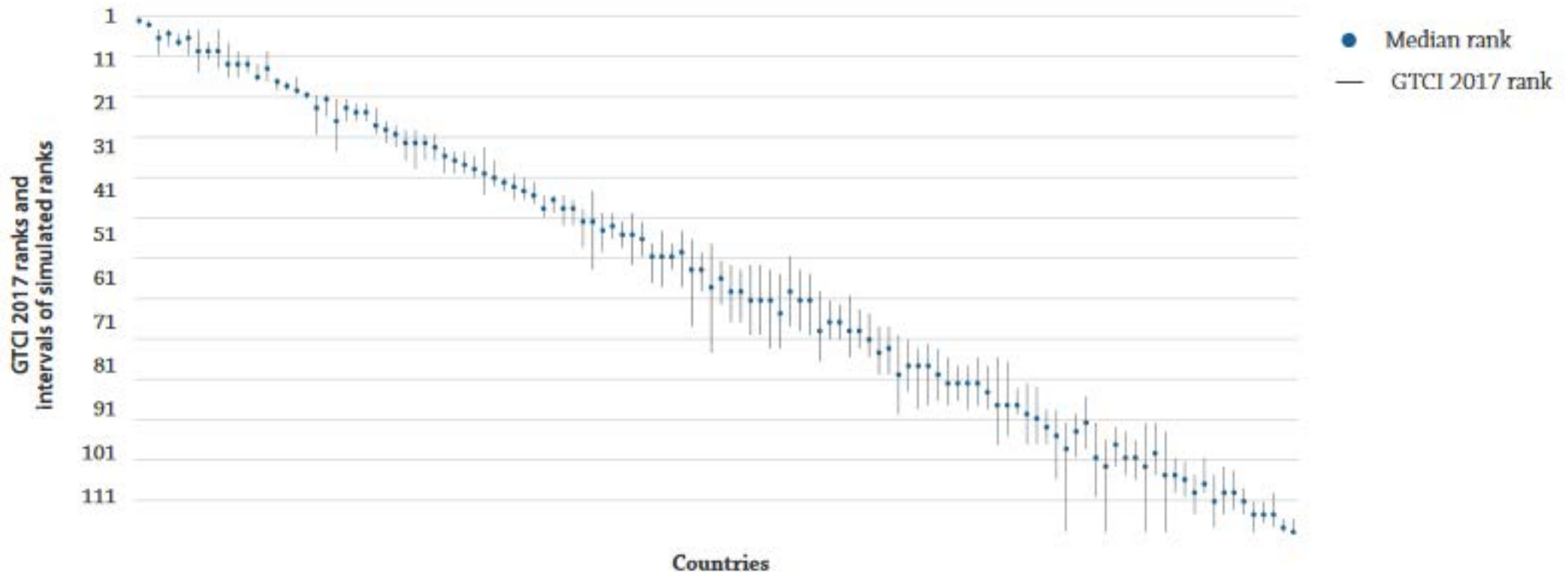
Uncertainty analysis for the GTCI 2017: Weights, missing data, aggregation, and normalisation

		REFERENCE	ALTERNATIVE
I. Uncertainty in the treatment of missing values		No estimation of missing data	Expectation Maximisation (EM)
II. Uncertainty in the aggregation formula at pillar level		Arithmetic average	Geometric average
III. Uncertainty in the method of normalisation		Partial normalisation	Full normalisation
IV. Uncertainty in the weights			
GTCI sub-index	Pillar	Reference value for the weight (within the sub-index)	Distribution assigned for robustness analysis (within the sub-index)
Input	Enable	0.25	U[0.15,0.35]
	Attract	0.25	U[0.15,0.35]
	Grow	0.25	U[0.15,0.35]
	Retain	0.25	U[0.15,0.35]
Output	Vocational and Technical Skills	0.50	U[0.40,0.60]
	Global Knowledge Skills	0.50	U[0.40,0.60]

Source: Statistical audit of the Global Talent Competitiveness Index (2017), JRC-COIN

Example: Global Talent Competitiveness Index

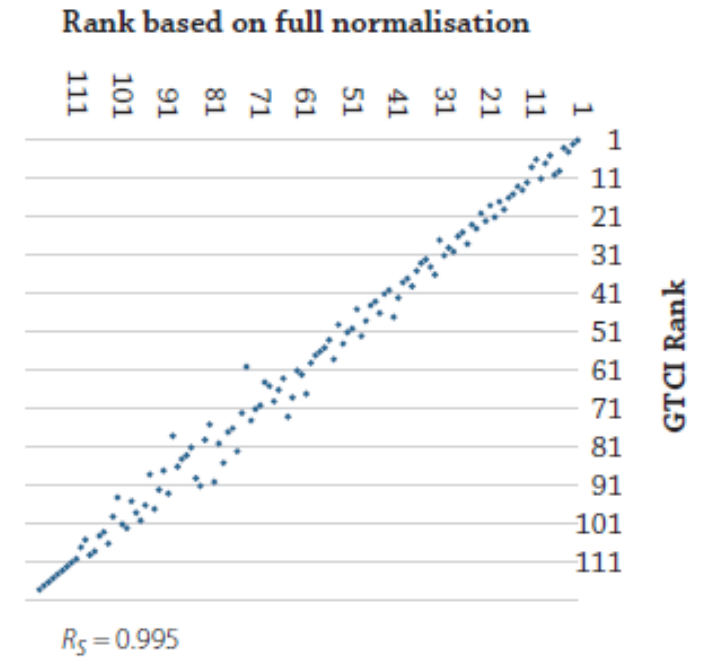
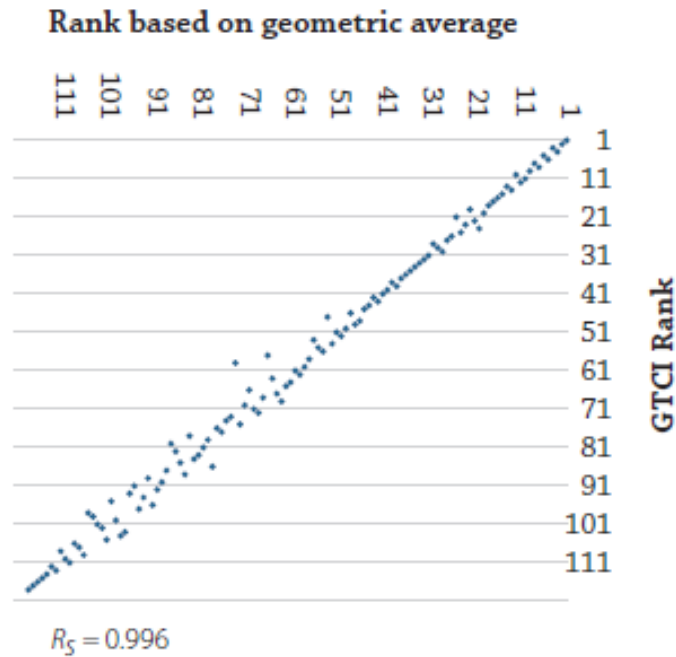
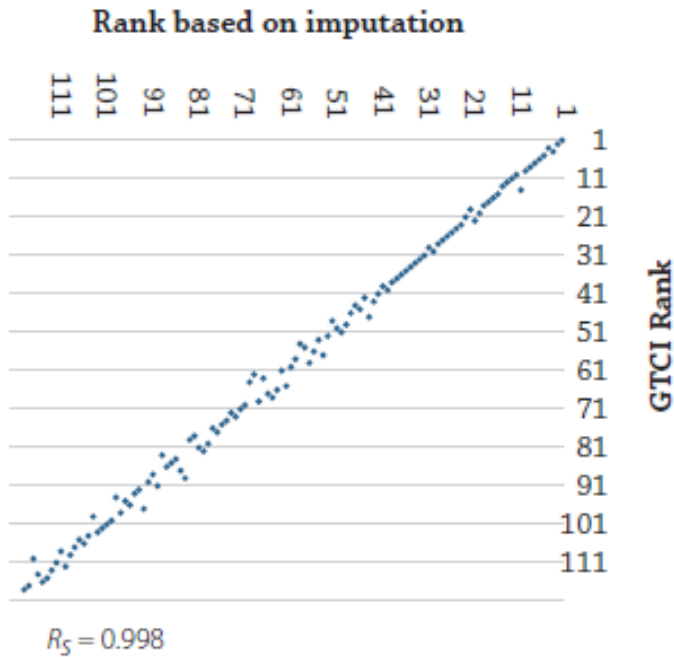
Robustness analysis (GTCI rank vs. median rank, 90% confidence intervals)



Source: Statistical audit of the Global Talent Competitiveness Index (2017), JRC-COIN

Sensitivity Analysis

Which assumptions are driving the uncertainty in the rankings?



From irresponsible musings to serious efforts

Composite indicators

- Are very widely used
- Fill a demand for which there is no other alternative

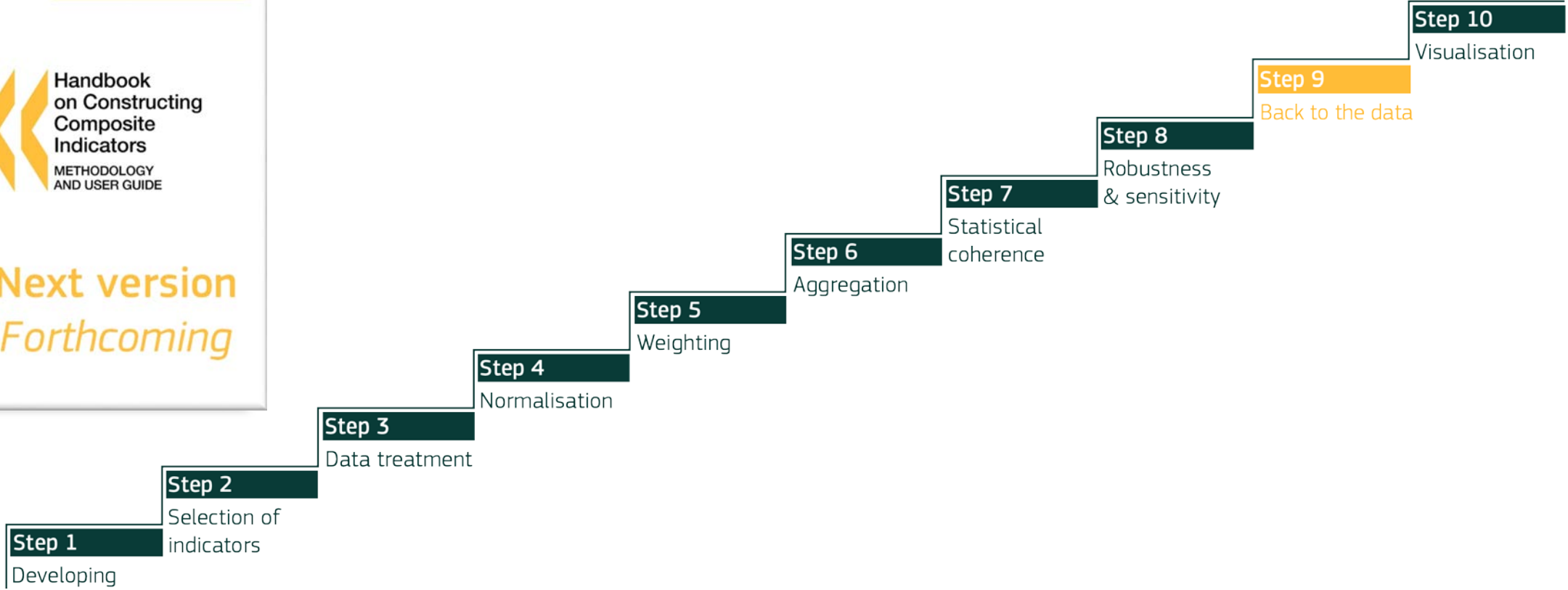
Therefore we should use available tools to **increase robustness and credibility**:

1. **Transparency**—detailed description of methodology, data sources, assumptions
2. **Statistical soundness**—analysis of correlations, data structure, effects of weights, etc.
3. **Uncertainty and sensitivity analysis**—check effect of alternative but plausible assumptions. Honestly acknowledge uncertainty.

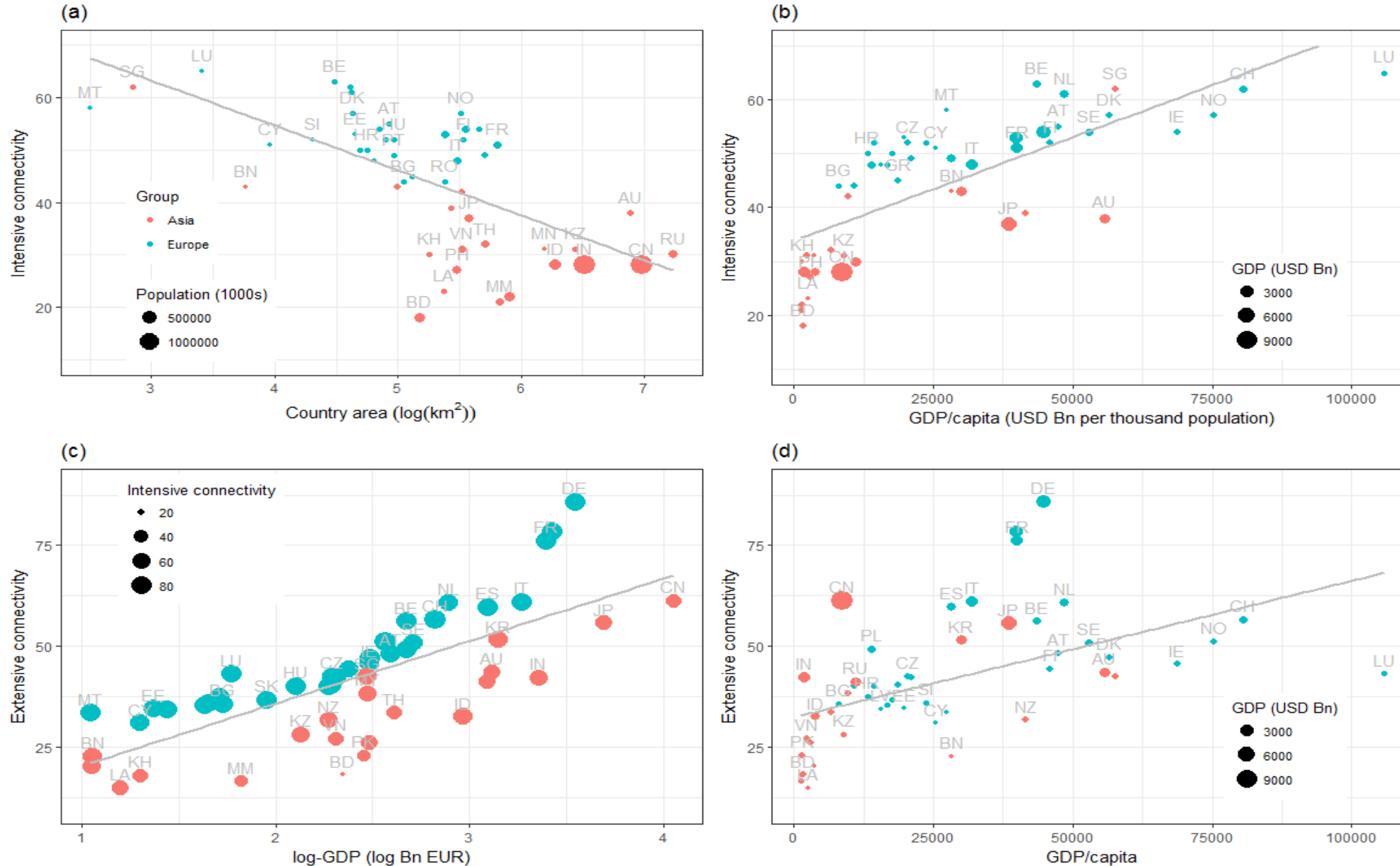
Suggested readings and references (Uncertainty and Sensitivity analysis)

- Aitbach, P. (2015). The dilemmas of ranking. *International Higher Education*, (42)
- Becker, W., Saisana, M., Paruolo, P., & Vandecasteele, I. (2017). Weights and importance in composite indicators: Closing the gap. *Ecological Indicators*, 80, 12-22.
- Booyesen, F. (2002). An overview and evaluation of composite indices of development. *Social indicators research*, 59(2), 115-151
- Saisana, M., Saltelli, A., & Tarantola, S. (2005). Uncertainty and sensitivity analysis techniques as tools for the quality assessment of composite indicators. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 168(2), 307-323.
- Saisana, M., d'Hombres, B., & Saltelli, A. (2011). Rickety numbers: Volatility of university rankings and policy implications. *Research policy*, 40(1), 165-177.

Make sense of the data

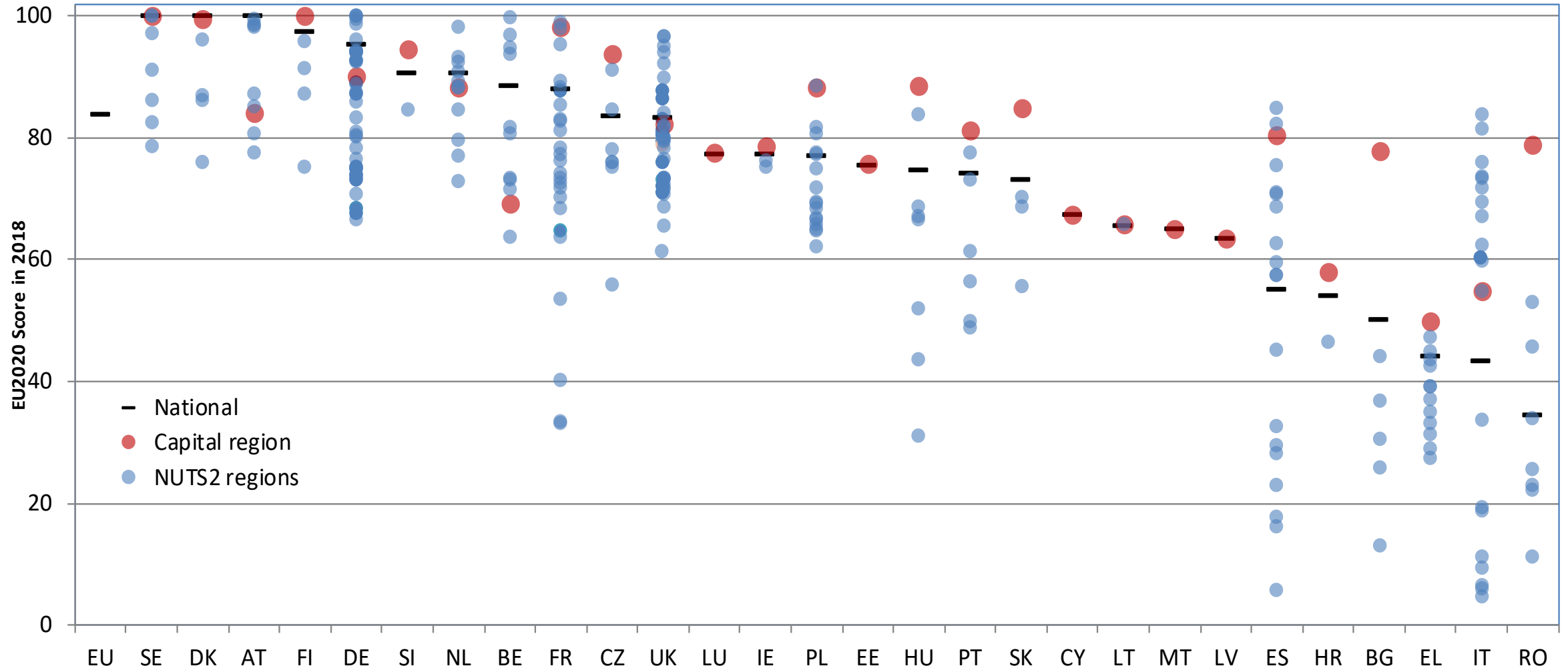


Story-telling: Examples



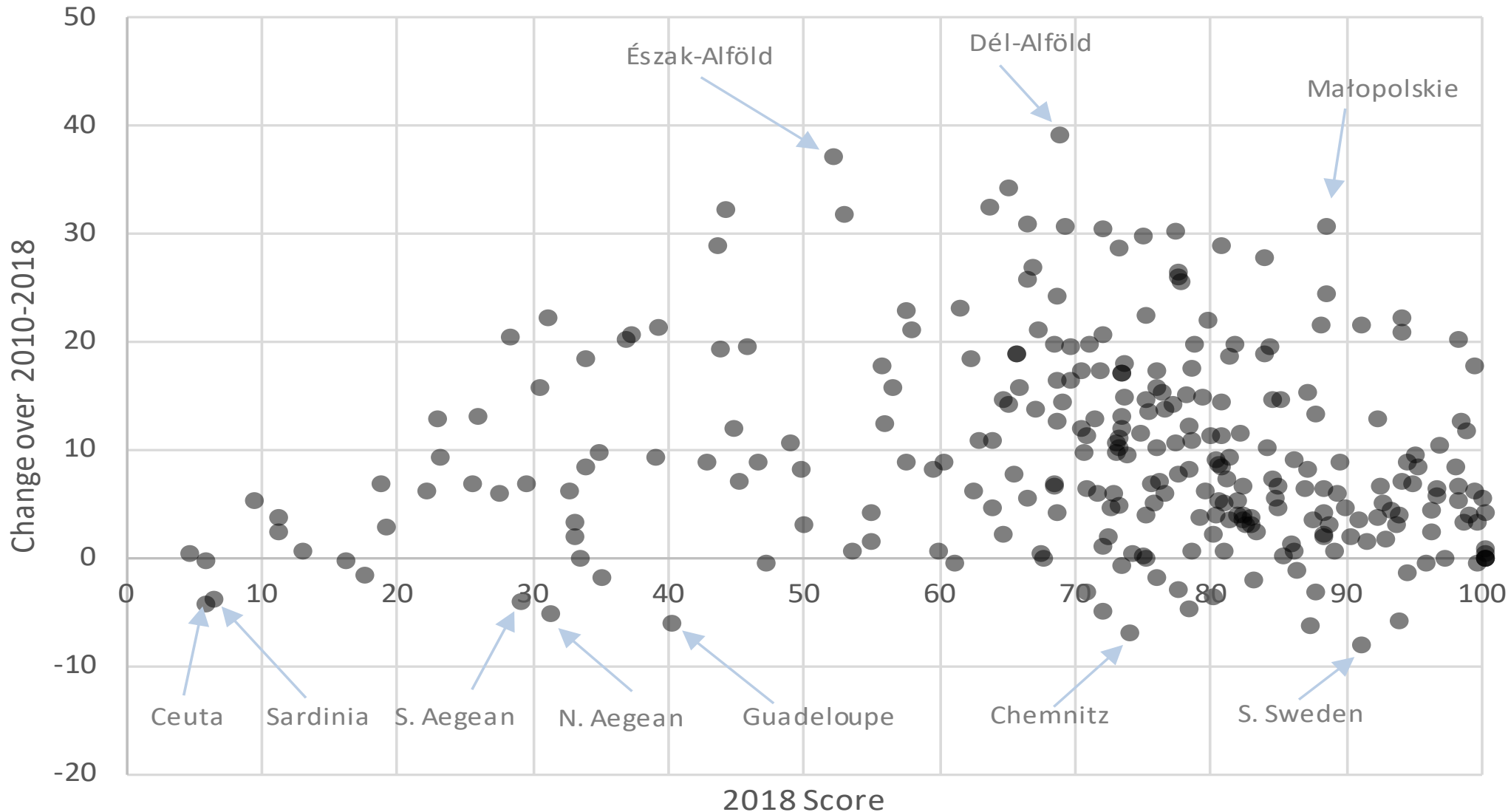
Story-telling: Examples

EU2020 Index



Story-telling: Examples

EU2020 Index



Thank you



© European Union 2020

Unless otherwise noted the reuse of this presentation is authorised under the [CC BY 4.0](https://creativecommons.org/licenses/by/4.0/) license. For any use or reproduction of elements that are not owned by the EU, permission may need to be sought directly from the respective right holders.

The European Commission's science and knowledge service

Joint Research Centre



The COIN Tool

William Becker

COIN 2019 - JRC Annual Training on Composite Indicators & Scoreboards
04-06/11/2019, Ispra (IT)

Download COIN Tool:

bit.ly/jrcweekslides

Pass: 2019JRCweek!

Alternative download
<http://tiny.cc/cointool>

What is it?

1

2

A free **Microsoft Excel-based tool** designed to help users in the process of building and analysing composite indicators.

Developed by the European Commission's **Competence Centre on Composite Indicators and Scoreboards** (COIN), at the Joint Research Centre.

Download at our website:

<https://composite-indicators.jrc.ec.europa.eu/>

Versions

Full Version

- With example data
- Without example data

Lite Version (reduced features, runs faster)

- With example data
- Without example data

Also please consult full **User Guide** (featuring general tips on composite indicators).

What does it do?

Given a data set, a structure, and methodological settings, the COIN Tool:

- Helps to visualise your data, including error-checking
- Checks and corrects for **outliers** (data treatment)
- **Imputes** missing data
- **Normalises** and **aggregates** indicators
- Calculates and visualises index **scores and ranks**
- Analyses **correlations** and helps balance your index
- Gives **alternative aggregations** (Borda, Copeland)
- Investigates **uncertainty** due to methodological assumptions
- **And many other things!**

Up to 300 units, 99 indicators, four aggregation levels (33 sub-pillars, 11 pillars, 3 sub-indexes).

What does it do?

Step 1: Concept definition

Step 2: Indicator selection

Step 3: Data treatment and analysis

Step 4: Normalisation

Step 5: Weighting

Step 6: Aggregating indicators

Step 7: Statistical and conceptual coherence

Step 8: Uncertainty and sensitivity analysis

Step 9: Making sense of the data

Step 10: Visualisation

COIN Tool

Before you start

Before you begin using the COIN Tool, you should have performed the following steps:

- Clearly defined the **concept** to be measured
- Clearly defined the **purpose** of the index, its proposed added value, and who the end users should be (this can help guide methodological decisions)
- Outlined an initial **conceptual framework**
- Populated the conceptual framework with **indicators**, and gathered the data
- **Denominated any indicators** where needed, by size-related quantities such as GDP, population, country area
- Made **initial decisions on methodology**, e.g. on weighting. Equal weights can be used as a starting point (and may indeed be retained in the final version)

Of course, **these steps can be returned to and modified later on**, but you should have initial ideas in place. Once you are at that point, you are ready to begin using the COIN Tool. Good luck!

Tool Rules

The COIN Tool is built in Excel to make it as accessible as possible to the widest range of users. This stretches Excel a bit, so you should follow these rules:

1. To make the CT experience as fast as possible, **disable automatic workbook calculation**. This is done as follows:
 - Go to the “File” menu
 - Go to Options -> Formulas -> Workbook calculation and set to MANUAL
 - Press F9 to recalculate the workbook when needed.
2. The COIN Tool is **friendly-user** software. This means:
 - Do not change cells that are not blue or turquoise
 - You have to enter data in the way the CT expects: e.g. correct structure codes and only ‘n/a’ for missing data.
 - See the FAQ and Troubleshooting sections in the User Guide if errors occur.
3. It might be easier to work in **full-screen mode** on a laptop (View -> Full Screen)



THANK YOU

Welcome to email us at: jrc-coin@ec.europa.eu

COIN in the EU Science Hub

<https://ec.europa.eu/jrc/en/coin>

COIN tools are available at:

<https://composite-indicators.jrc.ec.europa.eu/>

The European Commission's
Competence Centre on Composite
Indicators and Scoreboards

