BAC on Bibliometrics
ABC on Bibliometrics

Benvinguts...

BAC Research Group on bibliometrics

http://bac.fundaciorecerca.cat/

THOMSON REUTERS Bibliometric Partners

Institució CERCA Centres de Recerca de Catalunya

FCRi Fundació Catalana per a la Recerca i la Innovació

Contents:

I Research activity

- Aim, phases, timings

II Research & Development (R&D) systems

- Basic notions, actors

III Bibliometrics

- Key points, Aim and validity

IV Methods in bibliometrics

- Sources, cleansing, indicators

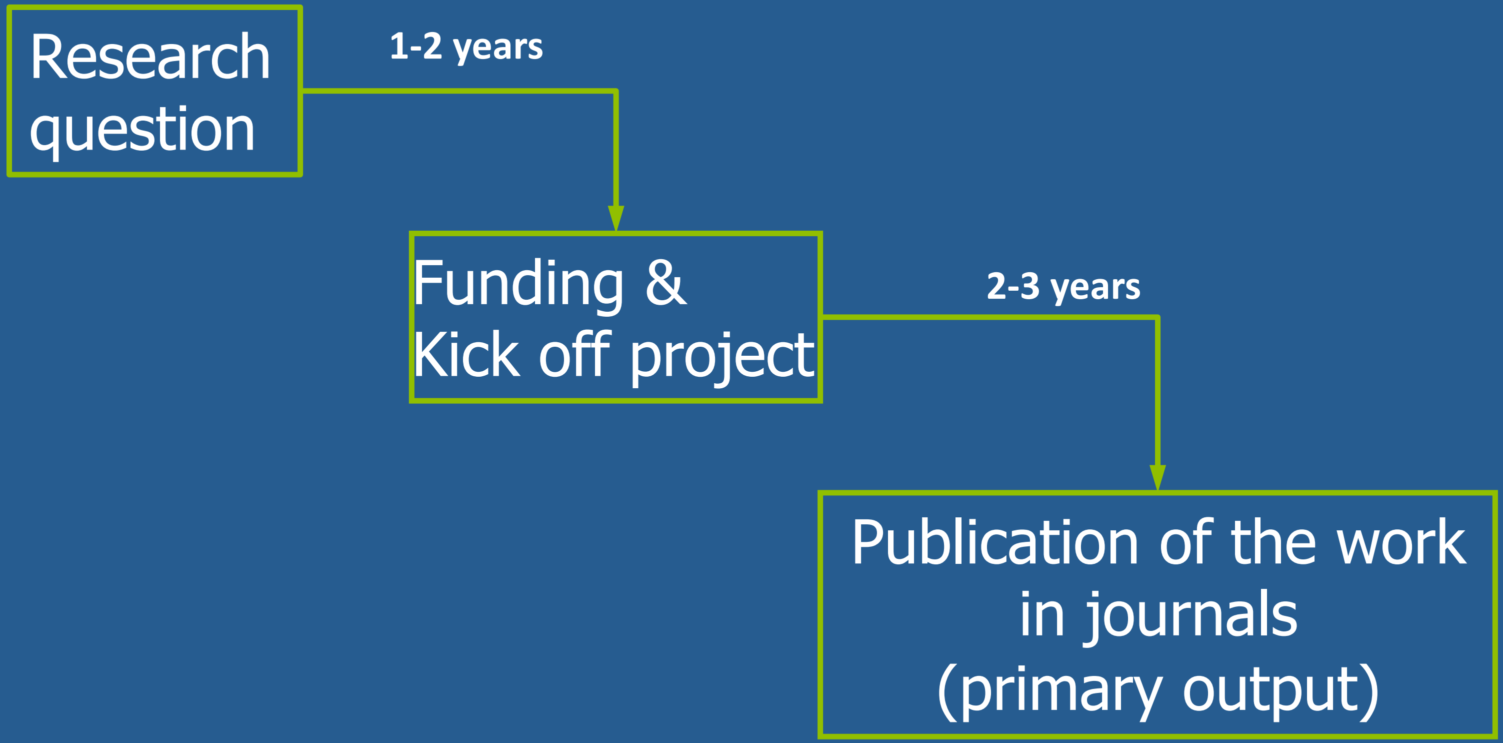
V Bibliometric analysis

- types, general schema, comparability

I Research activity: aim

- To increase our knowledge of “everything”. Scientists have been investigating systematically and sharing their findings in the form of reports since the 17th century.
- The reports have a common structure: introduction and aim, methodology, results, discusion and conclussion.
- About 1,000,000 publications are added to the body of knowledge of the planet each year.

I Research activity: phases & timings


Research question
1-2 years
Funding & Kick off project
2-3 years
Publication of the work in journals (primary output)

Different types of publications show different publication timings

II R&D systems: basic notions

- The concept of R&D system is a framework to understand / model the process of innovation.
- Innovation is the result of turning an idea into a process, product or service that (potentially) have value in the market.
- The notion of a system emphasizes the idea that the interactions between their components are of capital importance.

II R&D systems: actors, guidance

- Institutions, enterprises and governmental bodies involved with research are the actors most commonly referred as to the components of R&D systems (triple helix model).
- Governments guide their respective R&D systems with the help of Research Programs (RP).
- RP are the main instruments governments have to coordinate and integrate different research initiatives and priorities, resource allocation, etc.

III Bibliometrics: key points

- Originally, it was limited to collecting data on numbers of scientific articles and other publications, classified by author and/or by institution, field of science, country, etc., in order to construct simple "productivity" indicators for academic research.
- **1**) Collecting data
- **2**) Classifying data according different criteria
- **3**) Constructing indicators

1, Frascati Manual 2002: The measurement of scientific and technological activities, OECD, Paris, 2002

III Bibliometrics: object of study

- "Bibliometric analysis uses data on numbers and **authors** of scientific **publications** and on articles and the citations therein (as well as the citations in patents) to measure the "**output**" of individuals/ research teams, institutions and countries, to identify national and international networks, and to map the development of new (multidisciplinary) fields of science and technology."

Publications are the basic units of analysis and those actors involving in producing them become the object of study (of analysis) in bilbiometrics.

III Bibliometrics: validity indicators

- As publications on journals are the most basic unit in bibliometrics, bibliometric indicators should be used only in the study of fields of science and/or technology in which the results of research are shared in reports published in journals (articles basically).

IV Methods in bibliometrics: Sources

- Bibliometric reports are most commonly descriptive, and drawing conclusions further than the scope of the source data is a mistake.
- Source studies with large scope datasets
- Completeness, the required pieces of information are present in every, or almost all, records of the dataset.

IV Methods in bibliometrics: Sources

- Global
  - Web of Science (WOS), Thomson-Reuters
  - Scopus is produced by Elsevier
- Specialized
  - Medline, National Library of Medicine USA
  - Archive
  - REPEC
  - CITESEER
- Other: Google Scholar (it is not a source)

IV Methods: cleansing & classifying

- Information on publications is codified in different ways on journals.
- Entropy, which is visible in all data sources
- Analysing requires extensive
  1) Cleansing, purging
  2) Classifying

IV Methods: cleansing & classifying

- Classsification: grouping publications acccording to any of their attributes: year, authors, etc.
- Attribution: a publication can be (fully or partially) attributed to different entities: authors, centers, regions, etc.
- Precision and recall: used in the assessment of the quality of information retrieval processes.

IV Methods: classification & errors

- During the classification process we search the source dataset for subsets of publications sharing a specific attribute.
- i.e. attributable publications to a researcher
  - During this process we retrieve true positive publications, but also false positive pubs. Also we failed to retrieve some publications (true and false negative pubs).

False positive and false negative cases are classification "errors". Errors occcur…

IV Methods: classification

- Scientific and technical disciplines (fields)
- Organizations (centers) / sectoral groups /regions
- Authors

IV Methods: classification, fields

- Since different communities -> different publication patterns -> different citation rates, different tendency to international cooperation.
- Examples
  - 1) UNESCO
  - 2) Field of Science and Technology (FOS) OECD
  - 3) Journal Citation Report of Thomson-Reuters
  - 4) Medical Subject Headings thesaurus (MeSH) NLM
  - 5) SCOPUS

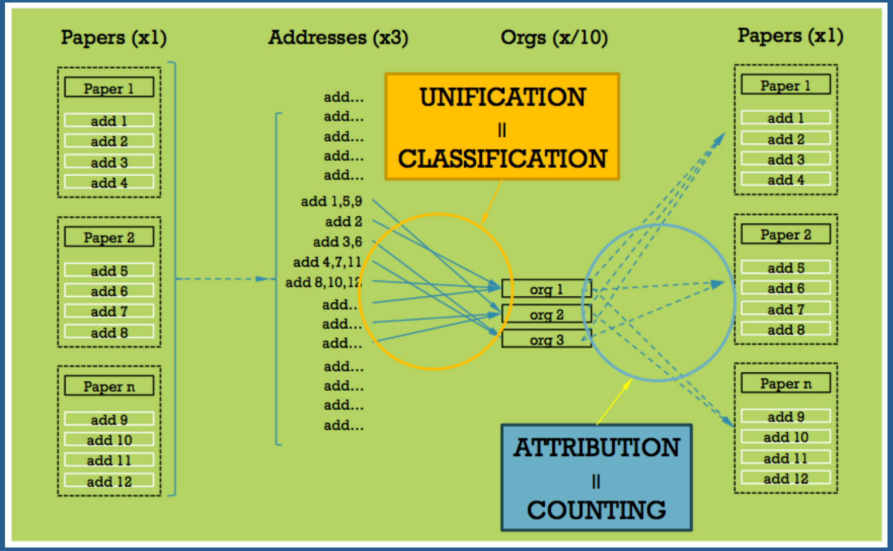
IV Methods: classification, centers

- The attribution of publications using the address field becomes a complex process as the focus of studies move below the national (macro) level.
- The normalization of addresses can be divided into two different
  - 1) Unification of addresses
  - 2) attribution of publications.

The precision of this process determines the quality of the studies

IV Methods: classification, centers

- The main challenge we face during the unification is dealing with entropy (disorder).
- The main challenge we face during the attribution is to get a good picture of reality.



IV Methods: classification, centers

- The complexity of this classification process increases as the structure of organizations become more and more complex
- As organization behave like living things the changes they experience along their “life cycle” add even more complexity to this process

IV Methods: classification, authors

- The problem
  - The lack of connection between authors and their publications.
- Common practice using a nickname instead of actual names.
- Nicknames or bibliographic names are created putting together the family name and the initial of the first name of authors.

IV Methods: classification, authors

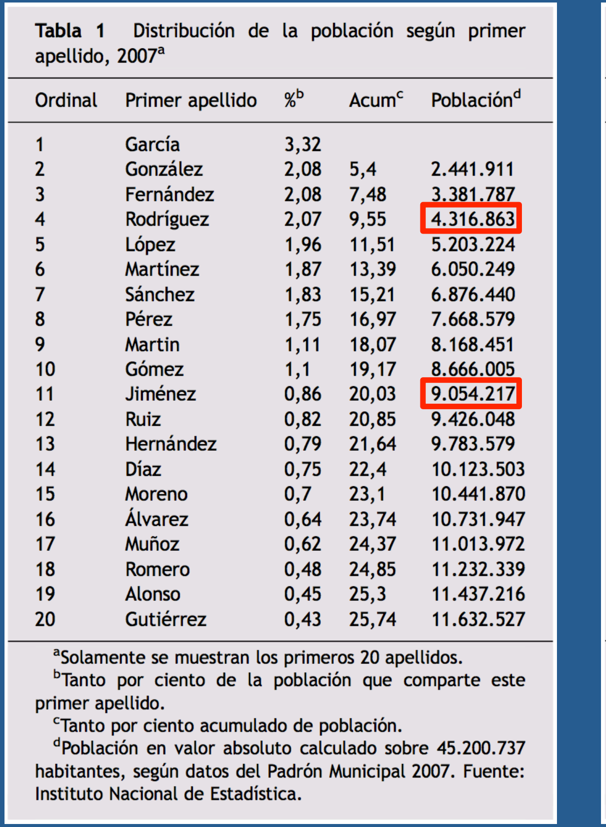


Tabla 1 Distribución de la población según primer apellido, 2007[a]

Ordinal	Primer apellido	%[b]	Acum[c]	Población[d]
1	García	3,32		
2	González	2,08	5,4	2.441.911
3	Fernández	2,08	7,48	3.381.787
4	Rodríguez	2,07	9,55	4.316.863
5	López	1,96	11,51	5.203.224
6	Martínez	1,87	13,39	6.050.249
7	Sánchez	1,83	15,21	6.876.440
8	Pérez	1,75	16,97	7.668.579
9	Martin	1,11	18,07	8.168.451
10	Gómez	1,1	19,17	8.666.005
11	Jiménez	0,86	20,03	9.054.217
12	Ruiz	0,82	20,85	9.426.048
13	Hernández	0,79	21,64	9.783.579
14	Díaz	0,75	22,4	10.123.503
15	Moreno	0,7	23,1	10.441.870
16	Álvarez	0,64	23,74	10.731.947
17	Muñoz	0,62	24,37	11.013.972
18	Romero	0,48	24,85	11.232.339
19	Alonso	0,45	25,3	11.437.216
20	Gutiérrez	0,43	25,74	11.632.527

[a] Solamente se muestran los primeros 20 apellidos.

[b] Tanto por ciento de la población que comparte este primer apellido.

[c] Tanto por ciento acumulado de población.

[d] Población en valor absoluto calculado sobre 45.200.737 habitantes, según datos del Padrón Municipal 2007. Fuente: Instituto Nacional de Estadística.

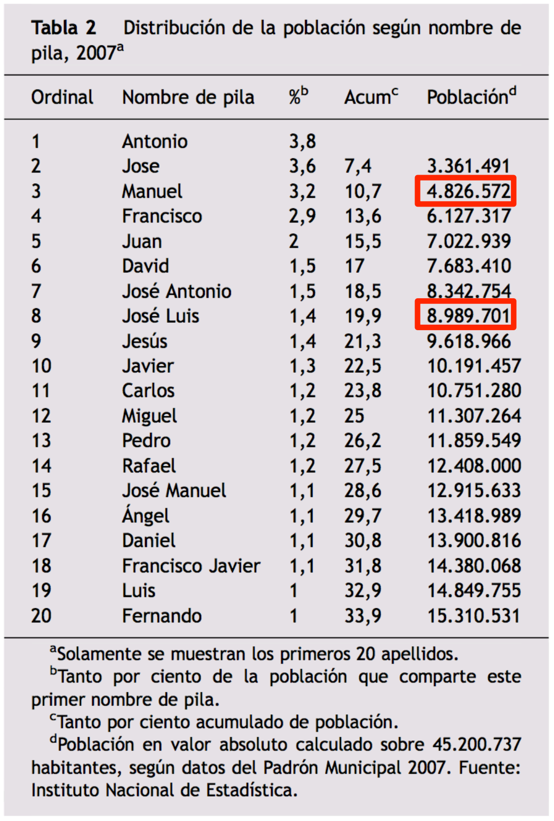


Tabla 2 Distribución de la población según nombre de pila, 2007[a]

Ordinal	Nombre de pila	%[b]	Acum[c]	Población[d]
1	Antonio	3,8		
2	Jose	3,6	7,4	3.361.491
3	Manuel	3,2	10,7	4.826.572
4	Francisco	2,9	13,6	6.127.317
5	Juan	2	15,5	7.022.939
6	David	1,5	17	7.683.410
7	José Antonio	1,5	18,5	8.342.754
8	José Luis	1,4	19,9	8.989.701
9	Jesús	1,4	21,3	9.618.966
10	Javier	1,3	22,5	10.191.457
11	Carlos	1,2	23,8	10.751.280
12	Miguel	1,2	25	11.307.264
13	Pedro	1,2	26,2	11.859.549
14	Rafael	1,2	27,5	12.408.000
15	José Manuel	1,1	28,6	12.915.633
16	Ángel	1,1	29,7	13.418.989
17	Daniel	1,1	30,8	13.900.816
18	Francisco Javier	1,1	31,8	14.380.068
19	Luis	1	32,9	14.849.755
20	Fernando	1	33,9	15.310.531

[a] Solamente se muestran los primeros 20 apellidos.

[b] Tanto por ciento de la población que comparte este primer nombre de pila.

[c] Tanto por ciento acumulado de población.

[d] Población en valor absoluto calculado sobre 45.200.737 habitantes, según datos del Padrón Municipal 2007. Fuente: Instituto Nacional de Estadística.

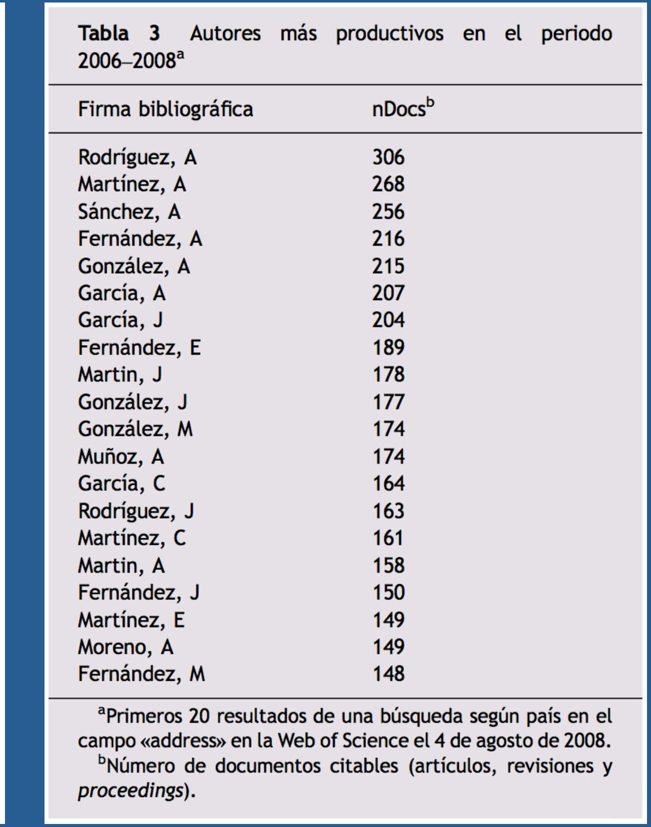


Tabla 3 Autores más productivos en el periodo 2006–2008[a]

Firma bibliográfica	nDocs[b]
Rodríguez, A	306
Martínez, A	268
Sánchez, A	256
Fernández, A	216
González, A	215
García, A	207
García, J	204
Fernández, E	189
Martin, J	178
González, J	177
González, M	174
Muñoz, A	174
García, C	164
Rodríguez, J	163
Martínez, C	161
Martin, A	158
Fernández, J	150
Martínez, E	149
Moreno, A	149
Fernández, M	148

[a] Primeros 20 resultados de una búsqueda según país en el campo «address» en la Web of Science el 4 de agosto de 2008.

[b] Número de documentos citables (artículos, revisiones y *proceedings*).

Mendez-Vasquez RI. Estar o no estar en el asunto: la evaluación individual del rendimiento científico. Aten Primaria. 2009;41(2):63–66

IV Methods: classification, authors

- Using available information (priority of application)
    1. Email address
    2. Correspondence address
    3. Rareness of the bibliographic name
    4. Main coauthors (most frequent)
    5. Host organizations (not correspondence add)
    6. Field of study: most frequent JCR disciplines

IV Methods: classification, authors

- Discrepancies in the number of publications
  - coverage of the source (journals and period)
  - Document type
  - Changes in first and family names (languages, marriage)
  - Mobility (change of host institution)
  - Changes in the research field.

Accuracy depends entirely on the amount of information available during this process.

IV Methods: Indicators, counting

- When we are counting publications we are actually saying: "I am giving organization X (o auhtor x) credit for these publications", or "these publications belong to organization X (or authro x)"
- Two methods:
  - Full credit: a unit per publication
  - Fractional credit: a fraction per publication

Bibliographic data -----------

Atomtronics with holes: Coherent transport of an empty site in a triple well potential

A. Benseny,1 S. Fernández-Vidal,1 J. Bagudà,1 R. Corbalàn,1 A. Picón,1,2 L. Roso,3 G. Birkl, 4 and J. Mompart1

1 Grup d'Òptica, Departament de Física, Universitat Autònoma de Barcelona, E-08193 **Bellaterra**, Spain
2 JILA, University of Colorado, **Boulder** 80309-0440, USA (present address)
3 Centro de Láseres Pulsados (CLPU), E-37008 **Salamanca**, Spain and
4 Institut für Angewandte Physik, Technische Universitä¨t Darmstadt, Schlossgartenstr. 7, D-64289 **Darmstadt**, Germany (Dated: June 16, 2010)

Subject Category: Optics; Physics according to Web of Science
----

```
Output of the
counting
Methods -----
```

8 authors

Entity	Whole count	Fractional count
Author 1	1	1/8
Author 2	1	1/8
Author 3	1	1/8
Author 4	1	1/8
Author 5	1	1/8
Author 6	1	1/8
Author 7	1	1/8
Author 8	1	1/8

4 organizations

Entity	Whole count	Fractional count
Center 1	1	1/4
Center 2	1	1/4
Center 3	1	1/4
Center 4	1	1/4

4 locations

Entity	Whole count	Fractional count
Bellaterra, Catalunya, Spain	1	1/4
Boulder, Colorado, USA	1	1/4
Salamanca, Castilla y León, Spain	1	1/4
Darmstadt, Hesse, Germany	1	1/4

2 JCR disciplines

Entity	Whole count	Fractional count
Discipline 1	1	1/2
Discipline 2	1	1/2

IV Methods: Indicators, counting

Feature	Full credit	Fractional c.
Counting (data management)	Easier	Complex
Resulting figures	Easy to understand	Not easy
Calculation of percentages	Not correct!	Correct!
Detection of errors	Possible	Not possible

IV Methods: Indicators, meaning

- The number of publications provides an estimation of the size and level of activity of an unit.
- Use this indicatror to group units according to their size (to stratify / or segment a population), and then analyze within groups.

The number of publications is the most basic indicator and most be included always in bibliometric reports.

IV Methods: Indicators, citations

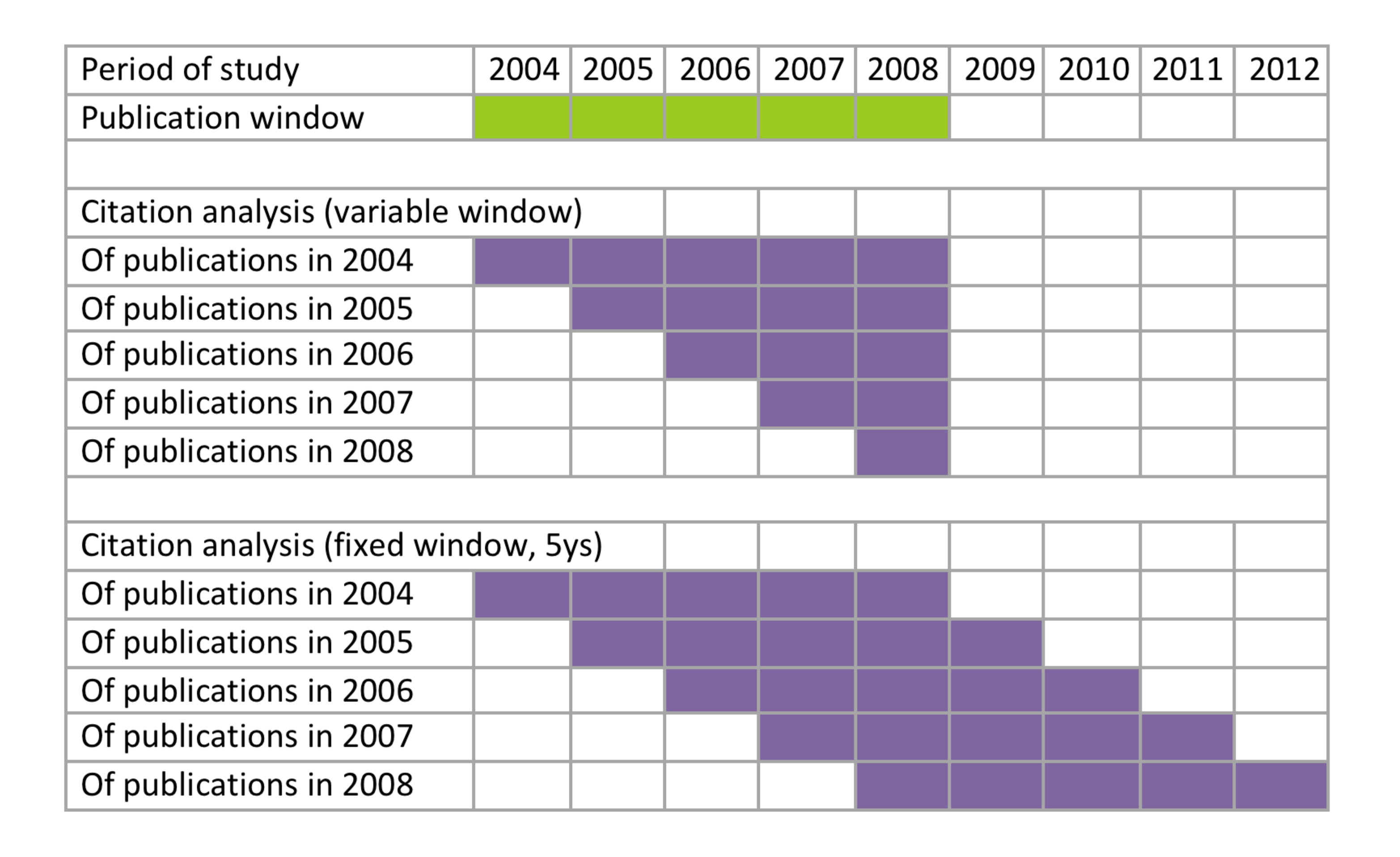
Three factors can modify the number of citations

1. **Time**
   - The number of citations increases with time
2. **Research field**
   - Different fields show different citation rates and tendency to international cooperation
3. **Type of document**
   - Articles, reviews and proceedings receive most of the citations recorded in a dataset

IV Methods: Indicators, citations

- In general, the number of citation reach a plato in 5 years in some discipline in natural science and biomedicine, while it may take 10 year in some disciplines in social science.
- Deviations from this pattern
  - Mayflowers
  - Sleeping Beauty
  - Hot papers

Period of study and citation windows



Period of study	2004	2005	2006	2007	2008	2009	2010	2011	2012
Publication window	🟩	🟩	🟩	🟩	🟩				
Citation analysis (variable window)									
Of publications in 2004	🟪	🟪	🟪	🟪	🟪				
Of publications in 2005		🟪	🟪	🟪	🟪				
Of publications in 2006			🟪	🟪	🟪				
Of publications in 2007				🟪	🟪				
Of publications in 2008					🟪				
Citation analysis (fixed window, 5ys)									
Of publications in 2004	🟪	🟪	🟪	🟪	🟪				
Of publications in 2005		🟪	🟪	🟪	🟪	🟪			
Of publications in 2006			🟪	🟪	🟪	🟪	🟪		
Of publications in 2007				🟪	🟪	🟪	🟪	🟪	
Of publications in 2008					🟪	🟪	🟪	🟪	🟪

IV Methods: Indicators, citations

- Analysis of citation in the extreme
  - Indicators based on extremly high number of citations are increasingly used in bibliometrics as proxis of excellence.
  - i.e, percentage of publications in the Top 10% most cited in the world
  - Top 1% or in the Top 1‰ (1 per thousand) also

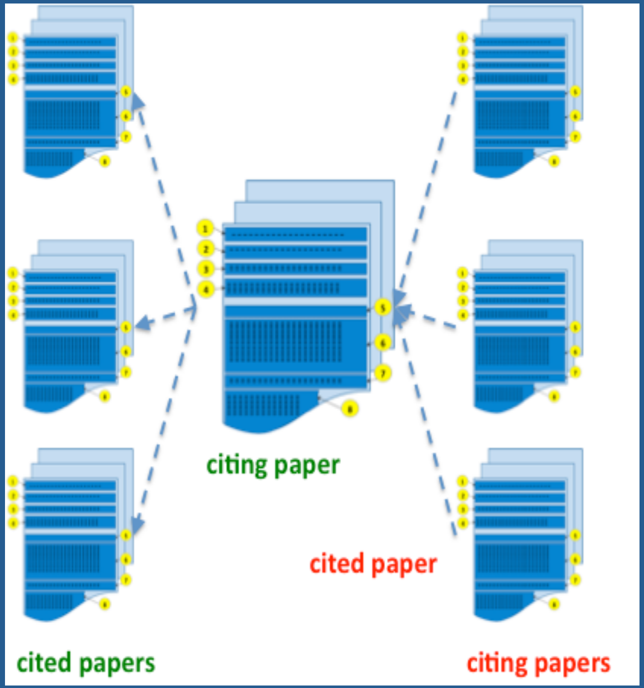
IV Methods: citations, meaning

- Research results in publications generate reactions of colleagues working inside and outside a specific field.

- These reactions are manifested in subsequent publications in different ways:
  - paying homage to pioneers
  - giving credit for related work (homage to peer)
  - Identifying methodology, equipment, etc

IV Methods: citations, meaning

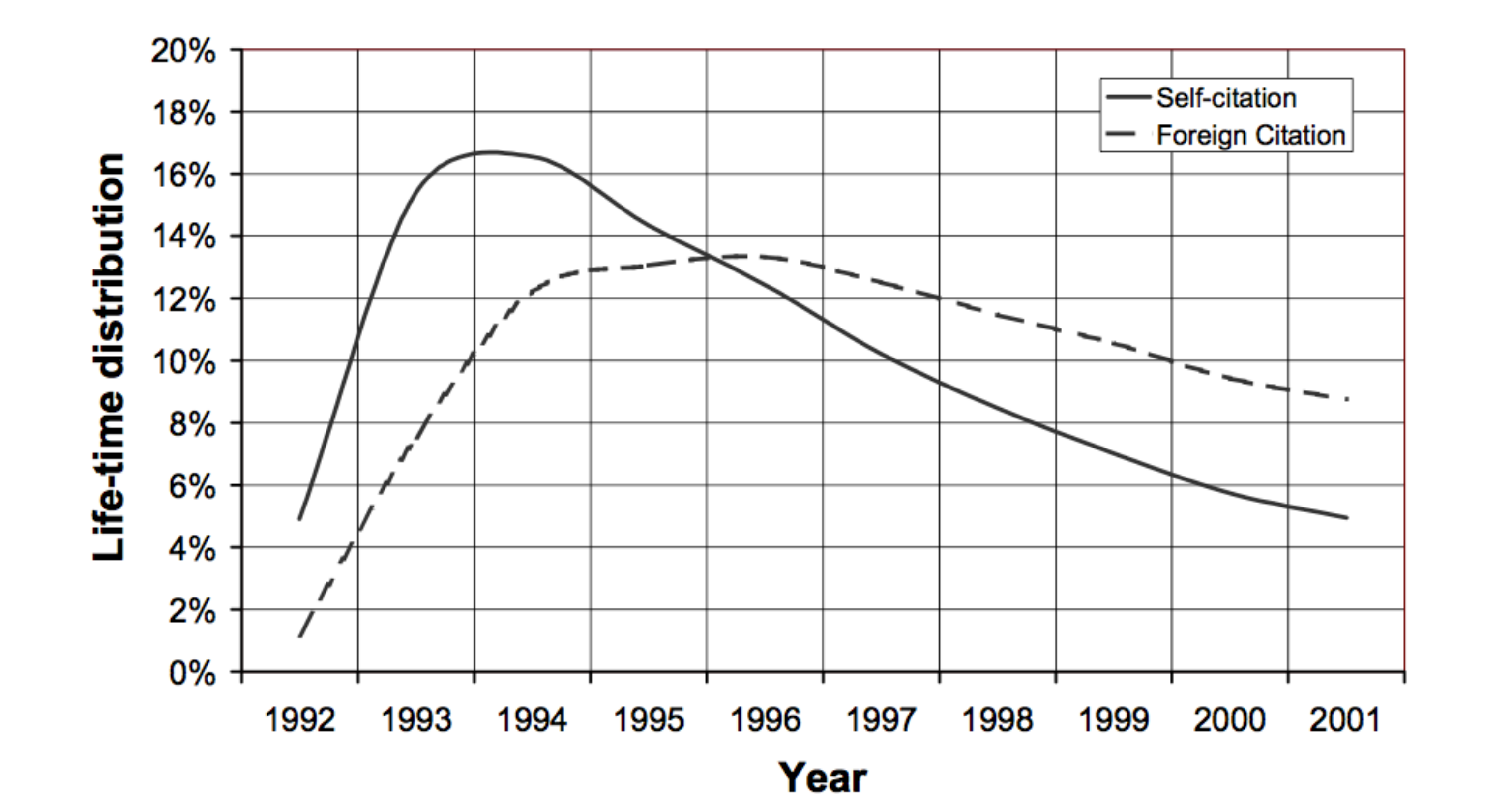
- Citations do not provide an 'ideal' monitor on scientific performance.
- However, its analysis enables assessing the impact of a work on colleagues.
- In general the more citations the more impact
- Howeve , it is not recommended at all using it as standalone indicator.



MacRoberts, M.H., MacRoberts, B.R. Problems of citation analysis. Scientometrics; 1996 36, 435–444.

IV Methods: citations, self-citations

Evolution of the share of self-citation (all fields combined)



source: Glänzel W. Bibliometrics as a research field. 2003

IV Methods: citations, normalized ind.

- The aim in constructing this type of indicators is counteracting the effects of time, research field and document type.
- This type of indicators enable comparing the impact of researchers devoted to different fields.
- There are 2 kinds of normalized indicators
  - Item oriented: high precision, low suceptibility biase
  - Field oriented: suceptible to biase

IV Methods: citations, normalized ind.

- Item oriented normalized indicator: an indicator that is calculated for every publication
  - Relative Citation Index (RCI)
  - CWTS field normalized citation score (crown ind.)
- Field oriented normalized indicator: all publications are categorized in an unique field (biase)
  - Impacto Normalizado (IN) Min. Econ. Comp.

IV Methods: citations, normalized ind.

- Item oriented normalized indicator: an indicator that is calculated for every publication
  - Relative Citation Index (RCI)
  - CWTS field normalized citation score (crown ind.)
- Field oriented normalized indicator: all publications are categorized in an unique field (biase)
  - Impacto Normalizado (IN), Min. Econ. Comp.

IV Methods: Cooperation indicators

- Cooperation can be assessed based on the addresses reported in publications, but also based on their number of authors.
  - Addresses of host institutions enable analyzing cooperation between territories
  - Authors (concurrence) enable detenting research groups

IV Methods: Coop. ind, meaning

- The higher their value, the better, as internactional cooperation associates with high impact.
- As for the indicators based on the number of coauthors, it should be used with caution, as this indicator is highly filed dependent.

IV Methods: excellence indicators

- Initally Top cited papers were defined as those included in the Top 10% most cited papers in the world.
- With time other definitions appeared in the bibliography, and currently Top 1% (HCP) and Top 1‰ most cited papers are also used as indicators of excellence.
- The implication of this definition is that the authors of this subset of publications have influenced thought, theory, and practice in world science and technology according to Westney.

Westney, l. C. H., National Science Board. (2010). Science and engineering indicators 2010, Arlington, VA, USA: National Science Foundation (NSB 10-01).

IV Methods: excell ind., meaning

- As this indicators are stimated using normalized reference values, they are realible.
- Use them to categorize researchers with similar profiles.

IV Methods: H index

- The H index combines measures of both the productivity and impact of the papers published by a researcher.
- An H index of 10 means that a resercher has published 10 papers, each of which has been cited at least 10 in other papers.

IV Methods: H index, limitations

- This index grows as citations accumulate and thus it depends on the 'academic age' of a researcher.
- This indicator should not be used to compare junior and senior researchers.
- This index is highly field dependent

Use the H index exclusively to compare researchers (not center, nor regions) with similar ages working on the same, or closely related research fields.

IV Methods: Journal Impact Factor (JIF)

- The JIF was developed by Eugene Garfield as an indicator to assist in the selection of journals during the creation of catalogues of sources.
- The JIF is the average (mean value) of citations to publications in a specific journal in the last 2 previous years.

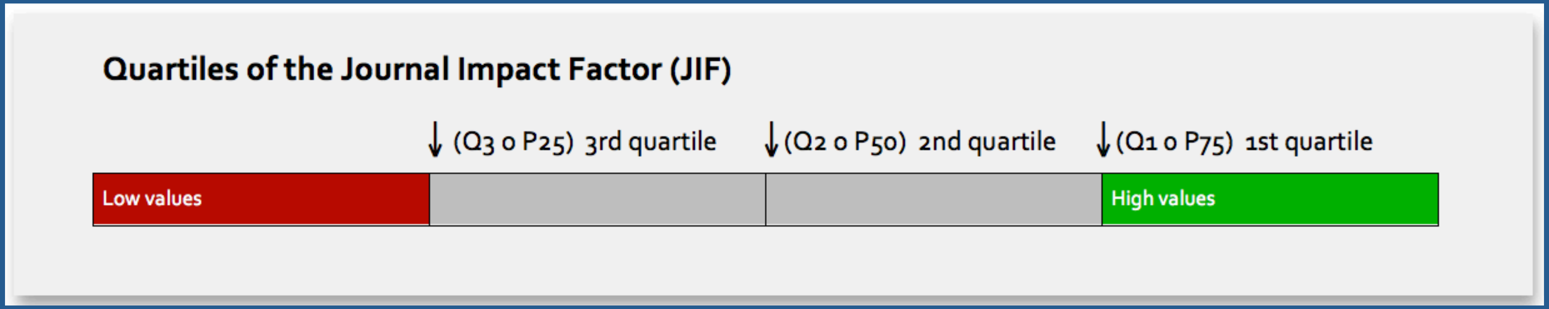
IV Methods: JIF, limitations

- The average is sensitive to extreme values
- Individual publications contribute unevenly to the JIF, specially highly cited publications.
- The most cited 15% of the articles account for 50% of the citations, and the most cited 50% of the articles account for 90% of the citations.

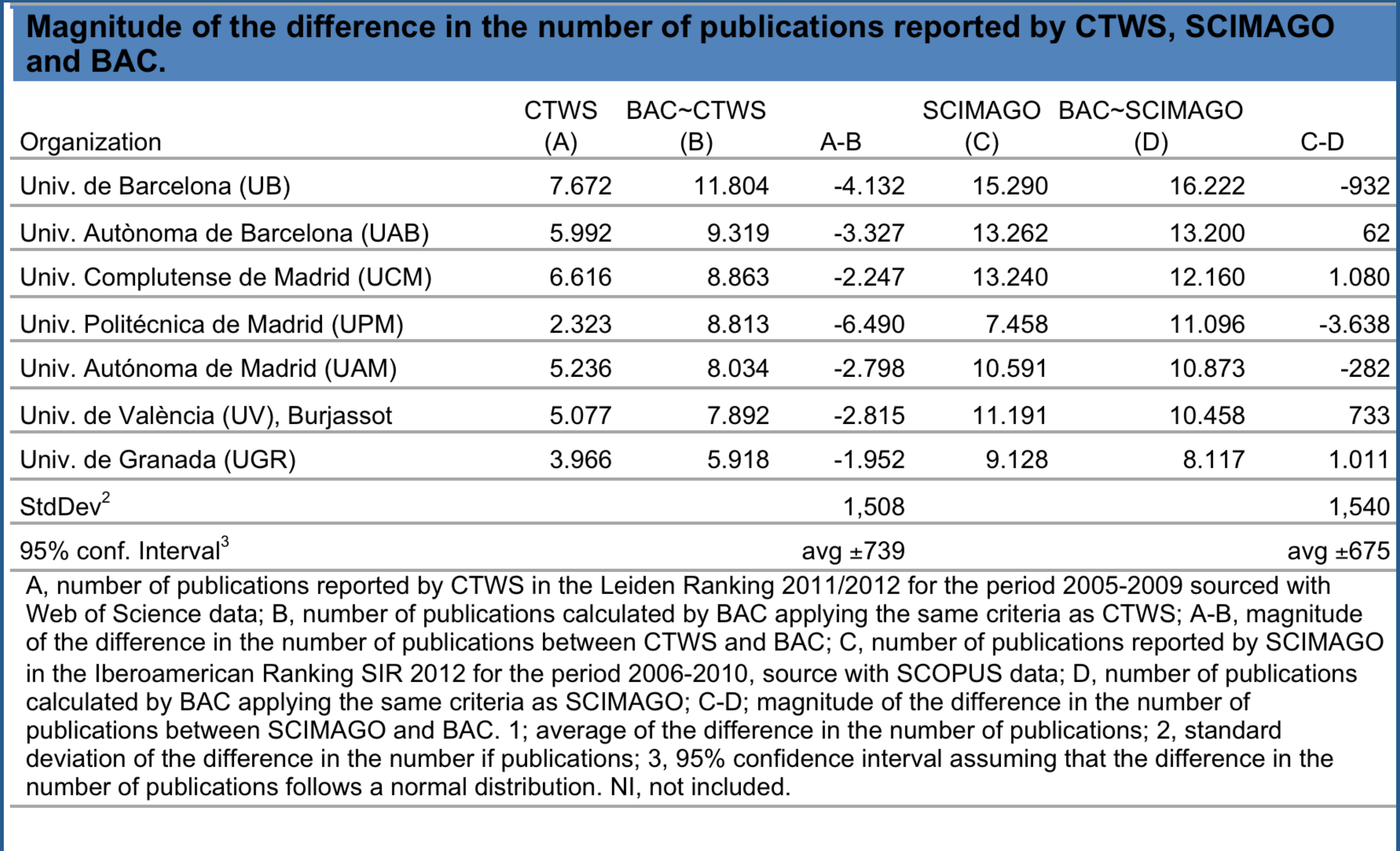
IV Methods: JIF, uses

- Calcualtion of the percentage of publicaitons in the Q1 (JIF ≥ P75 in respective JCR categorires).
  - This indicator aproximation ability of the researcher to overcome specific editorial filters

- Calculation of the sum of the JIF attributed to a center. This indicator is meaningless!



IV Methods: discrepancies figures



Magnitude of the difference in the number of publications reported by CTWS, SCIMAGO and BAC.

Organization	CTWS (A)	BAC~CTWS (B)	A-B	SCIMAGO (C)	BAC~SCIMAGO (D)	C-D
Univ. de Barcelona (UB)	7.672	11.804	-4.132	15.290	16.222	-932
Univ. Autònoma de Barcelona (UAB)	5.992	9.319	-3.327	13.262	13.200	62
Univ. Complutense de Madrid (UCM)	6.616	8.863	-2.247	13.240	12.160	1.080
Univ. Politécnica de Madrid (UPM)	2.323	8.813	-6.490	7.458	11.096	-3.638
Univ. Autónoma de Madrid (UAM)	5.236	8.034	-2.798	10.591	10.873	-282
Univ. de València (UV), Burjassot	5.077	7.892	-2.815	11.191	10.458	733
Univ. de Granada (UGR)	3.966	5.918	-1.952	9.128	8.117	1.011
StdDev^2			1,508			1,540
95% conf. Interval^3			avg ±739			avg ±675

A, number of publications reported by CTWS in the Leiden Ranking 2011/2012 for the period 2005-2009 sourced with Web of Science data; B, number of publications calculated by BAC applying the same criteria as CTWS; A-B, magnitude of the difference in the number of publications between CTWS and BAC; C, number of publications reported by SCIMAGO in the Iberoamerican Ranking SIR 2012 for the period 2006-2010, source with SCOPUS data; D, number of publications calculated by BAC applying the same criteria as SCIMAGO; C-D; magnitude of the difference in the number of publications between SCIMAGO and BAC. 1; average of the difference in the number of publications; 2, standard deviation of the difference in the number if publications; 3, 95% confidence interval assuming that the difference in the number of publications follows a normal distribution. NI, not included.

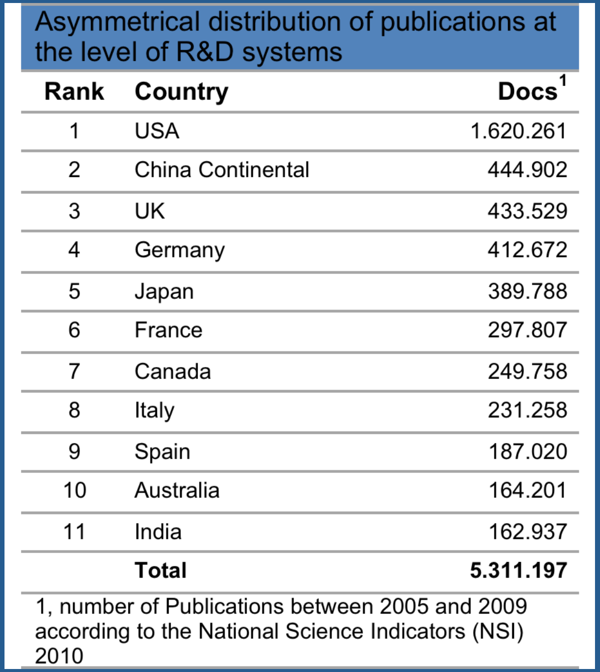
IV Methods: discrepancies figures

There are a number of factor that could explain such differences

- Source data (coverage and completeness)
- Period of the study
- Deepness of the normalization
- Percentage of error in the normalization
- Structure of the center and propagation rules
- Regional peculiarities
- Miss-location of addresses
- Missing addresses (full, partial)
- Types of document taken into account
- Counting method itself

V Bilbiometric analysis: distributions

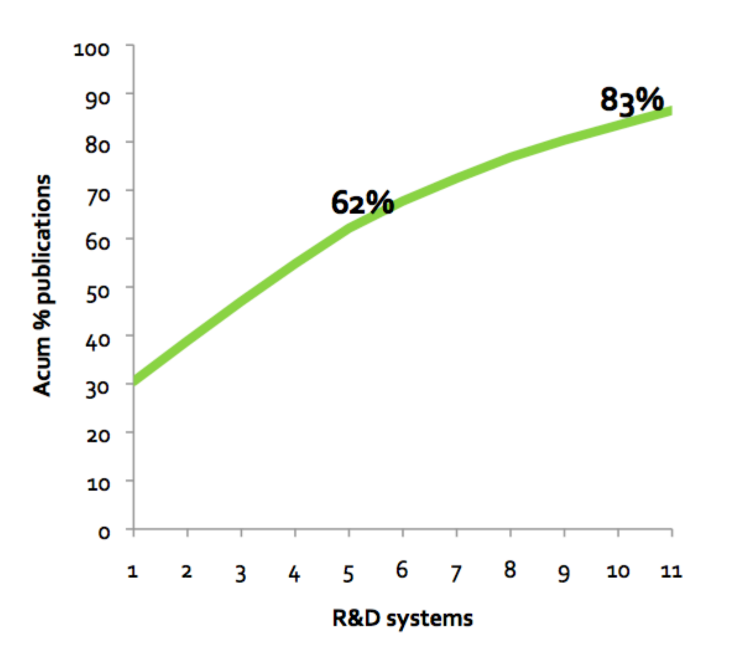
- Publications show asymmetric distributions at all levels in bibliometrics, as few actors account for the major part of the publication output (and citations).



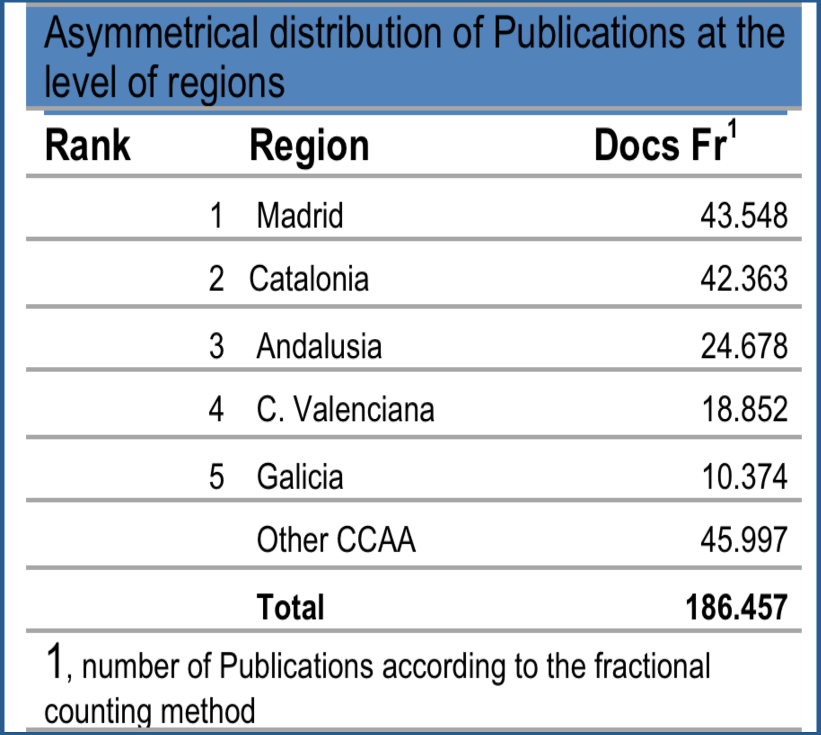
Asymmetrical distribution of publications at the level of R&D systems

Rank	Country	Docs[1]
1	USA	1.620.261
2	China Continental	444.902
3	UK	433.529
4	Germany	412.672
5	Japan	389.788
6	France	297.807
7	Canada	249.758
8	Italy	231.258
9	Spain	187.020
10	Australia	164.201
11	India	162.937
	Total	**5.311.197**

[1] 1, number of Publications between 2005 and 2009 according to the National Science Indicators (NSI) 2010



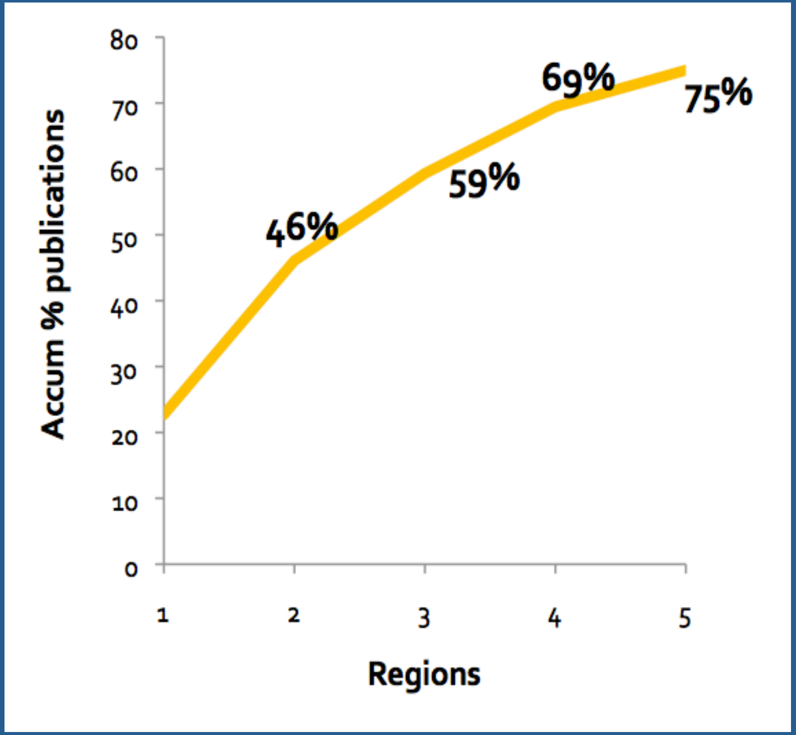
V Bilbiometric analysis: distributions



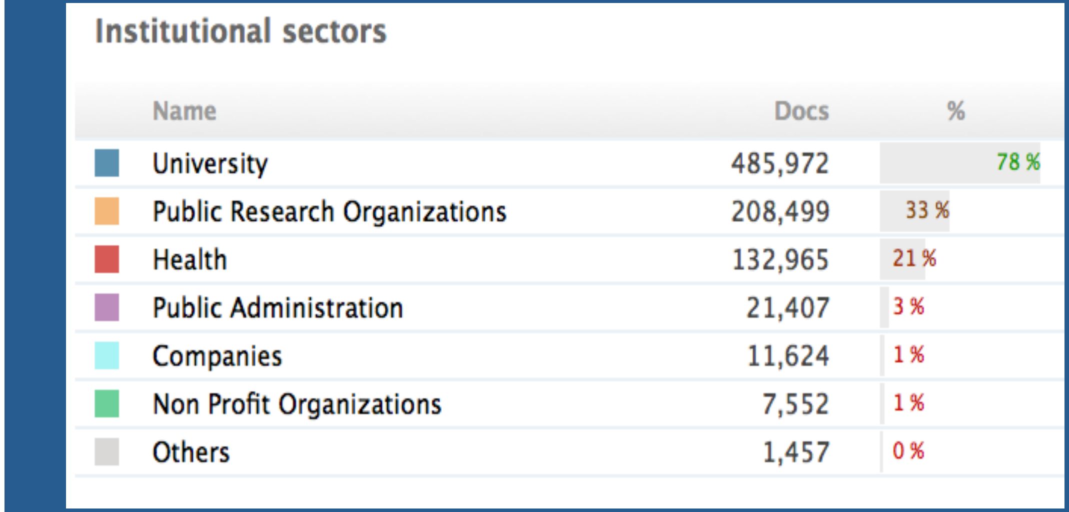
Asymmetrical distribution of Publications at the level of regions

Rank	Region	Docs Fr[1]
1	Madrid	43.548
2	Catalonia	42.363
3	Andalusia	24.678
4	C. Valenciana	18.852
5	Galicia	10.374
	Other CCAA	45.997
	Total	**186.457**

1, number of Publications according to the fractional counting method

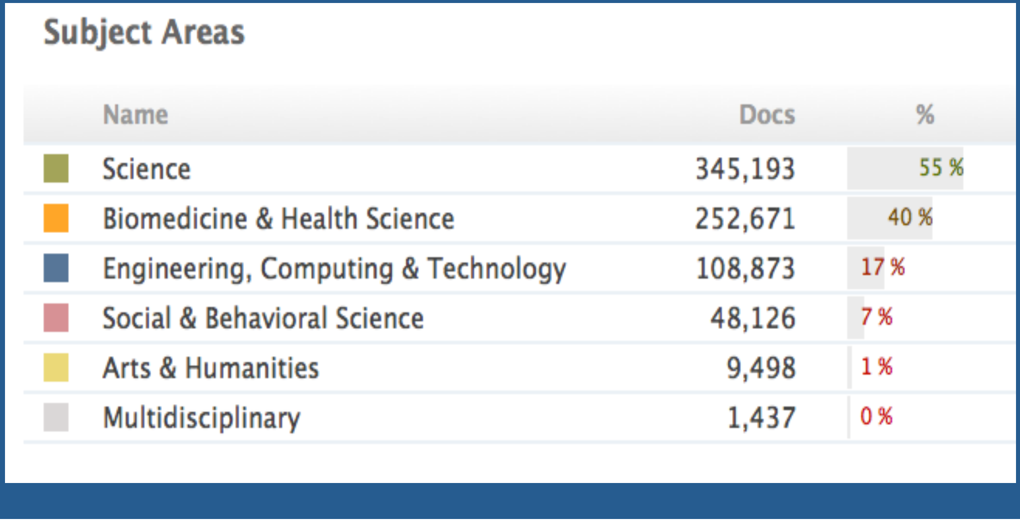


V Bilbiometric analysis: distributions



Institutional sectors

	Name	Docs	%
🟦	University	485,972	78 %
🟧	Public Research Organizations	208,499	33 %
🟥	Health	132,965	21 %
🟪	Public Administration	21,407	3 %
🟦	Companies	11,624	1 %
🟩	Non Profit Organizations	7,552	1 %
⬜	Others	1,457	0 %



Subject Areas

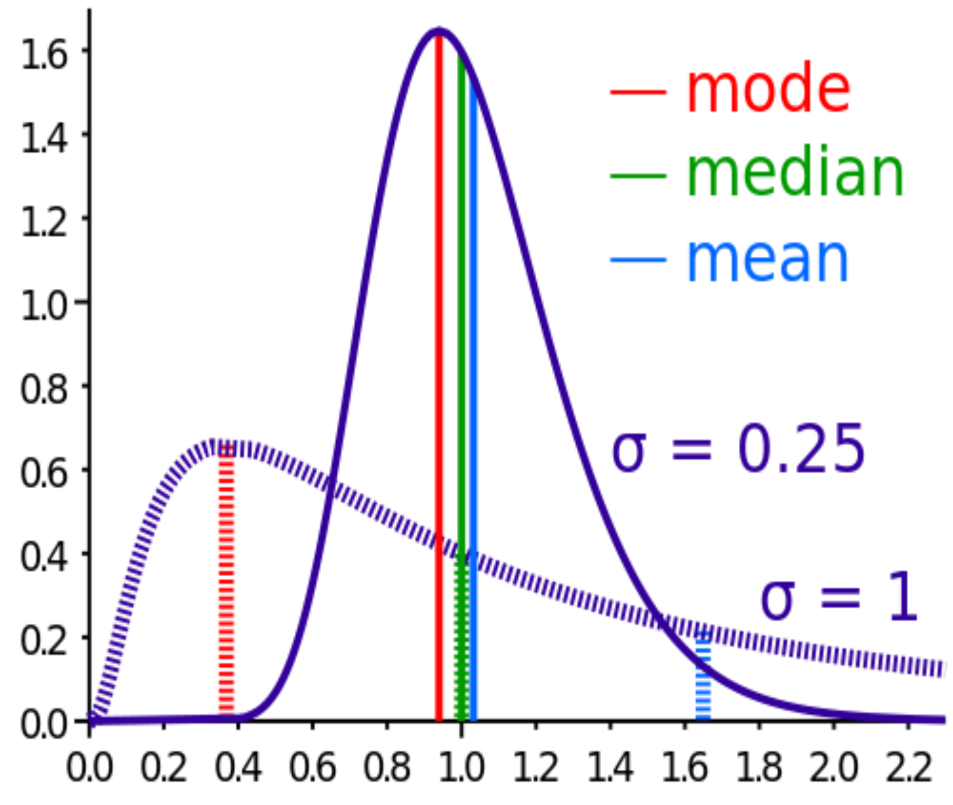
	Name	Docs	%
🟩	Science	345,193	55 %
🟧	Biomedicine & Health Science	252,671	40 %
🟦	Engineering, Computing & Technology	108,873	17 %
🟥	Social & Behavioral Science	48,126	7 %
🟨	Arts & Humanities	9,498	1 %
⬜	Multidisciplinary	1,437	0 %

V Bilbiometric analysis: statistics

- Given that the observations distribute asymmetrically so frequently, it is recommended using these 5 statistics:
  - Minimum
  - percentile 25 (p25)
  - median or percentile 50
  - percentile 75
  - interquartilic range (p75-p25)
  - Maximum

V Bilbiometric analysis: statistics

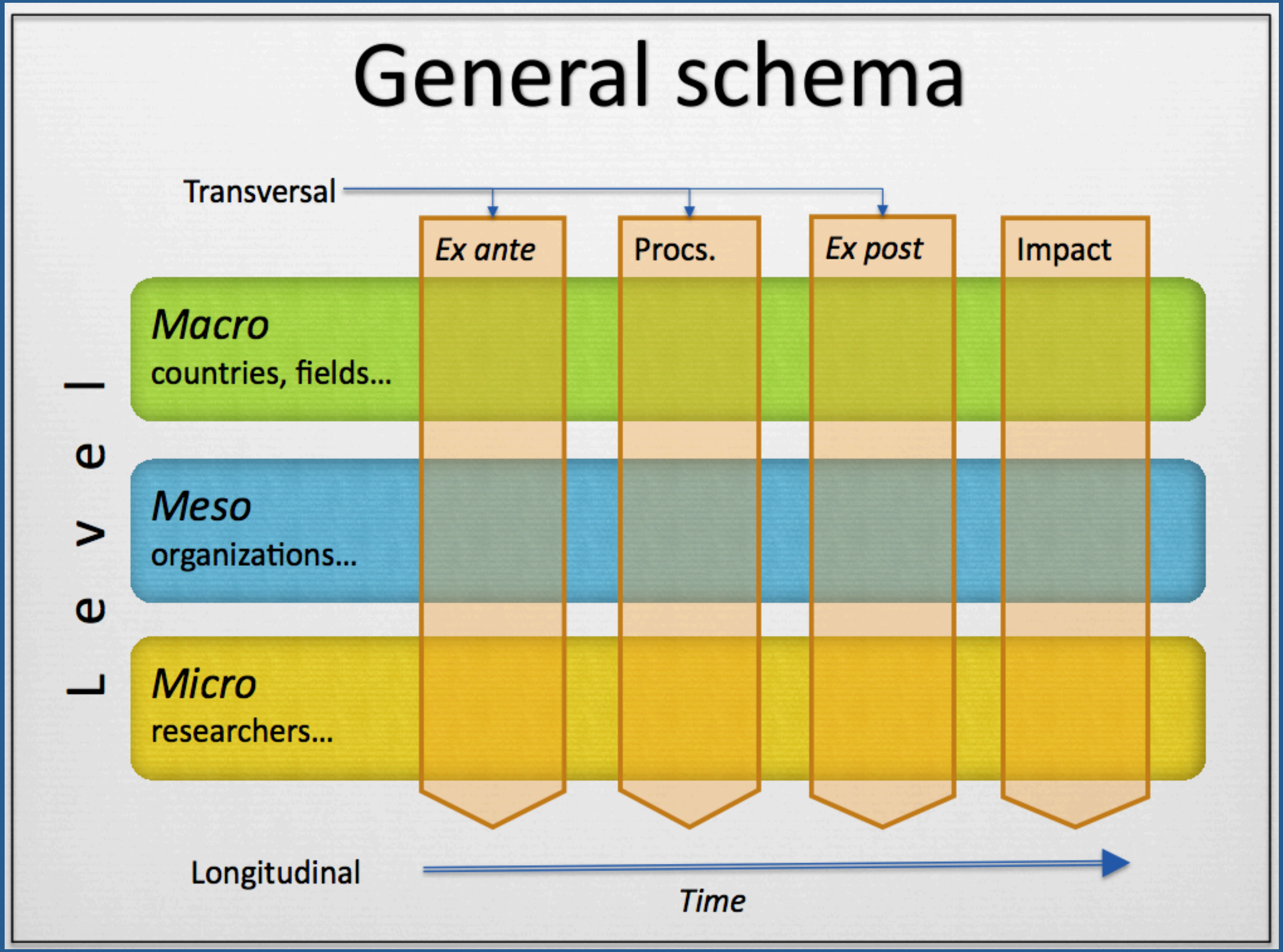
Effect of the distribution of the observations on different statistics



V Bilbiometric analysis: types

- Retrospective vs. prospective
- Univariate vs. multivariate
- Descriptive vs. Inferential
- Size / location
  - Micro
  - Meso
  - Macro
- Time (moment) the analysis
  - Ex ante
  - Process
  - Ex post
  - Impact/outcome
- Scope
  - Transcersal
  - Longitudinal

V Bilbiometric analysis



General schema

V Bilbiometric analysis: reporting

- Design a strategy of analysis aimed to provide answers to those objectives
- During the exploratory phase
  - Design a general schema of analysis and apply it systematically to all levels and actors in order to get the same indicators for all actors.
  - Check for completeness of data
  - Check distributions

V Bilbiometric analysis: reporting

- Explore the data going from general to specific points always, i.e. set the frame in which the unit(s) exists (the environment), and subsequently dive into lower level units to describe them in detail.
- i.e. when analyzing a university we should describe the region (CCAA or country of location) first, in order to set the frame/context for further comparisons.

V Bilbiometric analysis: reporting

- Defining dimensions of analysis makes things a lot easier.
- In bilbiometrics almost every attribute of a publication can be used as a dimension of analysis.

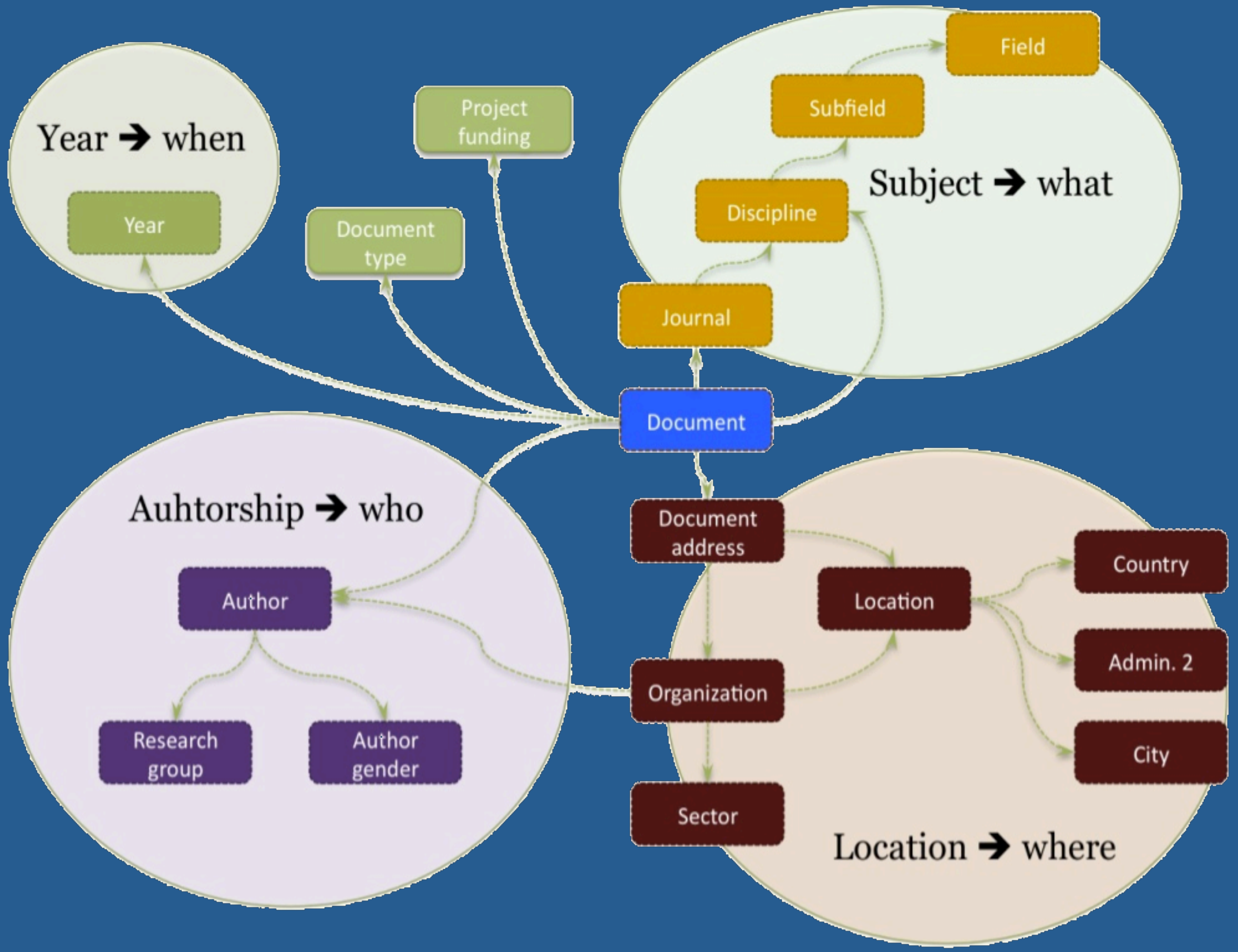
V Bilbiometric analysis: reporting

- The most common bibliographical attributes 4 categories:
- **What**? (matter): research fields, disciplines and journals, keywords.
- **When?** year of publication.
- **Where?** Here we include locations, and organizations, which are normally group into institutional sectors.
- **Who?** authors (and gender studies), as well as research groups.

V Bilbiometric analysis: reporting

- Notice that these attributes are dimensions and units of analysis at the same time.
- Select one dimension and calculate the indicators of the units for the rest of dimensions, and so on whenever it makes sense.
- As several indicators are normaly included in a regular bibliometric study, ordering tables will provide different views of the same phenomenon,

V Bilbiometric analysis: reporting



V Bilbiometric analysis: comparability

- Comparability is one of the most important issues in bibliometrics, since it assures fair assessment/ evaluation.
- However, comparing apples with apples is not always possible, as the components of R&D systems often show peculiarities.

V Bilbiometric analysis: comparability

- Select a (fair enough) classification system that enables grouping apples with apples, and so on.
- Compare the bibliometric indicators of the units inside every homogeneous groups.
- First, apply activity indicators to create subgroups according to size.
  - 3 groups: big, medium and small size units.
  - Order the units, within each size group, by other ind.

V Bilbiometric analysis: Ref. values

- Reference values serve as standards with which comparing a specific indicator of a particular unit is fair.

- Units showing higher values than the reference in a particular indicator are thought to be performing above the average in the specific dimension measured by the indicator.

V Bilbiometric analysis: Ref. values

- Limitations
  - Only reference values for citation rate are available currently
  - There are no widely accepted reference values for activity or cooperation indicators.
- Scope of reference values

V Bilbiometric analysis: Ref. values

- Scope
  - Global: when they are calculated over a wide range of values, let's say, countries, research fields, etc. (world league), or widely accepted
  - Local : when they are calculated on data of local actors (regional league).

VI Sources of bibliometric indicators

- USA: National Science Foundation (NSF)
- Europe: Cordis
- *Research Groups*: CTWS, SCIMAGO, BAC
- Companies: Evidence, london; Science-Metrix, Canada

Gràcies per la vostra atenció…

BAC

Raül Méndez-Vásquez

Research group on bibliometrics

http://bac.fundaciorecerca.cat/