Graphical Methods in Statistics — Source link

Stephen E. Fienberg

Institutions: University of Minnesota

Published on: 01 Nov 1979 - The American Statistician (Taylor & Francis Group)

Topics: Statistical graphics and Statistical theory

GRAPHICAL METHODS IN STATISTICS

by

Stephen E. Fienberg

Technical Report No. 304

October 31, 1977

Department of Applied Statistics

School of Statistics

University of Minnesota

St. Paul, Minnesota 55108

ABSTRACT

Graphical methods have played a central role in the development of statistical theory and practice. This presentation briefly reviews some of the highlights in the historical development of statistical graphics, and gives a simple taxonomy which can be used to characterize the current use of graphical methods. This taxonomy is used to describe the evolution of the use of graphics in some major statistical and related scientific journals.

Some recent advances in the use of graphical methods for statistical analysis are reviewed, and several graphical methods for the statistical presentation of data are illustrated, including the use of multi-colored maps.

KEY WORDS: Diagnostic plots; Graphical methods; History of statistics; Maps, statistical; Standards for graphics; Statistical graphics.

Author's Footnote:

Stephen E. Fienberg is Professor and Chairman, Department of Applied Statistics, School of Statistics, University of Minnesota, St. Paul, Minnesota 55108. This article is based on a General Methodology Lecture delivered at the 137th Annual Meeting of the American Statistical Association, Chicago, Illinois, August 18, 1977, and its preparation was supported in part by grant NIE-G-76-0096 from the National Institute of Education, U.S. Department of Health, Education, and Welfare. The opinions do not necessarily reflect the positions or policies of the National Institute of Education. The author wishes to acknowledge the assistance of Phillip Chapman and Stanley Wasserman in the preparation of this article, and the helpful comments and suggestions of Albert Biderman, Kinley Larntz, William Kruskal, and Howard Wainer.

1. INTRODUCTION

"----for no study is less alluring or more dry and tedious than statistics, unless the mind and imagination are set to work or that the person studying is particularly interested in the subject; which is seldom the case with young men in any rank in life."

These words were written 176 years ago by Wm. Playfair, one of the fathers of statistical graphics in his Statistical Breviary (1801). Playfair's purpose in developing his graphical representations of statistical data was to make the statistics a little more palatable.

We have come a long way since 1801. Charts and graphs now play an important role in data presentation. They are used in our textbooks and classrooms; they summarize data in our technical journals; they are playing an increasing role in government reports; they appear daily in our newspapers and popular magazines. In statistics, as a field, graphs and charts are used not only to summarize data, but also as diagnostic aids in analysis, to organize Monte Carlo results, and, of course, to display theoretical relations.

We have come far since the time of Playfair, but we have far to go. Actual practice in statistical graphics is highly varied, with good graphics being overwhelmed by distorted data presentation, cumbersome charts, and perplexing pictures. While advice on how and when to draw graphs is available, we have no theory of statistical graphics, nor, as Kruskal (1977) has noted, do we have a systematic body of experimental results to use as a guide. We have seen considerable innovation in graphics during the past 20 years, but the advances in statistical methodology have made room for even greater innovation in the future. These are the themes of this paper.

The qualities and values of charts and graphs as compared with textual and tabular forms of presentation have been succinctly summarized by Calvin Schmid(1954) in his Handbook of Graphic Presentation:

1. In comparison with other types of presentation, well-designed charts are more effective in creating interest and in appealing to the attention of the reader.
2. Visual relationships, as protrayed by charts and graphs, are more clearly grasped and more easily remembered.
3. The use of charts and graphs saves time, since the essential meaning of large masses of statistical data can be visualized at a glance.
4. Charts and graphs can provide a comprehensive picture of a problem that makes possible a more complete and better balanced understanding than could be derived from tabular or textual forms of presentation.
5. Charts and graphs can bring out hidden facts and relationships and can stimulate, as well as aid, analytical thinking and investigation.

This is, of course, what Playfair's work was all about. He said:

"---I have succeeded in proposing and putting in practice a new and useful mode of stating accounts,----------as much information may be obtained in five minutes as would require whole days to imprint on the memory, in a lasting manner, by a table of figures".

2. LANDMARKS IN THE HISTORY OF STATISTICAL GRAPHICS

A paper on graphical methods in statistics would be incomplete without some attention to historical development. This brief account owes much to the work of Beniger and Robyn (1977), and the social graphics project at the Bureau of Social Science Research (see also Fienberg and Franklin, 1975) led by Albert Biderman.

Although Sir Edmond Halley published the first known analysis of empirical data using a scatterplot (barometric pressure vs. elevation above sea level), it was only after the work of Crome and Playfair in the late 18th and early 19th centuries, that the use of graphs and charts for data display became accepted practice. Playfair gave us the bar chart in his Commercial and Political Atlas (1786), and the pie chart and circle graph in his Statistical Breviary (1801). His work provides excellent examples of good graphics; it conveys information and is pleasing to the eye.

Figures 1 through 4 are from Playfair's An Inquiry into the Permanent Causes of the Decline and Fall of Powerful and Wealthy Nations (1805), and they illustrate how far ahead of his time Playfair really was. The following material explains the information in each of the charts:

> Figure 1, "representing the rise and fall of all nations or countries, that have been particularly distinguished for wealth or power, is the first of the sort that ever was engraved, and has, therefore, not yet met with public approbation."
>
> "It is constructed to give a distinct view of the migrations of commerce and wealth in general. For a very accurate view, there are no materials in existence; neither would it lead to any very different

conclusion, if the proportional values were ascertained with the greatest accuracy.....I found the first rough draft gave me a better comprehension of the subject, than all that I had learnt from occasional reading, for half of my lifetime; and, on the supposition that what was of so much use to me, might be of some to others, I have given it with a tolerable degree of accuracy."

Figure 2 represents the increase of the annual revenues of England and France from the beginning of the 17th century to the present time. The top of the chart shows the monarchs of the two countries during this time span. The yellow line is the revenue of France (measured in pounds sterling), the red line is the revenue of England. The latter is broken into two parts, with the "interest of debt", shown in green, subtracted from total revenue to get the "free revenue", indicated in pink.

Figure 3 shows the amount of the exports and imports of England to and from all parts 1800 (sic) to 1805. Exports are indicated by the red line, imports by the yellow line, and the balance of trade is shaded green. Barely perceptable is a dotted line indicating "public revenue and expenditure."

Figure 4 gives the extent, population, and revenue of the principal nations in Europe in 1804. The circles are proportional to the areas of the countries or territories (the figures are on the chart as well). The red line on the left is the number of inhabitants; the yellow line on the right is the revenue in pounds. The scales for these lines are the same.

"The dotted lines, to connect the extremities of the lines of population and revenue, serve by their descent from right to left, or from left to right, to show how revenue and population are proportioned to each other.
The impression made by this chart is such that it is impossible not to see by what means Sweden and Denmark are of little importance, as to wealth or power; for, though population and territory are the original foundation of power, finances are the means of exerting it."

Playfair (1805, p. 190)

(Figures 1 - 4 go about here.)

Playfair recognized that, while charts save time, the idea of Schmid that they can allow large masses of data to visualized at a glance needs some qualification. At the beginning of the book from which the four figures were taken he notes:

"Opposite to each Chart are descriptions and explanations. The reader will find, five minutes attention to the principle on which they are constructed, a saving of much labour and time; but, without that trifling attention, he may as well look at a blank sheet of paper as at one of the Charts."

Playfair (1805, p. xvi)

Subsequent developments involved such famous names as, Bessel (graphic table), Fourier (cumulative frequency curve), and Quetelet (empirical mortality curves, graphs of frequency curves, plots of Histograms with limiting normal curves).

In 1849, Fletcher published the first statistical map (with tone wash) in a statistical journal, although so much had appeared elsewhere as early as 1819. Then, in 1857, Florence Nightingale introduced her "Coxcomb" chart to describe, by month, the causes of mortality in the British Army during the Crimean War. The "Coxcomb" is the forerunner of the modern-day Rose Chart, and other graphs used to show cyclic phenomena.

<u>The Statistical Atlas of the United States Based on the Results of the Ninth Census</u> (Walker 1874) contained the first examples of population pyramids and bilateral frequency polygons. The descendents of these graphical elders are among the most effective forms of graphical display.

Moving into the 20th century we find the Lorenz (1905) Curve published in <u>JASA</u>, which compares percentiles of cumulative distributions. Such a comparison of two cumulative distributions is the first example of what Wilk and Gnanadesikan (1968) have labelled as P-P plots.

3. PUBLISHED STANDARDS FOR GRAPHICS

With the rapid growth of graphic presentation came a professional concern for the need of standards. This concern was reflected in the proceedings of the International Statistical Congresses held in Europe from 1853 to 1876, and in abortive attempts at the sessions of the International Statistical Institute, at the beginning of the 20th century, to develop rules and standards for graphics.

Then, in 1914, as a result of invitations extended by the American Society of Mechanical Engineers, a number of national associations formed a joint committee on standards for graphic presentation. The committee's preliminary report, published in JASA in 1915, consisted of 17 basic rules of elementary graphic presentation, each illustrated by one or more figures. The rules are simple and direct, and several of them are just as relevant today as in 1915.

Ten of these rules pertain to the portrayal of time series data (with time on the horizontal axis), using arithmetic scales. We can thus see the kinds of charts and figures which dominated publications during the early part of this century.

The American Society of Mechanical Engineers continued this effort at standards, and has published various updates over the years. The emphasis, however, has remained on time-series charts, and there are few published lists of standards for other types of charts and graphs.

A rational set of graphical standards should be based on a theory for graphical presentation. Alas, we have no such theory, and the current prospects for its development remain dim.

4. CLASSIFICATION OF GRAPHICS

For a variety of reasons, it is useful to have a taxonomy of graphical forms. For example, to illustrate how the use of graphics has changed in our professional journals over the past 50 years, I wanted a means of dividing graphs and charts into groups of some sort. Being of scholarly persuasion I turned to the literature for some help in this matter.

Different authors have proposed different classification schemes over the years. Perhaps the most detailed scheme is due to Magill (1930), presented in his doctoral thesis in education at the University of Pennsylvania. His classification system contains 37 types of graphs with 3 types of subordinate features, and 12 types of maps with 33 subordinate and 62 associated features. He then proceeded to utilize this structure to describe the frequency of graphic forms used over the period of one year in a selection of advertisements, magazines, newspapers, and automobile instruction manuals. Unfortunately, his classification scheme was of little use for my purposes.

Schmid (1954) suggests that the basis for classifying charts and graphs must utilize one or more of the following criteria: (A) purpose, (B) circumstance of use, (C) type of comparison to be made, (D) form. Under purpose he lists the following three categories:

(1) Illustration,

(2) Analysis,

(3) Computation.

For the study of statistical journals described here, Schmid's lists of "forms" was far too restrictive and outdated, and to augment this list offered me the prospect of becoming a second Magill! The "types of comparisons" in Schmid's list were also not especially interesting for the task I had in mind. Finally, since the "circumstance of use" for the charts and graphs in my study was always to be the same, I was left with Schmid's list of three purposes.

I then turned to Tukey (1972) who suggests that there are three classes of graphs:

1. Those intended to show what has already been learned by some other technique (propaganda graphs).
2. Analytical graphs - to let us see what may be happening over and above what has already been described.
3. Those from which numbers are to be read off - substitutes for tables.

These are just the Schmid purposes, with Tukey's embellishments.

I continued my search, and was able to add only one further purpose, suggested by Tufte (1976):

4. Decoration - - graphs are pretty.

We have already seen the artistic beauty of Playfair's charts, and to describe them as decoration would almost be demeaning. Thus I chose not to include this purpose in my study.

5. USE OF GRAPHICS IN JASA AND BIOMETRIKA, 1920 - 1975

Phillip Chapman, a graduate student at Minnesota, and I have conducted a preliminary study of the evolution of the use of graphics in statistical journals subsequent to the 1915 standards report. Our purpose was not to assess the adherence to standards, but rather to determine whether the relative volume of usage of statistical graphics has changed over time and whether there has been a shift in the purposes to which the published graphs are being put, in particular a shift from illustration (and communication) to analysis (and exploration).

After some initial explorations we augmented the list of three purposes to six. Three of these were relevant for graphs not involving data:

I. Graphs depicting theoretical relationships, such as probability density functions, contours of multivariate densities, values of risk functions for different estimators, and theoretical descriptions of graphical methods.

II. Computational graphs and charts, used as substitutes for tables ---- e.g. Fox Chart's, nomograms, and especially charts with small detailed grid lines.

III. Non-numerical graphs and charts ---- e.g. maps, certain skull diagrams, Venn diagrams, flow charts.

For graphs involving data we focussed on the distinction between communication or summary, and analysis. For our purposes a graph displaying data or summarizing analyses was intended for communication, even when in addition to data summaries it contained a fitted theoretical curve. We interpreted analytical graphs to be ones actually involved in the analysis, and we required these to include, at a minimum, something beyond a straightforward examination of the traditional modes of data presentation.

The purpose of a graph involving data is not always apparent from the graph itself, and we spent many hours reading the accompanying text. Even then we had a large number of instances where the purpose was either both communication and analysis, or was difficult to determine precisely. Thus we finally decided to break the communications-analysis continuum into three categories:

IV. Graphs intended to display data and results of analysis --- e.g.; time series charts, histograms, results of Monte Carlo studies, scatter plots (even those with an accompanying regression line).

V. Plots and graphs with elements of both data display and analysis --- e.g., charts from older papers involving primitive forms of analysis; graphs of posterior distributions.

VI. Analytical graphs---residual plots, half-normal and other probability plots where conclusions are drawn directly from graph, graphical methods of performing calculations, spectrum estimates from time series.

We tried to make our classification of graphs consistent over time but, as the description of category V demonstrates, this was difficult. It is clear to us that what graphs were put into what categories is a function of our current perspective, and our personal biases.

Our preliminary study involved examing all graphs published in JASA and Biometrika during six five-year spans beginning with 1921-1925, and moving in ten-year increments up through 1971-1975. We chose these particular journals because they represent the two major English-speaking countries with substantial statistical professional groups, and because they have been published throughout the twentieth century. We chose five-year spans because of the distortions that could possibly result from idiosyncratic volumes or issues of journals. For example, one of the early issues of Biometrika which we examined contained primarily articles on skull measurements, and only very specialized graphs.

Tables 1 and 2 contain simple summaries of the relative volume and distribution of graphs and charts for both journals. These tables give two related measures of relative volume: (i) the number of graphs per 100 pages; (ii) the actual space taken up by the graphics as a percentage of total space. Special care is needed when interpreting these data because of changes in journal page size and format. The major change to be wary of is the JASA shift from 9" x 6" single column pages to 11" x 8½" double column pages in 1971. The second measure seems to handle this shift in a reasonable way.

(Tables 1 and 2 go about here.)

In Biometrika there has been roughly a constant volume of theoretical graphs over time, whereas in JASA there is a noticeable increase from 1941-45 to 1951-55. In both journals type II (computational) graphs play an important role mainly in the 1950's and 1960's, and the changes in the volume of non-numerical graphs over time is not especially interesting.

Since the focus of this paper is on graphical methods associated with data, I have prepared some bar charts contrasting the two journals in terms of graphs of type IV, V, and VI (communication, mixed, analysis). Figures 5 and 6 give a graphical display of these data. These graphs clearly show the decline in the use of statistical graphics during this century, at least within two of our major statistical journals.

(Figures 5 and 6 go about here.)

6. RECENT INNOVATIONS IN STATISTICAL GRAPHICS

Despite what may appear to be a prolonged decline in the use of graphics in statistical journals, the past 20 years has seen an almost astonishing increase in innovative graphical ideas for data display and analysis. The statistical groups at Princeton University and at Bell Telephone Laboratories have provided much of the leadership for this development of what might be called the "new statistical graphics". I would like to quickly review some of these innovations.

6.1 Graphs for Displaying Multidimensional Data

The rapid spread of the use of computers for statistical analysis in the early 1960's led to an upsurge in work involving multivariate analysis. This, in turn, led to various proposals for representing multidimensional data in only two dimensions.

Anderson (1960) developed his method of using "glyphs" and "metroglyphs", which are circles of fixed radius with rays of various lengths representing the values of different variables. When the glyphs are plotted as points in a two-dimensional scatter plot we get a representation of (K + 2) - dimensional data, where K is the number of possible rays. There are many variants to the glyph technique, involving the plotting of triangles (Pickett and White, 1966), k-sided polygons (Siegel, Goldwyn, and Friedman, 1971) and weathervanes (Cleveland and Kleiner, 1974). Figure 7 shows a set of STARS (a version of the k-sided polygons) produced using TROLL, a computer system developed by the National Bureau of Economic Research (NBER) Computer Research Center for Economics and Management Science. Welsch (1976) describes the standard TROLL graphic capabilities, as well as a series of experimental graphic devices including STARS.

The data in Table 3 are taken from Ashton, Healy, and Lipton (1957), who used graphical techniques to compare measurements on the teeth of fossils and different "races" of men and apes. Andrews (1972) also used an excerpt of these data to produce a plot that we consider below. Here we use the same data-set as Andrews involving 8 measurements on the permanent first lower premolar. The values displayed are not the original measurements, but rather are the 8 canonical variables produced from the data on the men and apes, in order to maximize the between sum of squares relative to the within. In Table 3 we have the group means of the values of the canonical variables for the men and the apes.

(Table 3 and Figure 7 go about here.)

The STARS corresponding to the 9 "observations" are given in Figure 7. Each canonical variable is located along one of the 8 rays, beginning with variable 1 located at 3 o'clock, and running counter-clockwise. The polygon links the actual values of the coordinates for the observation, the circle is included for reference purposes, and the barely visible tick marks indicate the means for the 9 observations. The rays, circle and ticks are thus the same in each STAR.

An examination of the first nine STARS suggests that 1, 2 and 3 (the humans) form one group, 3, 4, 5 and 6, (the gorillas and orangutangs) form a second, and 8 and 9 (the chimpanzees) form a third. One should note how most of the separation into groups is based on the values of the first two canonical variates (these are the ones with the largest eigenvalues).

Andrews (1972) suggested representing k-tuple, $\underset{\sim}{x} = (x_1, x_2, ---x_K)$, via the finite Fourier series

$$f_{\underset{\sim}{x}}(t) = x_1/\sqrt{2} + x_2 \text{Sin } t + x_3 \text{ Cos } t + x_4 \text{ Sin } 2t + x_5 \text{ Cos } 2t + ---.$$

He then plotted $f_{\underset{\sim}{x}}(t)$ over the range $-\pi \leq t \leq \pi$ for each point $\underset{\sim}{x}$, for the 9 points in Table 3, and produced the graph in Figure 8. The graph distinguishes different values for humans (here labelled A, B, C), the gorillas and orangutangs (D, E, F, G) and the chimpanzees (H, I). These groups are the same as those we arrived at using the STARS. Note that the humans have been separated from the apes, and that at t_2 and t_4 the humans have a precise value, whereas the apes converge into their two groups at t_1. In his article Andrews goes on to develop significance tests and confidence intervals to make comparisons on the plots.

(Figure 8 goes about here.)

Noting that people grow up studying and reacting to faces, Chernoff (1973) proposed representing a point in 18-dimensional space by drawing a face whose 18 characteristics (such as length of nose, shape of face, curvature of mouth, size of eyes, etc.) are determined by the coordinates or position of the point.

Both Chernoff's faces and Andrew's Fourier plots are affected by interchanging coordinates. Thus a variety of displays may need to be tried before one can arrive at the best one for a given data-set. Chernoff and Rizvi (1975) report on an experiment involving random permutations in the assignment of coordinates to the 18 facial features, in a problem involving 36 observations from two multivariate normal populations, with approximately

18 observations from each population. They concluded that random permutations tend to affect the error rate in a classification task by a factor of about 25 percent. Their study did not, however, evaluate the efficacy of specific features, e.g. the eyes or the mouth.

To check on the implications of the Chermoff-Rizvi study, and to see how Chernoff's faces work in practice, I used the FACES program in TROLL with the same data as for the STARS and for Andrew's Fourier plot ---, the 9 observations on 8 variables from Table 3.

The use of the FACES program in TROLL turns out to be somewhat complicated when you have less than 18 variables. Unless you explicitly instruct the program to the contrary it assigns multiple features to each variable even though you assign only one. In two abortive attempts to draw FACES for the 3 human and 6 ape groups, the program internally assigned 13 characteristics to the 8 variates, even though I specified the assignment of only 8 characteristics.

Several additional attempts at the construction of faces with only 8 characteristics led to groupings of faces quite different from those I expected. One particular one was strongly influenced by the 8th canonical variate and led to two groups of apes, the females and the males! The final set of FACES I produced is included here as Figure 9. The 8 canonical variables are represented by the following facial characteristics: (1) face shape, (2) jaw shape, (3) eye size, (4) eye position, (5) pupil position, (6) forehead shape, (7) eyebrow slant, (8) mouth shape. Figure 9 does a moderately good job of producing the same three groups as we identified with the other graphic methods. Whether we would have stumbled across this grouping had we not been explicitly looking for it is another matter. This one experience with FACES suggests that their use requires considerable skill and experience.

(Figure 9 goes about here.)

6.2 Graphical Aids to Analysis-Diagnostic Plots

The multidimensional data plots in Section 6.1 are examples of computer-generated graphics that would have been either impractical or totally impossible to draw without the aid of the computer. Another area where the availability of computer-generated graphics has provided the impetus for innovative developments is diagnostic plots. These plots typically involve some form of data transformation and rescaling so that comparisons and deviations can be measured from a straight line. Some examples of these diagnostic plots are:

(A) Residual plots of various sorts, following on the suggestions of Anscombe and Tukey (1963),

(B) C_P plots for choosing subset regressions as suggested by Colin Mallows (see figures in Gorman and Tomin, 1968),

(C) Q-Q (Quantile-Quantile) and P-P (Percent-Percent) plots for comparing two distribution functions (Wilk and Gnanadesikan, 1968), as well as various hybrid plots,

(D) Diagnostic displays for robust regression (Denby and Mallows, 1977), showing the effects of varing the trimming parameter on the adjustment of residuals, and the values of regression coefficients,

(E) Three-dimensional isometric plots of changing periodograms (Blackstone and Bingham, 1974).

The use of diagnostic plots such as these are direct aids in data-analysis. In the course of analysis of a given data-set one typically needs to look at several different diagnostic plots. The computer makes such examination a reasonable task.

6.3 Semi-Graphic Displays

Not all recent innovations in graphics require the availability of sophisticated computers. Indeed, Tukey (1972, 1977) has suggested several semigraphical displays, that attempt to blur the distinction between "Table" and "graph", and that are easily prepared by hand at home or on the commuter train. The most well-known of these are the stem-and-leaf display, which is an alternative to tallying values into frequency distributions, and box-and-whisker plots.

7. STATISTICAL MAPS

Colored statistical maps have been in widespread use since the mid-nineteenth century, but there have been some recent advances and innovations. I will describe one of these, pioneered by the U.S. Bureau of the Census (see Meyer, Broome, and Schweitzer, 1975), the two-variable colored "cross" map. This type of map is intended to convey the spacial distribution of two variables and the geographic concentration of their relationship. Only an example does the method justice (or injustice, depending on your point of view). Figures 10, 11, and 12 are from the August 1976 issue of STATUS, (a now-defunct monthly chartbook of social and economic trends produced by the U.S. Bureau of the Census).

(Figures 10, 11 and 12 go about here.)

We begin by examining the two variables of study separately. First we have, in figure 10, the death rate from cardiovascular disease among males age 35-74 (1968-1971)---dark blue is high, yellow low. The data are displayed by county. The low death rates are concentrated in the western half of the country. Next we have, in Figure 11, a measure of overcrowded housing, the percentage of units with 1.01 or more persons per room (1970)---dark red is high, and again, yellow is low. The bivariate map is created by an overlay process, and there are 16 resulting colors representing the combinations of the variables, as in Figure 12. How does one interpret this map? The instructions in STATUS note:

"If the geographic relationships were random, the resulting map would show no particular tendency toward an areal concentration of similar colors, but instead would exhibit a patchwork of small contrasting color blocks throughout the country.

Examination of the map shows that there is, indeed, a geographic variation in the distribution of male cardiovascular mortality and overcrowded housing. The 16 individual colors which make up the map appear to be concentrated in sizable groups of contiguous counties."

This statement is, of course, a half-truth. Just as independence in an RXC contingency table doesn't lead to expect cell values of the same size, because of marginal structure, so too here the marginal univariate structure leads to non-random patterns.

Do not feel dismayed if you are having trouble figuring out what's going on in Figure 12. It takes considerable practice to learn to discriminate among the 16 colors and to organize the spatial bivariate relationship, even for those of us not afflicted with colorblindness. There are a variety of issues associated with the use of such maps that need to be resolved:

(1) Choice of class intervals,

(2) Choice of colors---note the colors do have to be matched,

(3) The number of classes to be used,

(4) Is the two-color system superior to a single color system and geometric patterning?

(5) Can individuals extract additional information from the bivariate map, over and above that which they can extract from the two univariate maps put side by side?

Tukey (1975) has suggested that:

"So called "Statistical maps" do not deserve so honored a name. "Patch maps" is more appropriate. We can, and must do better by assigning values to centers rathern than areas, by learning to adjust for area compositions, by bringing in spatial smoothing."

Tukey goes on to give detailed suggestions for improvements.

8. STATISTICAL EXPERIMENTS WITH STATISTICAL GRAPHICS

In the late 1920's and early 1930's F.E. Croxton and others published a series of papers in JASA reporting on studies comparing the relative merits of circles, bars, squares, and cubes for certain types of displays. The conclusions from these early attempts at experimentation were inconclusive and contradictory. I have been unable to locate any further work on experimentation (except on maps) until recent years. Earlier I mentioned the recent Chernoff-Rizvi (1975) experiment with faces. Also, it has been reported that census is doing various experiments with their two-color maps.

Wm. Kruskal has drawn my attention to a very carefully done study of the use of dot area symbols in cartography by Castner and Robinson (1968). They thoroughly describe the characteristics of dot patterns and their perception, focussing on such features as form, size spacing, arrangement, orientation, and reflectance density. From their study of these characteristics they devise a series of tests to evaluate the effects of varying some of these characteristics. The actual experiments are not fancy in an experimental design sense, but the careful and almost systematic approach to the problem is worthy of study by anyone contemplating an experiment with graphical forms.

The other only recently published report on experiments with statistical graphics that I know of is Wainer and Reiser (1976), who studied the response time of subjects to questions about different graphical displays of the same data. Wainer and Reiser's experiment is sophisticated in many senses, but it too has severe shortcomings. What is clear to me is that the design of good experiments in this area will tax the minds of the best statisticians, and those well-versed in the psychology of perception.

9. WHERE DO WE GO FROM HERE?

We have come far since the time of Playfair, but we still have far to go. We know how to prepare some forms of statistical graphics well, yet in other areas we have much to learn. Where do we go from here?

Clearly one of the things we need in the area of statistical graphics is MORE. We need to educate our students and ourselves to make MORE and better use of known graphical devices. We also need MORE attempts at innovation; the examples I have shown do not suffice. Finally, we need MORE attempts at synthesis. Let me elaborate.

I have suggested to you that despite the recent flurry of graphical innovation, many of our statistical journals publish fewer graphs and charts than ever before. This must change. First, we must teach statisticians and others how and when to draw good graphical displays of data. Second, we must encourage them to use graphical methods in their work, and in the material they prepare for publication. Third, we must change the policies in our professional journals so that graphics are encouraged, not discouraged.

Many areas of statistical methodology and analysis could benefit from graphical innovations:

(1) We need further work on displaying multidimensional data,

(2) We have few effective display devices for aiding in the fitting of ANOVA models to measurement data, and loglinear models to categorical data. Special attention must be paid to devices that utilize the hierarchical structure of the parameters in these models.

(3) As Tukey has indicated, much more can be done with statistical maps to make them worthy of the name.

Before we can arrive at a theory for statistical graphs we need more attempts at synthesis, but before we can expect effective synthesis we need considerable experimentation. For a profession that gave rise to the design and analysis of experiments, we have done surprisingly little to foster careful controlled experimentation with graphical forms to aid us in arriving at an informed judgement on what constitutes good graphical presentation.

REFERENCES

Anderson, Edgar (1957). "A Semigraphical Method for the Analysis of Complex Problems," Proceedings of the National Academy of Science, 13, 923-7. Reprinted, with an appended note, in Technometrics 2 (1960) 387-391.

Andrews, D.F. (1972). "Plots of High Dimensional Data," Biometrics, 28, 125-36.

Anscombe, F.J., and Tukey, John W. (1963). "The Examination and Analysis of Residuals," Technometrics 5, 141-160.

Ashton, E.H., Healy, M.J.R., and Lipton, S. (1957). "The Descriptive Use of Discriminant Functions in Physical Anthropology," Proceedings of the Royal Society B 146, 552-572.

Bachi, Roberto (1968). Graphical Rational Patterns. A New Approach to Graphical Presentation of Statistics. Israel Universities Press, Jerusalem.

Beniger, James R. and Robyn, Dorothy L. (1976). "The History and Future of Graphics in Statistics." Proceedings of the Social Statistics Section Vol. I, American Statistical Association, 192-197.

Beniger, James R. and Robyn, Dorothy L. (1977). "Quantitative Graphics in Statistics: A Brief History," American Statistician (in press).

Blackstone, Eugene H. and Bingham, C. (1974). "Cleaning Time Series by Frequency Domain Regression on Noise," Eascon' 74, 586-591.

Castner, Henry W. and Robinson, Arthur H. (1969). Dot Area Symbols in Cartography: The Influence of Pattern on their Perception, American Congress on Surveying and Mapping, Cartography Division, Technical Monograph, No. CA-4.

Chernoff, Herman (1973). "Using Faces to Represent Points in K-dimensional Space Graphically," Journal of the American Statistical Association, 68, 361-368.

Chernoff, Herman and Rizvi, M. Haseeb (1975). "Effect on Classification Error of Random Permutations of Features in Representing Multivariate Data by Faces," Journal of the American Statistical Association, 70, 548-554.

Cleveland, Wm. S. and Kleiner, Beat (1974). "The Analysis of Air Pollution Data from New Jersey and New York." Paper presented at the American Statistical Association Meeting, St. Louis.

Croxton, Frederick E. and Stein, Harold (1932). "Graphic Comparisons by Bars, Squares, Circles and Cubes." Journal of the American Statistical Association 27, 54-60.

Croxton, Frederick, E. and Stryker, R.E. (1927). "Bar Charts Versus Circle Diagrams." Journal of the American Statistical Association 22, 473-482.

Denby, L. and Mallows, C.L. (1977). "Two Diagnostic Displays for Robust Regression Analysis." Technometrics 19, 1-13.

Eells, Walter Crosby (1926). "The Relative Merits of Circles and Bars for Representing Component Parts." Journal of the American Statistical Association 21, 119-132.

Fienberg, Barry M. and Franklin, Carolyn A. (1975). Social Graphics Bibliography. Bureau of Social Science Research, Washington, D.C.

Fletcher, Joseph (1849). "Moral and Educational Statistics of England and Wales." Journal of the Statistical Society of London, 12, 151-176 and 189-335.

Funkhouser, H. Gray (1937). "Historical Development of the Graphical Representation of Statistical Data." Osiris, 3, 269-405.

Funkhouser, H. Gray and Walker, Helen M. (1935). "Playfair and his Charts." Economic History, 3, 103-109.

Gorman, J.W. and Toman, R.J. (1966) "Selection of Variables for Fitting Equations to Data." Technometrics 8, 27-51.

Huhn, R. von (1927). "Further Studies in the Graphic Use of Circles and Bars I: A Discussion of the Eell's Experiment." Journal of the American Statistical Association, 22, 31-36.

Joint Committee on Standards for Graphic Presentation, "Preliminary Report." Journal of the American Statistical Association 14, 790-797.

Kruskal, William (1977). "Visions of Maps and Graphs." Proceedings of the International Symposium on Computer-assisted Cartography, 1975, 27-36.

Lorenz, M.O. (1905). "Methods of Measuring the Concentration of Wealth. Journal of the American Statistical Association, 9, 209-219.

Magill, Walter H. (1930). The Determination of the Graphical Forms and the Frequencies of the Forms Employed in the Current Reading Matter of the Non-Specialist. Ph.D. Dissertation, University of Pennsylvania.

Meyer, Morton A., Broome, Frederick R. and Schweitzer, Richard H. Jr. (1975). "Color Statistical Mapping by the U.S. Bureau of the Census." American Cartographer 2, (No. 2) 100-117.

Nightingale, Florence. Mortality of the British Army. London (1857).

Nightingale, Florence. Notes on Matters Affecting the Health, Efficiency and Hospital Administration of the British Army. London (1858).

Pickett, R. and White, B.W. (1966). "Constructing Data Pictures." Proceedings of the 7th National Symposium of Information Display, 75-81.

Playfair, William (1786). The Commercial and Political Atlas. London.

Playfair, William (1798). Lineal Arithmetic; Applied to Show the Progress of the Commerce and Revenue of England During the Present Century. London.

Playfair, William (1801). The Statistical Breviary. London.

Playfair, William (1805). An Inquiry in the Permanent Causes of the Decline and Fall of Powerful and Wealthy Nations. London

Playfair, William. A Letter on Our Agricultural Distresses, Their Causes and Remedies. London (1821).

Schmid, Calvin F. (1954). Handbook of Graphic Presentation. Ronald Press, New York.

Siegel, J.H., Goldwyn, R.M. and Friedman, H.P. (1971). "Pattern and Process of the Evolution of Human Septic Shock." Surgery 70, 232-245.

Tukey, John W. (1972). "Some Graphic and Semigraphic Displays." In Statistical Papers in Honor of George W. Snedecor (T.A. Bancroft, ed.), 293-316.

Tukey, John W. (1975). "Methodology, and the Statisticians Response for both Accuracy and Relevance." General Methodology Lecture, American Statistical Association Meetings, Atlanta, Georgia.

Tukey, John W. (1977). Exploratory Data Analysis. Addison-Wesley, Reading, Mass.

Tufte, Edward R. (1976). Unpublished notes.

U.S. Bureau of the Census (1976). STATUS, A Monthly Chartbook of Social and Economic Trends. (August issue).

Wainer, Howard (1977). "Data Display Graphical and Tabular: How and Why." Unpublished Manuscript.

Wainer, Howard and Reiser, Mark (1976). "Assessing the Efficacy of Visual Displays." Proceedings of the Social Statistics Section, Vol. I, American Statistical Association, 89-92.

Walker, Francis A. (1874). Statistical Atlas of the United States Based on the Results of the Ninth Census. Washington, D.C.

Welsch, Roy E. (1976). "Graphics for Data Analysis," Computers and Graphics 2, 31-37.

Wilk, Martin B., and Gnanadesikan, Ram (1968). "Probability Plotting Methods for the Analysis of Data." Biometrika 55, 1-17.

Table 1. Number of Charts per 100 pages in JASA and BIOMETRIKA

(a) JASA

	Purpose of Graph								
Years	I[a]	II	III	Non-Data Subtotals	IV	V	VI	Data Subtotals	Totals
1921-25	0.89	0.10	0.40	1.39	7.77	2.62	0.99	11.38	12.77
1931-35	1.19	0.15	0.58	1.92	6.74	1.08	0.96	8.78	10.70
1941-45	0.43	0.14	0.33	0.90	6.29	0.99	0.47	7.75	8.65
1951-55	2.46	1.42	0.41	4.29	2.62	0.96	0.19	3.77	8.06
1961-65[b]	2.84	0.34	0.32	3.50	1.88	0.64	1.03	3.55	7.05
1971-75	5.37	0.05	0.91	6.33	4.70	0.76	1.92	7.38	13.71

(b) BIOMETRIKA[d]

Years	I	II	III	Non-Data Subtotals	IV	V	VI	Data Subtotals	Totals
1921-25	1.75	0	0.71	2.46	8.25	0.66	0.09	9.00	11.46
1931-35	4.37	0.08	1.13	5.58	9.84	0.11	0.30	10.25	15.83
1941-45[e]	2.94	0.33	0.33	3.60	2.45	0.33	1.14	3.93	7.52
1951-55	2.62	0.44	0.53	3.59	1.41	0.78	0.53	2.22	5.81
1961-65	3.88	0.65	0.18	4.71	0.80	1.16	0.36	2.32	7.03
1971-75	3.55	0.06	0.35	3.96	1.32	0.41	0.53	2.26	6.22

[a] I = Theoretical curves; II = graphs for computation; III = non-numerical charts and diagrams; IV = data display and summary; V = graphs with mixture of display and analysis; VI = analytical graphs.

[b] Slight change in page size

no adjustments made.

[c] Change page size and format

[d] Biometrika has a different page size and format from those used by JASA.

[e] Slight increase in amount of text per page.

Table 2. Percentage of Space Devoted to Charts and Graphs in JASA and Biometrika

(a) JASA

Years	I	II	III	Non-Data Subtotals	IV	V	VI	Data Subtotals	Totals
1921-25	0.38	0.05	0.25	0.68	4.00	1.06	0.62	5.68	6.36
1931-35	0.48	0.14	0.23	0.85	3.48	0.61	0.37	4.46	4.69
1941-45	0.20	0.19	0.28	0.67	3.33	0.47	0.22	4.02	4.69
1951-55	1.39	1.22	0.23	2.84	1.71	0.70	0.11	2.52	5.36
1961-65	1.63	0.26	0.19	2.08	1.00	0.50	0.55	2.05	4.13
1971-75	1.03	0.02	0.20	1.25	1.02	0.18	0.37	1.57	2.82

(b) Biometrika

Years	I	II	III	Non-Data Subtotals	IV	V	VI	Data Subtotals	Totals
1921-25	0.63	0	0.44	1.07	5.62	0.48	0.04	6.14	7.21
1931-35	2.26	0.05	0.79	3.10	5.44	0.06	0.22	5.72	8.82
1941-45	1.37	0.23	0.08	1.68	2.45	0.20	0.52	3.17	4.85
1951-55	1.12	0.41	0.23	1.76	0.48	0.32	0.26	1.06	2.82
1961-65	1.74	0.43	0.04	2.21	0.32	0.49	0.16	0.97	3.18
1971-75	1.67	0.05	0.15	1.87	0.65	0.19	0.19	1.03	2.90

Table 3. Permanent First Lower Premolar
Coefficients of Canonical Variates for Means
of 8 Groups (Andrews, 1972)

A. West African	-8.09	+0.49	+0.18	+0.75	-0.06	-0.04	+0.04	+0.03
B. British	-9.37	-0.68	-0.44	-0.37	+0.37	+0.02	-0.01	+0.05
C. Australian aboriginal	-8.87	+1.44	+0.36	-0.34	-0.29	-0.02	-0.01	-0.05
D. gorilla: male	+6.28	+2.89	+0.43	-0.03	+0.10	-0.14	+0.07	+0.08
E. female	+4.82	+1.52	+0.71	-0.06	+0.25	+0.15	-0.07	-0.10
F. orang-outang: male	+5.11	+1.61	-0.72	+0.04	-0.17	+0.13	+0.03	+0.05
G. female	+3.60	+0.28	-1.05	+0.01	-0.03	-0.11	-0.11	-0.08
H. chimpanzee: male	+3.46	-3.37	+0.33	-0.32	-0.19	-0.04	+0.09	+0.09
I. female	+3.05	-4.21	+0.17	+0.28	+0.04	+0.02	-0.06	-0.06

LEGENDS FOR FIGURES

Figure 1. The Rise and Fall of All Nations (Playfair, 1805)

Figure 2. The Increase of Annual Revenues of England and France, 1600-1804 (Playfair, 1805)

Figure 3. Exports and Imports of England 1700-1804 (Playfair, 1805)

Figure 4. Extent, Population, and Revenue of the Principal Nations of Europe, 1804 (Playfair, 1805)

Figure 5. Graphs and Charts Per 100 Pages in JASA and Biometrika, 1921-1975.

Figure 6. Percentage of Pages Devoted to Charts and Graphs in JASA and Biometrika, 1921-1975

Figure 7. STARS for Measurements on Permanent First Lower Premolar of Various Groups of Men and Apes

Figure 8. Andrew's Fourier Plot for Measurements on Permanent First Lower Premolar of Various Groups of Men and Apes (Andrews, 1972)

Figure 9. Chernoff's FACES for Measurements on Permanent First Lower Premolar of Various Groups of Men and Apes

Figure 10. Death Rate from Cardiovascular Disease Among Males Aged 35-74, 1968-1971 (U.S. Bureau of the Census, 1976)

Figure 11. Percentage of Housing Units With 1.01 or More Persons Per Room, 1970 (U.S. Bureau of the Census, 1976)

Figure 12. Bivariate Color Map of Male Cardiovascular Disease by Percentage of Housing Units With 1.01 or More Persons Per Room (U.S. Bureau of the Census, 1976)

Figure 1.

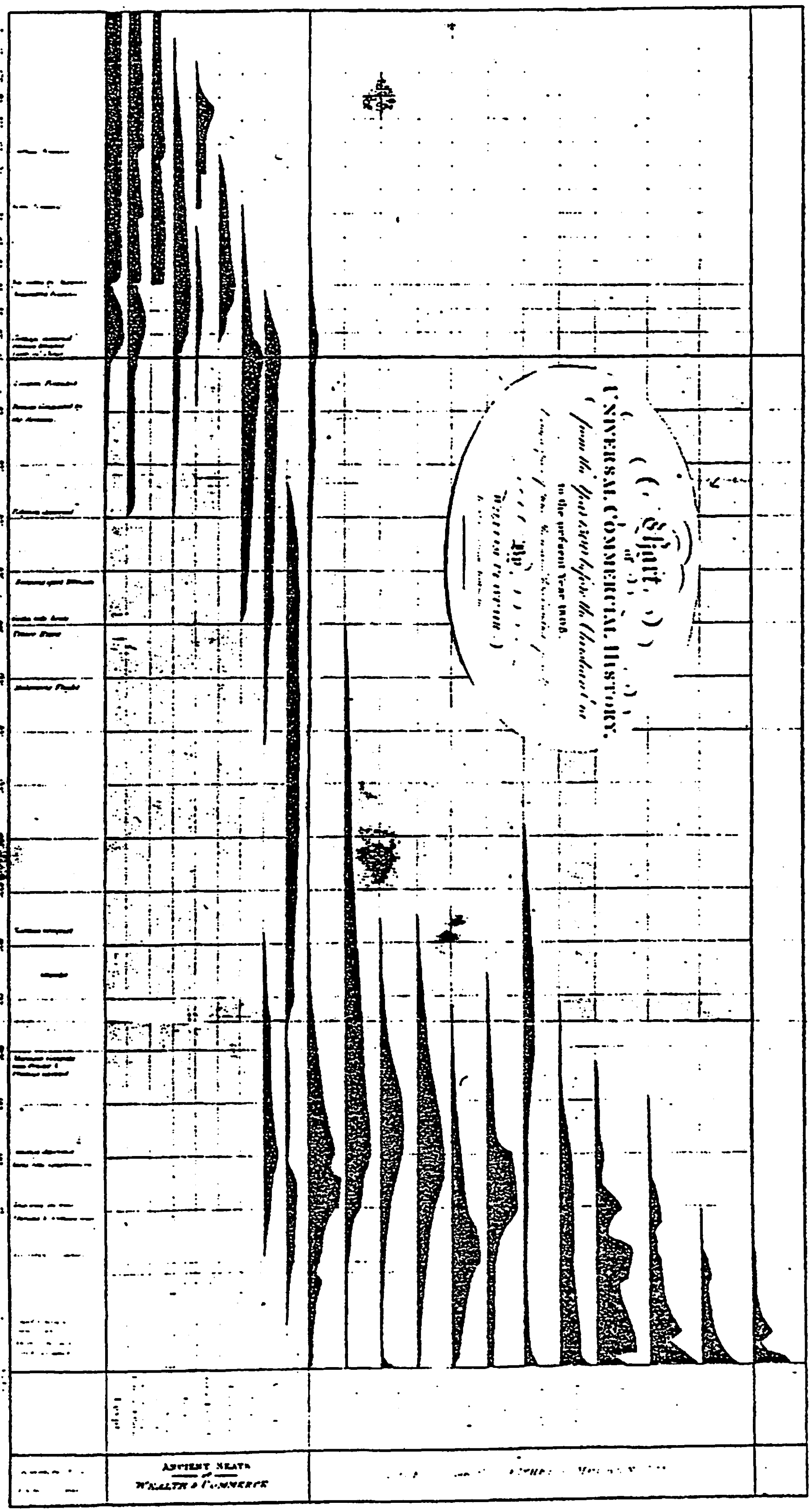
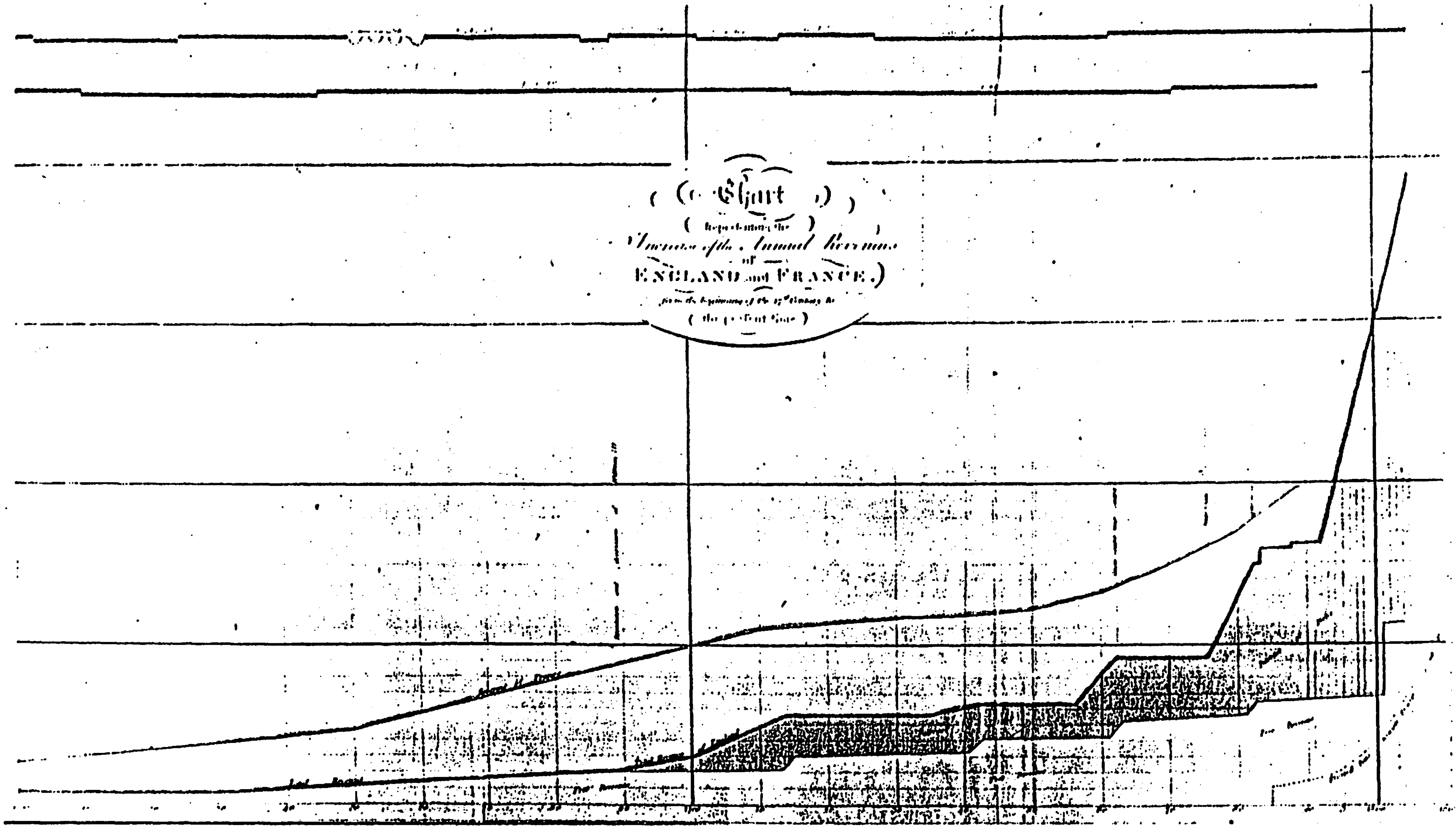


Chart
of
UNIVERSAL COMMERCIAL HISTORY,
ANTIENT SEATS
of
WEALTH & COMMERCE

Figure 2.



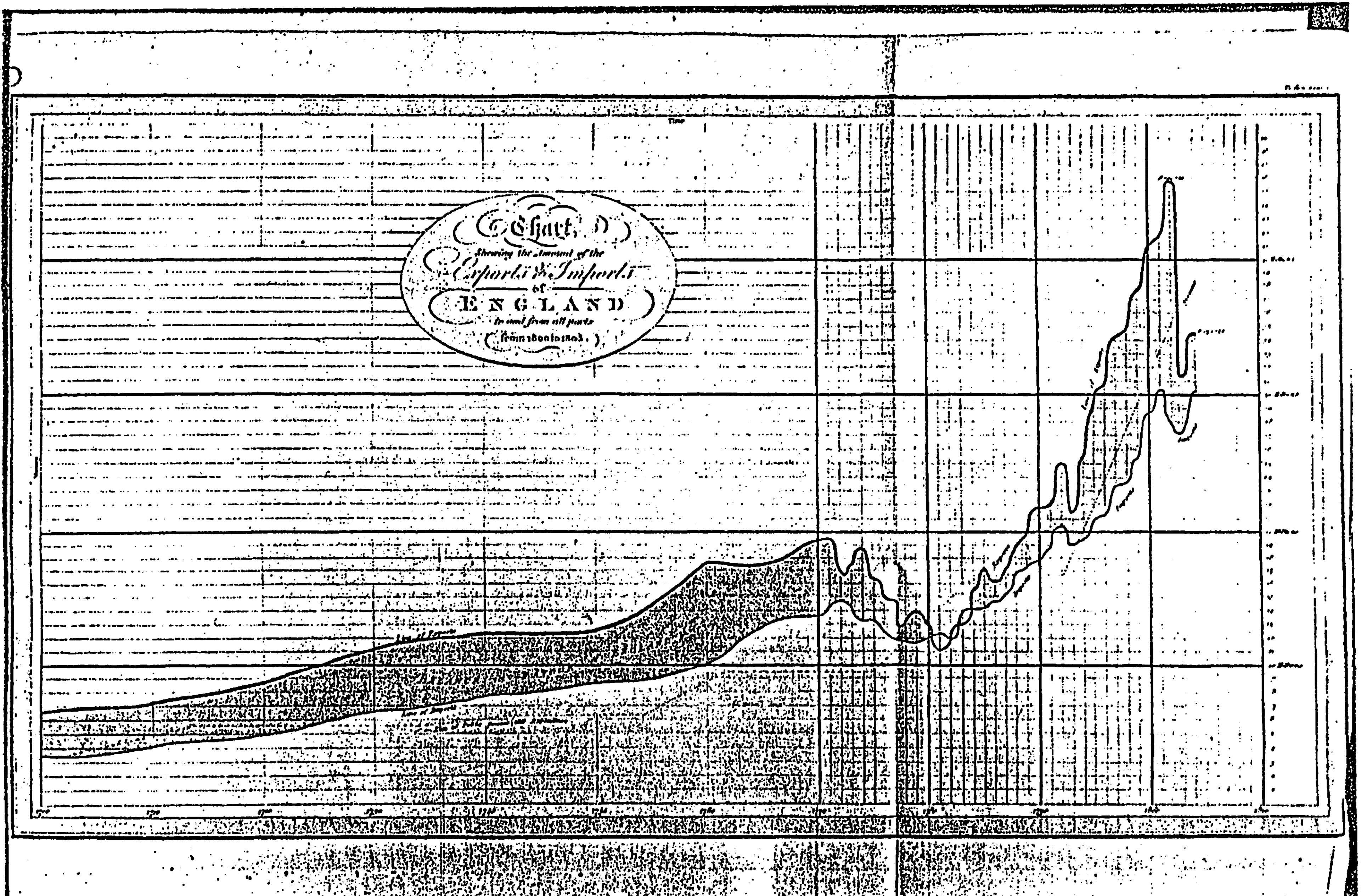

Chart,
Shewing the Amount of the
Exports & Imports
of
ENGLAND
to and from all parts

Figure 3

Figure 4.

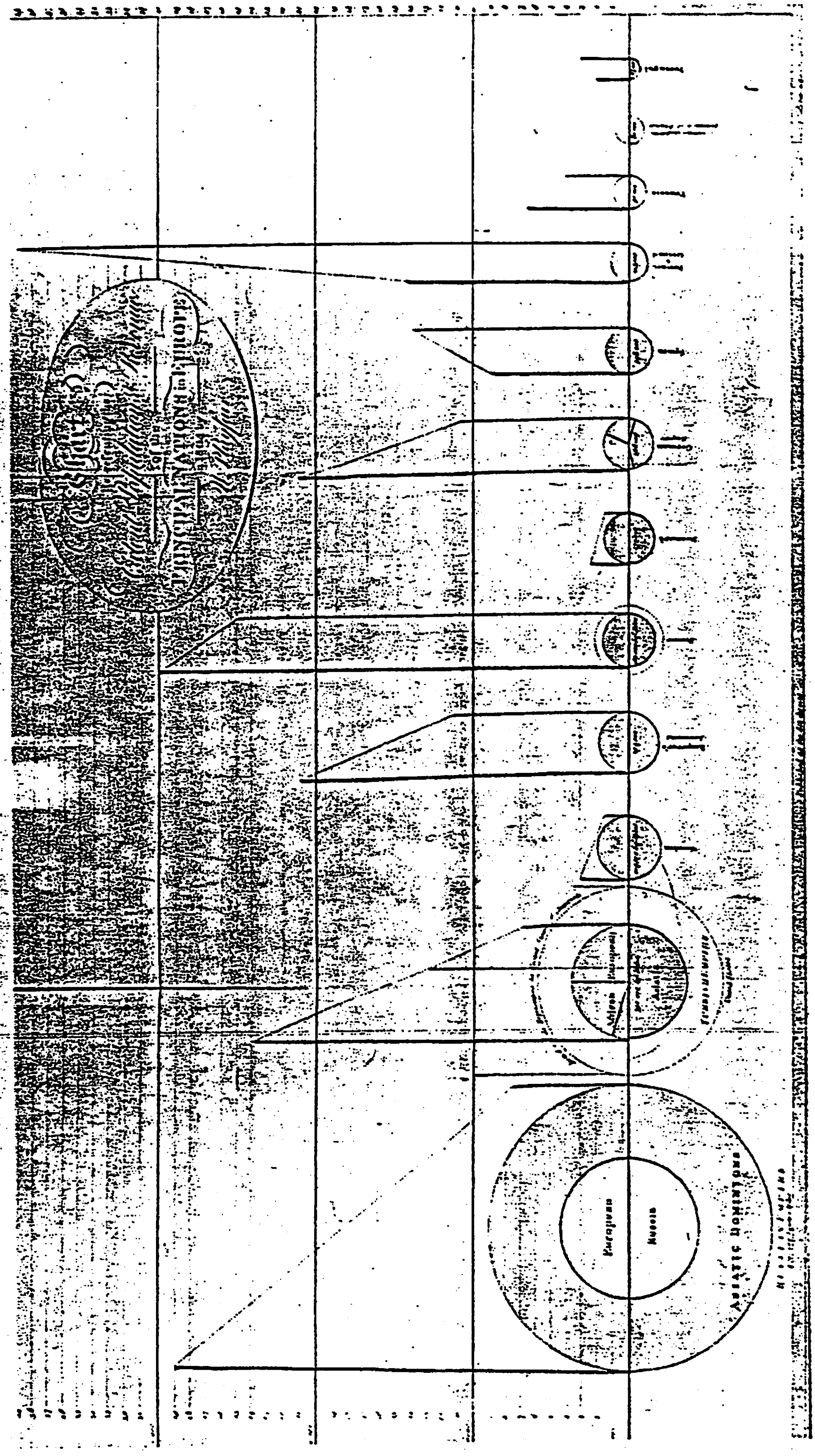


FIGURE 5. Graphs and Charts/100 Pages

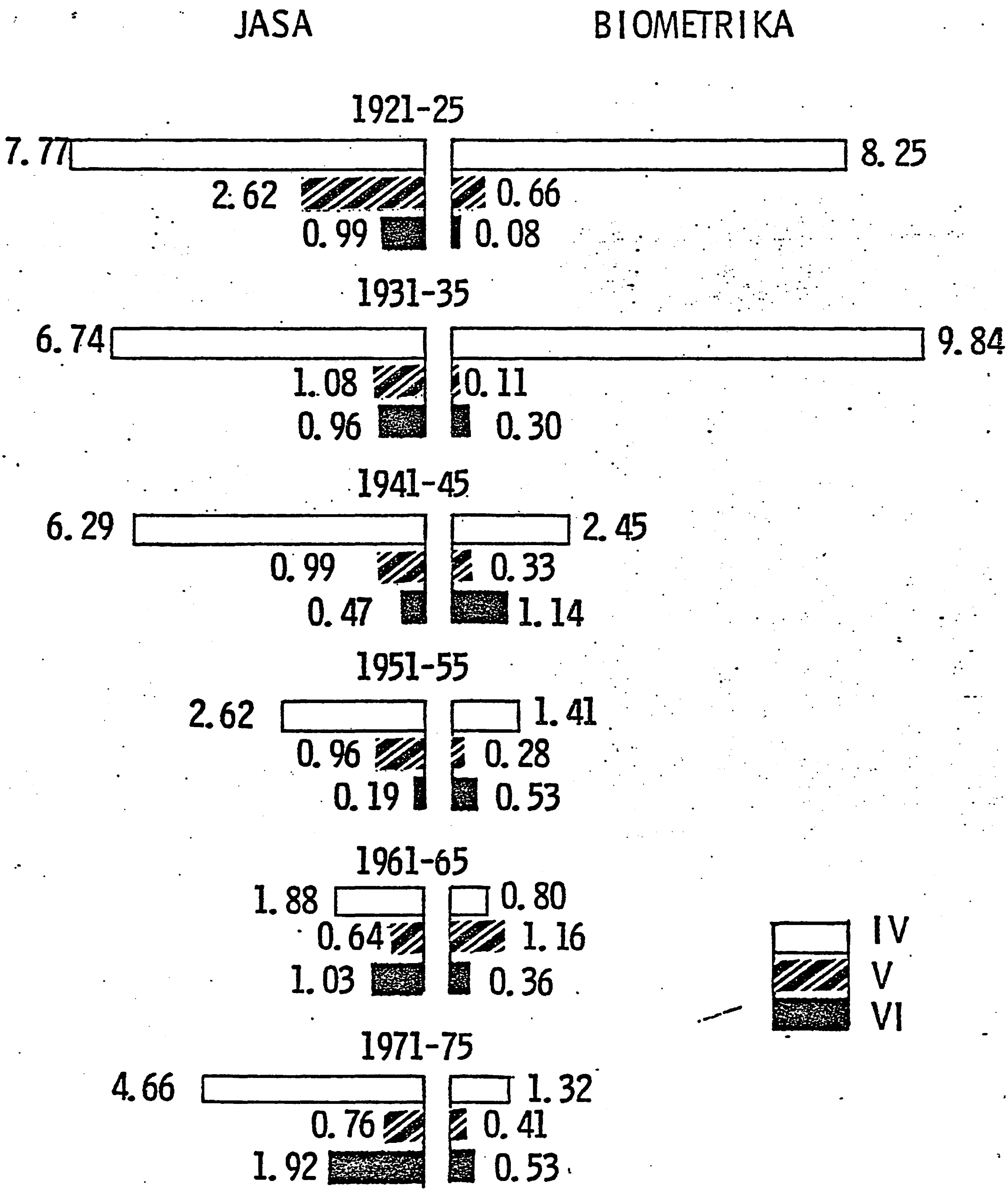


FIGURE 6. Percentage of Pages Devoted to Charts and Graphs

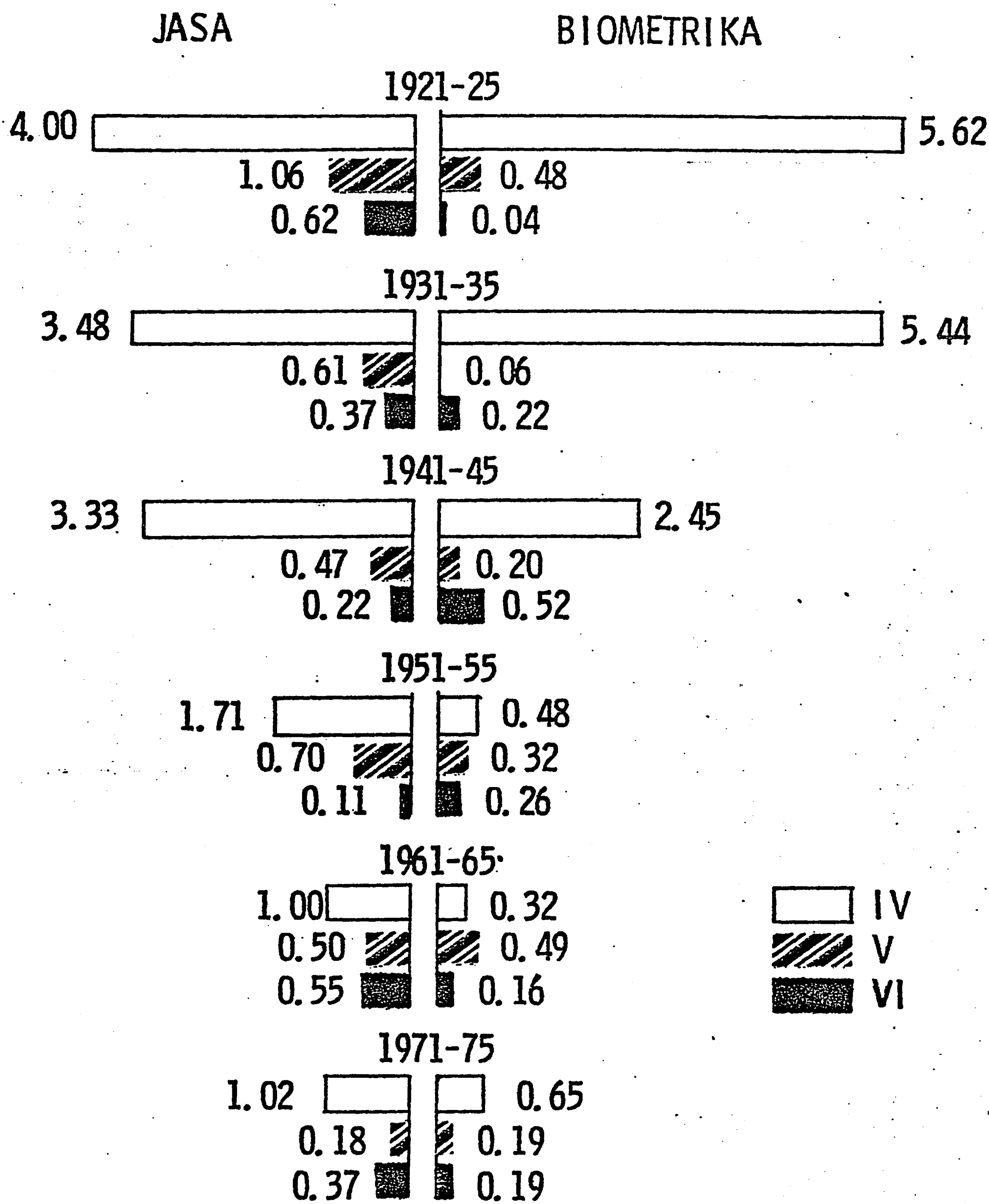
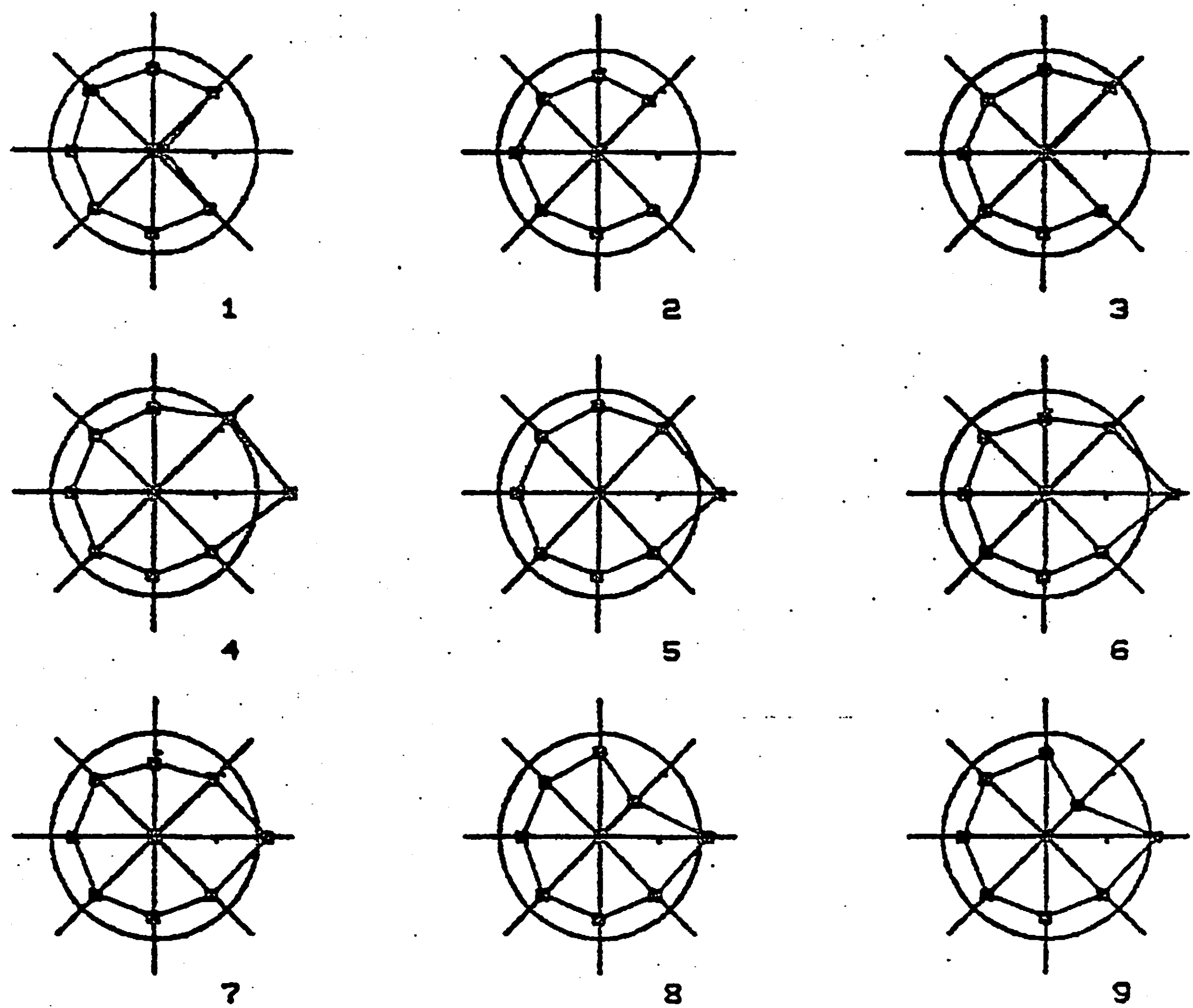


Figure 7



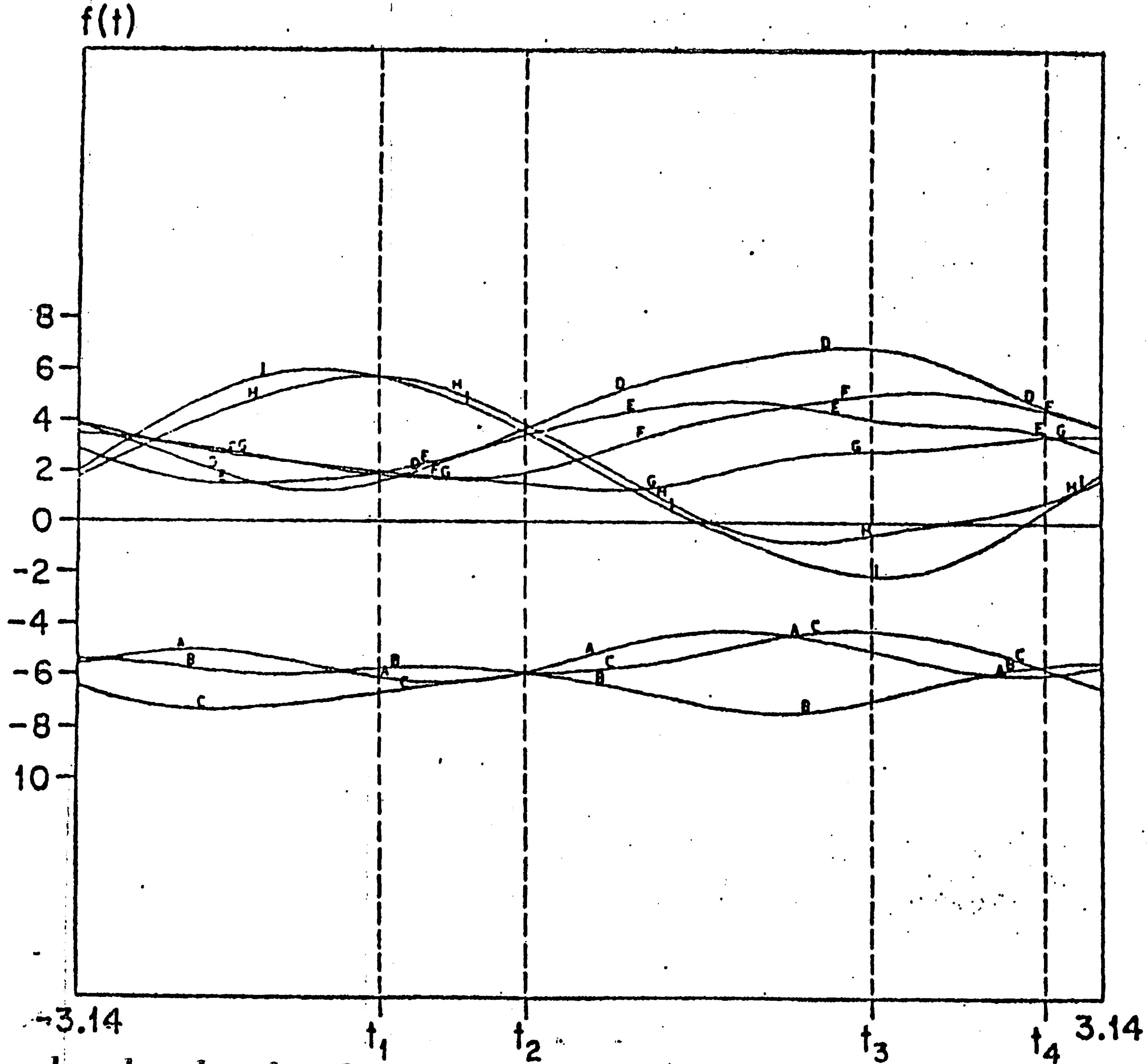


Figure 8

Figure 9

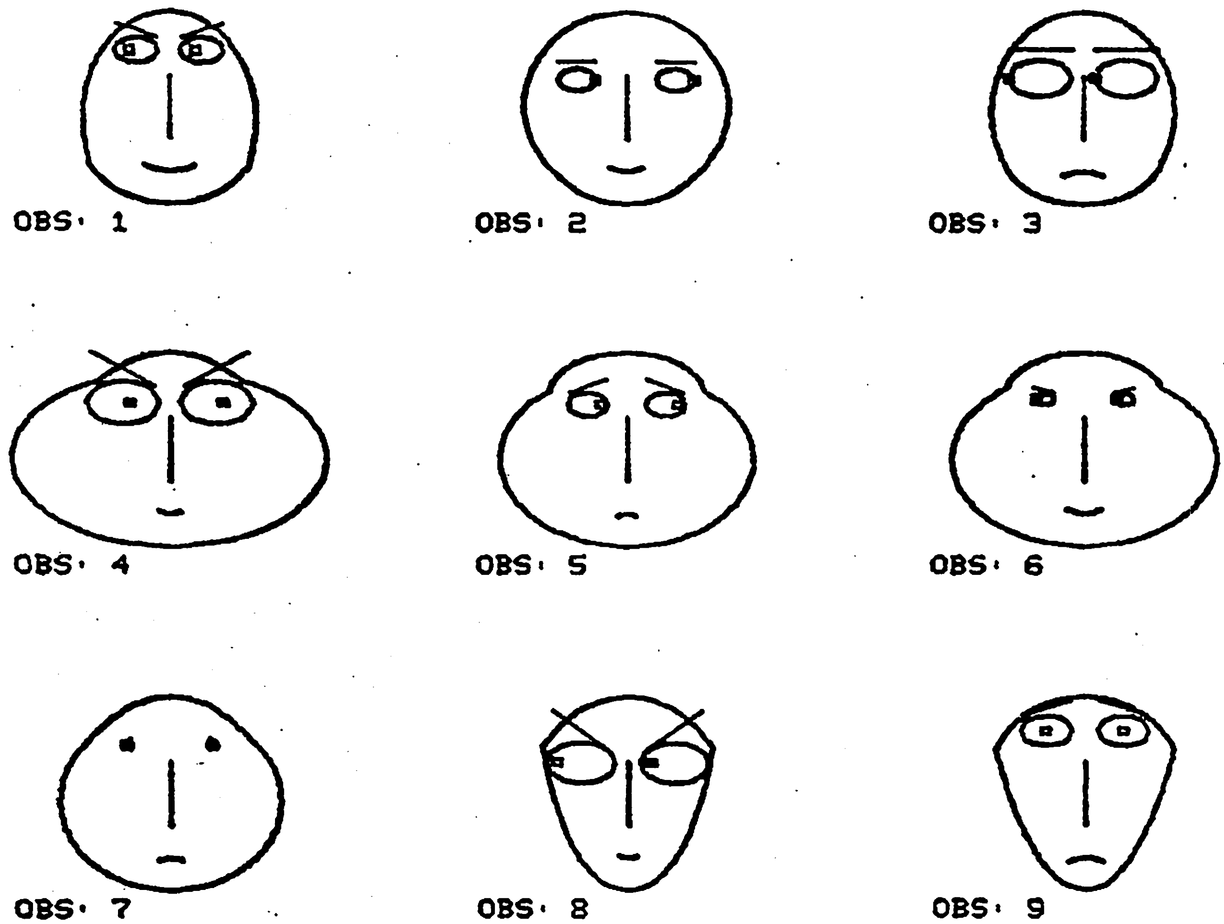
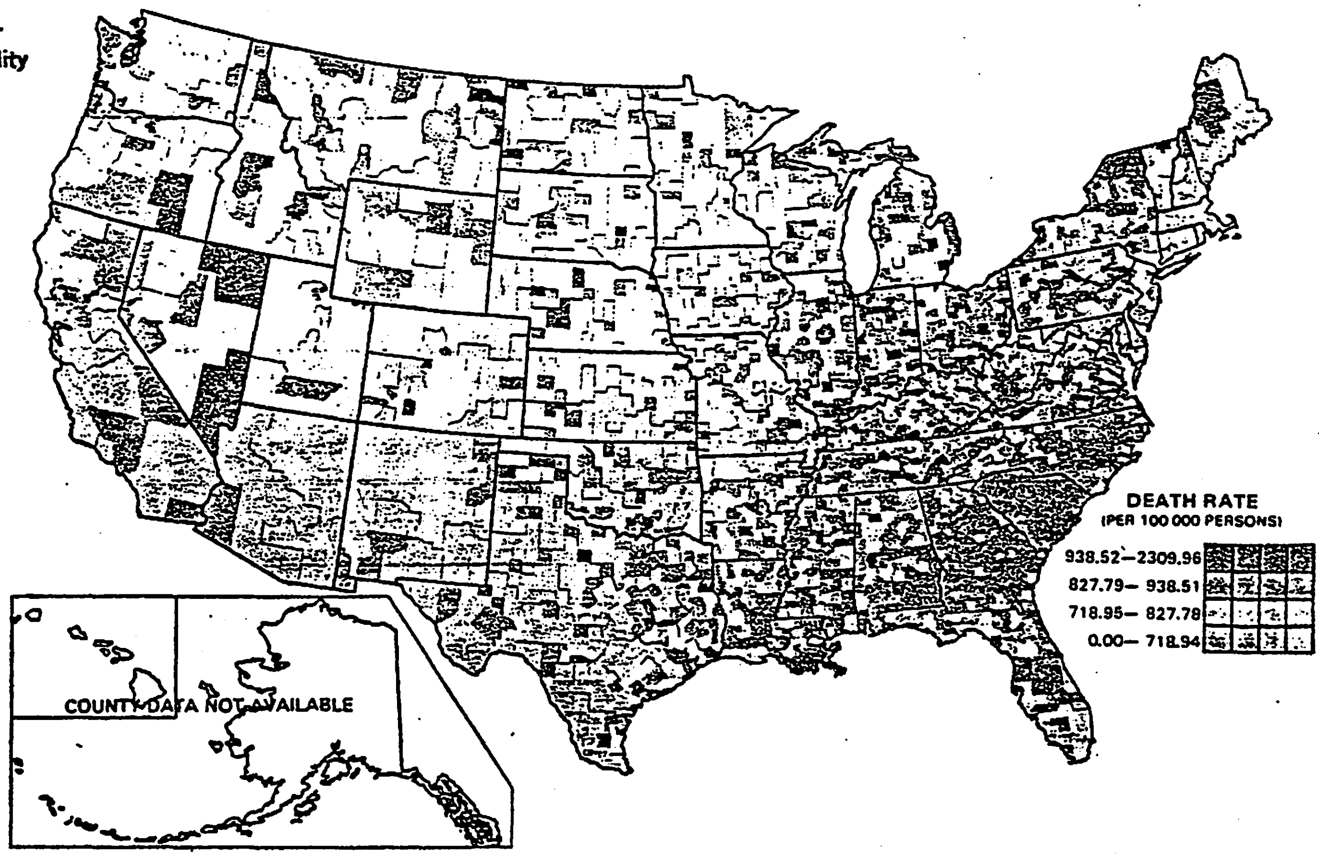


Figure 10

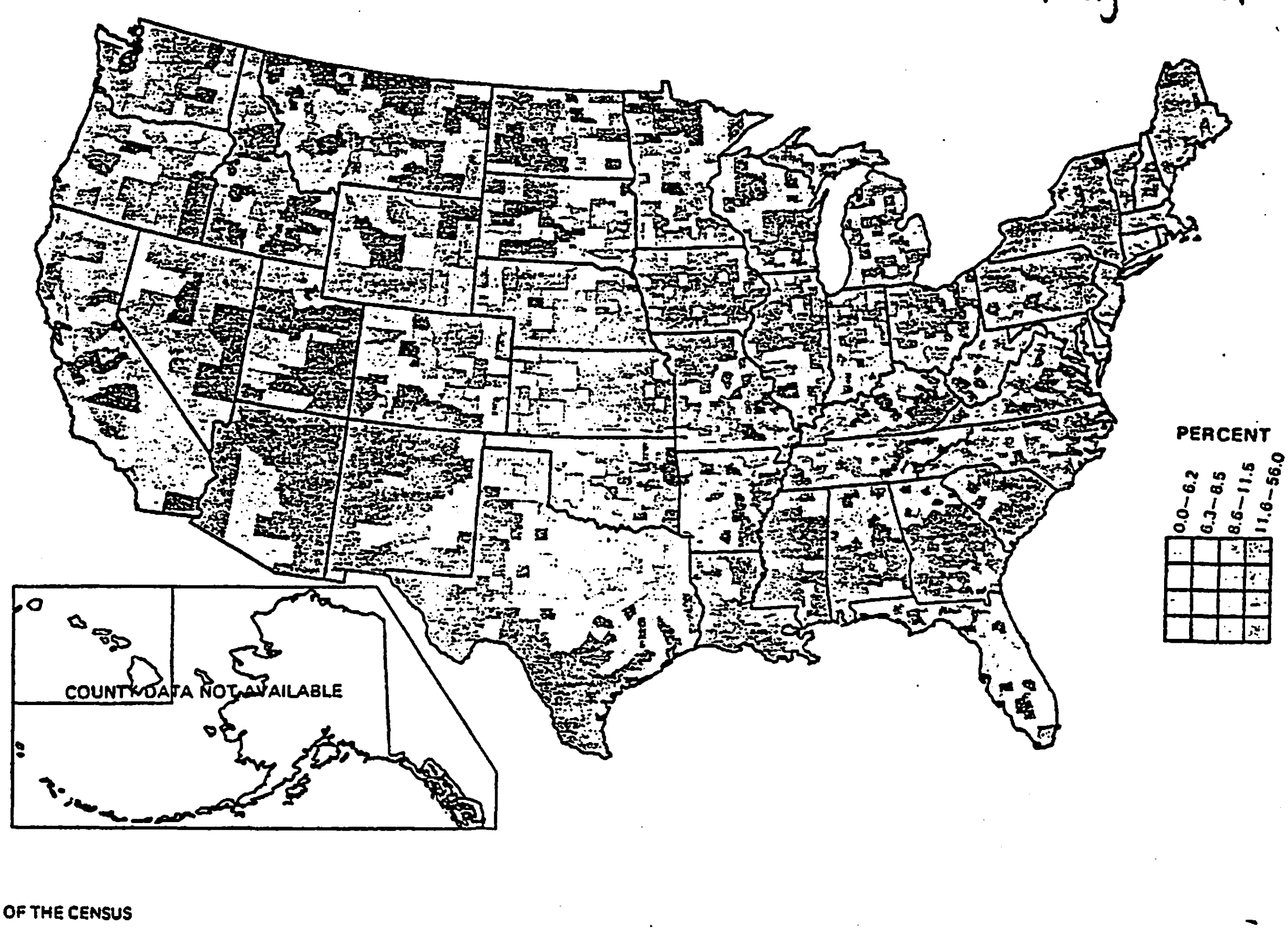
Cardiovascular Disease Mortality Males Aged 35-74 U.S. Counties 1968-1971



SOURCE U.S. DEAPRTMENT OF HEALTH, EDUCATION, AND WELFARE: PUBLIC HEALTH SERVICE, CENTER FOR DISEASE CONTROL; NATIONAL INSTITUTES OF HEALTH

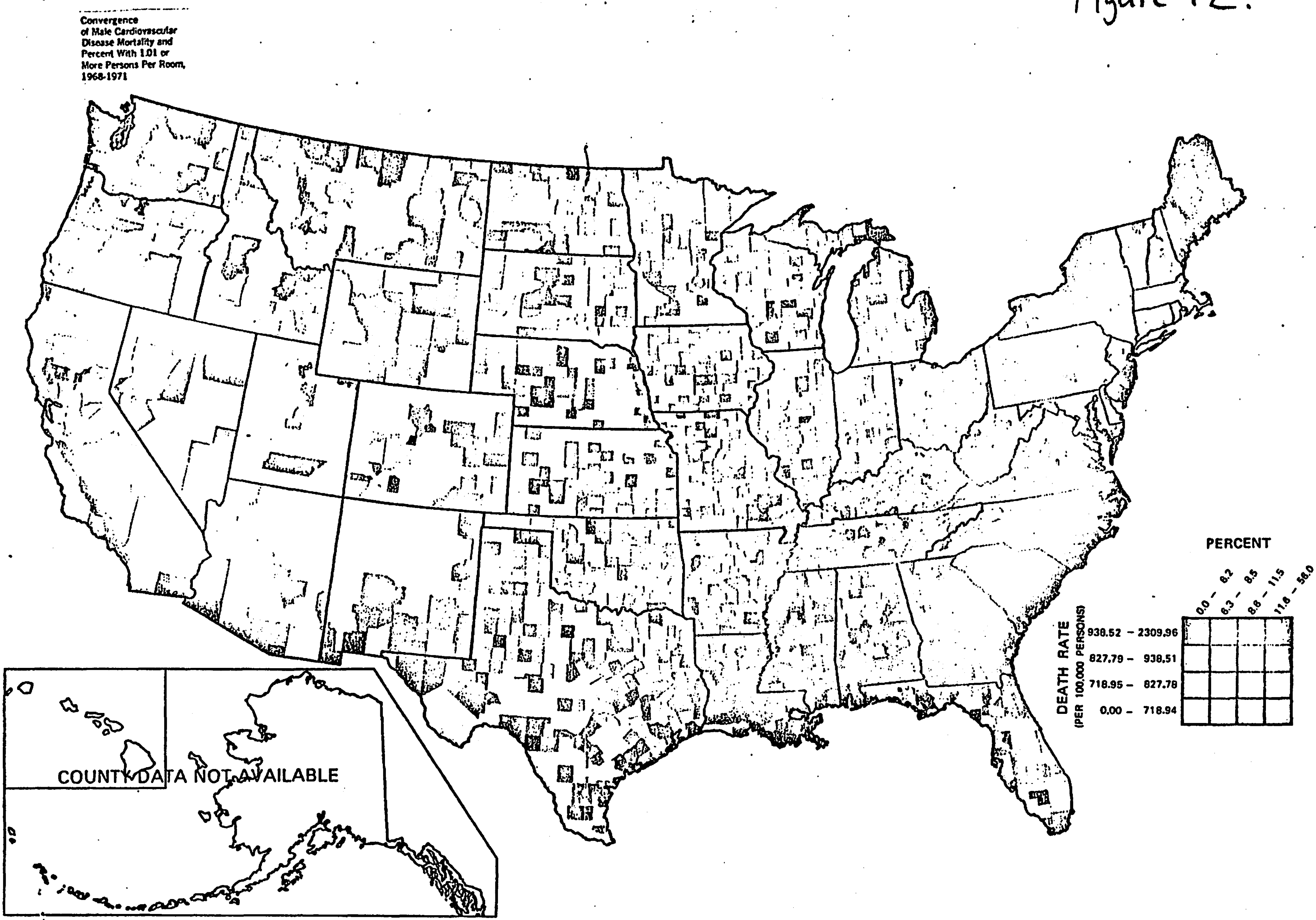
Figure 11

Percent With 1.01 or More Persons Per Room, by County, 1970



SOURCE BUREAU OF THE CENSUS

Figure 12.


Convergence of Male Cardiovascular Disease Mortality and Percent With 1.01 or More Persons Per Room, 1968-1971
COUNTY DATA NOT AVAILABLE
PERCENT
0.0 - 6.2
6.3 - 8.5
8.6 - 11.5
11.6 - 58.0
DEATH RATE (PER 100,000 PERSONS)
938.52 - 2309.96
827.79 - 938.51
718.95 - 827.78
0.00 - 718.94