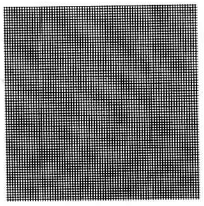
I want to reach that state of condensation of sensations which constitutes a picture.

Henri Matisse

8 *Data Density and Small Multiples*

Our eyes can make a remarkable number of distinctions within a small area. With the use of very light grid lines, it is easy to locate 625 points in one square inch or, equivalently, 100 points in one square centimeter.

Or consider how an 80 by 80 grid over a square inch—about 30 by 30 over a square centimeter—divides the space:[1]



25,281 distinctions

[1] A square grid formed on each side by n parallel black and n–1 parallel white lines contains n^2 intersections of two black lines (corners of squares), $(n-1)^2$ intersections of two white lines (white squares), and $2n(n-1)$ intersections of a black and white line (sides of squares), for a total of $(2n-1)^2$ line intersections or distinct locations.

With the help of considerable redundancy and context, our eyes make fine distinctions of this sort all the time. Measurement instruments used in engineering, architectural, and machine work are engraved with scales of 20 increments to the centimeter and 50 to the inch. Or consider the reading of fine print. The type in the U.S. *Statistical Abstract* is set at 12 lines per vertical inch, with each line running at about 23 characters per inch for a maximum density of 276 characters per square inch. The actual density, given the white space, is in this case 185 characters per square inch or 28 per square centimeter.

No. 1450. Steel Products—Net Shipments, by Market Classes: 1960 to 1978

[**In thousands of short tons.** Comprises carbon, alloy, and stainless steel. "N.e.c." means not elsewhere classified]

Market class	1960	1965	1970	1973	1974	1975	1976	1977	1978
Total [1]	**71,149**	**92,666**	**90,798**	**111,430**	**109,472**	**79,957**	**89,447**	**91,147**	**97,935**
Steel for converting and processing	2,928	3,932	3,443	4,714	4,486	3,255	4,036	3,679	4,612
Independent forgers, n.e.c.	841	1,250	1,048	1,213	1,339	1,098	952	998	1,192
Industrial fasteners [2]	1,071	1,234	1,005	1,278	1,331	675	912	848	870
Steel service centers, distributors	11,125	14,813	16,025	20,383	20,400	12,700	14,615	15,346	17,333
Construction, incl. maintenance	9,664	11,836	8,913	10,731	11,360	8,119	7,508	7,553	9,612
Contractors' products	3,602	5,018	4,440	6,459	6,249	3,927	4,502	4,500	3,480
Automotive	14,610	20,123	14,475	23,217	18,928	15,214	21,351	21,490	21,253
Rail transportation	2,525	3,805	3,098	3,228	3,417	3,152	3,056	3,238	3,549
Freight cars, passenger cars, locomotives	1,763	2,875	2,005	1,997	2,097	1,794	1,428	1,709	2,188
Rails and all other [3]	762	930	1,093	1,231	1,320	1,358	1,628	1,529	1,361
Shipbuilding and marine equip	622	1,051	859	1,019	1,339	1,413	969	869	845
Aircraft and aerospace	78	94	56	69	79	69	59	63	60
Oil and gas industries	1,759	1,936	3,550	3,405	4,210	4,171	2,653	3,650	4,140
Mining, quarrying, and lumbering	288	392	497	534	644	596	536	486	508
Agricultural, incl. machinery	1,003	1,483	1,126	1,772	1,859	1,429	1,784	1,743	1,805
Machinery, industrial equip., tools	3,958	5,873	5,169	6,351	6,440	5,173	5,180	5,566	5,992
Electrical equipment	2,078	2,985	2,694	3,348	3,242	2,173	2,671	2,639	2,811
Appliances, utensils, and cutlery	1,760	2,179	2,160	2,747	2,412	1,653	1,950	2,129	2,094
Other domestic commercial equip.	1,959	2,179	1,778	1,990	1,941	1,390	1,813	1,846	1,889
Containers, packaging, shipping	6,429	7,331	7,775	7,811	8,218	6,053	6,914	6,714	6,595
Cans and closures	4,976	5,867	6,239	6,070	6,349	4,859	5,290	5,173	4,950
Ordnance and other military	165	289	1,222	918	654	405	219	193	207
Exports (reporting companies only)	2,563	2,078	5,985	3,138	3,961	1,755	1,839	1,076	1,224

[1] Total includes nonclassified shipments, and, beginning 1970, data include estimates for a relatively small number of companies which report raw steel production but not shipments. [2] Bolts, nuts, rivets, and screws. [3] Includes railways, rapid transit systems, railroad rails, trackwork, and equipment.

U.S. Bureau of the Census, *Statistical Abstract of the United States: 1979* (Washington, D.C., 1979), p. 822.

Maps routinely present even finer detail. A cartographer writes that "the resolving power of the eye enables it to differentiate to 0.1 mm where provoked to do so. Clearly, therefore, conciseness is of the essence and high resolution graphics are a common denominator of cartography."[2] Distinctions at 0.1 mm mean 254 per inch.

[2] D. P. Bickmore, "The Relevance of Cartography," in John C. Davis and Michael J. McCullagh, eds., *Display and Analysis of Spatial Data* (London, 1975), p. 331.

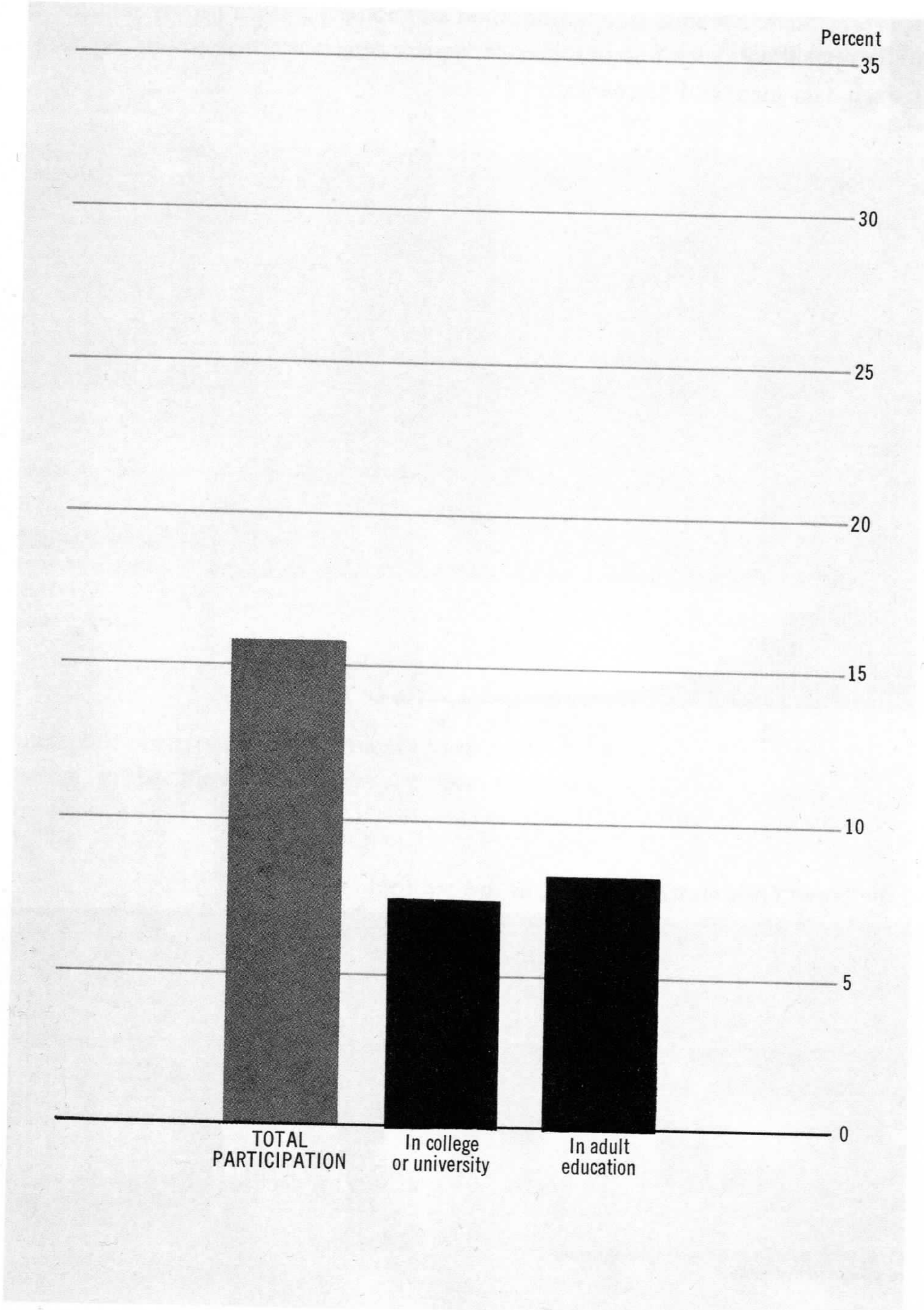
How many statistical graphics take advantage of the ability of the eye to detect large amounts of information in small spaces? And how much information should graphics show? Let us begin by considering an empirical measure of graphical performance, the data density.

Data Density in Graphical Practice

The numbers that go into a graphic can be organized into a data matrix of observations by variables. Taking into account the size of the graphic in relation to the amount of data displayed yields the *data density*:

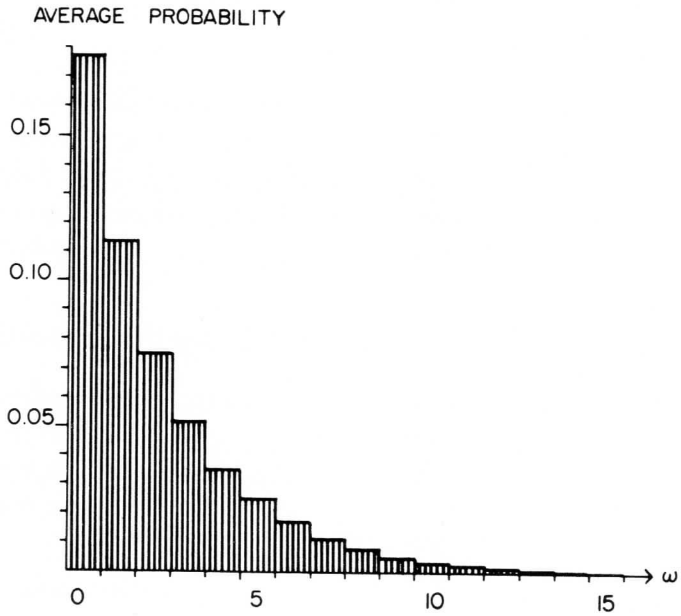
$$\text{data density of a graphic} = \frac{\text{number of entries in data matrix}}{\text{area of data graphic}}$$

Data matrices and data densities vary enormously in practice. At one extreme, this overwrought display (originally printed in five colors) presents a data matrix of four entries, the names and the numbers for the two bars on the right. The left bar is merely the total of the other two. The graph covers 26.5 square inches (171 square centimeters), resulting in a data density of .15 numbers per square inch (.02 numbers per square centimeter), which is thin indeed.

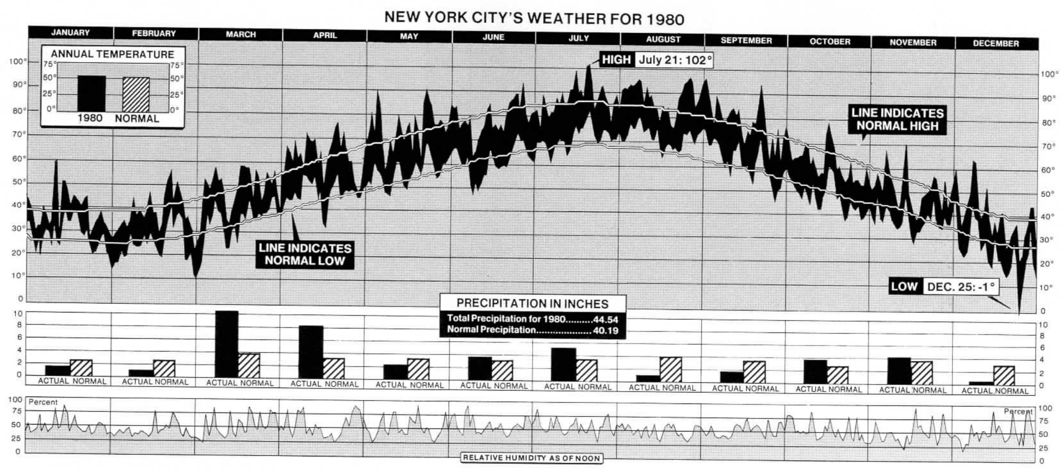


Executive Office of the President, Office of Management and Budget, *Social Indicators, 1973* (Washington, D.C., 1973), p. 86.

The exemplar from the JASA style sheet comes in at a lightweight 3.8 numbers per square inch (0.6 numbers per square centimeter) and a small data matrix of 32 entries:

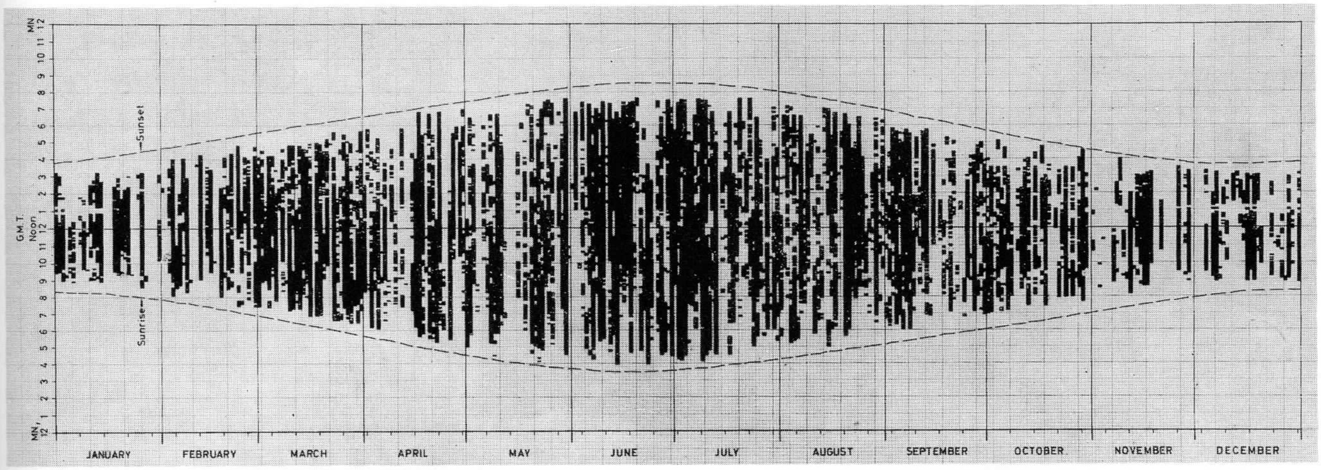


In contrast, the New York weather history, in this reduced version, does very well at 181 numbers per square inch (28 per square centimeter):

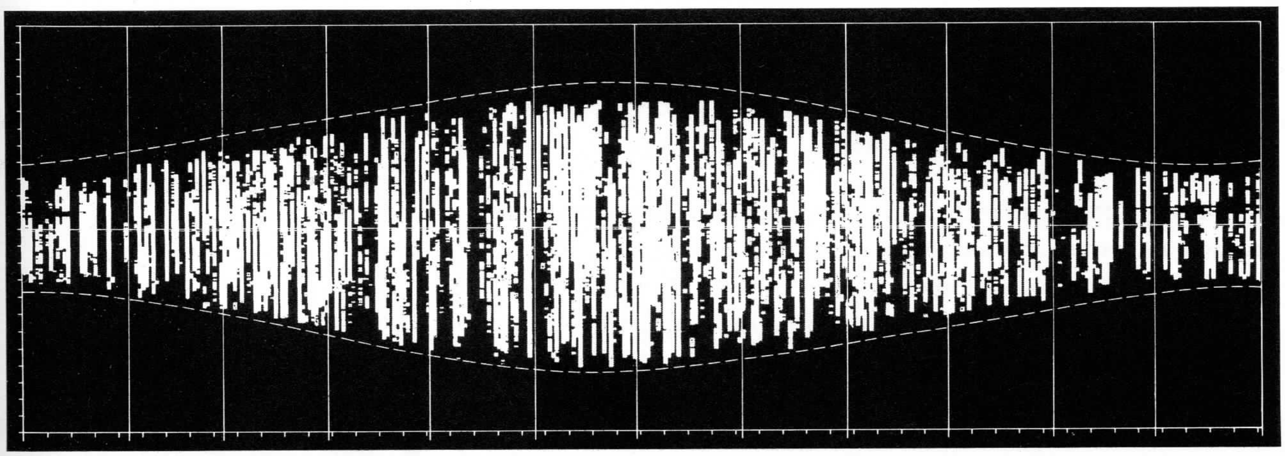


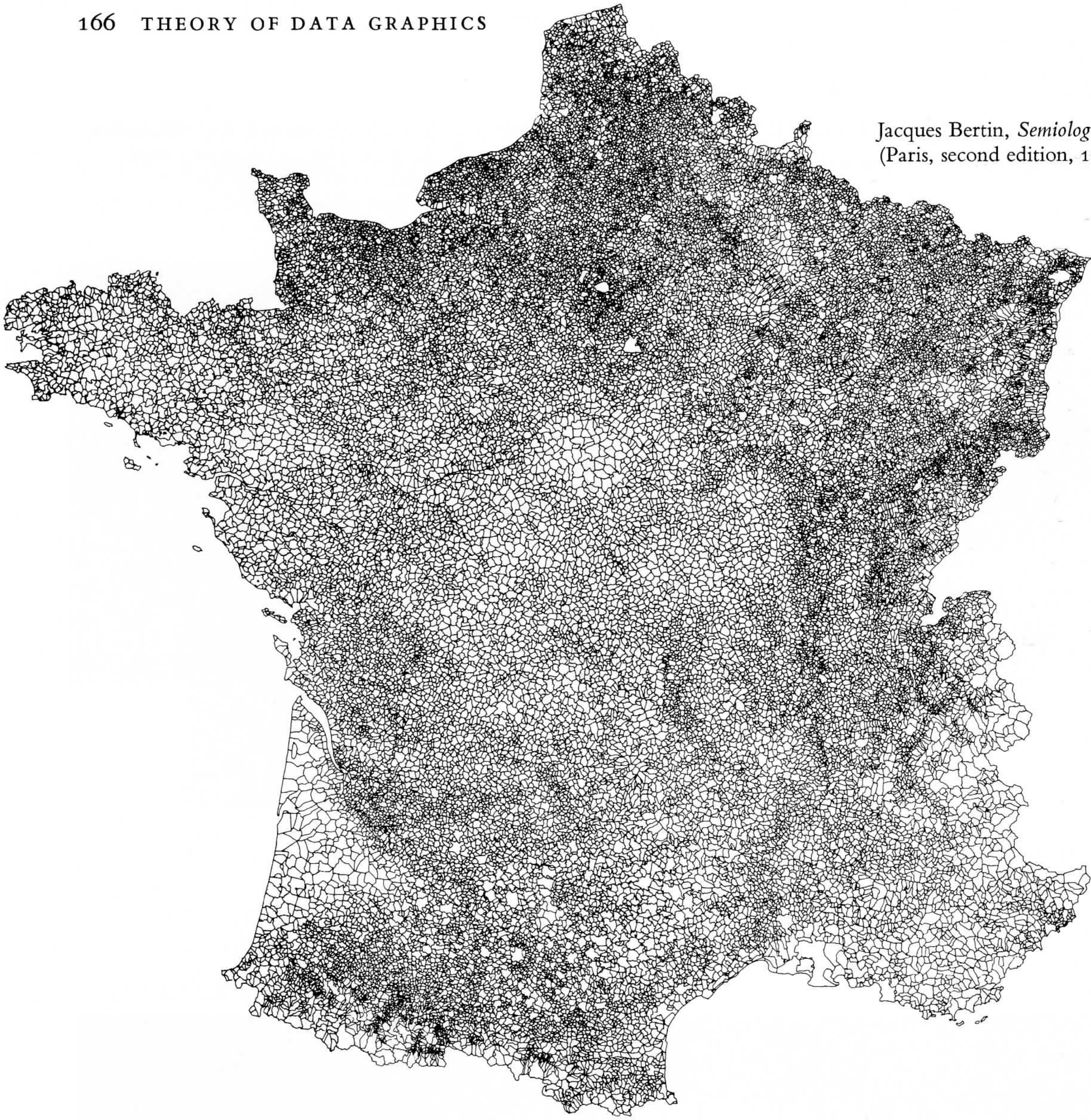
An annual sunshine record reports about 1,000 numbers per square inch (160 per square centimeter):

F. J. Monkhouse and H. R. Wilkinson, *Maps and Diagrams* (London, third edition, 1971), pp. 242–243.



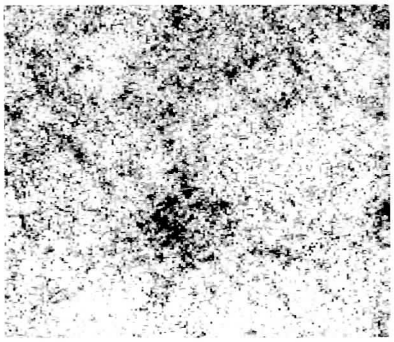
The visual metaphor corresponds appropriately to the data if the image is reversed, so that the light areas are the times when the sun shines:





Jacques Bertin, *Semiologie Graphique* (Paris, second edition, 1973), p. 152.

This map (27 square inches, 175 square centimeters) shows the location and boundaries of 30,000 communes of France. It would require at least 240,000 numbers to recreate the data of the map (30,000 latitudes, 30,000 longitudes, and perhaps six numbers describing the shape of each commune). Thus that data density is nearly 9,000 numbers per square inch, or 1,400 numbers per square centimeter.



The new map of the galaxies locates 2,275,328 encoded rectangles on a two-dimensional surface of 61 square inches (390 square centimeters). Each rectangle represents three numbers (two by its location, one by its shading), yielding a data density of 110,000 numbers per square inch or 17,000 numbers per square centimeter. That is the current record.

Data Density and the Size of the Data Matrix: Publication Practices

The table shows the data density and the size of the data matrix for graphics sampled from scientific and news publications. At least 20 graphics from each publication were examined.

The table records an enormous diversity of graphical performances both within and between publications. A few data-rich designs appear in nearly every publication. The opportunity is there but it is rarely exploited: the average published graphic is rather thin,

Data Density and Size of Data Matrix,
Statistical Graphics in Selected Publications, Circa 1979–1980

	Data Density (Numbers per square inch)			Size of Data Matrix		
	median	minimum	maximum	median	minimum	maximum
Nature	48	3	362	177	15	3780
Journal of the Royal Statistical Society, B	27	4	115	200	10	1460
Science	21	5	44	109	26	316
Wall Street Journal	19	3	154	135	28	788
Fortune	18	5	31	96	42	156
The Times (London)	18	2	122	50	14	440
Journal of the American Statistical Association	17	4	167	150	46	1600
Asahi	13	2	113	29	15	472
New England Journal of Medicine	12	3	923	84	8	3600
The Economist	9	1	51	36	3	192
Le Monde	8	1	17	66	11	312
Psychological Bulletin	8	1	74	46	8	420
Journal of the American Medical Association	7	1	39	53	14	735
New York Times	7	1	13	35	6	580
Business Week	6	2	12	32	14	96
Newsweek	6	1	13	23	2	96
Annuaire Statistique de la France	6	1	25	96	12	540
Scientific American	5	1	69	46	14	652
Statistical Abstract of the United States	5	2	23	38	8	164
American Political Science Review	2	1	10	16	9	40
Pravda	0.2	0.1	1	5	4	20

based on about 50 numbers shown at the rate of 10 per square inch. Among the world's newspapers, the *Wall Street Journal*, *The Times* (London), and *Asahi* publish data-rich graphics, with data densities equal to those of the *Journal of the American Statistical Association*. Most of the American papers and magazines, along with *Pravda*, publish less data per graphic than the major papers of other industrialized countries.

Very few statistical graphics achieve the information display rates found in maps. Highly detailed maps portray 100,000 to 150,000 bits per square inch. For example, the average U.S. Geological Survey topographic quadrangle (measuring 17 by 23 inches) is estimated to contain over 100 million bits of information, or about 250,000 per square inch (40,000 per square centimeter).[3] Perhaps some day statistical graphics will perform as successfully as maps in carrying information.

[3] Morris M. Thompson, *Maps for America* (Washington, D.C., 1979), p. 187.

High-Information Graphics

Data graphics should often be based on large rather than small data matrices and have a high rather than low data density. More information is better than less information, especially when the marginal costs of handling and interpreting additional information are low, as they are for most graphics. The simple things belong in tables or in the text; graphics can give a sense of large and complex data sets that cannot be managed in any other way. If the graphic becomes overcrowded (although several thousand numbers represented may be just fine), a variety of data-reduction techniques—averaging, clustering, smoothing—can thin the numbers out before plotting.[4] Summary graphics can emerge from high-information displays, but there is nowhere to go if we begin with a low-information design.

[4] Paul A. Tukey and John W. Tukey, "Summarization; Smoothing; Supplemented Views," in Vic Barnett, ed., *Interpreting Multivariate Data* (Chichester, England, 1982), ch. 12; and William S. Cleveland, "Robust Locally Weighted Regression and Smoothing Scatterplots," *Journal of the American Statistical Association*, 74 (1979), 829–836, are recent papers in the large literature.

Data-rich designs give a context and credibility to statistical evidence. Low-information designs are suspect: what is left out, what is hidden, why are we shown so little? High-density graphics help us to compare parts of the data by displaying much information within the view of the eye: we look at one page at a time and the more on the page, the more effective and comparative our eye can be.[5] The principle, then, is:

> Maximize data density and the size of the data matrix, within reason.

High-information graphics must be designed with special care. As the volume of data increases, data measures must shrink (smaller dots for scatters, thinner lines for busy time-series). The clutter of

[5] It is suggested in the analysis of x-ray films to "search a reduced image so that the whole display can be perceived on at least one occasion without large eye movement." Edward Llewellyn Thomas, "Advice to the Searcher or What Do We Tell Them?" in Richard A. Monty and John W. Senders, eds., *Eye Movements and Psychological Processes* (Hillsdale, N.J., 1976), p. 349.

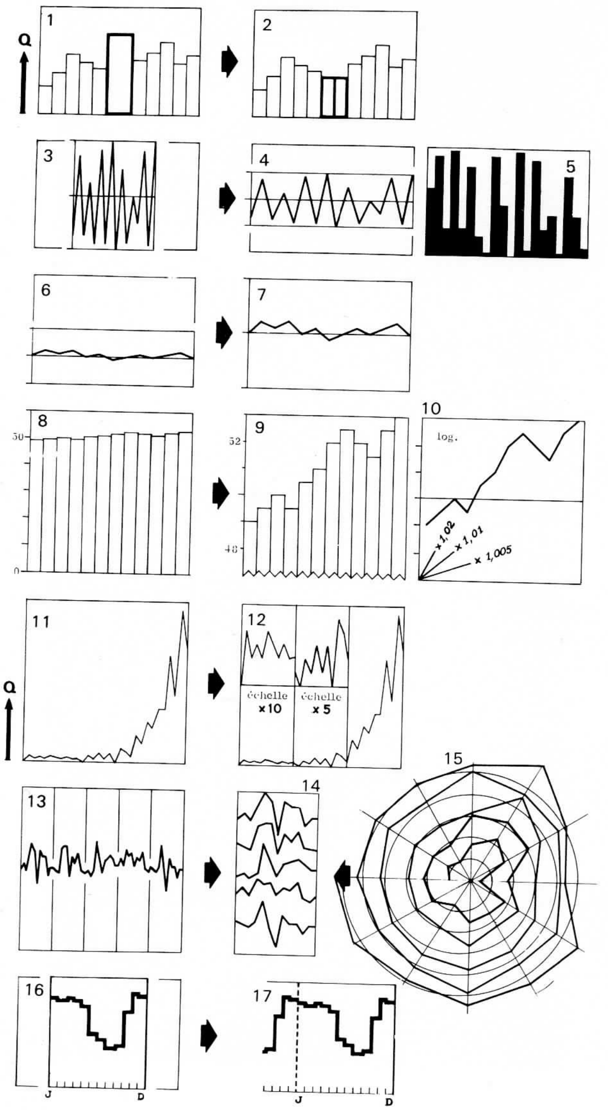
chartjunk, non-data-ink, and redundant data-ink is even more costly than usual in data-rich designs.

The way to increase data density other than by enlarging the data matrix is to reduce the area of a graphic. The Shrink Principle has wide application:

Graphics can be shrunk way down.

Many data graphics can be reduced in area to half their currently published size with virtually no loss in legibility and information. For example, Bertin's crisp and elegant line allows the display of 17 small-scale graphics on a single page along with extensive text. Repeated application of the Shrink Principle leads to a powerful and effective graphical design, the small multiple.

Jacques Bertin, *Semiologie Graphique* (Paris, second edition, 1973), p. 214.



PROBLEMES GRAPHIQUES POSES PAR LES CHRONIQUES

Un total sur deux cases (sur deux ans) doit être divisé par deux (1).
Un total pour six mois sera multiplié par deux dans des cases annuelles.

Courbes trop pointues, réduire l'échelle des Q; la sensibilité angulaire s'inscrit dans une zone moyenne autour de 70°.
Si la courbe n'est pas réductible (grandes et petites variations) employer les colonnes remplies (5).
Courbes trop plates : augmenter l'échelle des Q.

Variations très faibles par rapport au total.
Celui-ci perd de l'importance et le zéro peut être supprimé, à condition que le lecteur voit sa suppression (9). Le graphique peut être interprété comme une accélération si l'étude fine des variations est nécessaire (échelle logarithmique (10) (v. p. 240).

Très grande amplitude entre les valeurs extrêmes. Il faut admettre :
1°) Soit de ne pas percevoir les plus petites variations.
2°) Soit de ne s'intéresser qu'aux différences relatives (échelle logarithmique) sans connaître la quantité absolue.
3°) Soit admettre des périodes différentes dans la composante ordonnée et les traiter à des échelles différentes au-dessus de l'échelle commune (12).

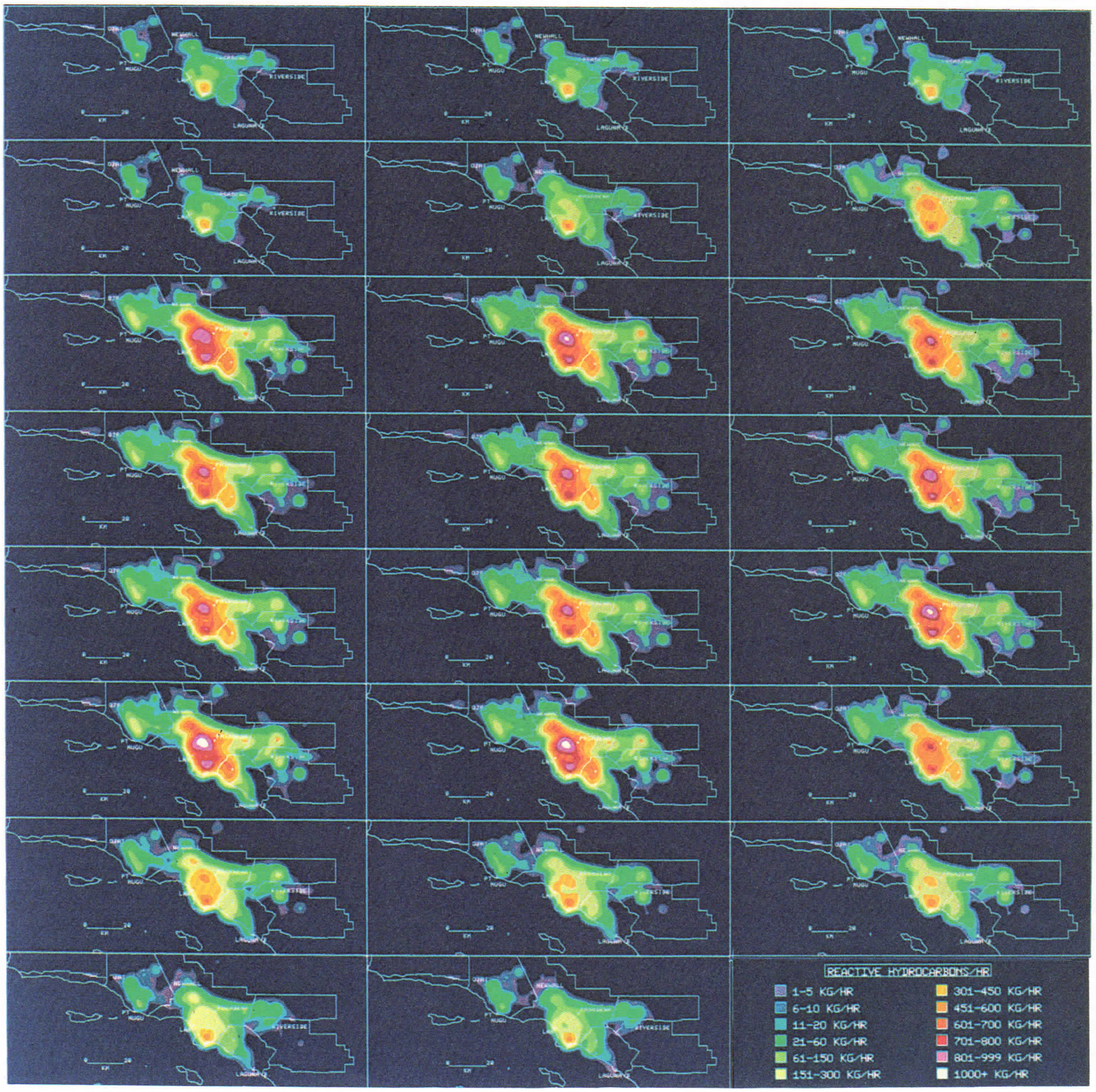
Cycles très marqués.
Si l'étude porte sur la comparaison des phases de chaque cycle, il est préférable de décomposer (13) de manière à superposer les cycles (14). La construction polaire peut être employée, de préférence dans une forme spirale (15) (ne pas commencer par un trop petit cercle); pour spectaculaire qu'elle soit, elle est moins efficace que la construction orthogonale.

Courbes annuelles de pluie ou de température. Un cycle possède deux phases (17), pourquoi n'en offrir qu'une à la perception du spectateur ? (16).

Small Multiples

Small multiples resemble the frames of a movie: a series of graphics, showing the same combination of variables, indexed by changes in another variable. Twenty-three hours of Los Angeles air pollution are organized into this display, based on a computer generated video tape. Shown is the hourly average distribution of reactive hydrocarbon emissions. The design remains constant through all the frames, so that attention is devoted entirely to shifts in the data:

From video tape by Gregory J. McRae, California Institute of Technology. The model is described in G. J. McRae, W. R. Goodin, and J. H. Seinfeld, "Development of a Second-Generation Mathematical Model for Urban Air Pollution. I. Model Formulation," *Atmospheric Environment*, 16 (1982), 679–696.



These grim small multiples show the distribution of occurrence of the cancer melanoma. The sites of 269 primary melanomas are recorded, along with the distribution between men and women. Note the data graphical arithmetic, similar to that of the multiwindow plot.

Arthur Wiskemann, "Zur Melanomentstehung durch chronische Lichteinwirkung," *Der Hautarzt*, 25 (1974), 21.

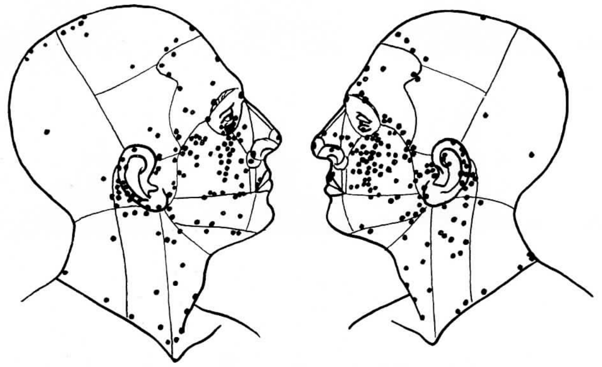


Abb. 1. Verteilung von 269 primären Melanomen auf Kopf und Hals

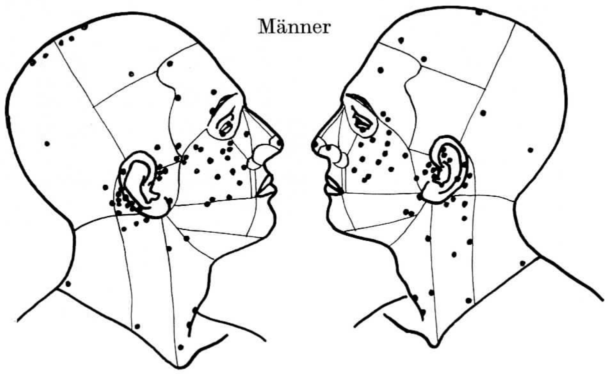


Abb. 2

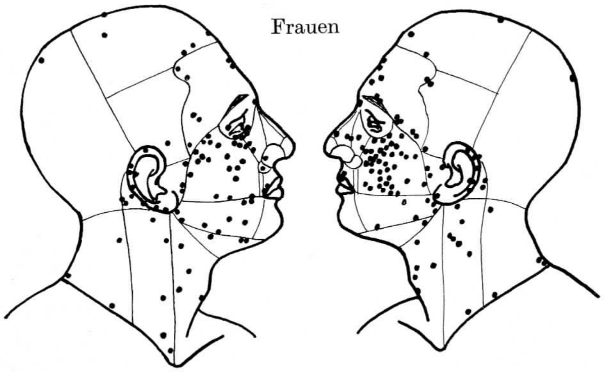
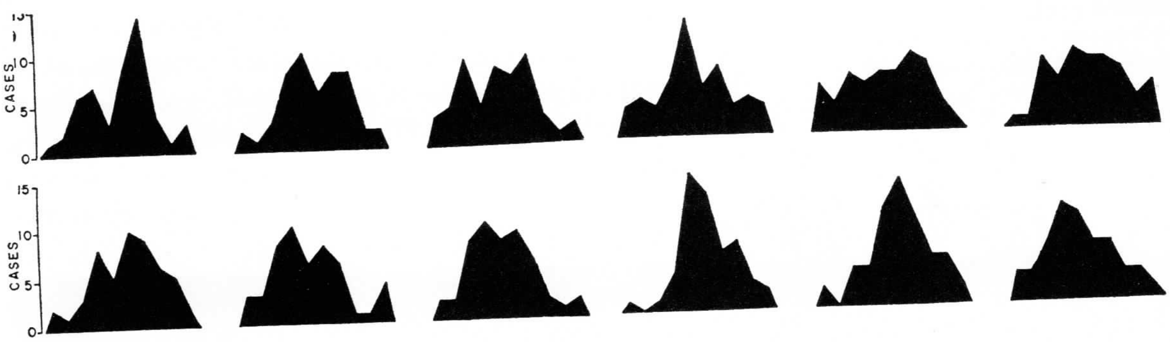


Abb. 3

Abb. 2 u. 3. Differenzierung der Melanomverteilung nach Geschlechtern

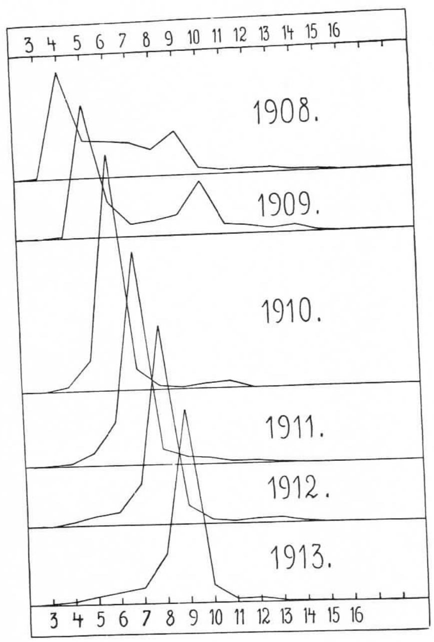
The effects of sampling errors are shown in these 12 distributions, each based on a sample of 50 random normal deviates:

Edmond A. Murphy, "One Cause? Many Causes? The Argument from the Bimodal Distribution," *Journal of Chronic Diseases*, 17 (1964), 309.



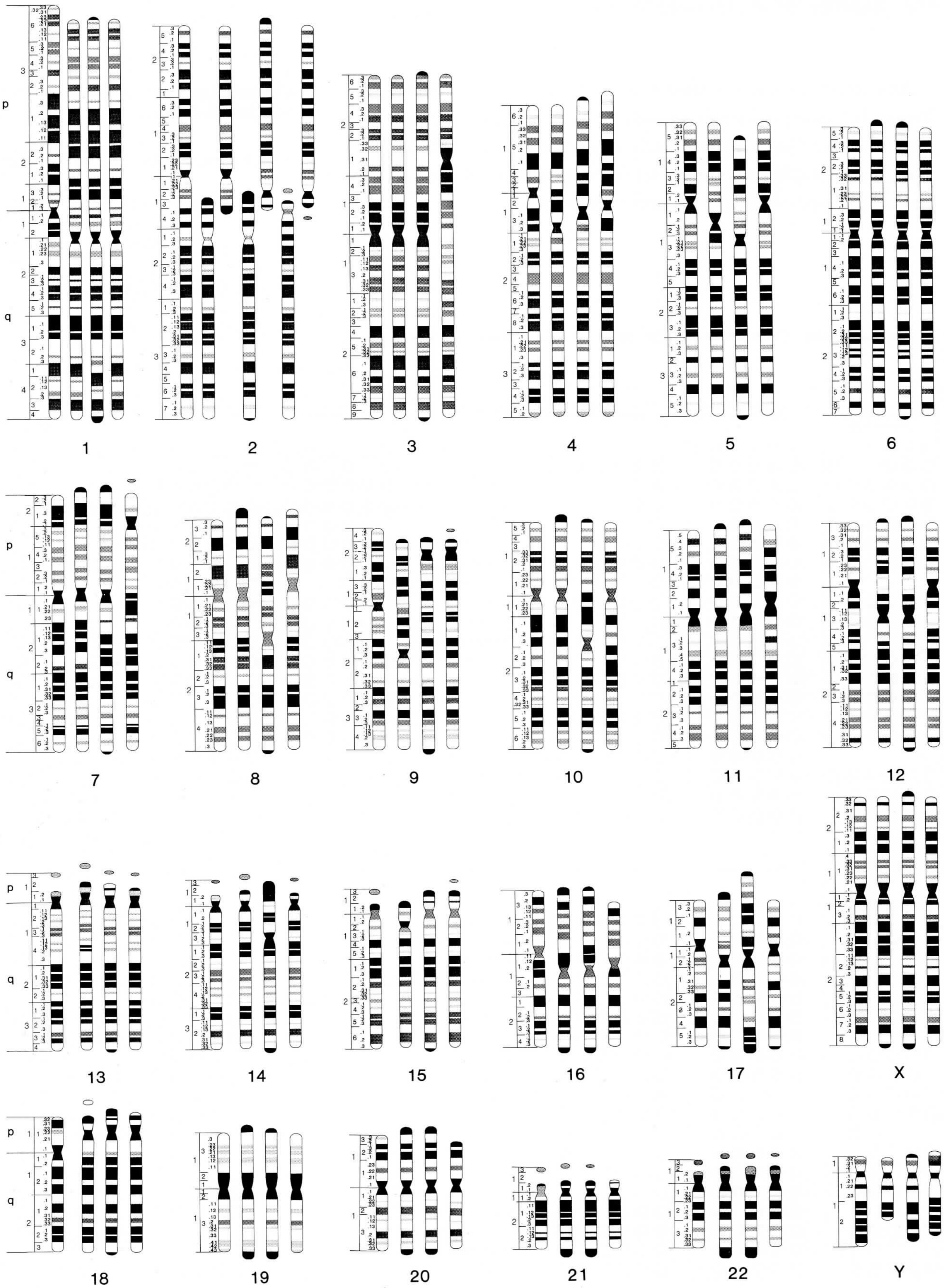
These six distributions show the age composition of herring catches each year from 1908 to 1913. A tremendous number of herring were spawned in 1904, and that class began to dominate the 1908 catch as four-year-olds, then the 1909 catch as five-year-olds, and so on:

Johan Hjort, "Fluctuations in the Great Fisheries of Northern Europe," *Rapports et Proces-Verbaux*, 20 (1914), in Susan Schlee, *The Edge of an Unfamiliar World* (New York, 1973), p. 226.



This next design compares a complex set of data: shown are the chromosomes of (from left to right) man, chimpanzee, gorilla, and orangutan. The similarities between humans and the great apes are to be noted.

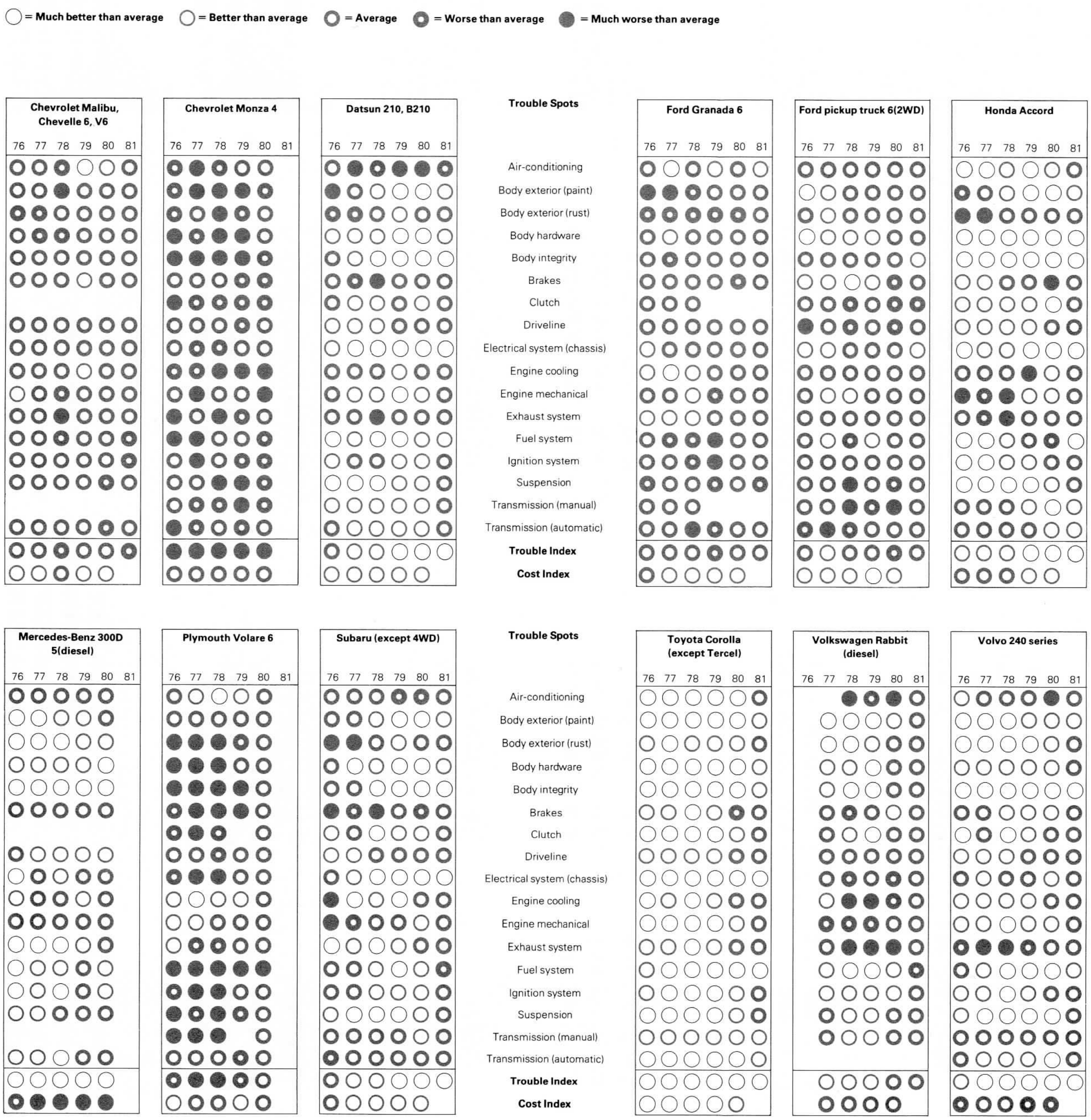
Jorge J. Yunis and Om Prakash, "The Origin of Man: A Chromosomal Pictorial Legacy," *Science*, 215 (March 19, 1982), 1527.



p
q
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
X
18
19
20
21
22
Y

And, finally, a visually similar small multiple, the *Consumer Reports* frequency-of-repair records for automobiles built from 1976 to 1981. This is a particularly ingenious mix of table and graphic, portraying a complex set of comparisons between manufacturers, types of cars, year, and trouble spots.

Consumer Reports, 47 (April 1982), 199–207. Redrawn.



Conclusion

Well-designed small multiples are

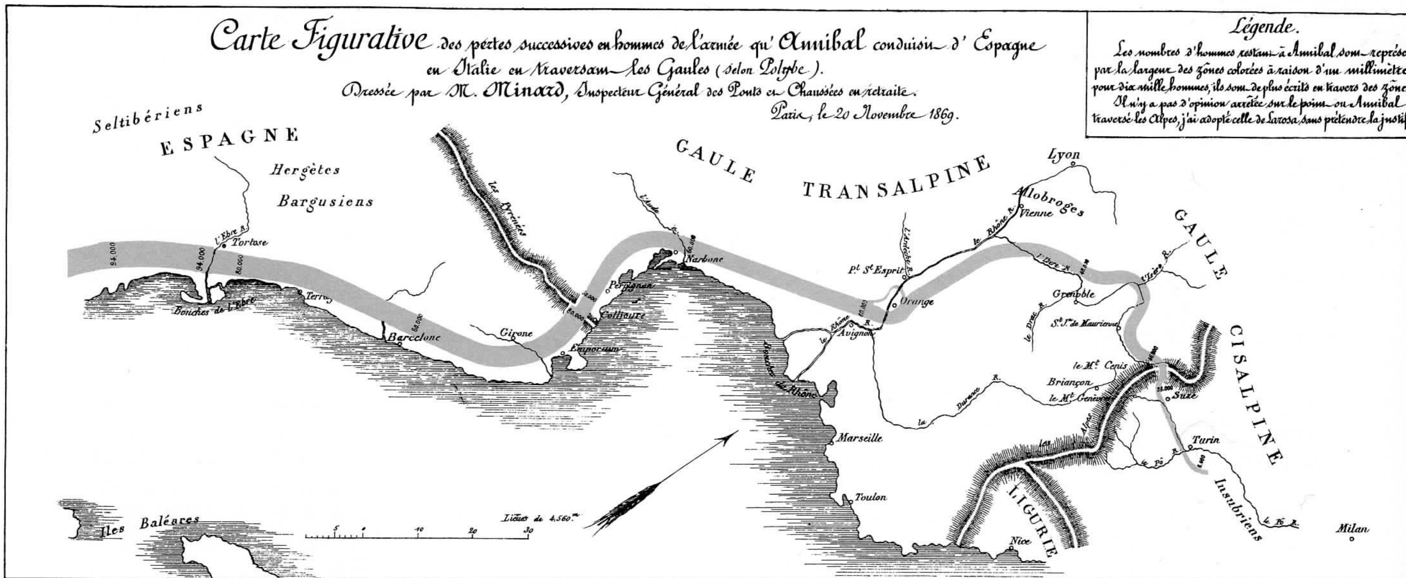
- inevitably comparative
- deftly multivariate
- shrunken, high-density graphics
- usually based on a large data matrix
- drawn almost entirely with data-ink
- efficient in interpretation
- often narrative in content, showing shifts in the relationship between variables as the index variable changes (thereby revealing interaction or multiplicative effects).

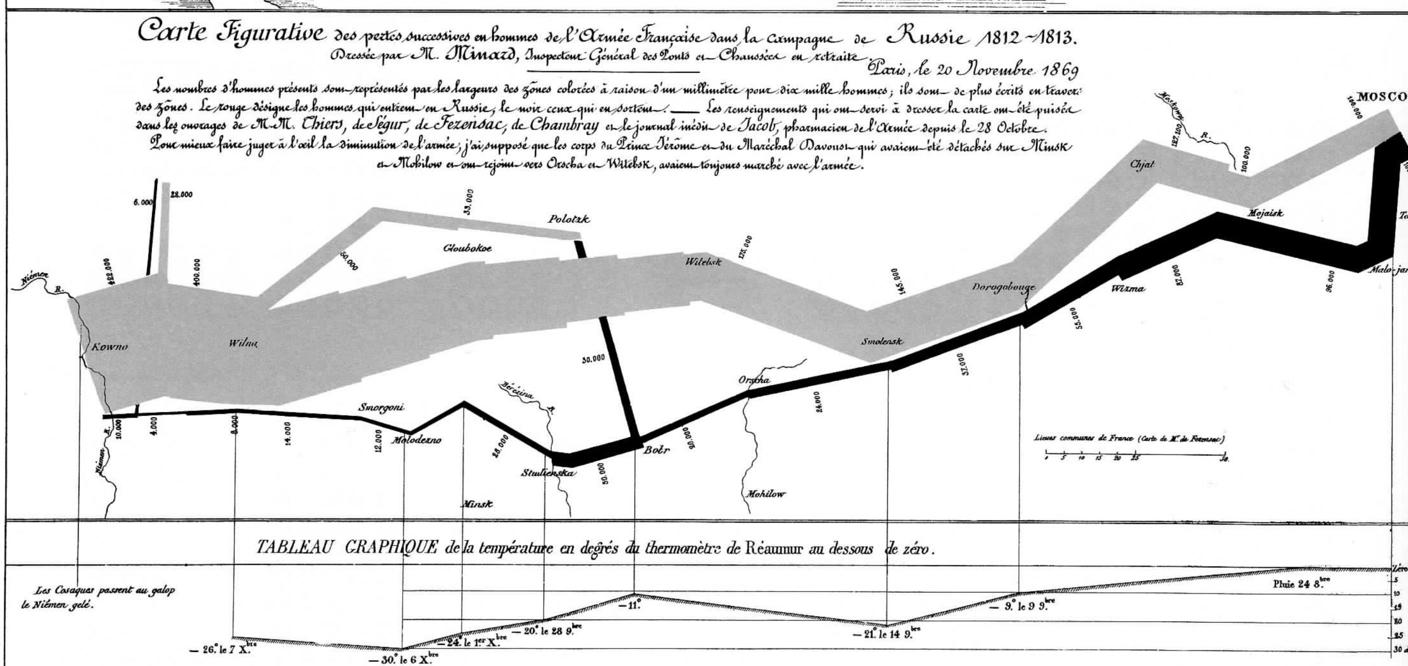
Small multiples reflect much of the theory of data graphics:

For non-data-ink, less is more.

For data-ink, less is a bore.[6]

[6] The two aphorisms on the meaning of "less" are, respectively, credited to Ludwig Mies van der Rohe and to Robert Venturi, *Complexity and Contradiction in Architecture* (New York, second edition, 1977), p. 17.





Autog. par Regnier, 8. Pas. Ste Marie St Gain à Paris.

Imp. Lith. Regnier et

9 *Aesthetics and Technique in Data Graphical Design*

Minard drew at least two versions of Napoleon's march to Moscow, the second in color with additional text describing the data sources. Another "Carte Figurative" was added in the 1869 plate, this one portraying Hannibal's campaign in Spain, Gaul, and Northern Italy. Minard's refined use of color contrasts with the brutal tones often seen in current-day graphics.

Charles Joseph Minard, *Tableaux Graphiques et Cartes Figuratives de M. Minard, 1845–1869*, a portfolio of his work held by the Bibliothèque de l'École Nationale des Ponts et Chaussées, Paris.

What makes for such graphical elegance? What accounts for the quality of Minard's graphics, of those of Playfair and Marey, and of some recent work, such as the new view of the galaxies? Good design has two key elements:

> Graphical elegance is often found in
> simplicity of design and complexity of data.

Visually attractive graphics also gather their power from content and interpretations beyond the immediate display of some numbers. The best graphics are about the useful and important, about life and death, about the universe. Beautiful graphics do not traffic with the trivial.

On rare occasions graphical architecture combines with the data content to yield a uniquely spectacular graphic. Such performances can be described and admired but there are no compositional principles on how to create that one wonderful graphic in a million. As Ben Shahn once said, "Aesthetics is for the artist like ornithology is for the birds."

What can be suggested, though, are some guides for enhancing the visual quality of the more routine, workaday designs. Attractive displays of statistical information

- have a properly chosen format and design
- use words, numbers, and drawing together
- reflect a balance, a proportion, a sense of relevant scale
- display an accessible complexity of detail
- often have a narrative quality, a story to tell about the data
- are drawn in a professional manner, with the technical details of production done with care
- avoid content-free decoration, including chartjunk.

The Choice of Design: Sentences, Text-Tables, Tables, Semi-Graphics, and Graphics

The substantive content, extensiveness of labels, and volume and ordering of data all help determine the choice of method for the display of quantitative materials. The basic structures for showing data are the sentence, the table, and the graphic. Often two or three of these devices should be combined.

The conventional sentence is a poor way to show more than two numbers because it prevents comparisons within the data. The linearly organized flow of words, folded over at arbitrary points (decided not by content but by the happenstance of column width), offers less than one effective dimension for organizing the data. Instead of:

> Nearly 53 percent of the type A group did something or other compared to 46 percent of B and slightly more than 57 percent of C.

Arrange the type to facilitate comparisons, as in this *text-table*:

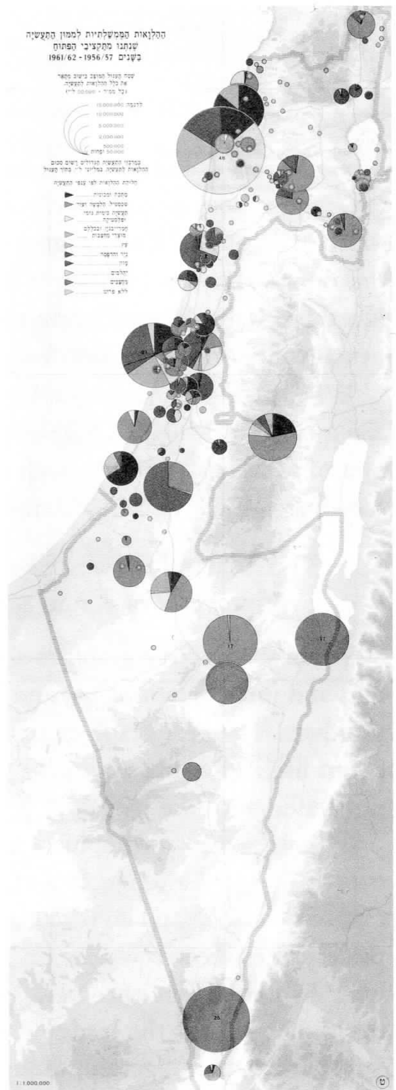
> The three groups differed in how they did something or other:

Group A	53%
Group B	46%
Group C	57%

There are nearly always better sequences than alphabetical—for example, ordering by content or by data values:

Group B	46%
Group A	53%
Group C	57%

Tables are clearly the best way to show exact numerical values, although the entries can also be arranged in semi-graphical form. Tables are preferable to graphics for many small data sets.[1] A table is nearly always better than a dumb pie chart; the only worse design than a pie chart is several of them, for then the viewer is asked to compare quantities located in spatial disarray both within and between pies, as in this heavily encoded example from an atlas. Given their low data-density and failure to order numbers along a visual dimension, pie charts should never be used.[2]



Department of Surveys, Ministry of Labour, *Atlas of Israel* (Jerusalem, 1956–), vol. 8, p. 8.

[1] On the design of tables, see A.S.C. Ehrenberg, "Rudiments of Numeracy," *Journal of the Royal Statistical Society*, A, 140 (1977), 277–297.

[2] This point is made decisively in Jacques Bertin, *Graphics and Graphic Information Processing* (Berlin, 1981). Bertin describes multiple pie charts as "completely useless" (p. 111).

Tables also work well when the data presentation requires many localized comparisons. In this 410-number table that I designed for the *New York Times* to show how different people voted in presidential elections in the United States, comparisons between the elections of 1980 and 1976 are read across each line; within-election analysis is conducted by reading downward in the clusters of three to seven lines. The horizontal rules divide the data into topical paragraphs; the rows are ordered so as to tell an ordered story about the elections. This type of elaborate table, a *supertable*, is likely to attract and intrigue readers through its organized, sequential detail and reference-like quality. One supertable is far better than a hundred little bar charts.

How Different Groups Voted for President

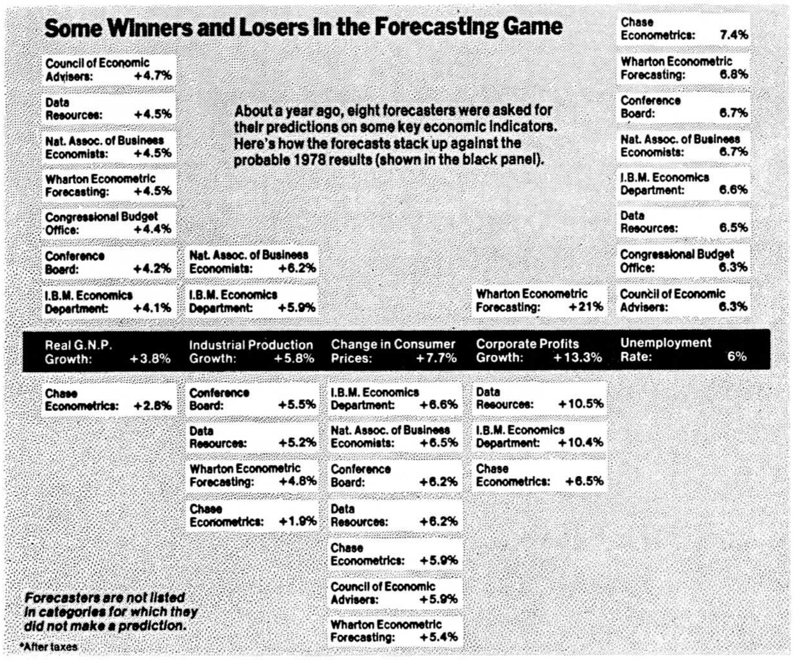
Based on 12,782 interviews with voters at their polling places. Shown is how each group divided its vote for President and, in parentheses, the percentage of the electorate belonging to each group.

	CARTER	REAGAN	ANDERSON	CARTER-FORD in 1976
Democrats (43%)	66	26	6	77 - 22
Independents (23%)	30	54	12	43 - 54
Republicans (28%)	11	84	4	9 - 90
Liberals (17%)	57	27	11	70 - 26
Moderates (46%)	42	48	8	51 - 48
Conservatives (28%)	23	71	4	29 - 70
Liberal Democrats (9%)	70	14	13	86 - 12
Moderate Democrats (22%)	66	28	6	77 - 22
Conservative Democrats (8%)	53	41	4	64 - 35
Politically active Democrats (3%)	72	19	8	—
Democrats favoring Kennedy in primaries (13%)	66	24	8	—
Liberal Independents (4%)	50	29	15	64 - 29
Moderate Independents (12%)	31	53	13	45 - 53
Conservative Independents (7%)	22	69	6	26 - 72
Liberal Republicans (2%)	25	66	9	17 - 82
Moderate Republicans (11%)	13	81	5	11 - 88
Conservative Republicans (12%)	6	91	2	6 - 93
Politically active Republicans (2%)	5	89	6	—
East (32%)	43	47	8	51 - 47
South (27%)	44	51	3	54 - 45
Midwest (20%)	41	51	6	48 - 50
West (11%)	35	52	10	46 - 51
Blacks (10%)	82	14	3	82 - 16
Hispanics (2%)	54	36	7	75 - 24
Whites(88%)	36	55	8	47 - 52
Female (49%)	45	46	7	50 - 48
Male (51%)	37	54	7	50 - 48
Female, favors equal rights amendment (22%)	54	32	11	—
Female, opposes equal rights amendment (15%)	29	66	4	—
Catholic (25%)	40	51	7	54 - 44
Jewish (5%)	45	39	14	64 - 34
Protestant (46%)	37	56	6	44 - 55
Born-again white Protestant (17%)	34	61	4	—
18 - 21 years old (6%)	44	43	11	48 - 50
22 - 29 years old (17%)	43	43	11	51 - 46
30 - 44 years old (31%)	37	54	7	49 - 49
45 - 59 years old (23%)	39	55	6	47 - 52
60 years or older (18%)	40	54	4	47 - 52
Family income				
Less than $10,000 (13%)	50	41	6	58 - 40
$10,000 - $14,999 (14%)	47	42	8	55 - 43
$15,000 - $24,999 (30%)	38	53	7	48 - 50
$25,000 - $50,000 (24%)	32	58	8	36 - 62
Over $50,000 (5%)	25	65	8	—
Professional or manager (40%)	33	56	9	41 - 57
Clerical, sales or other white-collar (11%)	42	48	8	46 - 53
Blue-collar worker (17%)	46	47	5	57 - 41
Agriculture (3%)	29	66	3	—
Looking for work (3%)	55	35	7	65 - 34
Education				
High school or less (39%)	46	48	4	57 - 43
Some college (28%)	35	55	8	51 - 49
College graduate (27%)	35	51	11	45 - 55
Labor union household (26%)	47	44	7	59 - 39
No member of household in union (62%)	35	55	8	43 - 55
Family finances				
Better off than a year ago (16%)	53	37	8	30 - 70
Same (40%)	46	46	7	51 - 49
Worse off than a year ago (34%)	25	64	8	77 - 23
Family finances and political party				
Democrats, better off than a year ago (7%)	77	16	6	69 - 31
Democrats, worse off than a year ago (13%)	47	39	10	94 - 6
Independents, better off (3%)	45	36	12	—
Independents, worse off (9%)	21	65	11	—
Republicans, better off (4%)	18	77	5	3 - 97
Republicans, worse off (11%)	6	89	4	24 - 76
More important problem				
Unemployment (39%)	51	40	7	75 - 25
Inflation (44%)	30	60	9	35 - 65
Feel that U.S. should be more forceful in dealing with Soviet Union even if it would increase the risk of war (54%)	28	64	6	—
Disagree (31%)	56	32	10	—
Favor equal rights amendment (46%)	49	38	11	—
Oppose equal rights amendment (35%)	26	68	4	—
When decided about choice				
Knew all along (41%)	47	50	2	44 - 55
During the primaries (13%)	30	60	8	57 - 42
During conventions (8%)	36	55	7	51 - 48
Since Labor Day (8%)	30	54	13	49 - 49
In week before election (23%)	38	46	13	49 - 47

Source: 1976 and 1980 election day surveys by The New York Times/CBS News Poll and 1976 election day survey by NBC News.

New York Times, November 9, 1980, p. A–28.

For sets of highly labeled numbers, a wordy data graphic—coming close to straight text—works well. This table of numbers is nicely organized into a graphic:



New York Times, January 2, 1979, p. D–3.

Making Complexity Accessible: Combining Words, Numbers, and Pictures

Explanations that give access to the richness of the data make graphics more attractive to the viewer. Words and pictures are sometimes jurisdictional enemies, as artists feud with writers for scarce space. An unfortunate legacy of these craft-union differences is the artificial separation of words and pictures; a few style sheets even forbid printing on graphics. What has gone wrong is that the techniques of production instead of the information conveyed have been given precedence.

Words and pictures belong together. Viewers need the help that words can provide. Words on graphics are data-ink, making effective use of the space freed up by erasing redundant and non-data-ink. It is nearly always helpful to write little messages on the plotting field to explain the data, to label outliers and interesting data points, to write equations and sometimes tables on the graphic itself, and to integrate the caption and legend into the design so that the eye is not required to dart back and forth between textual material and the graphic. (The size of type on and around graphics can be quite small, since the phrases and sentences are usually not too long—and therefore the small type will not fatigue viewers the way it does in lengthy texts.)

The principle of *data/text integration* is

> Data graphics are paragraphs about data and should be treated as such.

Words, graphics, and tables are different mechanisms with but a single purpose—the presentation of information. Why should the flow of information be broken up into different places on the page because the information is packaged one way or another? Sometimes it may be useful to have multiple story-lines or multiple levels of presentation, but that should be a deliberate design judgment, not something decided by conventional production requirements. Imagine if graphics were replaced by paragraphs of words and those paragraphs scattered over the pages out of sequence with the rest of the text—that is how graphical and tabular information is now treated in the layout of many published pages, particularly in scientific journals and professional books.

Tables and graphics should be run into the text whenever possible, avoiding the clumsy and diverting segregation of "See Fig. 2," (figures all too often located on the back of the adjacent page).[3] If a display is discussed in various parts of the text, it might well be printed afresh near each reference to it, perhaps in reduced size in later showings. The principle of text/graphic/table integration also suggests that the same typeface be used for text and graphic and, further, that ruled lines separating different types of information be avoided. Albert Biderman notes that illustrations were once well-integrated with text in scientific manuscripts, such as those of Newton and Leonardo da Vinci, but that statistical graphics became segregated from text and table as printing technology developed:

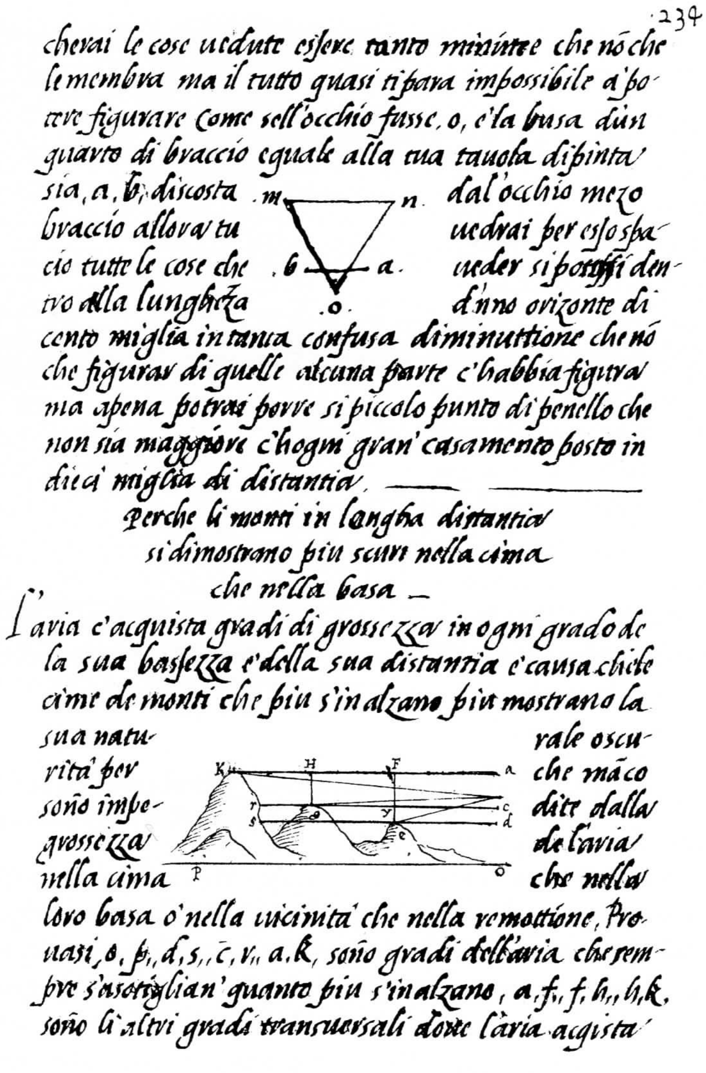
> The evolution of graphic methods as an element of the scientific enterprise has been handicapped by their adjunctive, segregated, and marginal position. The exigencies of typography that moved graphics to a segregated position in the printed work have in the past contributed to their intellectual segregation and marginality as well. There was a corresponding organizational segregation, with decisions on graphics often passing out of the hands of the original analyst and communicator into those of graphic specialists—the commercial artists and designers of graphic departments and audio-visual aids shops, for example, whose predilections and skills are usually more those of cosmeticians and merchandisers than of scientific analysts and communicators.[4]

3 "Fig.," often used to refer to graphics, is an ugly abbreviation and is not worth the two spaces saved.

4 Albert D. Biderman, "The Graph as a Victim of Adverse Discrimination and Segregation," *Information Design Journal*, 1 (1980), 238.

Page after page of Leonardo's manuscripts have a gentle but thorough integration of text and figure, a quality rarely seen in modern work:

Leonardo da Vinci, *Treatise on Painting* [*Codex Urbinas Latinus 1270*], vol. 2, facsimile (Princeton, 1956), p. 234, paragraph 827.


·234·
chevai le cose uedute essere tanto minute che nó che
le membra ma il tutto quasi ti para impossibile a po-
tere figurare Come sell'occhio fusse. o, c'la busa d'un
quarto di braccio equale alla tua tauola dipinta
sia, a, b, discosta .m n. dal'occhio mezo
braccio allora tu uedrai per esso spa-
cio tutte le cose che .b a. ueder si possi den-
tro alla lunghezza .o. d'uno orizonte di
cento miglia in tanta confusa diminuttione che nó
che figura di quelle alcuna parte c'habbia figura
ma apena potrai porre si piccolo punto di penello che
non sia maggiore c'hogni gran casamento posto in
dieci miglia di distantia.
Perche li monti in longha distantia
si dimostrano piu scuri nella cima
che nella basa
L'aria c'acquista gradi di grossezza in ogni grado de
la sua bassezza e della sua distantia e causa che le
cime de monti che piu s'inalzano piu mostrano la
sua natu- rale oscu-
rita per che mãco
sono impe- dite dalla
grossezza de l'aria
nella cima che nella
loro basa o nella uicinita che nella remottione, Pro-
uasi, o, p,, d, s,, c, r,, a, k, sono gradi dell'aria che sem-
pre s'assottiglian quanto piu s'inalzano, a, f,, f, h,, h, k,
sono li altri gradi transuersali doue l'aria acquista

Finally, a caveat: the use of words and pictures together requires a special sensitivity to the purpose of the design—in particular, whether the graphic is primarily for communication and illustration of a settled finding or, in contrast, for the exploration of a data set. Words on and around graphics are highly effective—sometimes all too effective—in telling viewers how to allocate their attention to the various parts of the data display.[5] Thus, for graphics in exploratory data analysis, words should tell the viewer *how* to read the design (if it is a technically complex arrangement) and not *what* to read in terms of content.

[5] Experiments in visual perception indicate that word instructions substantially determine eye movements in viewing pictures. See John D. Gould, "Looking at Pictures," in Richard A. Monty and John W. Senders, eds., *Eye Movements and Psychological Processes* (Hillsdale, N.J., 1976), 323–343.

Accessible Complexity: The Friendly Data Graphic

An occasional data graphic displays such care in design that it is particularly accessible and open to the eye, as if the designer had the viewer in mind at every turn while constructing the graphic. This is the *friendly data graphic.*

There are many specific differences between friendly and unfriendly graphics:

Friendly	Unfriendly
words are spelled out, mysterious and elaborate encoding avoided	abbreviations abound, requiring the viewer to sort through text to decode abbreviations
words run from left to right, the usual direction for reading occidental languages	words run vertically, particularly along the Y-axis; words run in several different directions
little messages help explain data	graphic is cryptic, requires repeated references to scattered text
elaborately encoded shadings, cross-hatching, and colors are avoided; instead, labels are placed on the graphic itself; no legend is required	obscure codings require going back and forth between legend and graphic
graphic attracts viewer, provokes curiosity	graphic is repellent, filled with chartjunk
colors, if used, are chosen so that the color-deficient and color-blind (5 to 10 percent of viewers) can make sense of the graphic (blue can be distinguished from other colors by most color-deficient people)	design insensitive to color-deficient viewers; red and green used for essential contrasts
type is clear, precise, modest; lettering may be done by hand	type is clotted, overbearing
type is upper-and-lower case, with serifs	type is all capitals, sans serif

With regard to typography, Josef Albers writes:

> The concept that "the simpler the form of a letter the simpler its reading" was an obsession of beginning constructivism. It became something like a dogma, and is still followed by "modernistic" typographers. . . . Ophthalmology has disclosed that the more the letters are differentiated from each other, the easier is the reading. Without going into comparisons and details, it should be realized that words consisting of only capital letters present the most difficult reading—because of their equal height, equal volume, and, with most, their equal width. When comparing serif letters with sans-serif, the latter provide an uneasy reading. The fashionable preference for sans-serif in text shows neither historical nor practical competence.[6]

[6] Josef Albers, *Interaction of Color* (New Haven, 1963, revised edition 1975), p. 4.