I. Analysis of the information

A rigorous definition of the components of the information, specifying their number, length, and level, must precede any graphic construction.

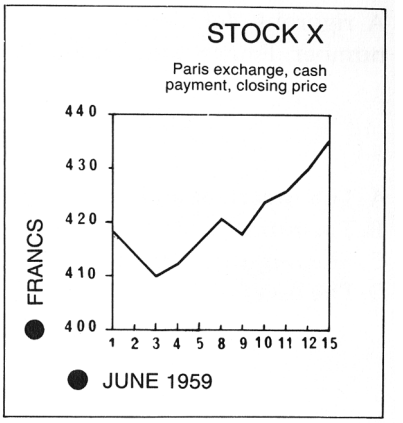
A. The invariant and the components
B. The number of components
C. The length of the components
D. The level of organization of the components

A. The invariant and the components

DEFINITION

Information is a series of correspondences observed within a finite set of variational concepts of "components." All the correspondences must relate to an invariable common ground, which we will term the "invariant." A precise analysis of these terms is the only means:
- to understand complex information
- to determine the best graphic representation of it
- to word its title and legend.

The following examples will enable us to specify the nature of the two concepts: invariant and components.



1

EXAMPLES

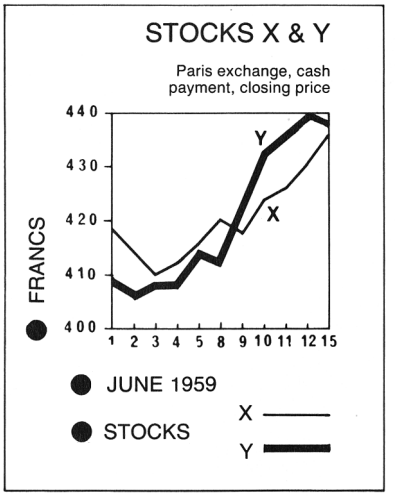
Example 1: The trend of stock X on the Paris exchange.

The INVARIANT is the complete and invariable notion common to all the data. In figure 1, it is the "quotation in new francs for stock X, cash payment, closing price on the Paris exchange."

Indeed, we cannot mix together within this information, cash and time payments, old and new francs, stocks X and Y, the London and Paris exchanges.

The COMPONENTS are the variational concepts. In this case the variational concepts are:
- quantities (of francs)
- time (in days)

The information has two components, and the drawing must thus utilize two visual variables: the two dimensions of the plane (figure 1).



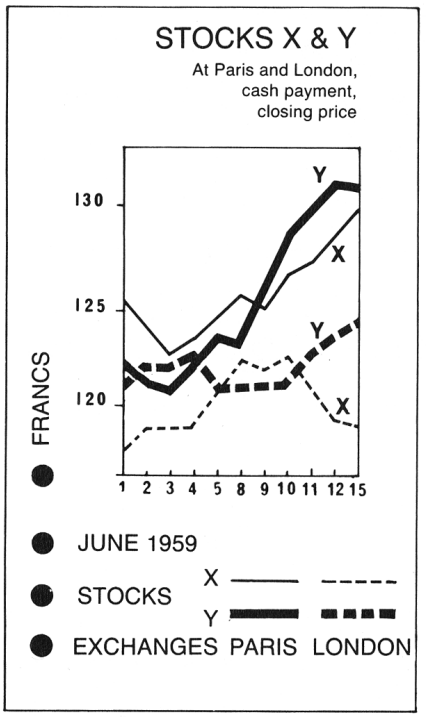
2

Example 2: Comparison of the trends of stocks X and Y.

INVARIANT –*quotation in francs, cash payment, closing price, on the Paris exchange*

COMPONENTS –*quantities of francs, according to*
–*the date*
–*different stocks (X and Y)*

This information has three components; the drawing must employ three visual variables. A size variation can be used to distinguish X from Y, as in figure 2.



3

Example 3: Comparison of the trends of stocks X and Y in London and Paris.

INVARIANT –*quotation, cash payment, closing price*
COMPONENTS –*quantities (indexed), according to*
–*the date*
–*different stocks (X and Y)*
–*different cities (London, Paris)*

Here the information has four components, and the drawing must employ four visual variables. In figure 3, for example, texture is used in addition to the three preceding variables.

Note that the definition of the invariant simplifies as the number of components increases.

In these three cases, the information is sufficiently familiar to be designated by a specific word: TREND. This word can serve as a title, since it summarizes the informational situation, and its immediate comprehension takes the place of a more lengthy analysis. However, modern information often correlates components of great diversity, which means that the invariant will be less familiar, its identification more difficult.

Example 4: Population residing in the Paris area, by department of birth (not including departments constituting the Paris area), given in absolute quantities.

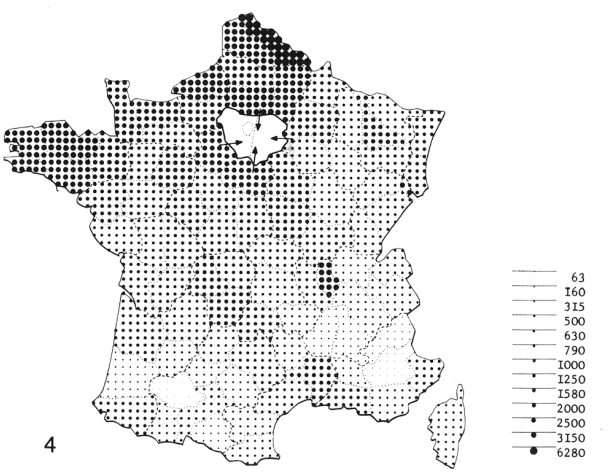
Example 5: Distribution per 100 persons born outside of the Paris area but residing there in 1962, according to their department of birth. There is no single word to capsulize either of these examples; there is no known title, since the situations are too unusual. A thorough analysis becomes imperative if the designer is to understand what must be expressed and the reader to understand what is being represented; if not, each risks serious error. For instance, what is the difference between the data in examples 4 and 5?

Further analysis reveals that there is none. These are simply two verbal formulae which express the same content, as shown in figure 4. Both cases are, in fact, constituted by:

INVARIANT –*a person, living in the Paris area, born in the provinces and counted in the department of birth*
COMPONENTS –*population according to*
–*departments*

The data are simply expressed in one case by the observed numbers, and the total corresponds to the number of observations (3 034 700); in the other case by numbers calculated per 100. All the numbers from the first example have been multiplied by the fraction 100/3 034 700 to produce the second. This merely involves changing the numerical scale, which in no way modifies the observed correspondences.

In such cases, it is only by seeking the precise definition of the invariant and the components that we can come to understand the true nature of the information.



4

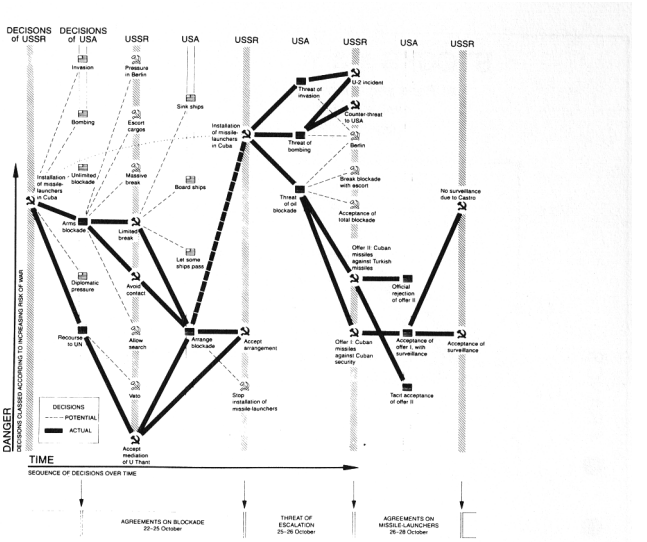
Example 6: The Cuban Missile Crisis—Principal elements of decision-making during the missile crisis of 1962 (see page 264). This is a nonquantitative problem; analysis alone permits a clear representation of the information.

INVARIANT *–a decision at the summit (made by a chief of state). The decisions are differentiated according to*

COMPONENTS *–nationality (American or Russian)*
– (potential or actual)
–danger (increasing risk of war)
–date (made on such and such a date)
–nature (of such and such a nature)

This information has five components, necessitating at least five visual variables. It cannot be perceived in its totality in one immediate image, as we note from observing figure 1.

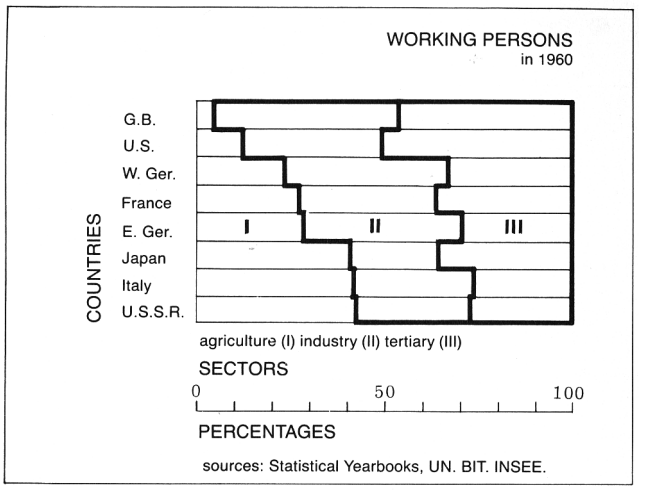
No word exists for identifying the second component; certain components are characterized only by listing their categories. Nor does any word exist for the whole of the information. The "title" is merely a paraphrase which orients the reader toward the subject but does not specify which components are involved. We are far from the word "trend" in the first examples.



1

THE ORDER OF THE COMPONENTS

When the data involve percentages, or when we decide to derive percentages from the raw data, this must be reflected in the analytic description. For example, in comparing the extent of the three main sectors of the work force in several countries:



2

Figure 2:

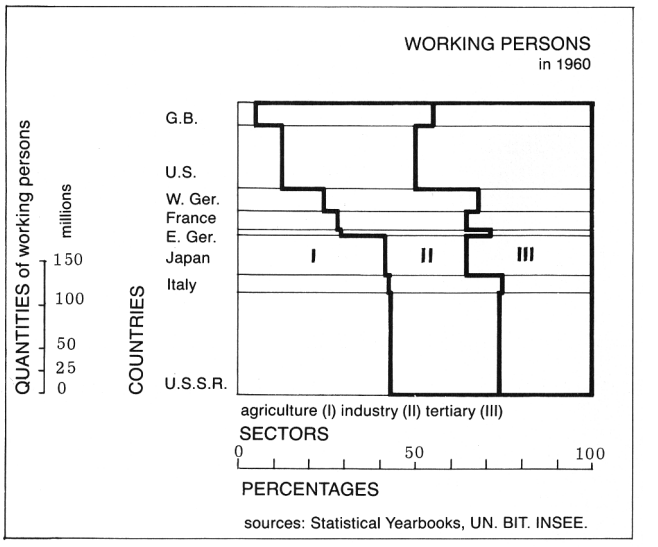
INVARIANT *–working persons (1960)*

COMPONENTS *–different countries*
–number (Q) per 100 working persons per country according to
–three main employment sectors

In this case, all the countries are considered as similar and equal to 100.

This situation is reflected in the analytic description by placing those components not affected by the quantities first, and by having the quantities followed by only those components which they govern.

This rule is also exemplified in the following analytic description:



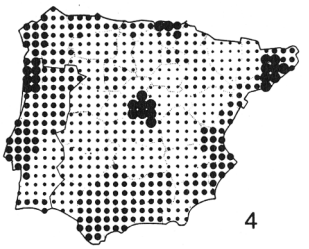
3

Figure 3:

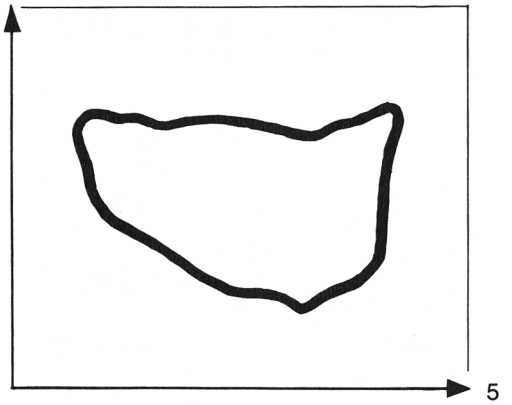
INVARIANT *–working persons (1960)*

COMPONENTS *–absolute Q according to*
–different countries
–Q per 100 working persons per country according to
–different employment sectors

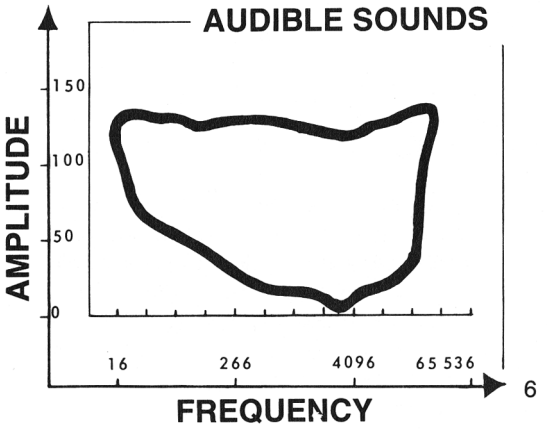
The first application of the notion of components is the following:



4



5



6

Wording of titles and legends

The two terms, "title" and "legend," are sometimes used interchangeably. A "legend," placed under a figure, will function as a title, while specialists will often seek the real "title" of a map in what some designers call the "legend."

Such confusion must be avoided. The headings (legend and title material) in a graphic image have two functions:

(1) to permit the reader to identify, *in the mind*, the invariant and components involved. This can be called EXTERNAL IDENTIFICATION, in the sense that it is independent from the graphic image itself.

(2) to identify, *in the drawing*, the visual variables corresponding to the components. This will be referred to as INTERNAL IDENTIFICATION.

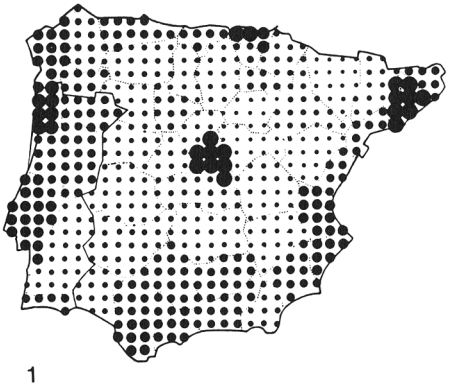
EXTERNAL IDENTIFICATION

External identification is independent of the graphic representation, because the drawing, in itself, cannot furnish all the elements necessary for identification. Figures 4 and 5 are not identifiable. Written or oral statements are indispensable for their external identification. To identify the content of figure 5 we must know

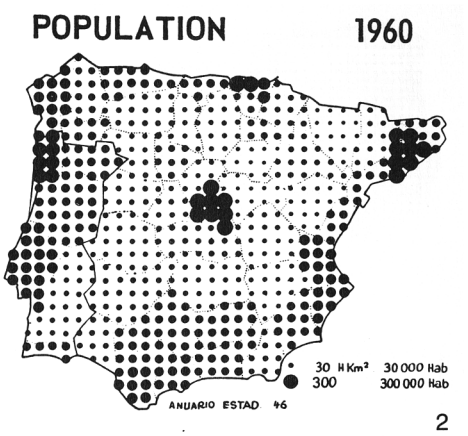
- the invariant: audible sounds
- the first component: the frequency of the sound (cycles per second)
- the second component: the amplitude of the sound (auditory level in decibels).

These factors are clearly identified in figure 6.

One cannot study a graphic intelligently without knowing the invariant and the components displayed in it. Titling a drawing speeds the acquisition of this knowledge and dispels potential ambiguity.



1



2

Identification of the components

In numerous cases, however, identification of the components can result from the drawing itself. For a given audience, figure 1 clearly involves:

– a map of the Iberian peninsula: a geographic component
– quantities: a quantitative component.

All that remains is to specify the invariant, as in figure 2.

In certain cases the drawing itself can provide the means of identifying the components due to the familiarity of the subject.

For the same reason, a simple expression describing the invariant can sometimes be sufficient to define the components. Terms such as trend, price, temperature of X, barometric pressure, can define the two components of the diagram as well. The reader can understand these verbal cues because the images involved are of relatively current and common usage.

But modern scientific research multiplies the occurrence of innumerable combinations which cannot be named concisely or where the specific term employed is familiar only to a limited number of individuals. The designer can thus encounter three types of cases:

– common figurations, where a simple expression suggests the components or where the image itself is sufficient (particularly in cartography) to make them recognizable.
– new combinations, involving recent terms which are not very familiar to the average reader (agroclimatic diagrams, concentration curves, "stemmas" [see page 279]…). Here, the term serves its purpose only for a limited number of "experts" in the given field.
– new combinations, which have no precise term to characterize them.

In the last two cases, a written description of the components is indispensable to external identification (see figure 6 on the preceding page). Thus:

In most cases, the written designation of the various components is included in the title.

Identification of the invariant

An image (e.g., figure 1) or a term (e.g., "trend") can serve to define the two components. But quantity of what? Trend of what?

In all cases, a written term is needed to define the invariant.

Trend of what? Of stock X, cash payment at closing price, in new francs, on the Paris exchange.

This involves category X of the component "different stocks."

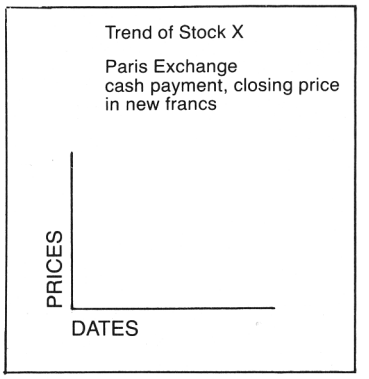
It involves the category "cash payment" of the component "different types of purchase."

It involves the category "Paris" of the component "different cities." . . .

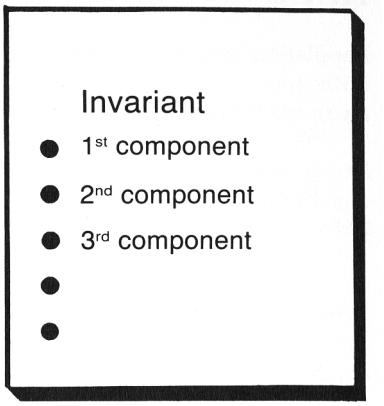
The invariant specifies the common ground for different components that extend throughout the larger informational set being investigated or the work being consulted.

It can therefore be worded as a function of the information immediately related to that being considered. Thus, in a grouping of trends "on the Paris exchange," this term need not be repeated for each diagram; in a chapter on "cash payments" this term could be eliminated. . . .

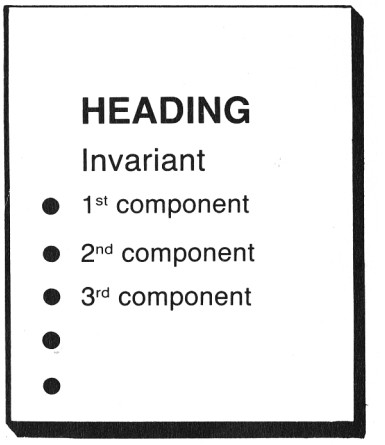
But omissions are often dangerous, since graphic information ought to be detachable from its immediate context in order to be related to all data having a common element. This is the case in analytic documentation.



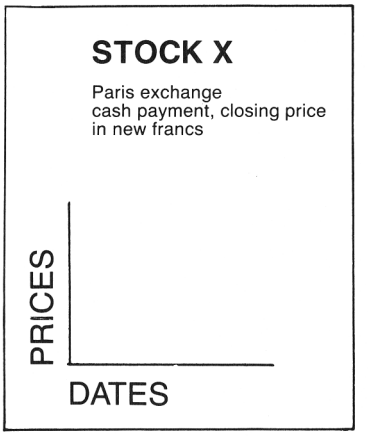
3



4



5



6

External identification: title composition

Consequently, in most cases external identification consists of writing, in visible characters and in a standard arrangement, (a) the invariant, and (b) all the components of the information. This is illustrated in figure 3.

The title will generally be set up as in figure 4, the order of the components following the rule outlined on page 18.

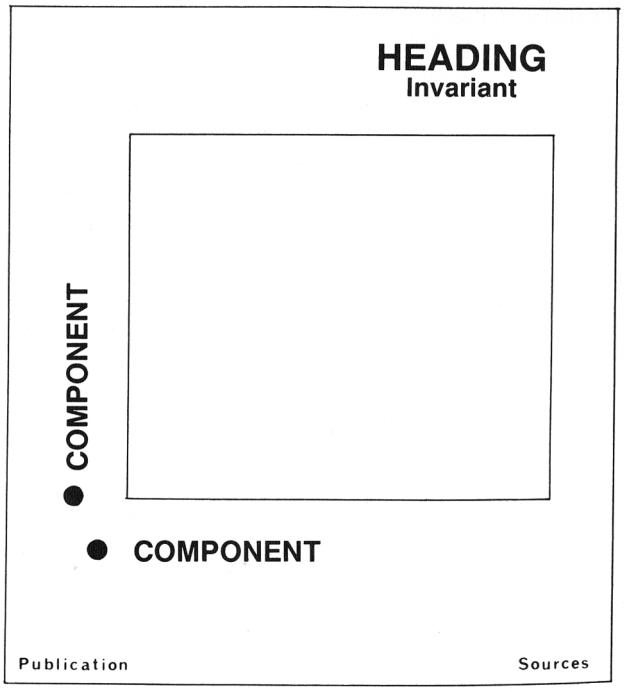
But the rigor and precision of this formula entails a relatively long title, which no longer fulfills the conditions of brevity and rapid comprehension, particularly necessary when studying a large number of drawings.

In order to avoid permanent omissions, it would seem that the most logical solution, the one meeting all the requisite conditions of external identification, is:
(1) to word the title according to the general formula outlined in figure 4;
(2) to cap it with a heading (see figure 5) whose wording depends on the larger informational set.

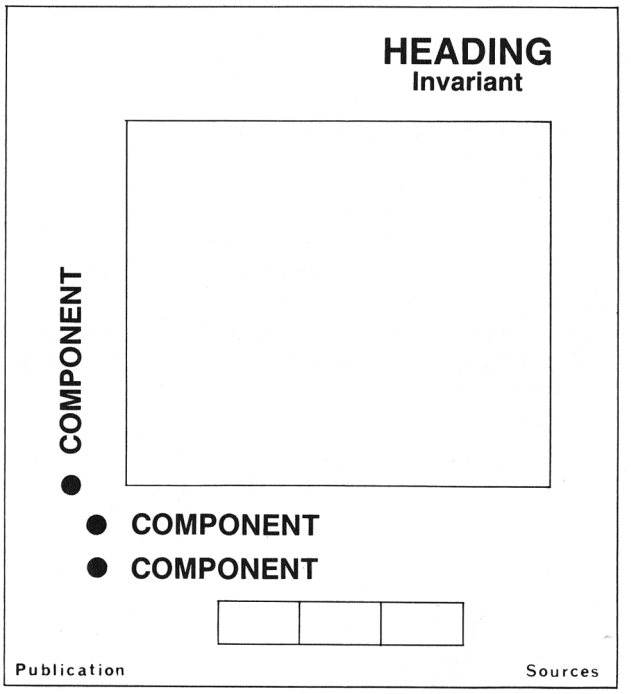
As we see in figure 6, the heading generally includes the name of the category (stock X) which belongs to the next higher informational set (different stocks). This solution affords the means of immediately perceiving the unique features of each representation, or, for the reader already acquainted with the documents, the means of rapidly locating a given representation. The following examples apply these principles and illustrate to what degree the current use of "title phrases" can be confusing. Indeed, there are a large number of title phrases which could fit a given set of information, while a given phrase could also apply to very different data.

TYPICAL TITLE PHRASES	SUGGESTED TITLES
Population residing in the Paris area, by department of birth (not including departments constituting the Paris area). Quantities in thousands.	MIGRATION TO PARIS Residents of the Paris area born in the provinces – absolute quantities according to – department of birth
Percentage distribution of persons born outside the Paris area but residing there in 1962, according to the department of birth.	
Number of persons residing in the Paris area (in 1962) per 100 persons born in each department (not including the departments of the Paris area).	RATE OF MIGRATION TO PARIS Residents of the Paris area born in the provinces – by department of birth – Q per 100 persons born in the department
Percentage of population living in rural communes [administrative subdivision of France] where 20–39.9% of the population is agricultural.	RURAL COMMUNES, 20–40% AGRICULTURAL Population living in rural communes where 20–39.9% of the population is agricultural – by department – Q per 100 persons living in all rural communes
Map of the distribution by ward of Parisian parents of students at the École Polytechnique	STUDENTS OF THE ÉCOLE POLYTECHNIQUE Residence of Parisian parents of students at the Ecole Polytechnique – quantity – by ward
Principal elements of decision-making during the Cuban "hot" crisis in 1962.	THE CUBAN CRISIS (1962) Decisions at the summit (made by chiefs of state) according to – whether the decisions were actual or only potential – date – danger of war – nationality (American or Russian) – nature of decision

TYPICAL TITLE PHRASES	POSSIBLE TITLES
Percentage of students enrolled in seventh grade, in private schools, by canton [administrative subdivision of France].	PRIVATE SCHOOLING IN THE SEVENTH GRADE Students enrolled in seventh grade, private schools – by canton – Q per 100 students in seventh grade, all schools SEVENTH GRADERS IN PRIVATE SCHOOLS Students enrolled in seventh grade, private schools – by canton – Q per 100 students enrolled in private schools
Distribution of the three main sectors (agriculture, industry, tertiary [other] of the work force, in percentage, by department.	MAIN SECTORS OF THE WORK FORCE Work force in 1954 – by department – Q per 100 working persons according to – three main sectors (agriculture, industry, tertiary) MAIN SECTORS OF THE WORK FORCE Work force in 1954 – by department – Q per 100 inhabitants according to – three main sectors (agriculture, industry, tertiary)
Changes in the population, aged 20 to 64, in France, between 1954 and 1962.	GROWTH OF ADULT POPULATION (1954–1962) Population aged 20 to 64. Difference between 1954 and 1962 – in absolute quantities – by department EVOLUTION OF ADULT POPULATION (1954–1962) Population aged 20 to 64 in 1962 – by department – Q per 100 persons aged 20 to 64 in 1954 VARIATION IN THE PROPORTION OF ADULTS (1954–1962) Difference between the 1954 percentage (Q of population aged 20 to 64 per 100 inhabitants) and the 1962 percentage – in quantities – by canton



1



2

INTERNAL IDENTIFICATION

Having identified the invariant and the components, the reader must still recognize which visual variables are being used to represent these components.

An "ombrothermal" diagram becomes meaningful when it is known that this word signifies:

INVARIANT *–planted areas (for a given type of vegetation) according to*

COMPONENTS *–the precipitation recorded in these areas (annual total)*
–the temperature recorded in these areas (annual average)

But we must still be able to recognize that on the drawing precipitation increases as we read, say, from bottom to top (ordinate), while temperatures increase from left to right (abscissa).

Constructions involving two components

Each of the two planar dimensions must be named.

To avoid repeating terms in the diagram, external and internal identification of the components can be combined. This leads to the standard arrangement displayed in figure 1. However, a different situation arises when there is a third component, and, in cartography, where the geographic order occupies the two planar dimensions.

Constructions involving more than two components

When three or more components are involved, we must utilize "retinal" variations: variations in the size of points or lines, in the value or color of the marks, etc. These variations are independent of position on the plane.

For each retinal variable, we must therefore draw a standard variation, indicate the name of the component which it represents, and relate its "steps" to the categories of the component. This is what we mean by the "legend."

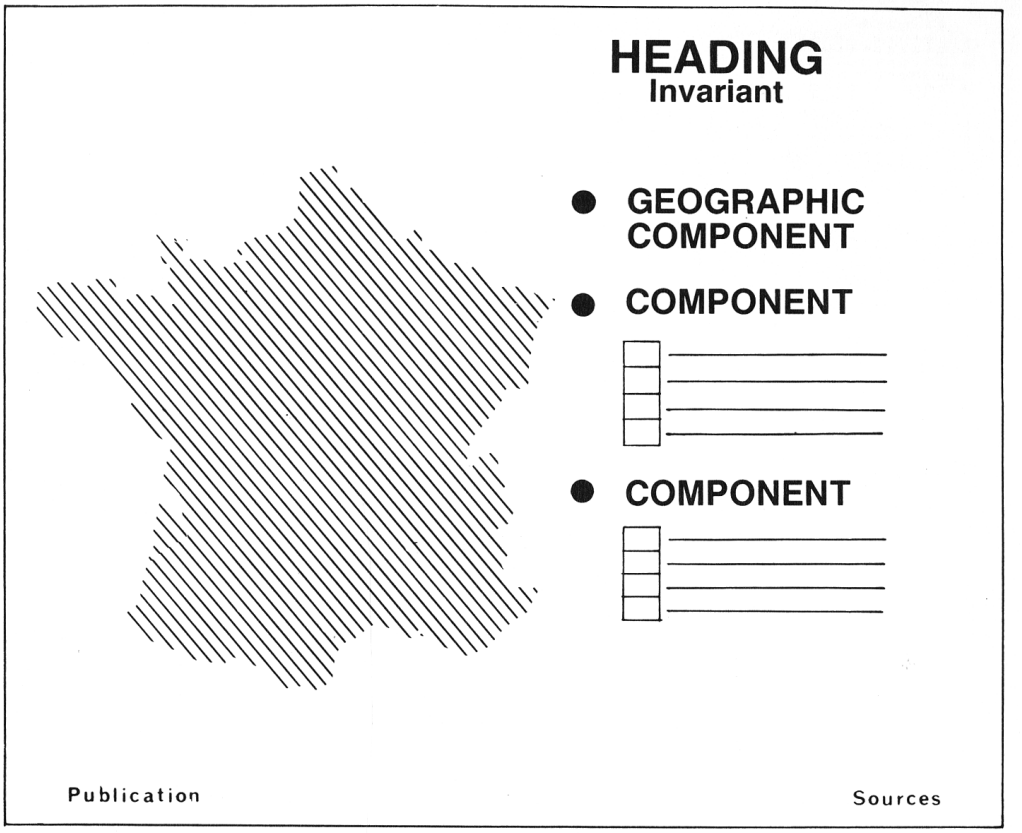
To avoid repetition, we arrive at the standard arrangement depicted in figure 2.

Maps

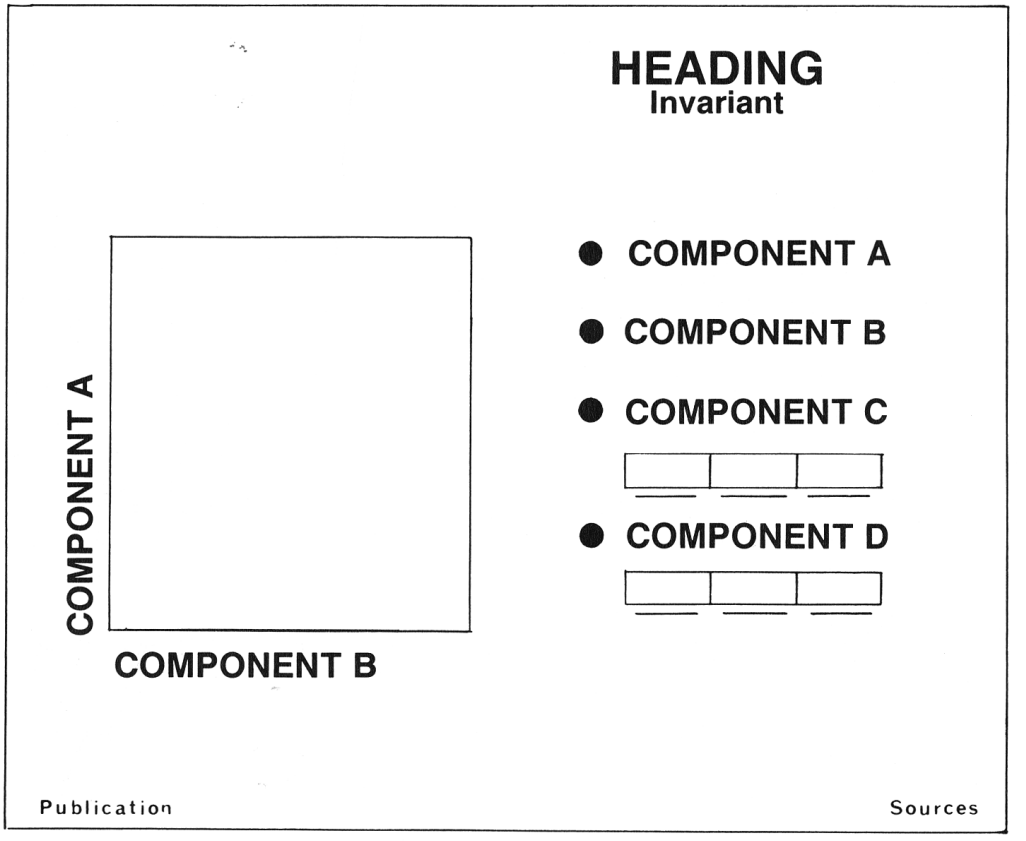
Here, two planar dimensions are occupied by the geographic entity, but any internal geographic categorization must still be specified. The enumeration areas can be departments, "agricultural zones," wards, cantons, communes, etc.

Furthermore, a map often involves several additional components.

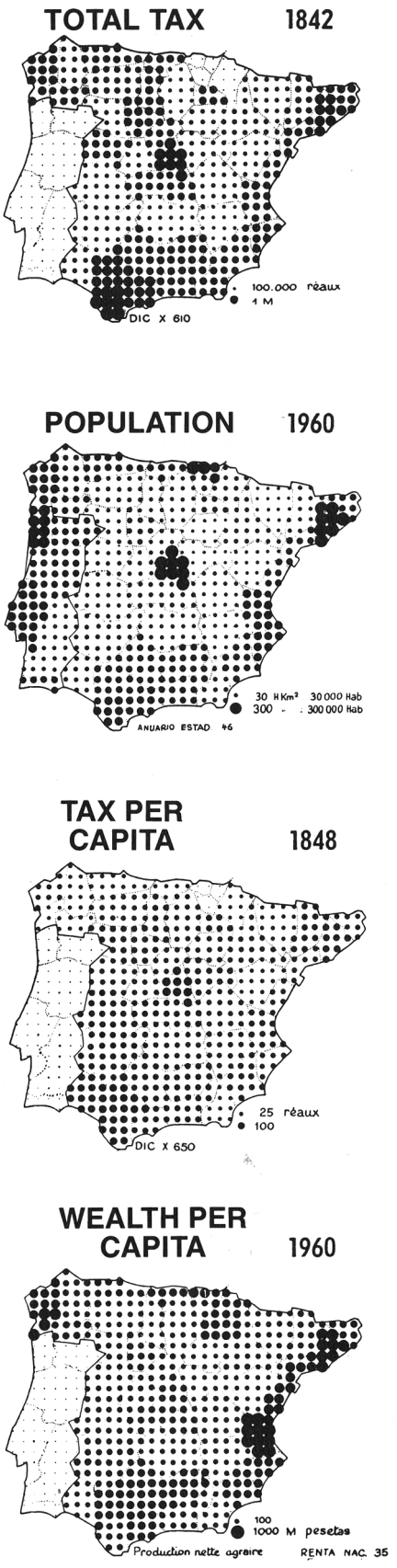
This results in the standard arrangement shown in figure 3. It can also serve as a model for those diagrams in which it is useful to group all the elements of identification (provided that the terms affiliated with the planar dimensions are repeated, as in figure 4).



3



4



1

IDENTIFICATION OF A HOMOGENEOUS SERIES

In a homogeneous series such as the collection of information on Spain portrayed in figure 1 (see also page 398) or the series of phonic analyses of folk songs in figure 2 (see also page 267), each map or diagram is identified by the invariant. It specifies a particular category from each of the components involved in the entire series of images:

– the category "1848" of the component "time"
– the category "population" in the component "different studies"

or, with the folk songs,

– the category "United States" from the geographic component
– the category "ballad" of the component "type"
– the category "first stanza" of the component "stanza and refrain."

The information itself is constituted by the entire series of maps or the whole collection of diagrams. The components extending throughout the series are constituents of the information, even though they are transcribed only by the writing, not by the images themselves. These components are the basis for the operation of classing and grouping, which constitutes the real objective of such representations. It is therefore important that:

– **The components extending throughout a homogeneous series are always transcribed in the same place.**
– **The categories of these components have maximum visibility and, if possible, are written in bold face and CAPITALS.**

Although these recommendations may appear relatively unimportant, they are nonetheless essential. The investigator who knows how to utilize all the properties of graphic representation and to derive the most from information processing will follow these rules and reserve sufficient space for an organized and efficient identification.

IDENTIFICATION OF SOURCES

From the perspective of generalized documentation, each graphic must be capable of being extracted from its original source and incorporated into any other set of information. However, the reader must be able easily to find the documentary sources and identify the original author (an expert whose field is similar or related to that of the reader).

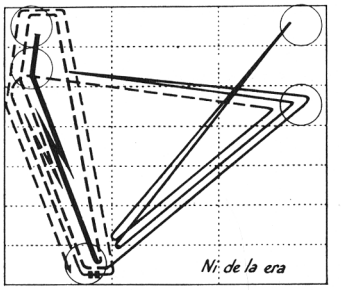
As a general rule, in or under each image or homogeneous group of images, the source, author, work, publisher, place of publication, and date should be indicated in such a way that these elements can be photographed with the drawing, as in figure 3.

COUNTRY	Region or Language	Type	Stanza or Refrain

2

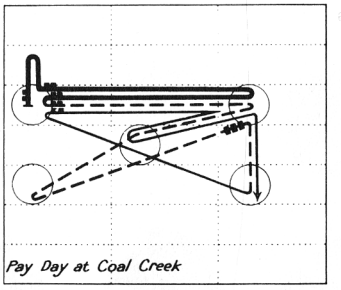
ESPAGNE ENV. SEVILLE CHT DE LABOUR S 1

3



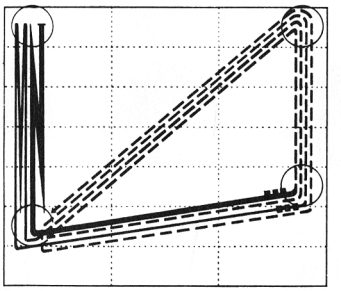
ETATS-UNIS KENTUCKY BLUE S 1

43



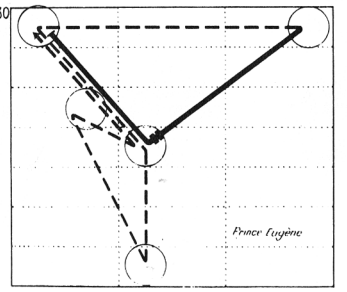
AFRIQUE BUSHMAN CHANT DE FEMMES

47



CANADA FRANÇAIS BALLADE R

30



3

A. LOMAX - Phonotactique du chant populaire - L'HOMME - Mouton, Paris 1964 Documents personnels

B. The number of components

Consequences of this notion

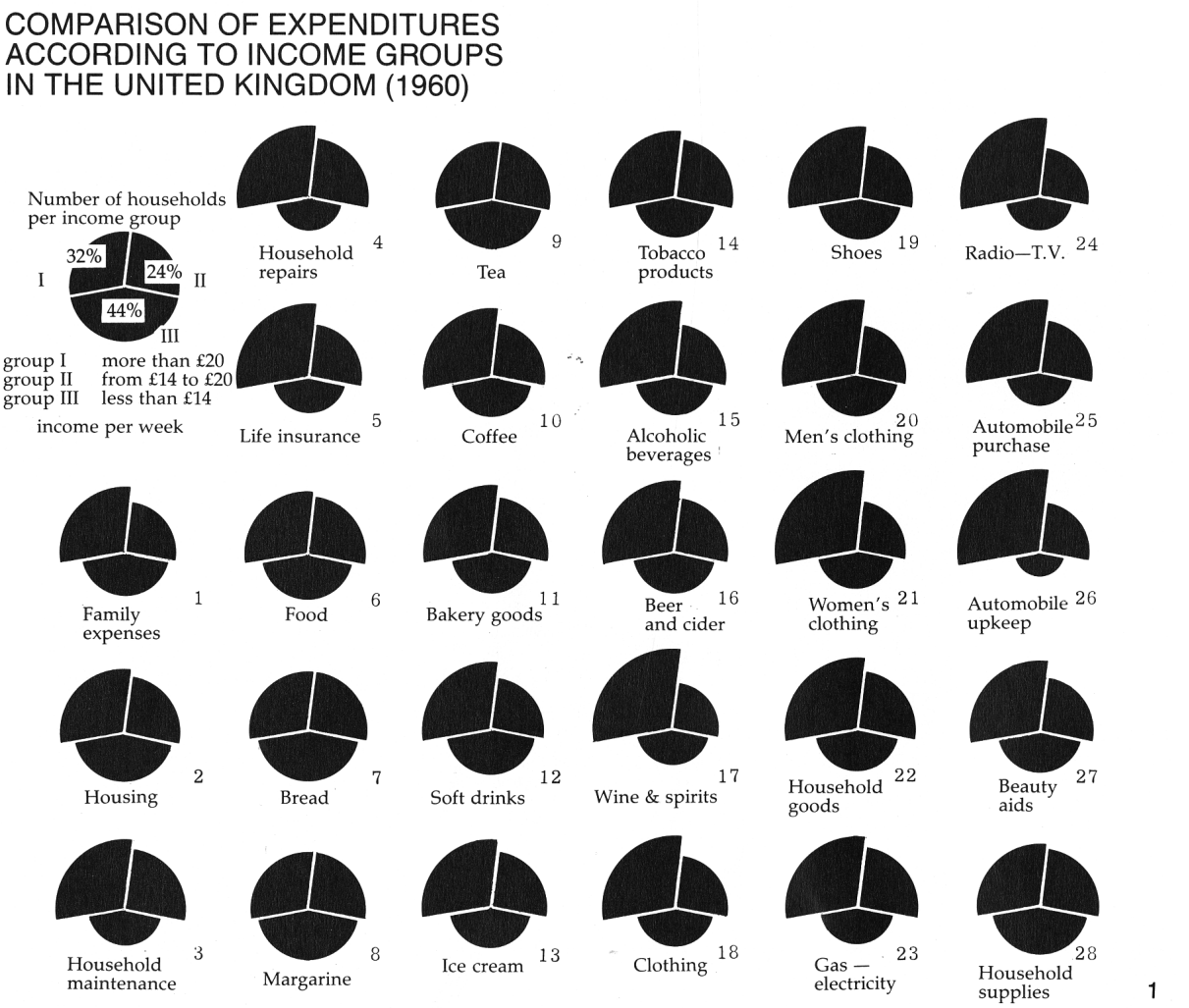
– The number of visual variables necessary for the representation is at least equal to the number of components in the information.
– With three components, the information can be perceived as a single image. Beyond that, the perception of several successive images is often necessary.
– There are at least as many types of possible questions as there are components.
– The number of components is the best basis for a classification of graphic constructions.

Take the following information:

INVARIANT *–expenditures by the British population per item*
COMPONENTS *–twenty-eight different items of expenditure*
–Q per 100 expenditures per item according to
–three income groups (upper, middle, lower)

Because the information involves only three components, it is possible to replace figure 1, necessitating the mental addition of numerous images, by figure 2, which presents a memorizable image.

It is a feeling of uniformity, of nonvariation, which results from figure 1, yet the distribution of expenditures could

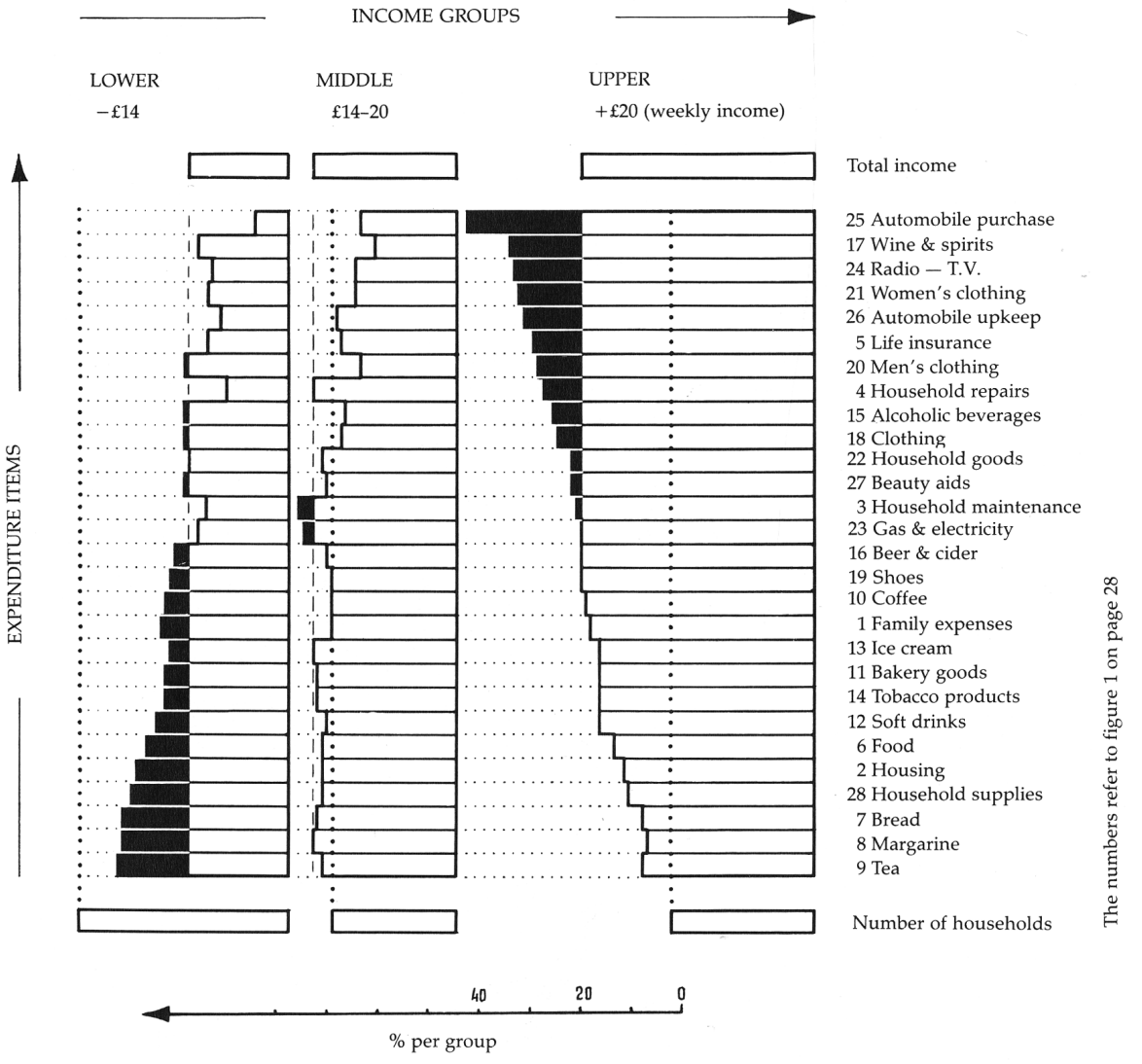


1

from Harry HENRY. Thomson Organisation Ltd.
Sources: Central Statistical Office. London 1961

hardly be similar for all three income groups.

In figure 2 the reader is struck and guided by the visible differences (underscored, incidentally, by the use of black) and is able to concentrate on them. The reader can rapidly perceive the logical order of the image: From left to right are the groups, from top to bottom are the items, whose order constitutes the very purpose of the information. The reader can pose questions about the characteristics of a group or an item, or about the order of these items, and feel confident of obtaining an answer.



The numbers refer to figure 1 on page 28

2

TRAFFIC ACCIDENT VICTIMS, in France, in 1958

VEHICLE (or pedestrian)
QUANTITIES

V	Pedestrians	Bicycles	Motorcycles	Four-wheeled vehicles
Q	28 957	17 247	74 887	63 071

VEHICLE
SEX
QUANTITIES

V	Pedestrians		Bicycles		Motorcycles		Four-wheeled vehicles	
S	M	F	M	F	M	F	M	F
Q	16 702	12 249	13 009	4 238	61 609	13 278	39 732	23 339

VEHICLE
SEX
CONSEQUENCES (dead, injured)
QUANTITIES

V		Pedestrians		Bicycles		Motorcycles		Four-wheeled vehicles	
S		M	F	M	F	M	F	M	F
C	Q								
d		1232	570	701	126	2 664	322	1 817	694
i		15 470	11 679	12 308	4 112	58 945	12 956	37 915	22 645

VEHICLE
SEX
CONSEQUENCES
AGE
QUANTITIES

V			Pedestrians		Bicycles		Motorcycles		Four-wheeled vehicles	
S			M	F	M	F	M	F	M	F
C	A	Q								
d	50		704	378	396	56	742	78	513	253
i			5 206	5 449	3 863	1 030	8 597	1 387	7 423	5 552
d	30		223	49	146	24	889	98	720	199
i			3 178	1 814	3 024	1 118	18 909	3 664	15 086	7 712
d	20		78	24	55	10	660	82	353	107
i			1 521	864	1 565	609	18 558	4 010	9 084	4 361
d	10		70	28	76	31	362	54	150	61
i			1 827	1 495	3 407	1 218	12 311	3 587	3 543	2 593
d			150	89	26	5	6	6	70	65
i			3 341	1 967	378	126	181	131	1 593	1 362

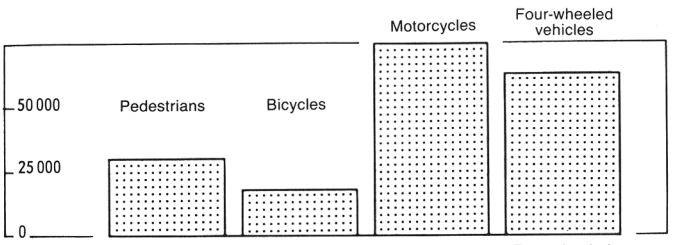
Source: Ministère des Travaux Publics

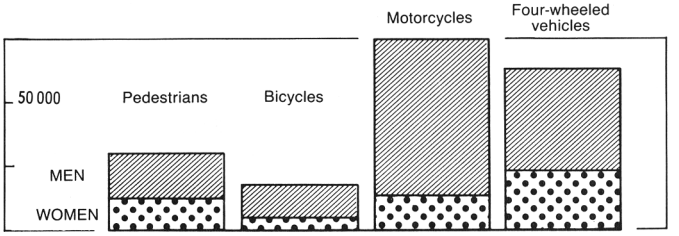
1

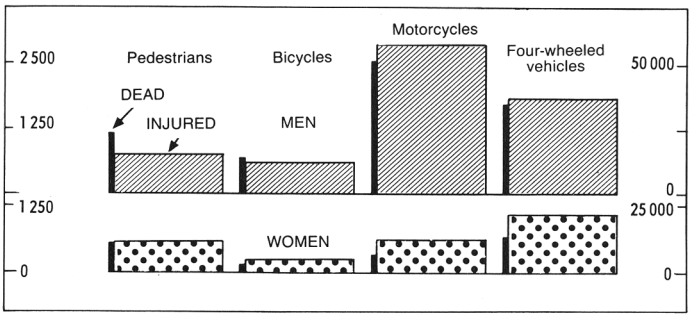
Information can have 2, 3, . . . *n*, components, and *n* can be quite large. It is sufficient that one of the components or the invariant be common to all the data.

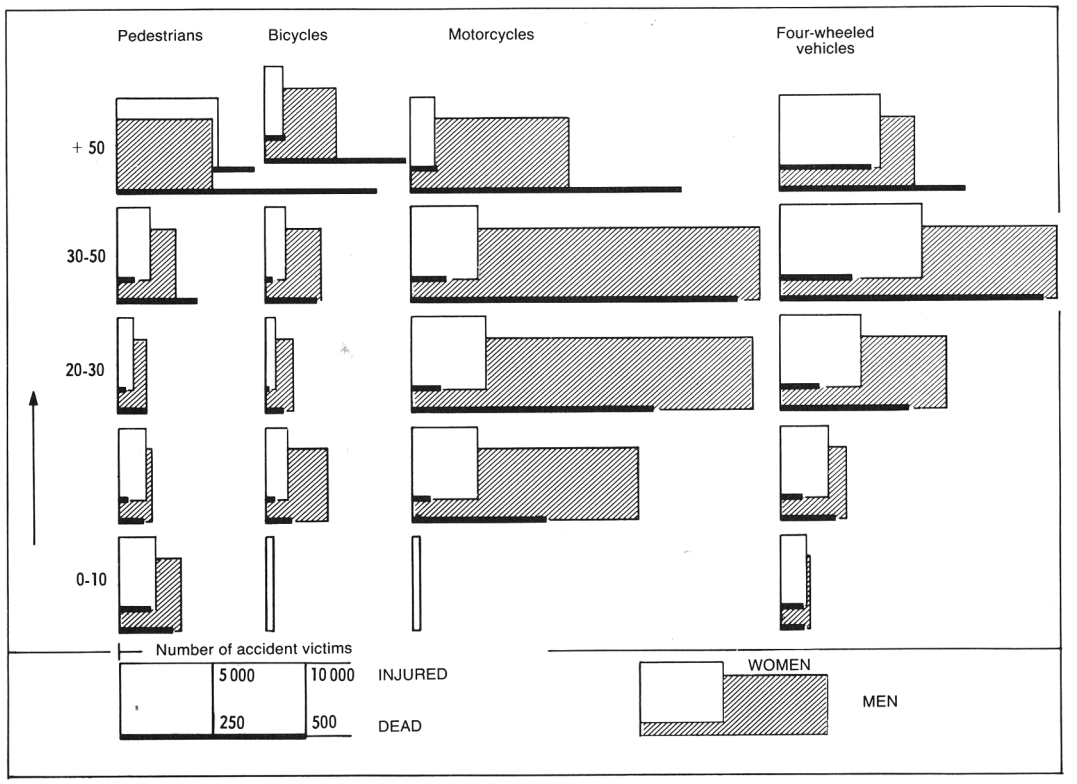
Consider the following example:
Analysis of highway accidents in France.
INVARIANT —*an accident victim*

This example can involve numerous components. As we see in figure 1 each additional component will generate new information. Furthermore, each additional component will also require a new visual variable leading to a different construction, as illustrated in figure 2 on the opposite page.

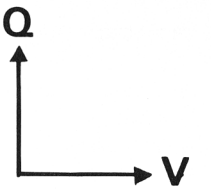


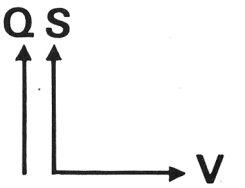


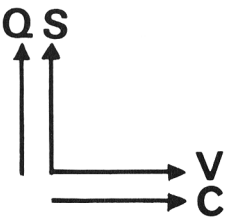


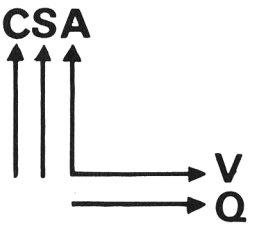


2









A set of information

Consider the distribution of the work force according to numerous concepts:

INVARIANT –*working persons in France*

COMPONENTS –*quantity according to*
–*geographic categories*
–*time*
–*age*
–*socioprofessional categories*
–*religious categories*
–*political categories*
–*various rates (mortality, fertility, birth, wealth, education …)*

When components are numerous, we will speak of a *set of information*. It is useful to analyze it as a whole and to consider the finite set of components in order to determine the most efficient and most economical processing system.

Regional studies, which have a defined geographic space as their common component; sociological surveys, whose common component is a group of individuals; historic research, whose common component is a period of time, are finite sets. To reduce them to their minimum constituents different systems of "information processing" are used. Depending upon which system is adopted, the graphic representation will lead to a series of diagrams, double-entry tables (see figure 1, page 30), maps, etc.

Likewise, at the moment of publication, the entire set of drawings must be conceived as a demonstrational unit. The layout must be considered a scientific problem, linked to the imperatives of reading and comparison, before it can be treated as an aesthetic problem (see, for example, page 401).

Information can have a single component

Figure 1: Conversational relationships among several individuals.

INVARIANT –*an exchange of conversation*

COMPONENT –*a series of individuals*

Figure 2: Map of developed areas in Poland, taken from F. Uhorczak (see also page 318).

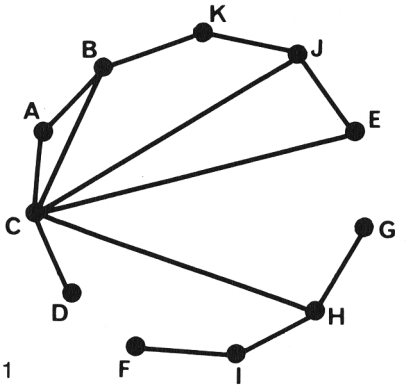
INVARIANT –*a site occupied by a building, whether a house, a factory, or a barn*

COMPONENT –*geographic space*

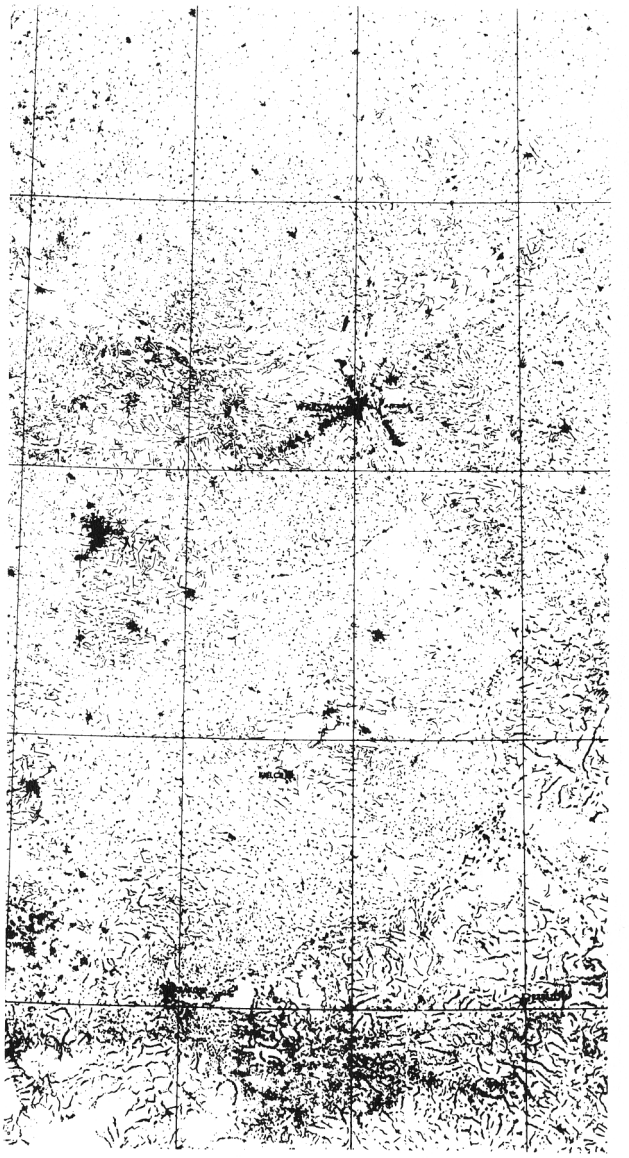
Definition of the number of components

In order to define the number of components in the information, the most convenient means consists of transcribing the data into a double-entry table (as in figure 1, page 30), before making any attempt at representing it graphically. There are as many components as entry-categories, plus one for quantities, if applicable.

Care must be taken, since such a table defines and clarifies the information but does not determine the most efficient graphic construction. This can sometimes be quite different from the layout of the table.



1


2

C. The length of the components

Consequences of this notion

– "Long" (extensive) components lead to basic or "standard" constructions.
– "Short" (limited) components lead to "special" constructions.
– The visual variables utilized must have at least the same length as the components which they are meant to represent.
– In a problem involving more than three components, the minimum number of images necessary is a function of the length of the components.

As variational concepts, components are, by definition, divisible. Their divisions bear different names according to circumstance and level.

One may speak of *elements* in the component "different persons forming a genealogical tree" or of *objects* in the component "different objects to be classed."

One may also speak of socioprofessional, geographic, or linguistic *categories*, and, in general, this term can be applied to all components.

One may speak of time, age, or income *classes*, and, in general, this term can be applied to all components which are ordered or quantitative. Finally, one may speak of *steps* of value, texture, or size for the visual variables (the components of the graphic sign-system). Each variable, each combination of variables, has a given length, which is most often quite limited. In fact, all these terms cover the same phenomenon, the useful and separable divisions of a component.

The term LENGTH of a component will be used to refer to the number of divisions that it enables us to identify.

The full significance of this notion can be appreciated when one comes to realize that the number of categories a person can grasp during the course of perception is quite small. This means that the visual variables can produce only a limited number of perceptible steps.

"Short" or limited components

The term *short components* will be used when length does not exceed four. Binary components have a length of two (sex, living or deceased, a decision which is actual or potential, etc.). Age is often divided into three main categories—youth, adult, elderly—as are the main employment sectors (agriculture, industry, tertiary). Short components are noteworthy in graphic problems. They simplify visual selection and enable us to use "special" constructions, which differ from the standard constructions.

"Long" or extensive components

The term *long components* will be used when length exceeds some fifteen divisions. Long components necessarily lead to "standard" constructions (see page 172).

The term *length* can be applied to a quantitative series when the latter is divided into steps or classes, or when it involves a "discrete" component (such as numbers of objects, inventories, or francs, which cannot be divided below a certain basic unit).

On the other hand, a series of numbers can be infinitely divisible; the term *length* is no longer applicable when the phenomenon is considered as continuous (such as speeds, altitudes, or temperature). Note, however, that the number of useful decimals is a finite number.

A phenomenon considered as continuous can nevertheless be expressed by graphic means, since the plane itself is continuous.

THE RANGE OF A QUANTITATIVE SERIES is the ratio of the smallest to the largest number.

This notion is quite different from that of length. A series whose extremes are 0.07 and 32 has the same range as a series whose extremes are 22 and 10 054; they both range from 1 to 457.

The practical range of a visual variation in size is limited; it cannot decrease below a ratio of 1 to 10 without losing the greater part of its efficacy. On the other hand, quantitative information can range from 1 to 1.2 (the size of individuals) as well as from 1 to 10 million (population maps). Thus it is easy to understand the importance of this notion and of the "range adjustments" which must often be introduced into quantitative representations in order to adapt the information to the faculties of visual perception (see page 357).

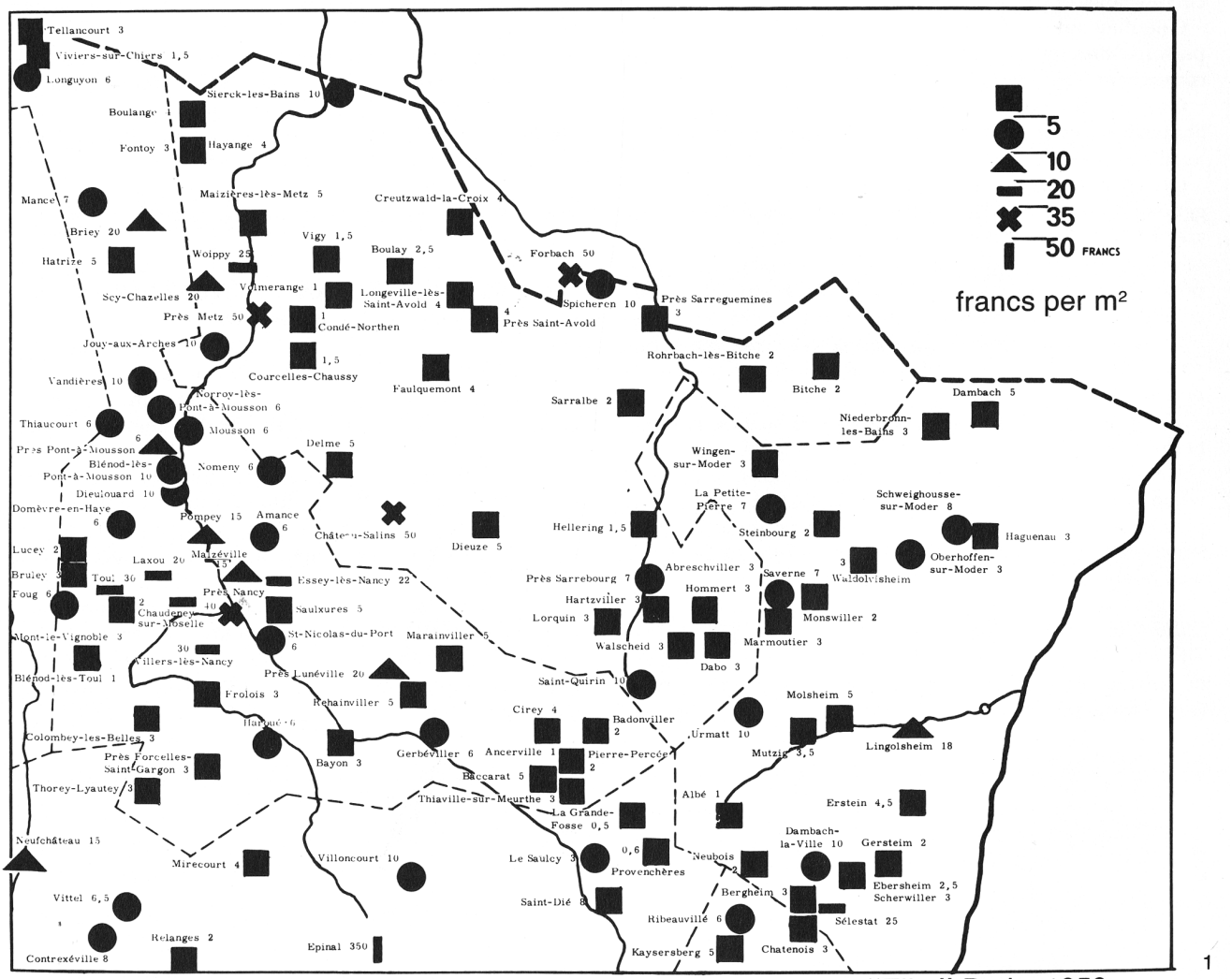
D. The level of organization of the components

The components of the information do not all involve the same intellectual approach. For purposes of information processing and/or display, the researcher will often attempt to order qualitative (nominal) categories such as trades, to compare ordered categories such as heat sensitivities, and to group neighboring quantitative values such as population densities.

A component can thus be characterized as *qualitative*, *ordered*, or *quantitative*; these are the three levels of organization. The visual variables which represent each component must permit parallel perceptual approaches. But, just like the components, the visual variables each have their own level of organization. An order will not be perceptible if the variable is not ordered; a ratio will not be perceptible if the variable is not quantitative. The notion of level of organization is thus of fundamental importance.

Consequences of this notion

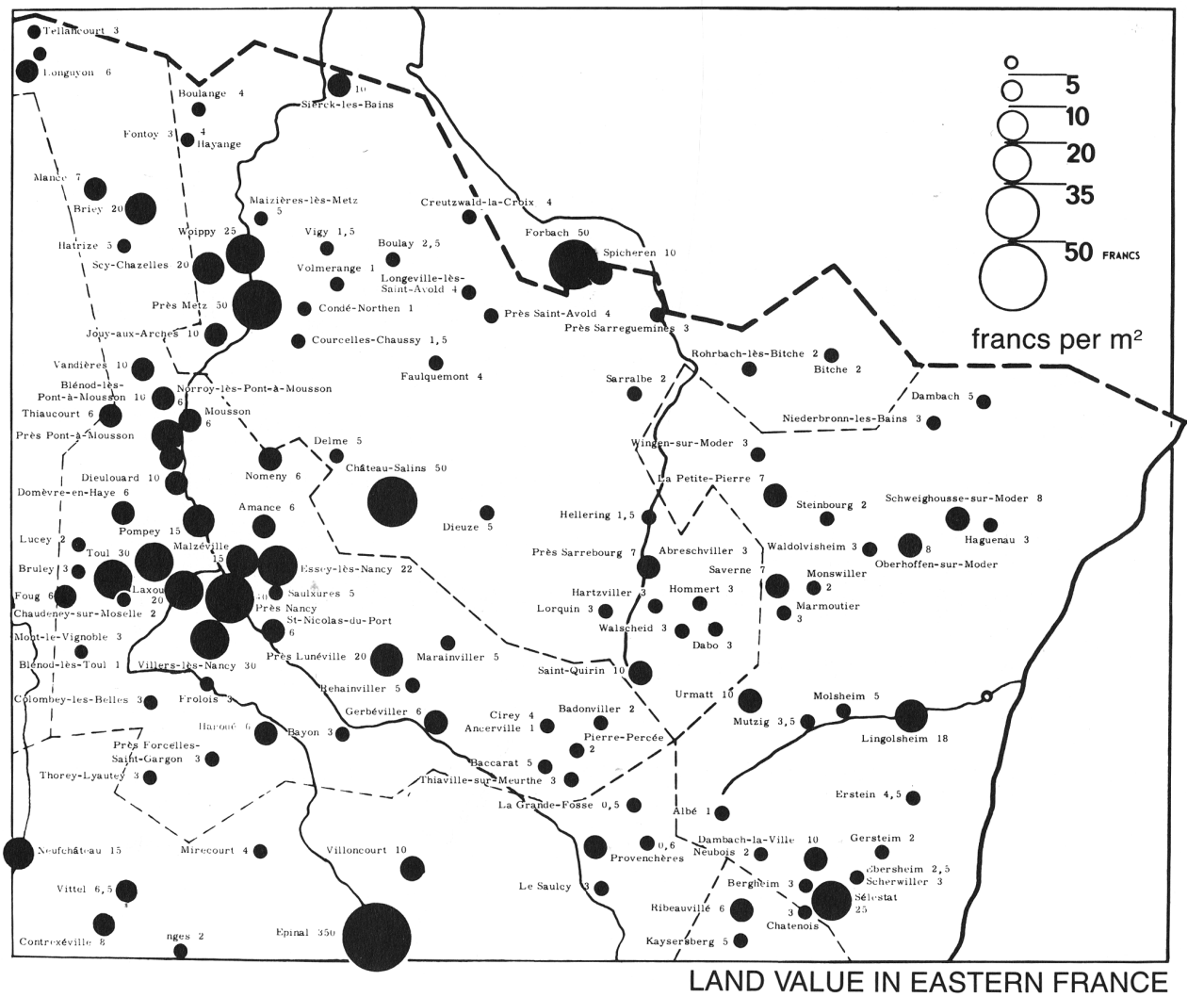
– The ordering of qualitative data, the comparison of ordered components, the groupings resulting from a quantitative component are the basis for the graphic processing of information.
– The visual variables must have a level of organization at least equal to that of the components which they represent.
– The three levels of organization lead to the first subclassification of graphic constructions.



1

LAND VALUE IN EASTERN FRANCE, from the weekly magazine, "Elle." Paris, 1959.

It is because the level of the visual variable utilized does not correspond to the level of the component represented by it that the map in figure 1 is inefficient and necessitates the burdensome reading of several successive images. When the levels correspond, as in figure 2, the map is visually retainable. It necessitates only one immediately perceived image.



2 LAND VALUE IN EASTERN FRANCE

THE QUALITATIVE LEVEL (OR NOMINAL LEVEL)

This notion includes all the innumerable concepts of simple differentiation: professions, products, languages, races, religions, leisure activities, diseases, colors, forms, social, ethnic, cultural or political traits . . .

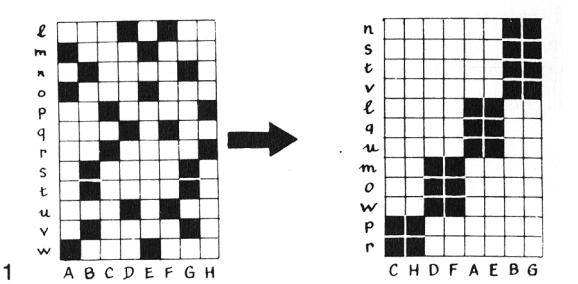
A component is qualitative when its categories are not ordered in a universal manner. As a result, they can be reordered arbitrarily, for purposes of information processing.

Qualitative categories are reorderable

In maritime commerce, for example, the categories coal, oil, wheat, wool, cotton, wine, wood, . . . of the component "merchandise" can be ordered in different ways, according to weight, total value, price per kilogram, volume, revenue production, fragility. . . .

Geographic groupings are also reorderable. Departments are commonly classed by alphabetic order, countries by their population, their production, their standard of living, their birth rate. . . .

The reciprocal ordering of two qualitative components (figure 1) or of a qualitative component in relation to an ordered component (figure 2) simplifies the images without diminishing the number of observed correspondences; indeed these operations of "permuting" and classing are the basis for graphic information-processing.



1

Qualitative categories are equidistant

As with ordered categories (discussed later), qualitative categories are, by definition, of equal importance, that is, "equidistant." Their graphic representation must not disturb this quality by highlighting a particular category or creating a priori groups of categories.

Two perceptual approaches

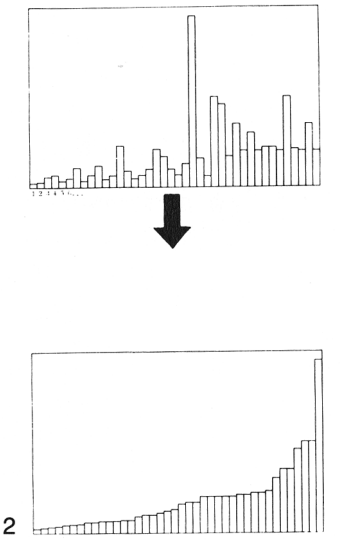
Faced with any qualitative concept, the observer can adopt two perceptual approaches:

This is different from that—a beech tree is different from an oak.

This is similar to that—beeches and oaks are similar—they are trees.

A *selective approach (difference)* is engendered by questions of an elementary or intermediate reading level. Where is a given category—the beech trees? When these questions are pertinent, it is important that the component be represented by a selective variable.

An *associative approach (similarity)* is engendered by questions of an intermediate or overall reading level. Where is a given component—the forest—all categories of trees combined? In order to reply to this question, the variable must permit equalizing and grouping all the categories during perception; it must be associative.


2

THE ORDERED LEVEL (ORDERING AND REORDERING)

This level groups all the concepts which are capable of ordering categories in a universally acknowledged manner. Each person will agree in the same way that this is more than that and less than the other.

Ordered concepts are always defined, more or less directly, in relation to:

- a temporal order: age, generation, matrimonial status, geologic era
- an order of sensory discrimination: heat, vision (black–gray–white, large–medium–small, here–near–far), weight (heavy–medium–light), health
- an order of intellectual or moral discrimination (good–mediocre–bad)
- certain social structures, such as military or administrative hierarchies.

A component is ordered, and only ordered:

- **when its categories are ordered in a single and universal manner**
- **when its categories are defined as equidistant**.

Ordered categories cannot be reordered

More precisely we can say that the reclassification of ordered categories is generally a source of confusion in the process of communication.

Consider the following example:

Propensity to theft according to age and amount of theft (based on V. V. Stanciu, Theft in Department Stores, unpublished study).

INVARIANT –*theft in department stores*
COMPONENTS –*age groups*
–*quantities (per 100 persons per age group) according to*
–*classes of amount of theft*

The components "age" and "amount of theft" are ordered, producing figure 3: In black is the highest percentage for each column, that is, the age group with the strongest observed tendency.

It can be of interest to reorder the component "amount of theft" for the purposes of constructing a linear relationship (figure 4), which permits one to reflect on theft psychology. Interpretation is more delicate, however, because reading from left to right no longer has an ordered meaning.

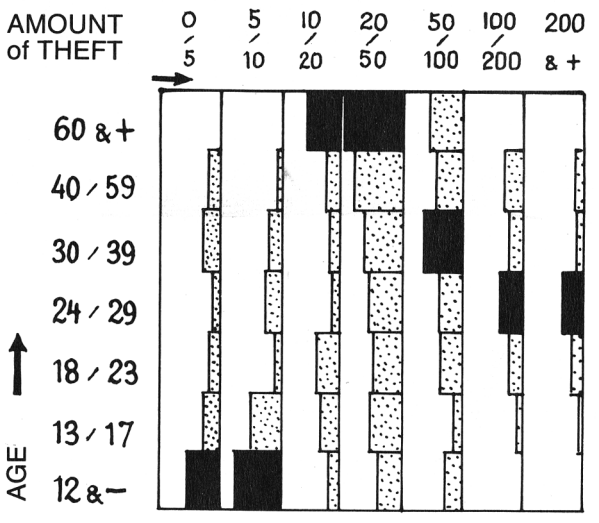
These examples demonstrate that the amount of theft is not ordered in a direct way by age, nor age by amount of theft.

Ordered categories are defined as equidistant

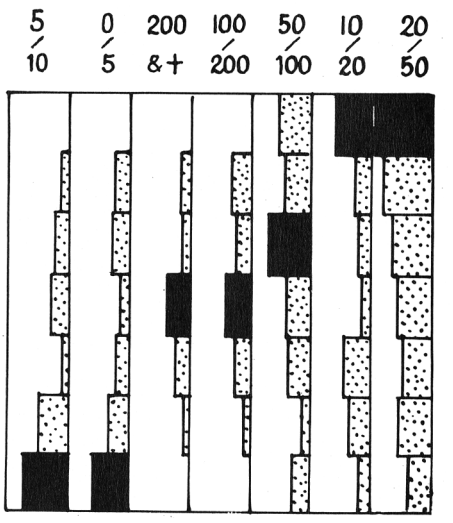
This characteristic distinguishes an ordered component from a quantitative component. The series: bachelor, husband, widower, deceased, constitutes a universal order. But there is no reason, a priori, to bring together any two categories or to form groups. These categories are ordered and at equal distances from each other.

The same is true for the component "amount of theft," even though its categories are defined by *numbers*. These are *ordinal numbers* which merely serve to rank the categories.

In any graphic transcription involving an ordered component, particularly when using "retinal" visual variables, the designer must try to preserve this equidistance. A priori visual groups must be avoided, since the very purpose of graphic processing is to discover, a posteriori, the groupings which result from the information.

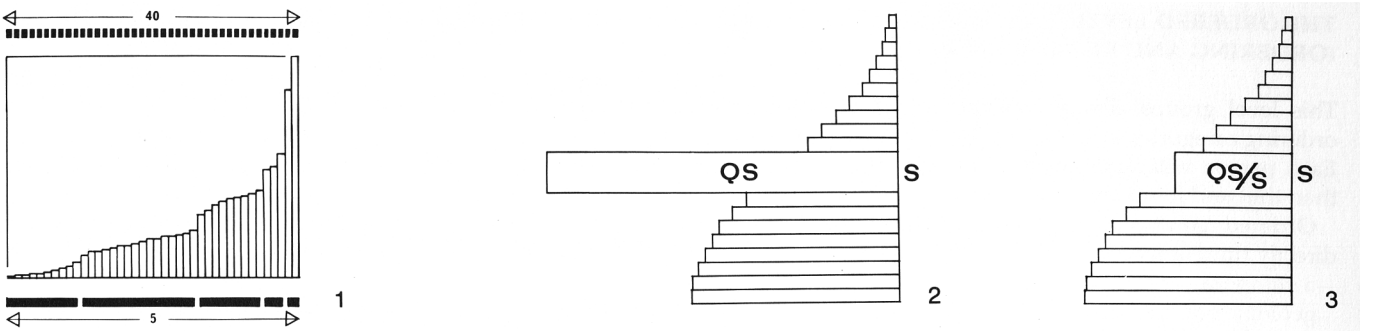


3



4

THEFT IN DEPARTMENT STORES
Distribution of delinquents according to age and amount of theft



THE QUANTITATIVE LEVEL (INTERVAL-RATIO LEVEL)

This level is attained when there are countable units, leading us to say: this is double, half, four times that. . . .

A series of numbers is quantitative when its object is to specify the variation in distance among the categories.

With a quantitative series of numbers, Q, it is possible to represent a variation in the length of columns, as in figure 1, and from this to derive groups, characterized by slight differences in length (slight "distances").

Relations among quantities and enumeration units

Before representing given quantities, any graphic must first depict the units (geographical areas, time periods, age groups, etc.) within which these quantities are being enumerated. A population map by commune is, first of all, a map of communes.

When the enumeration units are unequal:

- the representation of these units on the plane can result in figures which are equal (points on a diagram) or unequal (a map of communes);
- the representation of the quantities can be independent of the inequality of these figures (a single point per area) or dependent on it (color over the entire area).

But the quantities themselves can be independent of the inequality of the enumeration units (death rate in a commune) or dependent on it (total population). Therefore, graphic representation necessarily leads to an initial analysis of any quantitative series in terms of these relations.

Quantities dependent on the enumeration units (or QS)

In comparing communes, the geographer must calculate population density in order to take the unequal areas of the communes into account. This calculation is necessary because the quantities of population *are not independent of the unequal areas* (S)* of the communes. The same is true for the historian who uses quantities of immigrants enumerated over unequal periods (S) of time or for the demographer who uses quantities of persons counted within unequal age groups (S). These quantities are not independent of the dimension (S) of the enumeration units.

*S will be used frequently to signify area, since it comes from the French equivalent *surface* (translator's note).

Quantities of the form QS are:

absolute quantities (Q) counted according to variable units (S), whether these quantities are expressed:

by the observed numbers:

Q of tons of milk per department (S), Q of persons per period (S) . . .

in hundredths (or in thousandths of the total):†

Q of milk per department (S), Q of persons per age-group, expressed in hundredths (or in thousandths) of the total of the series, that is, QS × 100/total of the series.

Test: For a hundred what? For a total of the series equal to a hundred;‡

by an index:†

Q of milk consumed per period (S), per 100 liters consumed in 1950 (Qi), Q of milk produced per department (S), per 100 liters produced in the department of Calvados (Qi), that is, QS × 100/Qi.

Test: For a hundred what? For a hundred liters produced in Calvados;‡

by a ratio based on a variable independent of S:

Q is monthly average amount of milk produced per department (S), or Q is communal average of expenditures per period (S), that is, QS total/number of units (months, communes). When S is represented by lines or by areas, the graphic transcription of QS can lead to serious errors (page 45). It is generally necessary to transform the data by making the calculation QS/S, as illustrated in figures 2 and 3.

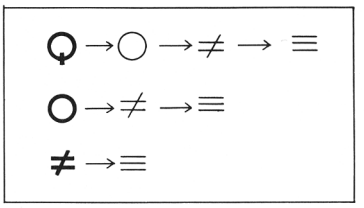
Quantities independent of the enumeration units (or Q)

The geographer looking for quantities of population independent of the area of a commune must calculate densities: Q of population/area, that is, QS/S = Q.

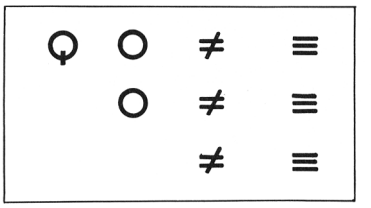
The demographer will reduce a class variation in the same manner: Q of persons/length (in years) of the class,

†Hundredths and indices merely involve a simple change of numerical scale, useful in verbal communication to the degree that all the numbers become intelligible when it is understood what one hundred or one thousand represents. Graphically, absolute Q, hundredths, or indices produce the same image of a series.

‡The test, for a hundred what?, is indispensable. It permits the elimination of false percentages and the comprehension of what is in question. It obliges one to furnish the elements of an answer and it reveals, all too frequently, a series for which an answer is impossible, so that clarification or elimination of the series is required.



4



5

that is, QS/S = Q. But quantities Q are not all of this nature.

Quantities of the form Q are:
samples, altitude, temperature, commodity prices, number of workers per factory, etc. These are measurements or real values sampled at a point which is by definition without length or area. They thereby characterize an invariable enumeration unit;
reductions to a unitary class, densities, "absolute" frequencies, such as the examples cited earlier, which result from the operation QS/S = Q;
simple ratios, in which the variable unit (S) relates the two terms of the ratio:
Q of wheat produced per commune (S)/Q of hectares sowed per commune (S),
Q of emigrants per period (S)/Q of boats per period (S), that is, QaS/QbS = Qa/Qb = Q′;
"percentages" and "rates" which multiply the simple ratios by a hundred (or a thousand):
Q of deaths per commune (S) × 1000/Q of persons per commune (S),
Q of working persons per age group (S) × 100/Q of persons per age group (S), that is, QaS × 100/QbS = Qa/Qb × 100 = Q′ per hundred.

Test: For a hundred what? For a total of a hundred persons per commune.‡
The graphic transcription of absolute quantities (Q) is simpler. However, one must know how to avoid a confusion with graphic solutions suitable only to QS (figure 16, page 45).

INCLUSIVENESS OF THE LEVELS OF ORGANIZATION

Graphic conventions

In order to designate a component and, at the same time, specify its level of organization, the following signs will be used:
Q—a quantitative series measuring variations in distance among ordered categories.
O—a component whose categories arc equidistant and inscribed in a single, universally acknowledged order.
≠—a qualitative component whose categories are defined and equidistant.
≡—a qualitative component whose differential characteristics can be disregarded (i.e., which can be approached "all categories combined").

Inclusiveness of perceptual approaches

The level of organization determines the perceptual approaches that can be adopted toward a component. These approaches are ordered and inclusive. In effect, *for a quantitative component*, it is possible to adopt:
a quantitative perceptual approach and ask the question: What is the ratio between the two lengths, between the two populations, between the two areas. . . ?
an ordered perceptual approach and ask the questions: In what order are the lengths given? Does the order of the departmental population quantities correspond to the alphabetic order of the departments . . . ?
a selective perceptual approach and ask the question: Where are all the cities of 15 000 inhabitants?
an associative perceptual approach and ask the questions: What is the distribution of the "cities," disregarding any differentiation among them? Where is the forest, disregarding any differentiation of age, size, or kind of trees?

Thus:
– **All quantitative series can be considered as merely ordered.**
– **All the categories of an ordered series can be considered as merely differentiated.**
– **All the categories of a qualitative series can be considered as similar.**

But:
– **A series which is only qualitative is not ordered (although we can reorder it arbitrarily, page 36).**
– **A series which is only ordered is not quantitative (although it may be defined by ordinal numbers, page 37).**

The system of inclusion resulting from these statements is expressed in figure 4, a more readable version of which is given in figure 5. This enables us to identify the perceptual approaches which each component can generate, to choose a visual representation of at least an equal level, and to classify the visual variables among themselves.