

Politics depends more and more on global indicators. Which data are really important, though? This is the paradox of the future: *do-it-yourself* statistics for global comparisons. After insects, penguins and cloistered

# The power of number is growing: are they the right ones?

STATISTICS 1

by Donato Speroni

monks, the next big-screen stars could be numbers: the dry, dull data of world statistics presented in such an interesting and dynamic way that even the masses will be entertained

**A** Swedish foundation that presents its findings free of charge on its website [www.gapminder.org](http://www.gapminder.org) has been transforming numbers into entertainment for years. Its motto is: "Making sense of the world by having fun with statistics!" The trick is to make the graphic displays that describe the behaviour patterns of different countries move through time, by combining them with historical statistical series. For example, by intersecting the numbers from the pro-capita gross domestic product (GDP) with the infant mortality rate on a Cartesian axis, and projecting that data from the past decades as if it were an animated cartoon, the impression of economic development of certain parts of the world becomes very clear, but so too is the fact that in certain cases, despite the increase in average wealth, social conditions have not improved at all. The site is fun as well; moreover, it offers a graphic representation of the UNPD (the United Nations Programme for Development) for 2005, which explains many things about the distribution of well-being and wealth throughout the world. It is available in eight languages, but Italian is not one of them: this is not surprising though, given that the activity of this foundation, which has received support from the United Nations, and is quite well-

known internationally, has caught the attention of just a few wayward bloggers in Italy. Hans Rosling is a professor of international medicine who, with the collaboration of his computer scientist son, invented Gapminder in order to make the problems of world health more visible, and who, with help from the UN and a number of other agencies involved in matters of international development, has extended its use to all major global phenomena. He says he thought up this presentation method when he realised that a sample group of Swedish students, who were given an elementary test comparing living conditions in different countries, were only able to correctly answer two out of five questions. "Even a group of Nobel Prize winners was not able to answer more than 50% of the questions correctly on average, the same percentage that a sample of chimpanzees would answer correctly by chance", he explains on his website. Rosling's theory is that our ideas about the world are, in large part, the result of stereotypes; and in order to change them, you simply need to know a few facts, presented in an effective manner. Since our attention span is becoming more and more limited due to the number of messages we are inundated with, though, there is a need for



上  
テム  
リツ  
台金  
ナイ  
精  
工  
部  
株

45  
97  
318  
42  
291  
47  
34  
97  
11  
25  
13  
0  
3  
1  
3  
0  
1  
2  
7  
1  
5  
5  
5  
9  
7

7 1/2  
1 13/16  
6.57  
6.60  
6.62  
6.57  
6.61  
6.59  
6.61  
6.62  
6.59  
6.61  
6.50  
6.55  
6.58  
6.60  
6.60

+0.23  
+0.24  
+0.24  
+0.28  
+0.25  
+0.27  
+0.28  
+0.26  
+0.30  
+0.31  
+0.27

05 - 10110.10137  
06 - 11146.10140  
06 - 11148.31149  
07 - 12125.04125  
08 - 13121 5/8 30  
09 - 14150 3/8 50  
09 - 14k144.2 14  
09 - 14k139.1413  
15 k 142.17 142.2  
15 k 136.16 136.1  
15 k 128.29 129.0  
122.27 122.2  
102.0  
104.4  
117.3  
1.16  
1.17  
1.22  
1.23  
1.24  
6 1/2  
8 1/8  
14  
45 1/2  
33 3/8  
-1  
-1

1 1/2  
3 1/8  
1 1/2

+1.31  
+1.29  
+1.25  
+1.23  
1.16  
1.17  
1.22  
1.23  
1.24

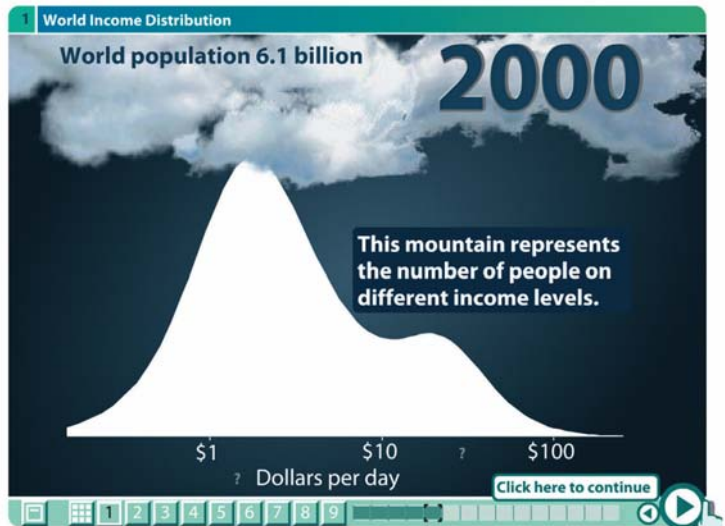
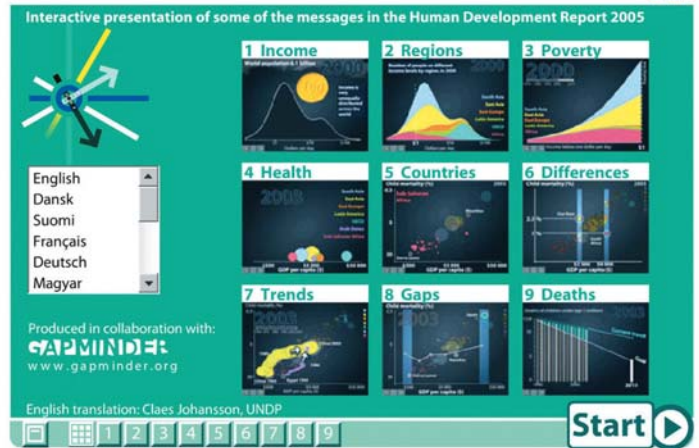


entertaining and dynamic techniques to be applied to the presentation of official numbers. The difficulty in presenting numbers on video is all too familiar to radio and television journalists; this in an age when information makes ample use of data to summarise any type of phenomenon, from levels of corruption to the results of government policies. The printed press is full of rankings and those in positions of power must keep up. Even if the rotten apples offered up by bogus institutions were eliminated from this barrage of numbers, an enormous amount of good-quality information is still available to assist international public opinion: numbers which are worked up by public institutions and private subjects according to methodologies that have received wide scientific consensus. A mass of revelations, questionnaires and analyses which result in enormous public investment, the involvement of hundreds and thousands of economists, statisticians and social scientists throughout the world, and the search for information in all aspects of human behaviour. The results end up, after some initial fireworks shot off in the press, by and large stored in data warehouses that are grossly under-utilised, even by the most expert of researchers. The world is like a patient who has at his disposal every possible examination offered through modern technology, and continuously refers to them, but hasn't yet found the right diagnostician able to effectively cross-reference the information in order to propose an adequate therapy.

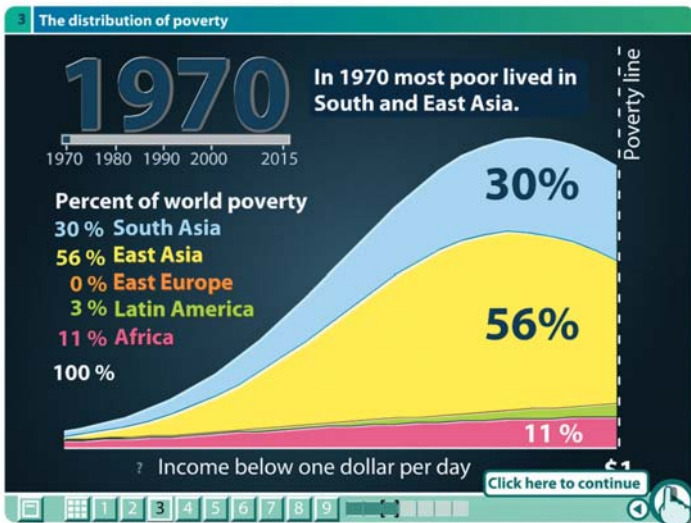
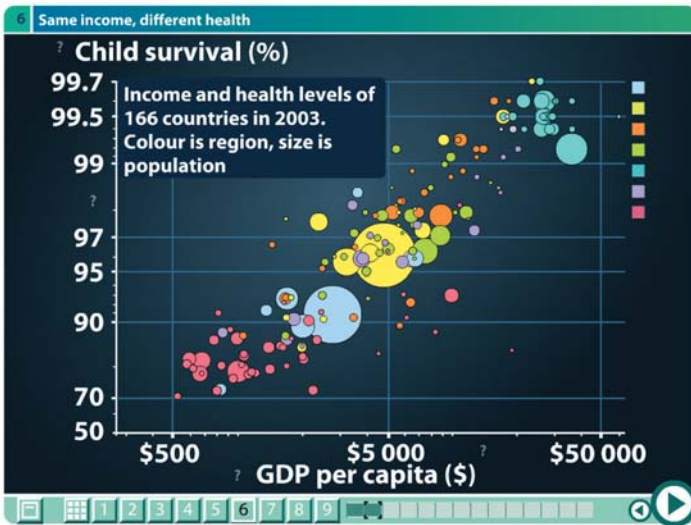
**How can we give political weight to this data?**

How can we best use this mass of numeric information? The roads that have been taken in recent years obviously lead to the construction of a consensus on the selection of the most important series, which must be presented with a format of easily readable indicators. However, with regard to the methods for elaborating data, there are a variety of techniques and the debate has become political. The first solution, preferred by the media, consists of synthetic indicators. A variety of different statistics are grouped together, sometimes given different weight depending on their importance, to make a synthetic indicator. It is, for example, the route taken by "Il Sole 24 Ore" in the construction of the indices of quality of life among the provinces in

**Human Development Trends 2005**



Italy, or by the World Economic Forum (Davos) for the global Index for competitiveness. This is similar to the way that the UNPD created its Human Development Index (HDI) which tracks the per-capita gross product, life expectancy and literacy rates in each country. Some years ago, it seemed that the HDI was destined to replace GDP in terms of importance, but now many experts wrinkle their noses at it, as would a physician, when faced with an index number which, despite being based on a number of different analyses, is meant to gauge the overall health of a patient on a scale from one to one hundred. The alternative to synthetic indicators is the concurrent examination of various historical series that are particularly significant. This



Statistical synopses taken from the Internet. The graph describing income distribution is particularly striking: the majority of the world’s population survives on daily earnings of around \$2 US per day

solution is more complicated, but also more serious, according to experts. The case of competitiveness, examined in detail in the interview with economist Roberto Monducci that follows, demonstrates that no single piece of data can provide an accurate overview of the strength of a nation’s position on the stage of global economics. Instead, a number of different factors must be analysed: productivity, the situation of the national currency, the

ability to sell to international markets; however, the mid-range resistance of the country’s social system must also be measured in some way. It is not easy to translate this combination of data into something more convincing than an analysis for field experts. And it is above all in the presentation of complex data that the new presentation techniques can provide for an increase in quality, on the condition that institutes of official statistics, as Gapminder hopes, in addition to producing numbers of acceptable quality, make them transparent and accessible free of charge, as well as comparative and homogenous over time. They must also take advantage of the possibility to utilise the most advanced web graphics techniques to present them, something many marketing agencies have been doing for some time.

**Which data really count?**

“Power from data”, the power that comes from knowledge of numbers, is the title of the online training programme of Statistics Canada which is available to both teachers and students. Statistics is a component that is becoming more and more important for reaching a consensus: not only among the willing elite who exchange ideas and analyse data, but also among the masses who are strongly influenced by the media. But its growing importance also requires solutions to some essential questions regarding democracy and development. The first regards the origin of data and the quality of elaboration techniques. east dealt with this subject (no. 9), in an interview with Enrico Giovannini, the economist who manages the office of statistics of the Organisation for Economic Cooperation and Development (OECD) and who deeply understands the problems inherent in the international system of statistics. No less important is another question that floats among those who are intolerant of these domineering numbers: what do these numbers really mean, when they tell the story of a human community, be it a region or a nation, and are compared with others? Are we certain that the numbers we discuss on a daily basis are truly worthy of influencing such major decisions?

A couple of examples show that it is not just a subject limited to field experts. Two years ago, east (no. 2) signalled the tendency of economists to search for alternatives to the



07 1007516 3502542 060  
70 4972772 4272701 34  
37 0303174 527487 20  
50 5467299 0201354 22  
20 0001627 4381780 13  
59 0675202 5733167 00  
37 5070513 5075054 28  
20 4616370 0551240 01  
53 3166937 3615457 59  
12 1707064 0973100 34  
00 0007923 5166497 00  
58 3154629 3702136 13  
04 1345207 5203164 40  
78 1546792 0252631 37  
22 3007154 0136159 00  
00 4792700 4623730 23  
56 3167927 1743465 09  
51 0034679 3764750 02  
64 2514569 0334047 18  
97 0601740 4575731 54  
17 1546455 0451642 00  
07 2004016 0605534 42  
52 4975679 0715973 39  
74 4313134 4370067 29  
55 3146970 5937017 40  
20 0924562 0001723 55

omnipresent GDP, perhaps by measuring the degree of happiness of individuals. Back then it seemed to be little more than a statistical curiosity, but in the meantime the debate has grown, with contributions made by important international organisations. And it has now become a political matter. It's no longer only King Jigme Singye Wangchuck of Buthan who has decided to calculate his annual growth objectives in terms of Gross Domestic Happiness, or GDH, refusing to recognize GDP value; no longer is it only the anti-global activists who challenge traditional measurements of wealth, even the new British Conservative leader David Cameron has proposed substituting the GDP (Gross Domestic Product) with the GWB (General Well Being). It is an issue that has even been addressed by "The Economist", with the cover of its 2006 year-end special edition reading: "Happiness – and how to measure it". In an enjoyable essay the British publication reports that the idea of measuring happiness has accompanied the history of the science of economics since the 1800s, but is now fashionable once again because it has become more and more evident that the increased availability of material goods is not what makes people happy: indeed, everyone feels that they are in continual competition with others. Be warned though: one must not "fossilise" society by renouncing competition. "Capitalism can make a society rich and keep it free. Don't ask it to make you happy as well." This is the magazine's theory, which is perfectly in line with its libertarian leanings. The debate goes on.

The comparison among different nations' systems is often vigorously discussed as well. From the measure of a nation's competitiveness comes its recipe for industrial politics; from its twin sister, the degree of productivity, come its work policies. But there are those who doubt the value of these data, when they are not limited to measuring the performance of single enterprises. For example, the economist Paul Krugman contends, "Competitiveness is a word without meaning when applied to national economies. The obsession with competitiveness is wrong as well as dangerous". Krugman first made these assertions in 1994, but the debate continues today and is an animating force in the culture of European economics, so much so that last November, during the Eighth National



Chiara Saraceno, a Professor of Sociology of the Family, proposed the integration of traditional indicators of competitiveness with those of social quality. Among others, the “indicators of empowerment”, which mirror the economic ideas of the Nobel Prize winner Amartya Sen

Conference on Statistics, Chiara Saraceno, a Professor of Sociology of the Family, proposed the integration of traditional indicators of competitiveness with those of social quality. Saraceno put forth among others the “indicators of empowerment”, which mirror the economic ideas of the Nobel Prize winner Amartya Sen, and are careful to also measure contextual aspects, or rather, as the Italian scholar explains, “the capacity of people to respond and influence change and receive benefits from social development in terms of activation and reinforcement of their own potential”: it’s a fascinating subject, still evolving greatly.

**Must we depend upon official statistics?**

In addition to the discomfort of those who think that up to this point too much importance has been given to economic rather than social indicators, there are those who show their intolerance for the fact that good statistics, given their cost, inevitably end up “falling from on high”. They come from institutions that should be immune to political influence (not always the case), but which in any case send their messages to public opinion in a “top down” manner, with the prescriptions for recipes which institutes habitually do not



share, and returning to the medical metaphor, the ratio between the diagnosis and the prognosis is minimal: for a single illness perhaps there may be three or four different cures, but in any case the treatment options are limited. Like alternative medicine, there should also be alternative statistics: it is possible, argue critics, that official statistical researchers, despite all their good intentions and skill, are able to see only certain items and no others, just as Western physicians ignored the force fields of the human body that are well-known in Chinese medicine. But statistics are not the same thing as kinesiology: it is almost impossible for opponents to organise a system of revelations. An example is the media hubbub that occurred in Italy with the juxtaposition of "perceived" and "official" inflation rates during the changeover from the Lira to the Euro. With the excuse that it was too expensive to create an alternative, adequate monitoring system to ISTAT, in the end opponents of the new currency gave room to the elaborations based on small numbers or on far-fetched assumptions that would have never been taken seriously in other countries.

So? Are we condemned to a monopoly of official institutions that can only guarantee us the quality of data? Up until yesterday, it looked as though it was a paradox without a solution, but thanks to new technology, it seems there is a way out. In Italy there has been talk, probably for the first time, at a seminar on the subject of data quality during the previously mentioned conference on statistics. The solution, proposed by Giovannini himself and already being studied at the OECD, consists of using official statistical questionnaires to create private statistics. To understand the scope of this novelty, the reader must be patient. We must begin from a premise: whether they are collected for a census, for a sampling or for administrative information, a large portion of the most important statistics are created by managing and combining millions of different pieces of information. Italy can be used as an example, but the problems are similar throughout the world: every three months, work force findings keep thousands of interviewers busy verifying the employment status of 175,000 individuals resident in 1,246 cities and towns in every province throughout the country. 700,000 interviews are conducted each year, with tens of millions of answers. ISTAT gets its quarterly

GLOBAL COMPETITIVENESS INDEX 2006 AND 2005 COMPARISONS

CONTRY/ECONOMY	GCI	GCI	GCI	CHANGES 2005/2006	POSITIONS EARNED OR LOST
	2006 RANK	2006 SCORE	2005 RANK		
Switzerland	1	5.81	4	ä	3
Finland	2	5.76	2	à	0
Sweden	3	5.74	7	ä	4
Denmark	4	5.70	3	æ	-1
Singapore	5	5.63	5	à	0
United States	6	5.61	1	æ	-5
Japan	7	5.60	10	ä	3
Germany	8	5.58	6	æ	-2
Netherlands	9	5.56	11	ä	2
United Kingdom	10	5.54	9	æ	-1
Hong Kong SAR	11	5.46	14	ä	3
Norway	12	5.42	17	ä	5
Taiwan, China	13	5.41	8	æ	-5
Iceland	14	5.40	16	ä	2
Israel	15	5.38	23	ä	8
Canada	16	5.37	13	æ	-3
Austria	17	5.32	15	æ	-2
France	18	5.31	12	æ	-6
Australia	19	5.29	18	æ	-1
Belgium	20	5.27	20	à	0
Ireland	21	5.21	21	à	0
Luxembourg	22	5.16	24	ä	2
New Zealand	23	5.15	22	æ	-1
Korea, Rep.	24	5.13	19	æ	-5
Estonia	25	5.12	26	ä	1
Malaysia	26	5.11	25	æ	-1
Chile	27	4.85	27	à	0
Spain	28	4.77	28	à	0
Czech Republic	29	4.74	29	à	0
Tunisia	30	4.71	37	ä	7
Barbados	31	4.70	—	n.a.	
United Arab Emirates	32	4.66	32	à	0
Slovenia	33	4.64	30	æ	-3
Portugal	34	4.60	31	æ	-3
Thailand	35	4.58	33	æ	-2
Latvia	36	4.57	39	ä	3
Slovak Republic	37	4.55	36	æ	-1
Qatar	38	4.55	46	ä	8
Malta	39	4.54	44	ä	5
Lithuania	40	4.53	34	æ	-6
Hungary	41	4.52	35	æ	-6
Italy	42	4.46	38	æ	-4
India	43	4.44	45	ä	2
Kuwait	44	4.41	49	ä	5
South Africa	45	4.36	40	æ	-5
Cyprus	46	4.36	41	æ	-5
Greece	47	4.33	47	à	0
Poland	48	4.30	43	æ	-5
Bahrain	49	4.28	50	ä	1
Indonesia	50	4.26	69	ä	19
Croatia	51	4.26	64	ä	13
Jordan	52	4.25	42	æ	-10
Costa Rica	53	4.25	56	ä	3
China	54	4.24	48	æ	-6
Mauritius	55	4.20	55	à	0
Kazakhstan	56	4.19	51	æ	-5
Panama	57	4.18	65	ä	8
Mexico	58	4.18	59	ä	1
Turkey	59	4.14	71	ä	12
Jamaica	60	4.10	63	ä	3
El Salvador	61	4.09	60	æ	-1
Russian Federation	62	4.08	53	æ	-9

Source: Global Competitiveness Report 2006 - 2007 of the World Economic Forum.



CONTRY/ECONOMY	GCI	GCI	GCI	CHANGES 2005/2006	POSITIONS EARNED OR LOST
	2006 RANK	2006 SCORE	2005 RANK		
Azerbaijan	64	4.06	62	æ	-2
Colombia	65	4.04	58	æ	-7
Brazil	66	4.03	57	æ	-9
Trinidad and Tobago	67	4.03	66	æ	-1
Romania	68	4.02	67	æ	-1
Argentina	69	4.01	54	æ	-15
Morocco	70	4.01	76	ä	6
Philippines	71	4.00	73	ä	2
Bulgaria	72	3.96	61	æ	-11
Uruguay	73	3.96	70	æ	-3
Peru	74	3.94	77	ä	3
Guatemala	75	3.91	95	ä	20
Algeria	76	3.90	82	ä	6
Vietnam	77	3.89	74	æ	-3
Ukraine	78	3.89	68	æ	-10
Sri Lanka	79	3.87	80	ä	1
Macedonia, FYR	80	3.86	75	æ	-5
Botswana	81	3.79	72	æ	-9
Armenia	82	3.75	81	æ	-1
Dominican Republic	83	3.75	91	ä	8
Namibia	84	3.74	79	æ	-5
Georgia	85	3.73	86	ä	1
Moldova	86	3.71	89	ä	3
Serbia and Montenegro	87	3.69	85	æ	-2
Venezuela	88	3.69	84	æ	-4
Bosnia and Herzegovina	89	3.67	88	æ	-1
Ecuador	90	3.67	87	æ	-3
Pakistan	91	3.66	94	ä	3
Mongolia	92	3.60	90	æ	-2
Honduras	93	3.58	97	ä	4
Kenya	94	3.57	93	æ	-1
Nicaragua	95	3.52	96	ä	1
Tajikistan	96	3.50	92	æ	-4
Bolivia	97	3.46	101	ä	4
Albania	98	3.46	100	ä	2
Bangladesh	99	3.46	98	æ	-1
Suriname	100	3.45	—	n.a.	—
Nigeria	101	3.45	83	æ	-18
Gambia	102	3.43	109	ä	7
Cambodia	103	3.39	111	ä	8
Tanzania	104	3.39	105	ä	1
Benin	105	3.37	106	ä	1
Paraguay	106	3.33	102	æ	-4
Kyrgyz Republic	107	3.31	104	æ	-3
Cameroon	108	3.30	99	æ	-9
Madagascar	109	3.27	107	æ	-2
Nepal	110	3.26	—	n.a.	—
Guyana	111	3.24	108	æ	-3
Lesotho	112	3.22	—	n.a.	—
Uganda	113	3.19	103	æ	-10
Mauritania	114	3.17	—	n.a.	—
Zambia	115	3.16	—	n.a.	—
Burkina Faso	116	3.07	—	n.a.	—
Malawi	117	3.07	114	æ	-3
Mali	118	3.02	115	æ	-3
Zimbabwe	119	3.01	110	æ	-9
Ethiopia	120	2.99	116	æ	-4
Mozambique	121	2.94	112	æ	-9
Timor-Leste	122	2.90	113	æ	-9
Chad	123	2.61	117	æ	-6
Burundi	124	2.59	—	n.a.	—
Angola	125	2.50	—	n.a.	—

employment and unemployment rates from these “statistical atoms” which also furnish the data for other employment-related publications and analyses throughout the year.

If one digs deeper there is even more to discover, though. Indeed, the material is already currently available to universities and private researchers in two different formats: in the form of “databases” containing “standard files”, or rather “molecules” obtained by grouping some responses in order to safeguard the privacy of those interviewed; or else by working directly with micro-data, or each individual response, but on the condition that experimentation of combinations occur only in places that are inaccessible to hackers.

Otherwise, entire fortunes of personal information could be stolen for commercial or criminal purposes. Something is lost, however, in the aggregation of standard files: external researchers complain because the “molecules” offered to their research “contaminate” the content of the information. On the other hand, combining micro-data in new elaborations is not easy because in Italy, at least for now, it can only happen in the “synchrotron of the statistical atoms”, that is the super-secure laboratory Alice at the ISTAT headquarters in Rome. Recently, however, technology has found a way to guarantee online consultation without the risk of micro-data being downloaded and stolen. For example, the micro-data of the Australian Bureau of Statistics can be queried and combined online from any place throughout that country’s vast territory. And OECD even thinks it will be able to connect the micro-data from different national statistics in one single database. In this way it will be possible, for example, to analyse the behaviours of global enterprises, which are today described as a kaleidoscope of territorially limited revelations.

We can therefore expect a future in which alongside good “top down” statistics, there will also be good “bottom up” statistics: any researcher will be able to use for free, or at very little cost, the wealth of information from statistics institutes. Returning to the example of the Italian work force, the Censis or for that matter, any university, will be able to provide employment and unemployment rates which, instead of focusing on the traditional age range from 15 to 64 years, will provide a snapshot from the ages of 18 to 67, a range that is much more reflective of modern society, without





having to wait for ISTAT's results. Additionally, the supply of statistical breakdowns by territory or employment category would increase. It will take years to adequately prepare networks and micro-data; but as confirmation of the trend of "do-it-yourself" statistics, or personalised elaborations on the basis of reliable data, there is also a new software from the British Institute of National Statistics, which allows anyone to calculate their own personal inflation rate, thereby modifying price indexes according to individual consumption habits. It is a solution which, if applied in Italy, would surpass by leaps and bounds the age-old debate of the multiplying effect that the official "baskets" have on the cost of living: in addition to the numbers that are currently reported for the categories of "factory and office workers", it will be possible to calculate the statistics of retired persons, young couples or even plumbers and journalists, each with their own consumption dynamics. However, these personal baskets do even more: they draw on the irreplaceable conclusions that come from official statistics at thousands of stores and supermarkets and they are combined based on individual behaviour.

### **Towards the data highway?**

"Bottom up" statistics is simply the application of a principle that is often used in other areas: the separation of

Each year, scholars and politicians from all over the world meet at the World Economic Forum in Davos, Switzerland. Recent years have seen an increase in criticisms of the Forum, which some feel has become a promenade for the powerful.

infrastructures from the subjects that use them in order to offer products and services. In this case the infrastructure would be offered up as official micro-data: perhaps imperfect, but it has the characteristic of being a natural monopoly. What better way to conduct a census? Information products, on the other hand, involve the elaboration of facts not just by an official institution, but also by a multitude of external subjects, in moderation and with a freedom that is well above what actually exists.

Even along the information highway of statistical data, one will encounter rumbling Ferraris as well as shoddy old bangers: indeed, the quality of the results will depend upon the criteria used for elaboration purposes, which must always be declared and will in any case be open to discussion. We are certain to be thunderstruck by the new discoveries that some researchers will put forth while others will be considered ridiculous for making such bad use of data: much like what already occurs in other fields of scientific debate.