

ON THE GRAPHICAL PRESENTATION OF
STATISTICS.

BY LEWIS M. HAUPT,
Member of the Institute.

While statistics are said to be uninteresting, they are, nevertheless, invaluable as indices to the progressive man of affairs who would keep in touch with the industrial development of his own or other countries. But masses of figures arranged in tabular form are confusing and often misleading, without great labor in classifying or in reducing them to units of comparison. Frequently only total quantities or values are given of imports, exports, output, product or prices, rendering a large amount of calculation necessary before unit values or percentages for various years, periods or commodities may be secured, and without which no comparison is possible.

To show the growth of any particular industry, it is evident that at least two factors must be represented, viz., time and quantity, or time and value. When all three are known, and the totals only are given, they must be reduced to the value of the unit for each time stated, to determine the actual returns. Since quantity may be expressed in many different denominations, as in pounds, tons, yards, bushels, bales, barrels, stones, cords, etc., they must also be brought to a common basis for comparison.

To avoid many of these difficulties, the practice of representing statistics by the aid of diagrams, plain, colored or shaded in conventional symbols, is rapidly growing in favor, but the existing conventions possess defects which render them almost useless as bases of comparison for commercial purposes, unless accompanied by abstract figures, in which case the former objections are not eliminated.

;
THE EXISTING DEFECTS.

For example, in comparing quantities of any particular kind, as of population, a column is built up of colored blocks having three dimensions, requiring the eye to

compare these volumes with one another. Often they are not even similar volumes, so that the cubes of their edges cannot be used for the purpose intended, and unless the figures be branded on the blocks, no idea can be obtained of even their relative volumes. In many other cases areas of various forms and sizes are used, especially rectangles and circles, but frequently sectors are employed in circles of different diameters, to indicate growth from year to year. Here, again, figures of two dimensions are therefore erroneously used to represent quantities of only one dimension. It is true that the circumference of the circle may be subdivided into arcs, representing proportional parts or percentages of traffic or product, but unless the ratios are indicated in the degrees of the central angle, or the relative lengths of arc, the sectors possess but little actual value for comparison.

Even where statistics are represented in columnar form, laid off on both sides of a central axis, the curve of increase is bisected, one-half of it being on either side of the axis, and, therefore, the correct curve is not shown on the diagram, but must be deduced from it by a mental operation.

Again, it frequently happens that the intervals of time when the record is made are quite variable, so that without interpolation a correct idea of the form of the curve cannot be obtained even with the abstract figures stated. Certainly the rectangles drawn to scale, unless taken in connection with their time intervals, become quite misleading. The absence of uniform scales, and the omission of any reference thereto, results in representing the parts as sometimes greater than the whole, a manifest absurdity.

These and many other incongruities arise from the promiscuous pictorial methods in vogue intended to educate the mind by unscientific, popular methods, resulting in "confusion worse confounded."

SUGGESTIONS FOR GENERAL USE.

All quantities, of whatever denomination, kind or extent, which are capable of being measured, may be expressed in abstract figures, and such figures may always be repre-

sented by linear distances drawn to scale by taking any number of such parts as a unit. As the number representing the quantity or price may vary from time to time, so the lengths of the lines will vary for the corresponding times, and these lines may, therefore, be compared, and the ratio of any one to any other be determined by projecting them upon a sheet of cross-section paper, starting always from a straight line or edge.

Thus, if the law of the increment of any variable quantity is to be represented, the times at which the variations occur may be laid off along a straight line, while the actual value of the variation may be platted in a series of lines at right angles to the first, constituting a system of rectilinear coördinates, of which the times may be the abscissa (x) and the quantities the ordinates (y). By drawing a line through the extremity of the ordinates, the curve showing the change in value of the quantity is at once revealed.

If the increments for equal times are equal, the curve will become a right line, and a prediction can safely be made as to future probabilities. If it bend upward, the increments are increasing; if downward, decreasing. They may oscillate violently, indicating lack of stability or great fluctuations in value, generally traced to some external cause affecting the value of the commodity used in exchanges, as a plethoric or stringent money market, abundance or paucity of crops, etc.

TO ILLUSTRATE DEVELOPMENTS.

The resulting diagrams, if platted as indicated above, will give a simple, accurate exhibit of the conditions existing at any particular date, whence they may readily be connected with correlative events, and the relation between cause and effect be largely determined with accuracy. As population is the basis for, and principal element in, industrial development, and as it is important to determine the probable increase as affecting questions of consumption or traffic, it may be instructive to illustrate the above method by plating the curve of the growth of population during the past century in the United States. Thus, on the diagram, it will

be seen that the axis of X is horizontal, having a scale of ten years to 1 inch, while that of Y is vertical, having a scale of 10,000,000 to an inch. (See *Plate I.*)

Taking the origin of coördinates at O , and the data from the census reports for each decade, beginning with 1790, it will be seen that the curve at first departs but slightly from the horizontal axis, but gradually increases its flexure upwards and must, in course of time, approach a tangent to the vertical, becoming, in short, an asymptote—this, too, notwithstanding the fact that the *percentage* of increase decreases from the beginning to the end of the century.

It is thus shown at once that the increment is not a simple ratio, as in an arithmetical progression, but is being compounded each decade. Hence it is that the population under normal conditions increases more rapidly each decade, although the percentage may decrease. Between 1790 and 1890, for instance, the population increased 1,491 per cent., or 14.91 per cent. on an average each year, if computed as simple interest upon the original principal. Yet the rate each decade before the civil war was varied from 33 to 36 per cent., or only $3\frac{1}{3}$ to $3\frac{2}{3}$ per cent. per annum.

During the war decade it was only 23 per cent., increasing to 30, and again decreasing during the last decade to 25 per cent.

The actual increments of population, however, were nearly 12,500,000 in 1880-90, as compared with less than 2,000,000 in 1800-10, when it was 36 per cent. of the whole.

What this rapid or compound increment portends in the near future for the trade prestige and influence of this nation may be more clearly understood by extending the curve at the constant ratio of that of the past decade, namely, 25 per cent., from which it appears that the population of 1890 will have doubled itself by 1920, or in thirty years, and by 1950, it will have again nearly doubled if the ratio remain constant and the government stable and liberal.

It is not to be expected, however, that the ratio will remain constant for so long a period, but it will not be many years before our population will reach 150,000,000 souls.

TO REPRESENT VARIATIONS.

Fluctuations in values may also be illustrated most forcibly by the use of the graphical method.

Take, for example, the relative values of gold and silver during the past century, and plot the curve of their ratios as per the diagram (see *Plate II*), and it will reveal at a glance the relation between the values of those commodities and legislation. The stability of their relative values, from the close of the last century, through three wars and several financial crises, up to 1873, the year of the passage of the demonetization act, is clearly shown, as well as the effects on the ratio due to subsequent acts. This diagram, therefore, is a sharply-defined representation of the fact that the value of silver as a medium of trade has depreciated rapidly, while that of gold has as rapidly appreciated, neither being stable nor a "standard," and, but for the providential development of large deposits and the reduction in the cost of extraction of gold, the discrepancies in values would have been much greater.

TO DETERMINE FRANCHISES.

Columns of figures are also confusing when it is desired to determine the values of a franchise from the tabulated annual revenues of any enterprise, and an average of the net returns is also misleading, since the profits will generally be less during the earlier stages of development than in the later year. Here, again, the superiority of the graphical curves plays an important part in revealing at once the actual condition of affairs.

To illustrate: take the case of the Monongahela River Improvement from 1840 to 1896, and plot the receipts and expenses, as shown in *Plate III*. These curves indicate readily the fluctuations due to internal and external causes far better than can be done by an array of figures. If a dam breaks, the expense account goes up and the revenues down: if enlargements or betterments are added, the operating expenses reach a higher plane, while the revenues are correspondingly increased.

But probably the most practical use of this diagram is

PLATE II.—CURVE SHOWING EFFECTS OF LEGISLATION ON RELATIVE VALUES OF GOLD AND SILVER.

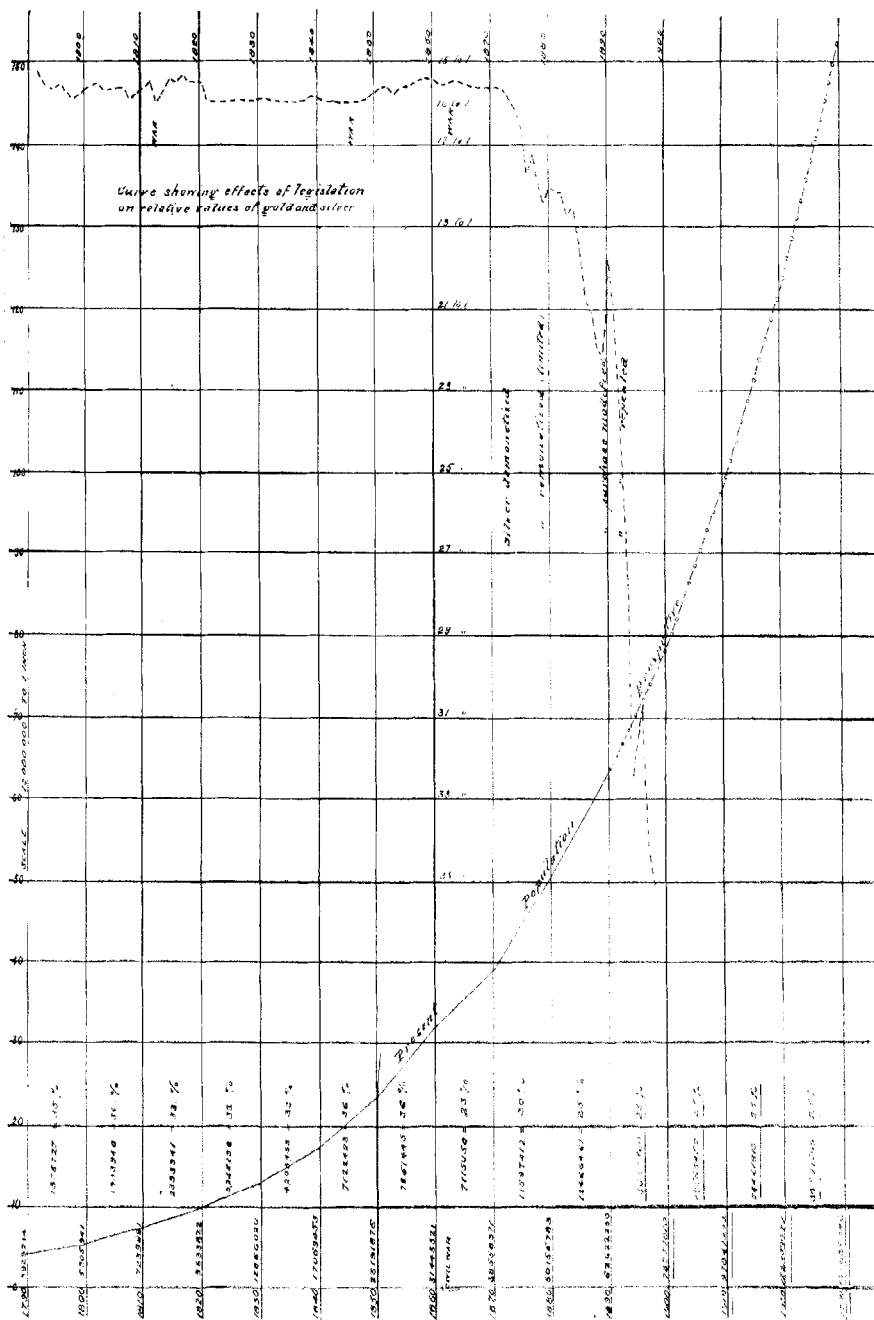


PLATE I.—CURVE SHOWING GRAPHICALLY THE INCREMENTS OF POPULATION

to determine the value of the company's franchise by the establishment of the average financial gradient. This is done, approximately at first, by a trial line with an assumed annual increment, and by then taking the algebraic sum of the ordinates of the revenue above and below this line. If

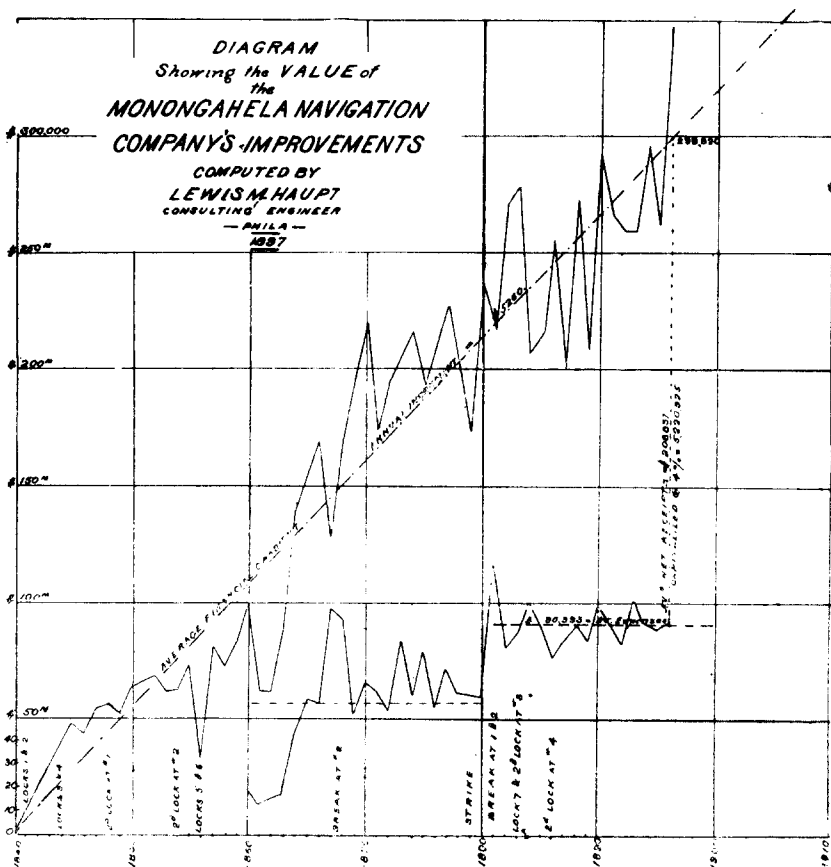


PLATE III.

this sum is not equal to zero, it should be further adjusted until it is so, when the true line becomes known. By capitalizing this increment at the current rate per cent., the value of the franchise is readily found.

These few illustrations will serve, it is hoped, to estab-

lish the great superiority of the linear methods of representing statistics as compared with the pictorial volumes and areas, which are almost useless to the statistician as well as to the public. If this brief article will aid in eliminating so unsatisfactory a method, and in substituting therefor the general use of the more scientific and accepted projections, it will have served its purpose.

SUMMARY OF THE REPORT OF THE COMMISSION ON THE EXTENSION AND IMPROVEMENT OF THE WATER SUPPLY OF THE CITY OF PHILADEL- PHIA.

Following is a summary of the conclusions presented in the report of Messrs. Rudolph Hering, Samuel M. Gray and Joseph M. Wilson, the experts named by the Mayor of the city of Philadelphia, with the consent of the Councils, to examine and report upon the questions involved in providing for the urgent requirements of the city in connection with its present and future water supply :

"The deplorable condition of the city's water supply, which it is sought to remedy, is due to the pollution of its sources, to the lack of effective pumping machinery, and to the insufficient capacity of the distributing system.

"The question of first importance is the source of supply, and to this nearly all of our thought and time has been devoted.

"Most of the water is now obtained from the Schuylkill River, within the city limits. Five pumping stations take from it about 200,000,000 gallons daily. One pumping station is located on the tidal estuary of the Delaware River at Lardner's Point, and supplies about 15,000,000 gallons daily.

"The Schuylkill water is being polluted at many points from its source down to the city line. Beginning with the mine waters, the coal dust and some sewage from the upper parts of the water shed, the pollution is increased below by the sewage of cities and villages situated along the river and its chief tributaries, by the manufacturing refuse and by the surface water from agricultural districts, all of which render the water sometimes turbid, unpalatable, impure and dangerous to health.

"The Delaware water at Lardner's Point is less turbid after rains than the Schuylkill water ; it is also softer and less polluted. Its flow is many times larger. While this water is, therefore, now somewhat better than the Schuylkill water, the growth of the city, the newly-built or projected sewers above and below the intake, and the tidal oscillation of the water, tend to a continually increasing pollution also of the water taken from the Delaware River.

"It, therefore, becomes imperative either to select a new source of supply or to improve the present one, so that it will become thoroughly satisfactory