"Capacity Building and Strengthening Institutional Arrangement / Data Yearbook"

Workshop: "How to produce an Environmental Data Year Book"

The method of presentation of environmental information

Ms. Mariaconcetta Giunta, Mr. Giovanni Finocchiaro,
Ms. Cristina Frizza, Mr. Luca Segazzi

APAT

Agency for Environmental Protection and Technical Services

Definition

A presentation to the mind in the form of an idea or image

A creation that is a visual or tangible rendering of someone or something

Representation

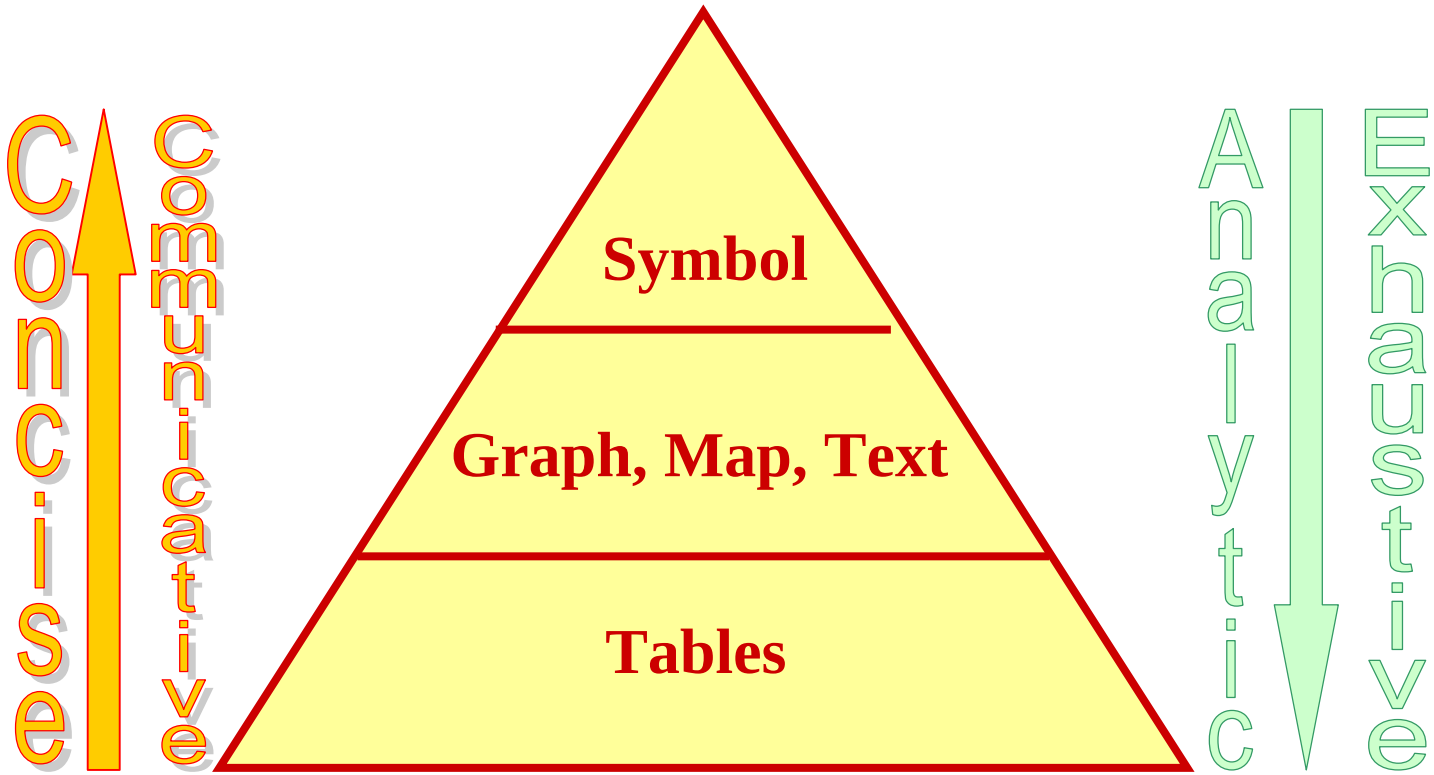
The data representation is an important phase of the development of information processing, in particular, about the transmission effectiveness.

Several modes to represent the results of an environmental analysis exist.

It is very important to choose that it is more easy and quick to understand by the stakeholders.

Some representations are more analytic and exhaustive, others are more communicative and concise.

Kind of Representation (1)

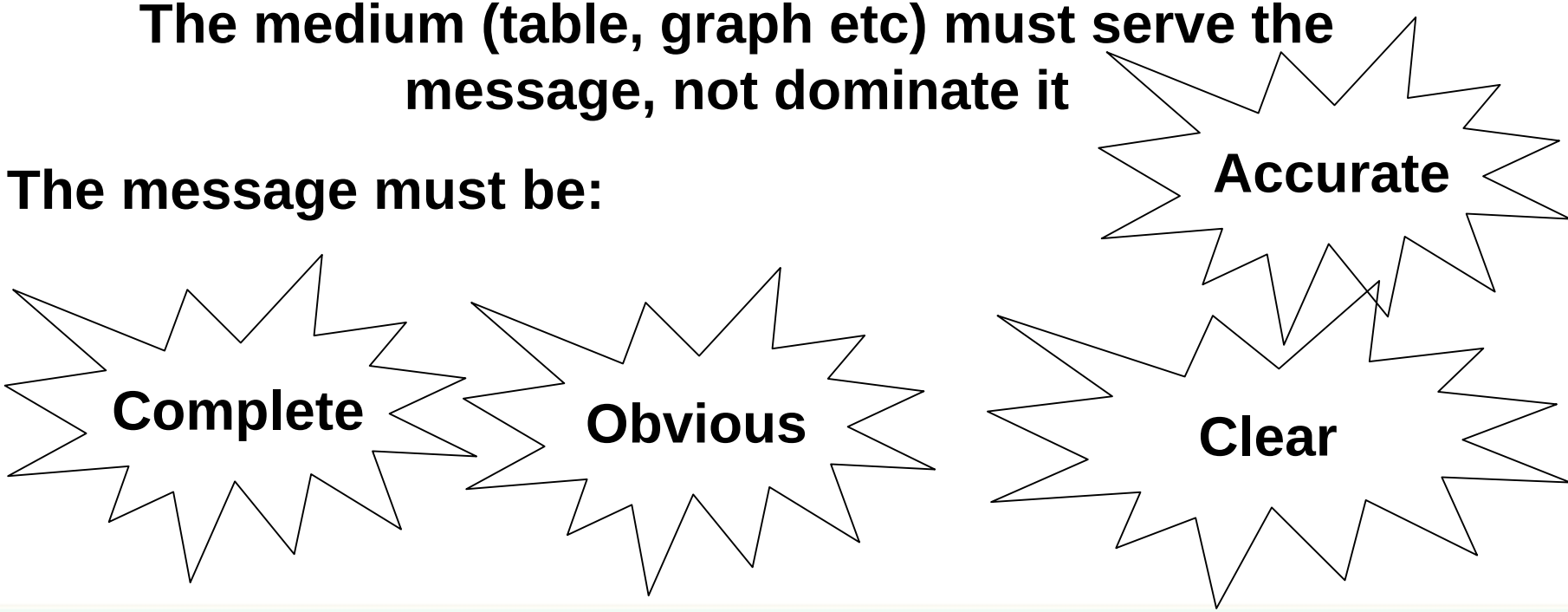


Kind of Representation (2)

Each kind of Representation, table, graph or map, text, and symbol, giving environmental statistical data, must be able to stand alone (as should each paragraph of text), with all necessary background information included.

The medium (table, graph etc) must serve the message, not dominate it

The message must be:



Stakeholders (1)

Want:

- **To be told why the author is telling them this information**
- **The conclusions that the author thinks it is important to know**
- **An idea of what sort of conclusions it is safe to draw from the data – including how reliable the data are**

Stakeholders (2)

Need:

- **A description of how and why the information was obtained**
- **Elaboration on any footnotes to graphs and tables,if needed**
- **Explanations of any inconsistencies in presentation**
- **References to other chapters where they would be helpful**

In short you should answer the questions you would ask if your presentation came from outside the Agency and you did not agree with the conclusions.

If you do this, you have done enough

Kind of representation

Text

Table

Graph

Map

Symbol

Text

- Conclusions
- Numbers
- Use the correct tense
- Conversion from other sources

Conclusion

Every conclusion must be backed by corroborative evidence

If you cannot refer to a graph or table you must give the data as you draw the conclusion. You may also need to give some intermediate calculation such as a percentage as well.

The statistics are there to help the reader draw the right conclusions, to give confidence in the conclusion drawn and to show why they were reached.

You will present your data/statistics in such a way that it easy to draw the right (ie true, not politically or commercially convenient) conclusion and difficult to draw the wrong one

Numbers

The text often needs to get across ideas like “most of” or “rapidly increasing”

If you do not give numbers, people will think that you are relying on guesses or feelings in your bones

If the numbers are not immediately obvious from graphs or tables, it is better to give them in the text

“most…..92 out of 120…..about three quarters…”

“rapidly increasing…by about 20% per year since 1990, (see Figure xxx)”

Use the correct tense

People frequently use the wrong tense when reporting statistics or data

The rule is:

- **past tense for describing the data (because you are describing the state of affairs when the date were collected)**
- **present or, rarely, future tense for drawing generalised conclusions (because you are using the data to make estimates about the present or future)**

Conversion from other sources

Many data, whether as graphs, maps, tables or as text, will have been quoted from other publications.

Agency authors must remember that is the agency house style that is paramount (it is very important).

The original publication was designed for its own purpose, aimed at its own readership and carrying its own baggage of assumptions.

Sometimes a graph, map or table includes information that is irrelevant to the purposes of your publication.

It should be reworked to exclude this information.

Tables

When to use the tables

- when the data are too complex (trying to get several ideas across at once or there are too many levels or categories in one variable);
- when you need to show the data behind your graphs;
- when the actual values are important (graphs only convey relationships and orders of magnitude – sometimes you may need to show both graphs and tables for the same data);
- when a lot of data are missing or below the limit of detection;
- when labels of categorical data are too complex;
- when any graph you try to draw looks silly

Kind of tables (1)

In general we can divided the tables in two groups: Informal Table (in-text table); Formal Table

Informal Table

This kind of table is positioned within the text, immediately following the text that it supplements.

Aim

To explain or to simplify facts in the surrounding text

Characteristics

It cannot stand alone

- **Does not receive a table number or table title**
- **It is not listed in the table of contents**
- **It is short with few rows and data (word) columns (usually only one horizontal rule)**

Example Informal Table

Response Status	Number of Agency
Yes	15
No	3
Missing	3

Kind of tables (2)

Formal Table

This type of table features a minimum of three horizontal rules.

The first rule separates the table title from the boxhead (area that spans across the top of the table and that labels the different categories of the data appearing in the vertical columns)

The second rule encloses the unit of measurement

The third rule encloses the boxhead

The forth rule indicates the end of the table body and separates it from the footnotes, notes, and sources notes

For ease of reading table adding additional horizontal and vertical lines

Example Formal Table (1)

Table 3.1: Total waste generation by type of waste

Year	Municipal waste	Non Hazardous waste	Hazardous waste	Construction & Demolition[E]	TOTAL
	t*1.000				
2000	28.959	51.913	3.896	27.291	112.059
2001	29.409	55.090	4.269	30.954	119.721
2002	29.864	49.374	4.991	37.346	121.575
2003	30.034	52.366	5.419	42.548	130.367

Source: APAT
NOTE:
E= Estimated data

Kind of tables (3)

For very large tables it may be helpful to the reader if a blank, vertical line space is inserted between logical groups of data.

To make a table visually more comprehensive, some of the data rows may be shaded a light tint, or data rows may be separated with fine lines to help the eye travel across the whole page and to continue to the next data line without dropping to a wrong line or to wrong data.

When a formal table is longer than one page, the word "Continued" (in parentheses) is added at the end of each table and in the word "following" is added at the opening of the second part of the table.

Footnotes in a multipage table appear only on the last page of the table

Example Formal Table (2)

Table 3.1: Total waste generation by type of waste

Year	Municipal waste	Non Hazardous waste	Hazardous waste	Construction & Demolition	TOTAL
	t*1.000				
2000	28.959	51.913	3.896	27.291	112.059
2001	29.409	55.090	4.269	30.954	119.721
2002	29.864	49.374	4.991	37.346	121.575
2003	30.034	52.366	5.419	42.548	130.367

Table numbering

In most publications, tables (and graphs) are numbered independently and sequentially throughout a report, beginning with the Arabic number numeral 1.

In other cases section or chapter numbering (2.4, 5.6 etc) may be appropriate

Table Titles

A table title, always preceded by the word "Table", should be short and concise. It is important to include three elements:

- **What (data illustrated)**
- **Where (geographic area represented)**
- **When (data date)**

Abbreviations in Tables (1)

- **Blank cells: A cell (intersection of a data line with a column head) is never left blank. Blank cells may cause confusion in the editing and transcription process and may be misinterpreted by the reader.**
- **No data reported: Use a single dash or hyphen (-) in tables cells for which no data are reported. This category includes situations where the data element exists on the survey form but no data was reported for that cell.**

Table 3.1: Total waste generation by type of waste

Year	Municipal waste	Non Hazardous waste	Hazardous waste	Construction & Demolition	TOTAL
	t*1.000				
2000		51.913	3.896	27.291	112.059
2001	29.409	-	4.269	30.954	119.721
2002	29.864	49.374	4.991	--	121.575
2003	30.034	52.366	5.419	42.548	130.367

Abbreviations in Tables (2)

- **Not applicable: Use a double hyphen with a space inserted between the two dashes (- -) in table cells for which data are not applicable. This category includes situations where it is not possible for data to exist for that particular cell in the table.**
- **Not meaningful: When a calculation of a percent change would not be meaningful, use NM. NM should be defined in endnotes as "NM=Not meaningful."**

Table 3.1: Total waste generation by type of waste

Year	Municipal waste	Non Hazardous waste	Hazardous waste	Construction & Demolition	TOTAL
	t*1.000				
2000		51.913	3.896	27.291	112.059
2001	29.409	-	4.269	30.954	119.721
2002	NM	49.374	4.991	--	121.575
2003	30.034	52.366	5.419	42.548	130.367

Abbreviations in Tables (3)

- **Not available: Use the code "NA" in the data cell if data are not available. This category includes situations where the data element exists on the survey form but for some reason no data is available for that reported table cell.**
- **Preliminary data: When the entry is a preliminary figure, enter P to the left of the number, leaving no space between the symbol and the data. P should be defined in endnotes as "P=Preliminary data."**
- **Revised: When the entry has been revised, enter R to the left of the number, leaving no space. R should be defined in endnotes as "R=Revised data."**

Year	Municipal waste	Non Hazardous waste	Hazardous waste	Construction & Demolition	TOTAL
	t*1.000				
2000	R 28.959	51.913	3.896	27.291	112.059
2001	29.409	P 55.090	4.269	30.954	119.721
2002	29.864	49.374	4.991	**NA**	121.575

Abbreviations in Tables (4)

- **Withheld: When data entries have been withheld for proprietary reasons, enter the abbreviation W in the data cell. W should be defined in endnotes as "W=Data withheld to avoid disclosure."**
- **Rounded to zero: When the entry has been rounded to zero, the asterisk symbol (*) should be placed in the cell and defined in the endnotes as "*Number less than 0.5 rounded to zero." When the typeface is small or difficult to distinguish, (*) may be used. Consistency is important within a publication.**

Year	Municipal waste	Non Hazardous waste	Hazardous waste	Construction & Demolition	TOTAL
	t*1.000				
2000	*	51.913	3.896	27.291	112.059
2001	29.409	55.090	4.269	30.954	119.721
2002	29.864	49.374	4.991	W	121.575

Abbreviations in Tables (5)

Zero: Only when the entry is known to be exactly zero, a 0 should be entered in the cell. It is placed in the first whole number position, left of where the decimal would appear in the column. The use of 0.0 is not necessary, even though other entries are decimals. Do not use a dash (hyphen) or leave a blank to mean zero.

•Estimated data: When a number is an estimate, enter E to the left of the number, leaving no space between the symbol and the data. E should be defined in the endnotes as "E=Estimated data."

Year	Municipal waste	Non Hazardous waste	Hazardous waste	Construction & Demolition	TOTAL
	t*1.000				
2000	0	51.913	3.896	27.291	112.059
2001	29.409	55.090	4.269	30.954	119.721
2002	29.864	49.374	4.991	E37.346	121.575

Figure (Graphs and Maps)

A figure is more than ten thousand words

Graph can take the form of line charts, bar charts, scatter diagrams, dot charts or pie charts.
A map can show a graph in each region, colour code regions according to some statistic, show contours of a statistic, or colour or symbol code spot data on the points to which they apply.
It is possible for a map to do more than one oh these things, but that is not recommended

Drawing graphs

- Orientation
- Scales and Legends
- Colours
- Kind of Graphs
- Time series

Orientation

Get it the right way up

Line charts, scatter diagrams and most bar charts show how the y-axis (vertical axis) variable changed when the x-axis (horizontal axis) variable changed. If you get it the other way round, people will misunderstand it.

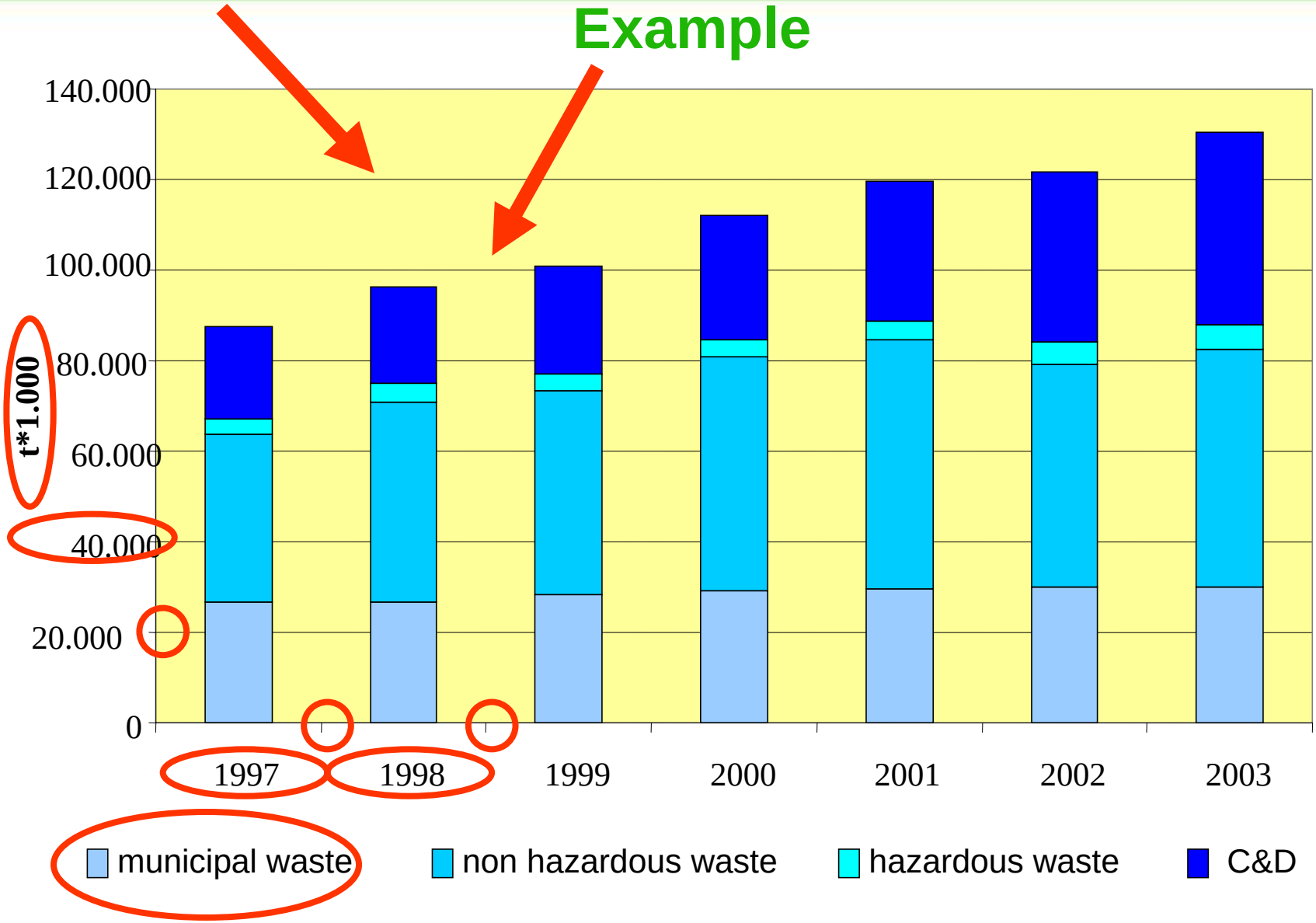
The only exception is a simple bar chart comparing the effects of one qualitative variable; if there are too many categories or their names are too long, you may find it easier to read if the bars are shown horizontally

Scales and Legends (1)

The scale of your graphs is indicated along the x- and y- axes by means of:

- **tick marks (little lines sticking out from the axis)**
- **numbers or category names by some of the tick marks**
- **grids (lines or dotted lines across or up and down the graph)**
- **explanatory legends (giving units etc) written above or along the axes**

Example


140.000
120.000
100.000
80.000
60.000
40.000
20.000
0
t*1.000
1997
1998
1999
2000
2001
2002
2003
municipal waste
non hazardous waste
hazardous waste
C&D

Scales and Legends (2)

Make scales and legends easy to read

The units used should be given in the axis legend

- **Decimal scaling**
- **Non-linear scales**
- **Ranges of data**

Colours

- **Colours must be informative**

  Properly used, colour can make your message easier to understand
- **Monochrome**

  is the cheapest method, you should make your patterns and shadings get denser as the values increases, but remember that black or dark patterns will tend to dominate the graph
- **Full colour**

  There is an obvious value in using colours that have a relevant meanings (eg red/black - grossly polluted; blue/green – clean; blue – cold; red – hot) Colour density (lighter colours – low numbers/less pollution; richer colours other case)
- **Colour for qualitative information**

  Colours should be simple, relevant and consistent. If there are groups of categories, use colour density to help bring this out

Graphs appropriate to the message (1)

The shape of a graph gives out a subliminal message of its own

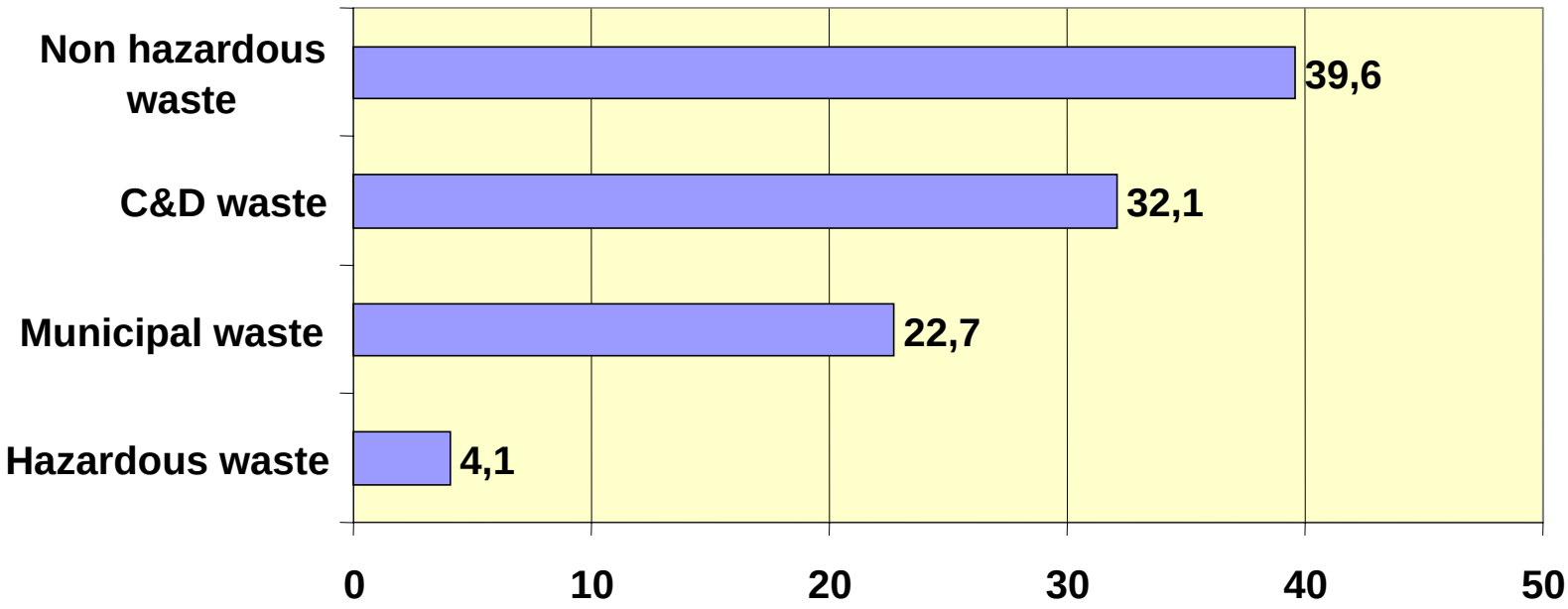
- **Bar charts**

They are most suitable for comparing amounts falling into different categories

There can be space between bars for amounts of different items such as emissions of different substances

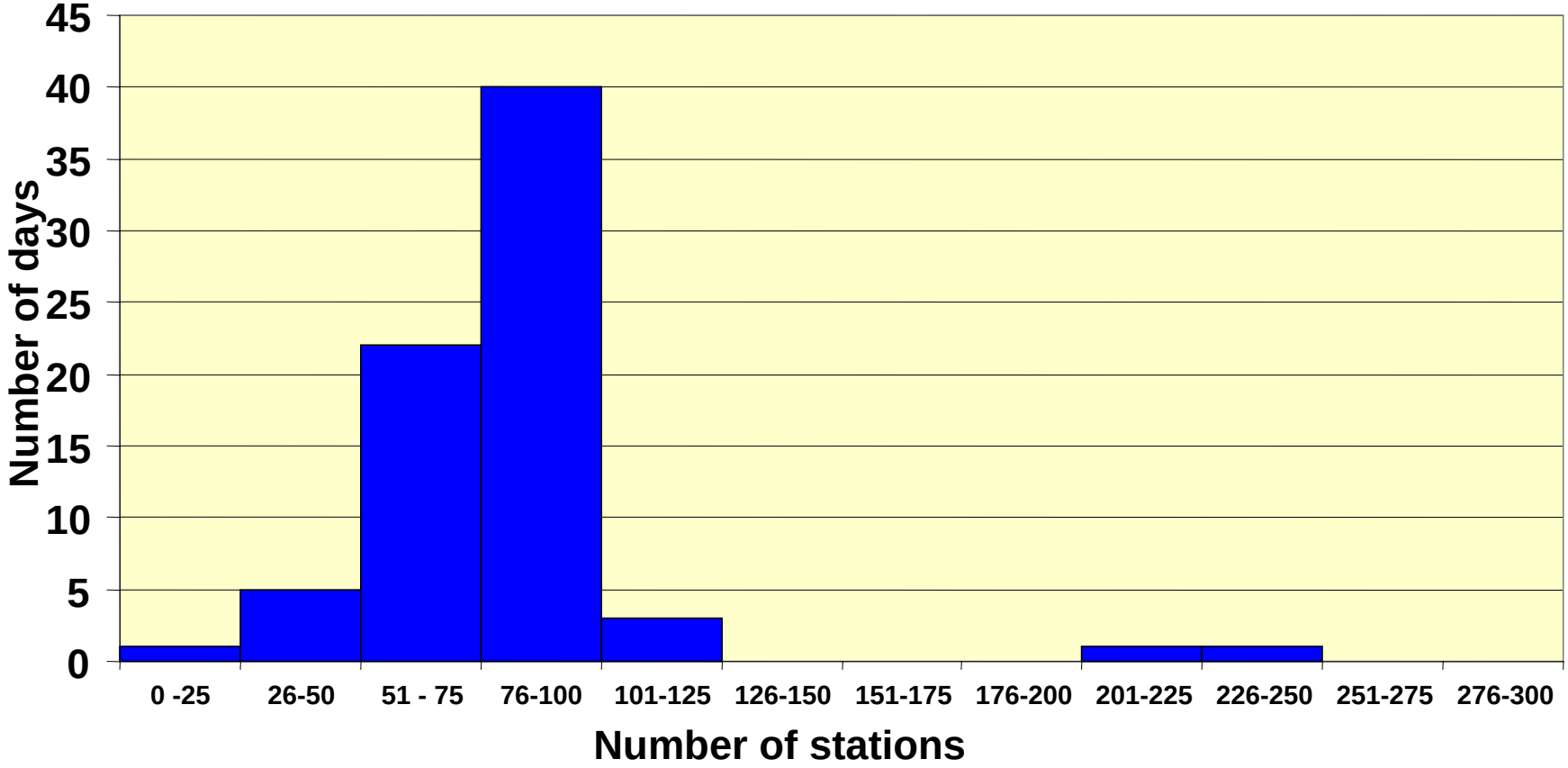
Bars can be adjacent for numbers of people exposed to given amounts of a pollutant (the gaps should be represent either zero or where are data missing) or for times series (but it is better to use a line chart)

Horizontal bar chart



The bar chart is drawn with the bars horizontal because the legends would be too crowded across the page under the axis. They are arranged in descending order, so that just reading down the descriptions gives information, without the reader even having to look at the far end of the bars.

Bar chart - Histogram



The x-axis is a continuous variable, so the bars are drawn touching each other

Graphs appropriate to the message (2)

- **Line charts**

Line chart can be used where a bar chart with adjacent bars can be used.

A bar chart is more appropriate when dividing the members of a sample into groups

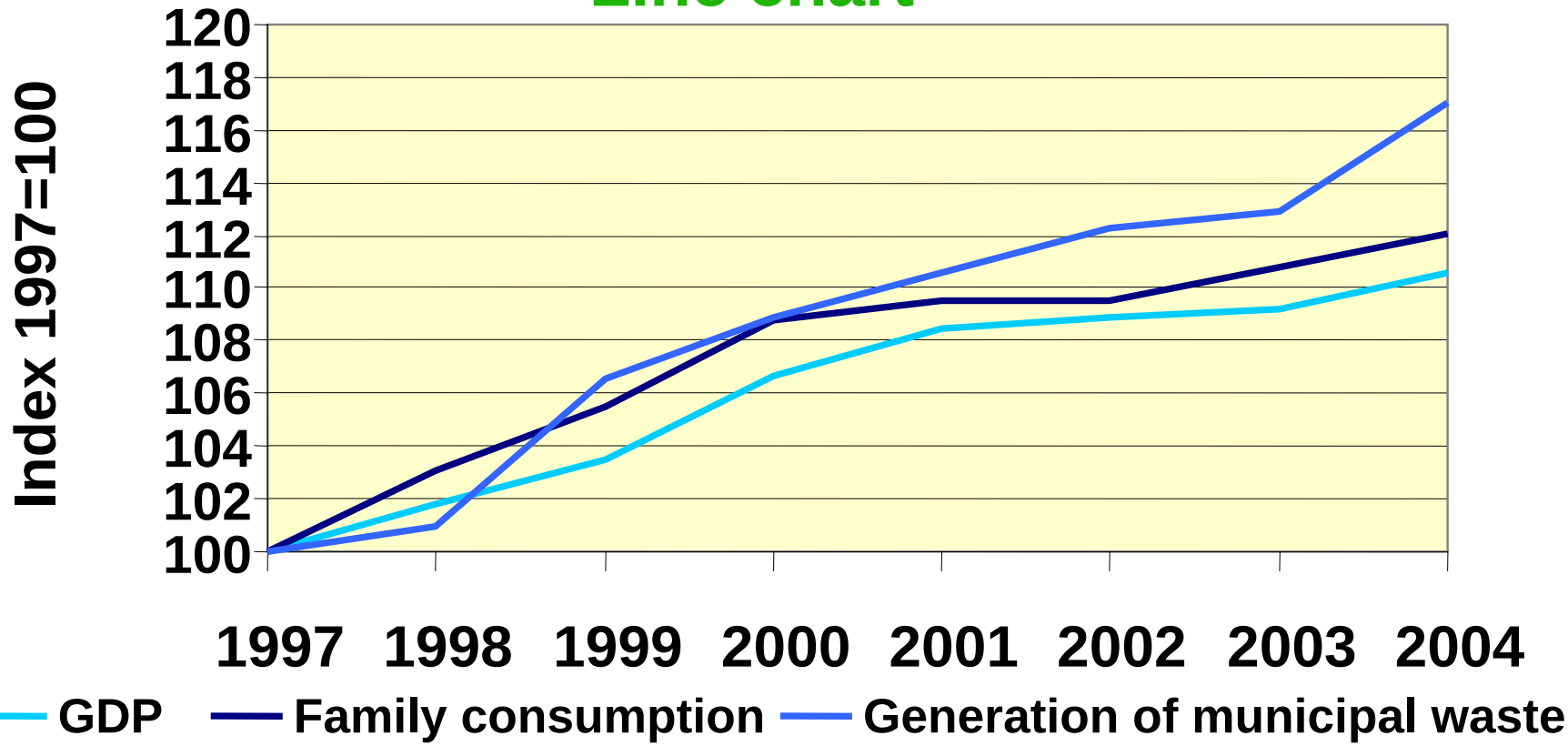
A line chart is better when illustrating gradual change, as in the time series.

A line chart can differentiate between missing data (a gap in the line) and a recorded zero (the line plunges to the baseline) whereas a bar chart cannot.

- **Multiple line charts**

These charts consist of more than one line chart sharing a zero. Draw the lines with different styles or colours. Remember that different line styles will have a different impact on the eye

Line chart



This time series also shows things that change gradually; so a line chart is more appropriate here as well

Graphs appropriate to the message (3)

- **Stacked bar and line charts**

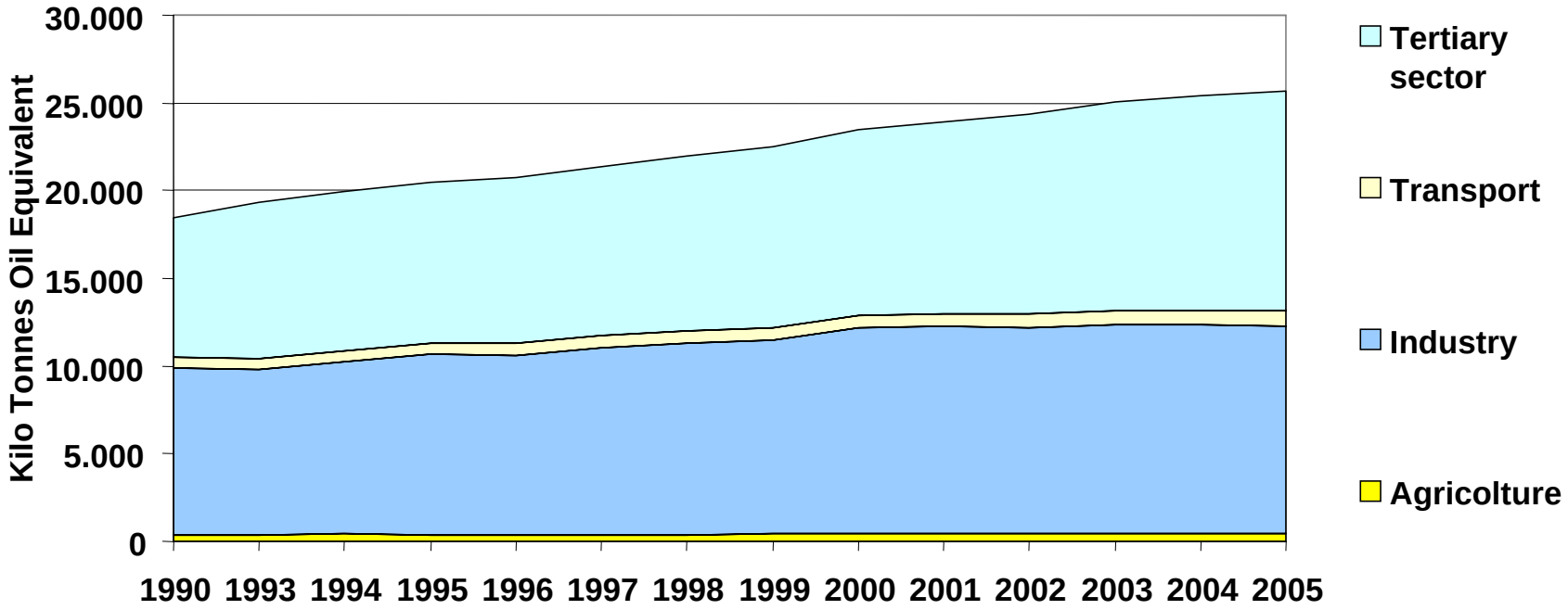
Stacked bars or lines imply the elements add up to something meaningful and important.

The emphasis is on the total and, perhaps, on the group at the bottom of the chart.

To colour or shade the spaces between the lines.

The key should show rectangles of the appropriate colours or shadings.

Stacked time series



This time series shows things that change, gradually, this is better represented by a line chart than by the unexpected changes of a bar chart.

Graphs appropriate to the message (4)

- **Multiple bar charts**

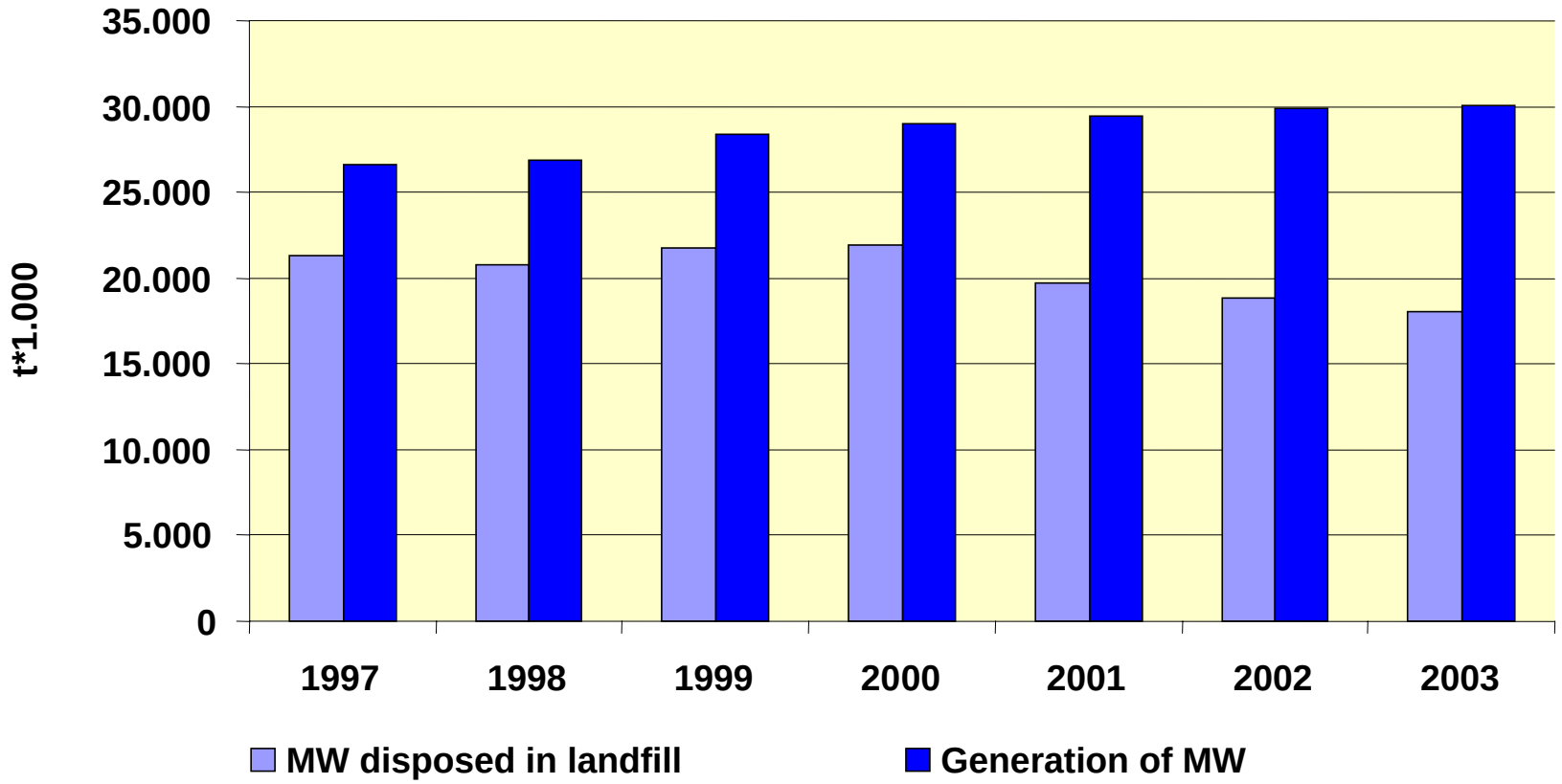
Multiple bar chart consist of repeated single bar charts arranged side by side or one above the other (if the bars are horizontal)

Separate the groups with larger spaces

These charts imply that the data do not add up

The key for a multiple bar chart will be written below the bars or in a separate area for a stacked bar chart

Multiple bar charts



The bars are shown as shades of blue to help give them equal weight

Graphs appropriate to the message (5)

- **Pie chart**

Pie charts imply that the elements add up to something meaningful but the relative proportions are what is interesting.

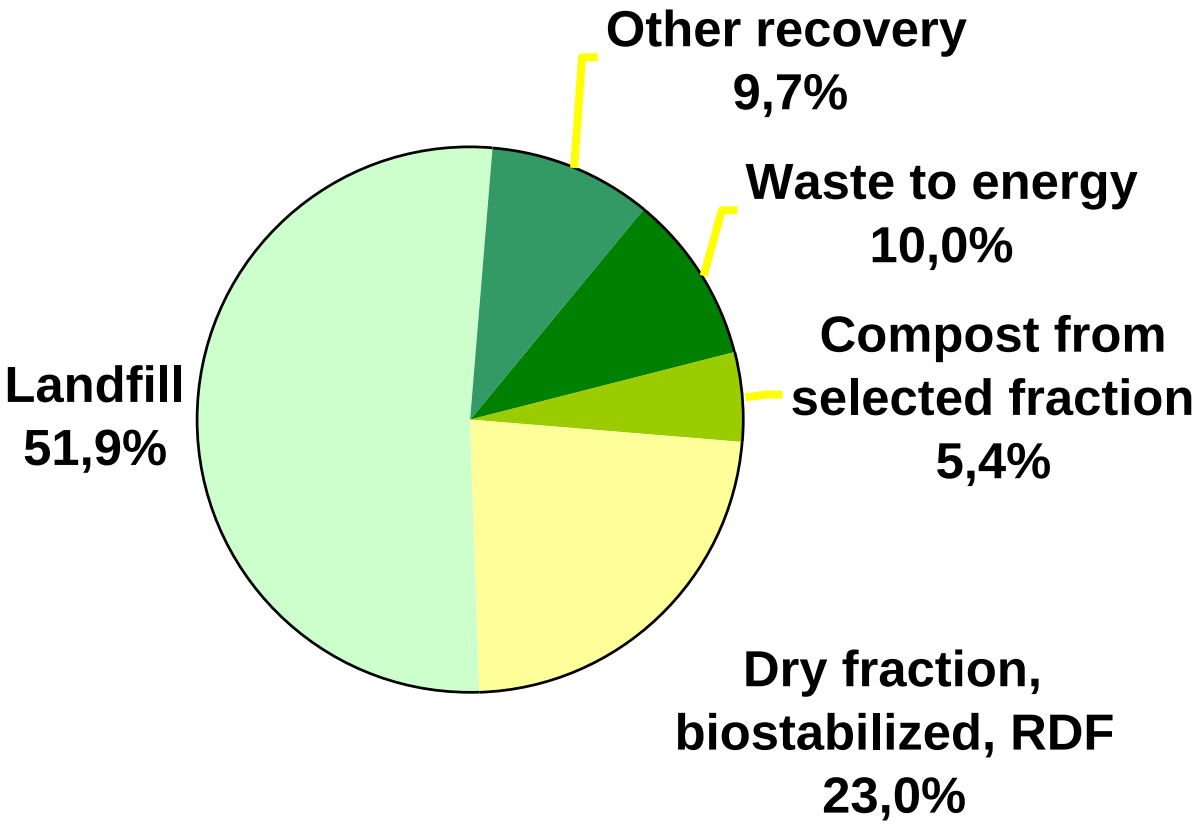
The reader will prefer to follow the segments clockwise round the pie and to have what the segments represent written against the segments themselves.

If that looks wrong, a separate key should show rectangles of the appropriate colour or shading.

Writing the amounts beside the segments of a pie is a good indicator that you should be considering some other way of presenting the information.

When comparing two or more pies make that the totals are represented by the area not the diameter of the pie

Pie charts



It is usually necessary to show either amounts or proportions represented by each slice of the pie

Graphs appropriate to the message (6)

- **Scatter Diagrams**

Scatter diagrams relate two variables by showing a point for each item of data on a two-dimensional area.

If you want to show the effect of one variable on another, the former (independent) variable should be on the x-axis and the latter (dependent) variable on the y-axis.

Mention the dependent variable first in the title.

If you show the grids, show both horizontal and vertical grids

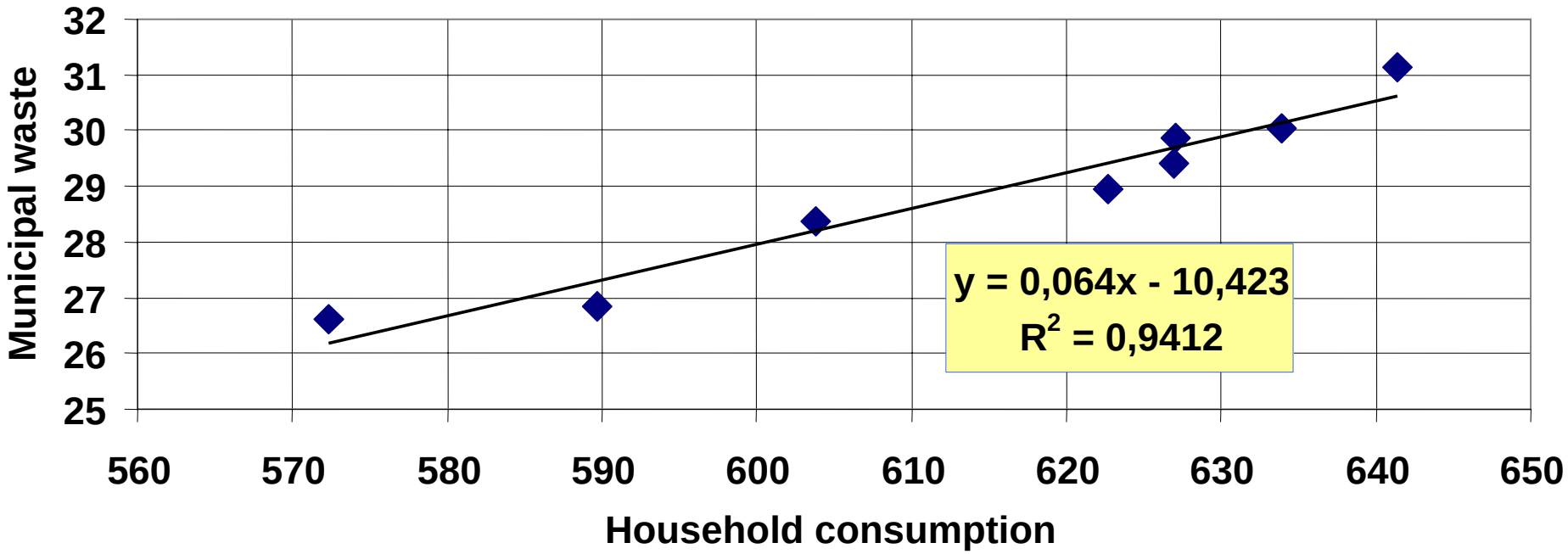
It is often helpful to show groups of data (eg different regions) by varying the colours or types of points, but be careful to make them have the same visual impact.

Yellow points are invisible

Explain point types in a separate key

Scatter Diagrams

Relationship between the municipal waste (dependent variable) and the household consumption (1997-2004)



The general shape of the points is the most important aspect of this graph, so the points stand out

Graphs appropriate to the message (7)

- **Box and whisker Plots**

If you have a range of y-values for each of a set of x-values, a scatter diagram can look silly, with too much information lost in a blot in the middle of each group.

It may be easier to show what the data mean if you use a box plot. That is a rectangle reaching from, say, the 5^{th} to the 95^{th} percentile, a line across the box at the mean (or median).

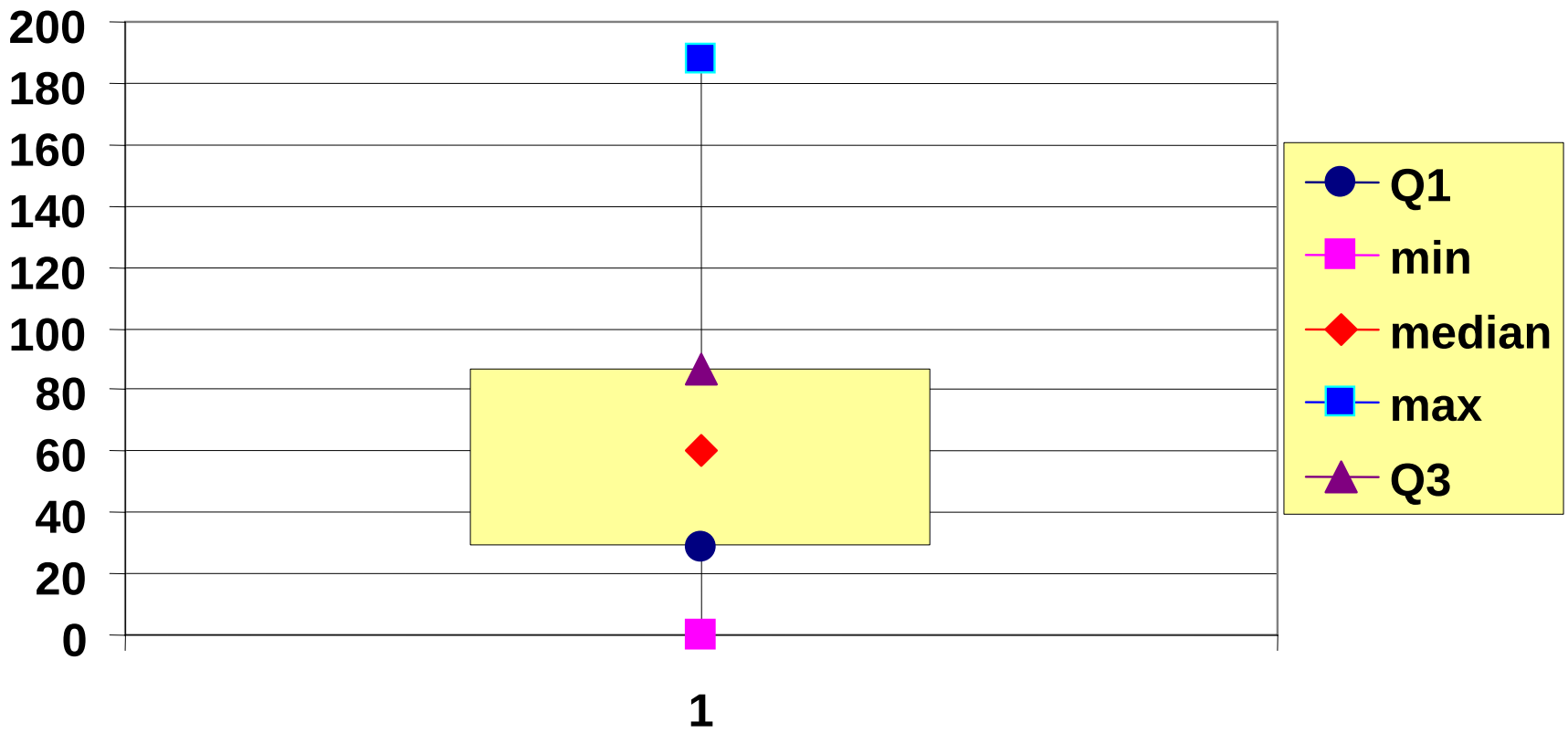
The ends of the vertical lines indicate the maximum and minimum data values.

The points outside the ends of the whiskers are outliers or suspected outliers.

The key should show an example box explaining each of the statistics included.

Box and whisker Plots

PM10 - Number of excedences days of daily limit value (50 mg/m3)



Time series

- Data in any time series should be the same in all respects except time. If they are not so, it affects the conclusions to be drawn and therefore needs explaining, along with the effect of such changes
- If the changes affect a single year, the data should be footnoted
- If they affect all subsequent years, draw a line on the graph and explain either in the axis legend or in a note
- It is often easier to omit seriously deviant years and explain that they were omitted because of lack of comparability

Maps

Special forms of graphs more suitable to represent the spatial variation of variables and indicators of environmental interest.

We have different kind of maps:

- **Pie charts on maps**
- **Bar charts on maps**
- **Colour coding or shading regions**
- **Time series on maps**

Maps

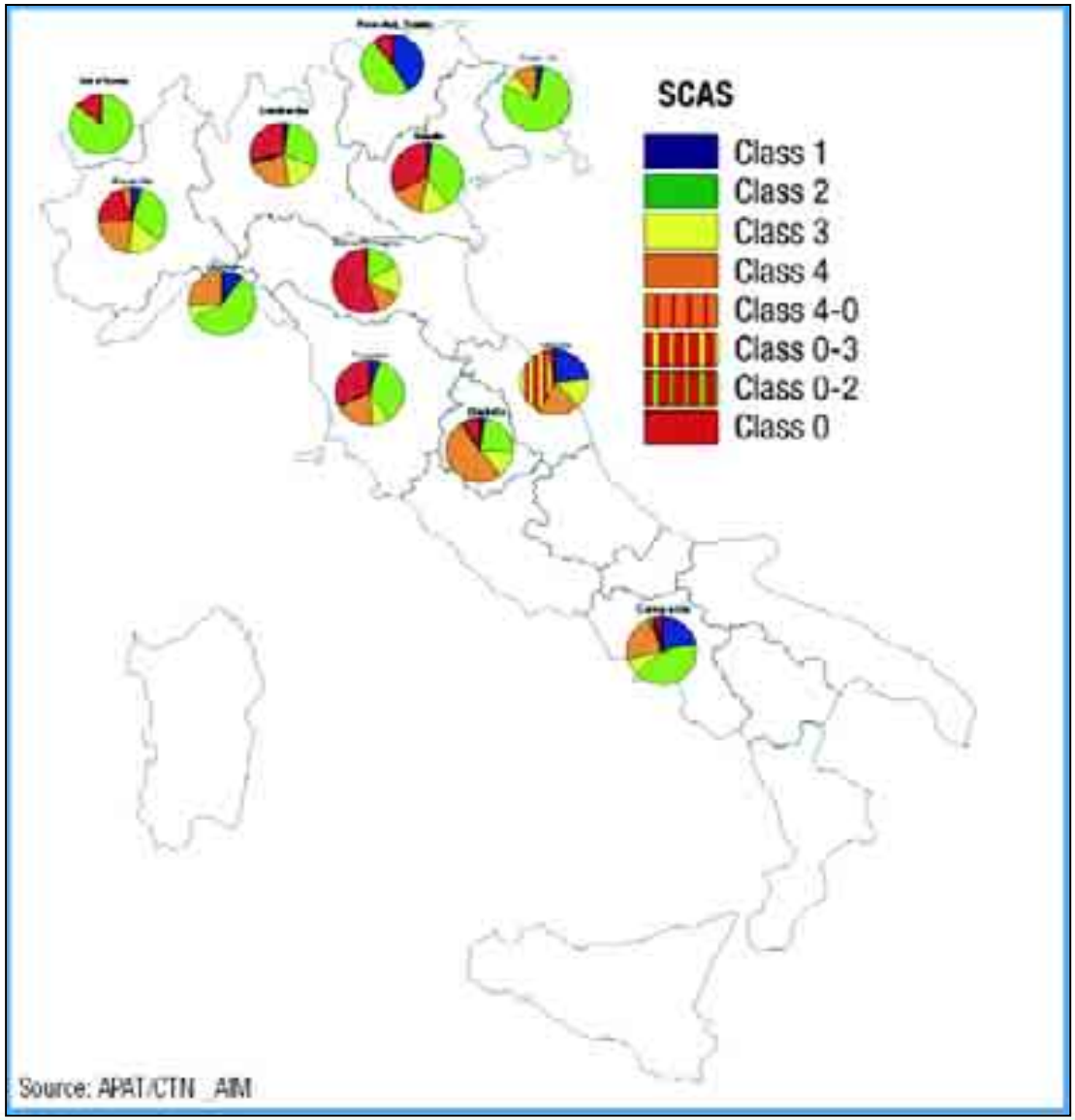
Regional maps will normally have a simple pie or bar chart in each region, or be colour or shading coded with respect to some variable.

Generally cartograms and cardiograms are used when the variability as regards administrative contests wants to be represented (municipal territories, provincial, etc).

The first obtains, usually filling with various colours the administrative borders of the chosen level; the second associating opportune graphs (linear or areal) to sayings border.

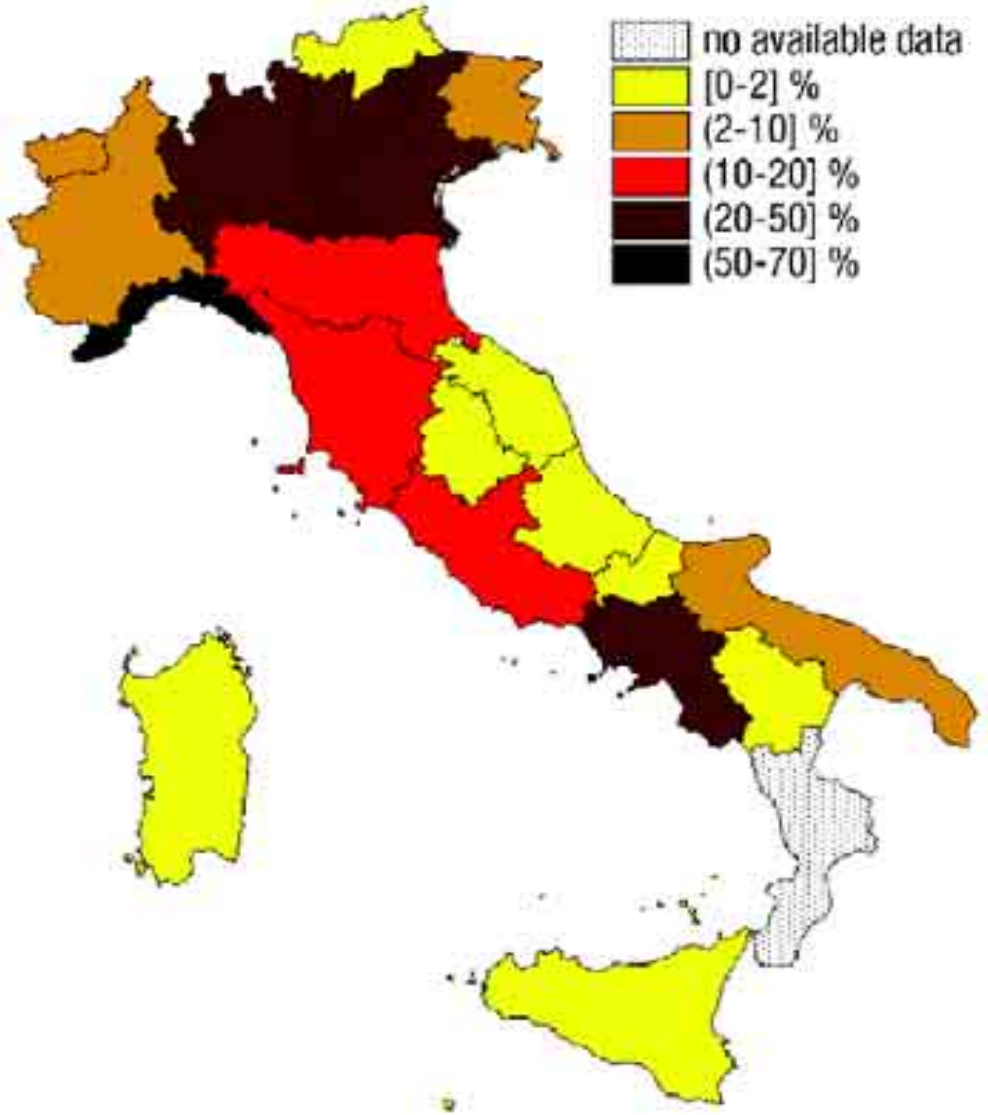
Pie charts

Chemical State of Underground Waters



Shading regions

Percentage of local authorities that have approved noise zoning guidelines (at 31/12/2003), compared to the total number of local authorities in each region/autonomous province



Symbolic representations

Pictograms do not have to be confused with the symbolic graphs, that even if with the approximations owed to the discrete value, objective data representations supply.

We use this kind of representation when we want to synthesise to the maximum the result of an analysis, that is the evaluation what of that result does the analyst.

Symbols which express the various judgement levels are generally used.

The number of levels of judgement more widely used for this type of representation is three: positive, negative, neutral.

Kind of Symbols

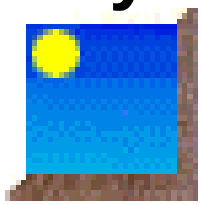
Chernoff Icons

The target will reasonably be achieved, based on the indicator trend

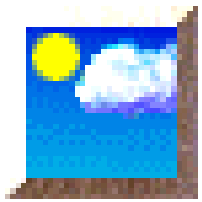
The indicator subject-matter is moving in the right direction, but the targets will hardly be achieved within the established timeframe

It's really distant from the target

Meteo symbols



Serene



Cloudy

Thunderstorm

Main References

- APAT (different years), *"Environmental data yearbook"*
- EUROSTAT (2000), *"Towards environmental pressure indicators for the EU," First Edition 1999.*
- Ebert U. and Welsch H. (2004), *"Meaningful environmental indices: a social choice approach"*, Journal of Environmental Economics and Management, vol. 47, pp. 270-283.
- Environmental Protection Agency (EPA), Council for Regulatory Environmental Modeling (CREM), *"Draft Guidance on the Development, Evaluation, and Application of Regulatory Environmental Models"*
- Environment Agency, England and Wales (1998), *"Presentation of Statistics"*
- Energy Information Administration, *"EIA Publishing Style Guide"*