THE ROYAL BANK OF CANADA
MONTHLY LETTER

Vol. 37, No. 1 HEAD OFFICE: MONTREAL, JANUARY 1956

The Uses of Statistics

ALL of us must work with figures and numbers every day. The simplest assertion in conversation, as well as the most profound business analysis, deals with quantities. Statistics is one way of bringing big numbers down to a size where they can be grasped with ease.

Men are fond of collecting figures, pushing them around into different combinations, drawing conclusions from them and quoting them. Many a man has built himself a reputation as a shrewd executive merely by demanding: "Give me the figures!"

What we do with figures once we have them is another story. There is an art in handling the information that figures give. This essay is concerned with the preliminaries: how to get the right figures for the purpose we have in mind, how to throw them into readily-understandable form, and how to protect ourselves against certain dangers.

There can be little argument against the need for knowledge of our number language. Lord Kelvin put it in a straightforward way: "When you can measure what you are speaking about, and express it in numbers, you know something about it; but when you cannot measure it, when you cannot express it in numbers, your knowledge is of a meagre and unsatisfactory kind."

However, we need to go about collecting the figures in the right way. Unassorted data, no matter in what great masses we gather them, are of little service until we arrange them, classify them, and relate them to one another and to other data. To achieve this, we turn to methods of abbreviating and summarizing the facts we have gathered. The size of a factory may be described by its average monthly output and the importance of a bank may be characterized by its average deposits. To compare the income value of two or more stocks we do not need to memorize long tables of rates of interest, but only to be able to say that the average rate of this stock is 6½ per cent and of that stock 6¾ per cent.

What statistics show

Statistics is not a new science. The earliest English book on the subject was published in 1661, while the first complete work was *The Statistical Account of Scotland*, edited by Sir John Sinclair in 1791. The term "statistics" is thought to have originated in the Latin word "status", meaning a state, and from that came "statistica" meaning one versed in matters of state. The word "statist" crept into the English language, and was used by Shakespeare in *Hamlet:* "I once did hold it as our statists do, A baseness to write fair."

As defined by Galton, the English anthropologist of the 19th Century, the object of statistical science is to discover methods of condensing information concerning large groups of allied facts into brief and compendious expressions suitable for discussion.

Upon such data there is a host of private and public decisions made every day. Business men make plans in the light of known facts in their field of interest, and check the progress of their business by reference to statistical material collected within their organizations. Government policy is built upon detailed knowledge about the nation's employment, production, and purchasing power. Legislation on conservation, crime suppression, public health, education, housing, industrial relations and economic stabilization must stem from accurate up-to-date information.

We do not value statistics for their own sake, but only as they pay for their keep by enabling us to make better plans, to check performance, to close budget gaps. Through statistics we enlarge our individual experience, obtain facts in definite form, simplify and classify numerical facts so as to make them readily comparable, and qualify ourselves to interpret conditions and foresee trends and tendencies. Without planning

we cannot control either business or domestic finance, and planning can only be done effectively when we have a knowledge of conditions such as we gain from perusal of statistical information.

Discretion needed

There is a cause and effect relationship between all phases and areas of business activity. The business man, however well versed he may be in the techniques of his own business, needs to know also the trends in major supplying, purchasing and competitive businesses. He needs to be able to forecast reasonably well from the record of the past what is ahead.

At the same time he must be careful to avoid over-indulgence in statistical control. Too much information may be gathered, adding to the expense of office work. Statistical indigestion may result when we fail to ask the purpose for which data are to be used, and to appraise the value of the information against the cost of obtaining it. Executive control is not gained by piling up reports, but by the use that is made of them.

Gathering of statistics should be tested by practicality: what purpose will be furthered by this statistical data? One could keep a staff busy for a long time averaging the numbers in a telephone directory, but it is hard to know what one would do with the resulting statistic. As Dr. J. A. Gengerelli puts it in lighter mood in an article in *The Scientific Monthly:* "Consider for example, the blades of grass on a lawn; what a capital fact-finding project they would make! We could use the very best in stratified sampling devices to calculate an unbiased estimate of the number of blades of grass contained in the defined area". And then what would you have?

It will pay office managers and executives to look at the statistical efforts being made by their staffs with a view to finding those that are counting blades of grass. Collection of figures, started for a good purpose which was served in a short time, has a habit of going on and on, and even spreading into new areas.

Sources of information

Most of the information basic to the compiling of statistics in a business enterprise will come from two sources: government departments and the firm's office.

Day after day the Dominion Bureau of Statistics and the statistical departments of provincial and municipal governments issue detailed reports about all facets of Canada's production, distribution and consumption of goods; about population, migration and earnings; about electricity, coal, oil, transportation, agriculture, forestry, fisheries, mining, and, in short, every activity that enters into our lives.

These reports range from the very comprehensive decennial census to monthly, weekly, and daily reports. For $2 a year you may have laid on your desk every week a mimeographed summary, about 16 pages, of the latest information collected by the federal government. It is the *D.B.S. Weekly Bulletin,* obtainable from the Queen's Printer, Ottawa.

Since the Bureau of Statistics was created in 1918 to centralize the statistical work of the government, it has won world-wide acclaim for the quality and comprehensiveness of its output. It has developed a well-rounded and well-balanced system of statistics adequate to meet the needs of the public, business people, students and the government.

But even in government activity a warning given earlier has some application: statistics should not be collected needlessly. Governments at various levels require large numbers of statistical figures from business and individuals, thus imposing a financial and labour burden on them. It is part of the statistician's duty to look critically once in a while at the resulting mass of data, to keep it within reasonable bounds.

Within our own spheres of activity — business, educational, social — most of our statistical data will come from our accounting departments, to which reports are made by other departments like purchasing, producing, and marketing. Other information will be gleaned from newspapers, trade journals, investment services, research bureaus, and the confidential exchange of figures between members of trade associations.

Principles of statistics

Most people will agree that if we can measure what we are dealing with, even roughly, it is far better to make some measurement than no measurement at all. In almost all events connected with business we are able to do better than that: we can come very close to accuracy.

The basis of statistics is the law of large numbers, sometimes called the law of average. We may formulate it in this way: "A reasonably large number of items, chosen at random from a large group, have the characteristics of the group."

When we are dealing with statistics of people, goods, finances, and so forth, we may be able to predict the probable course of the whole galaxy without being able to tell what course will be taken by any particular person, parcel or dollar.

All statistical information resolves itself into simple judgments of magnitude, comparisons between this and that or between similar things at different times.

The principal comparisons based upon statistics are: the same thing at different times; something in relation to a larger thing of which it may be a part; and one thing in its relation to something else which is supposed to influence it.

Averages

An average is simply a way of combining a number of numbers so as to obtain a single number that for our purpose can be used to represent the entire set, or to summarize certain of its properties. It gives us a quick understanding of the general size of the individuals in a group.

When properly arrived at, a series of simple averages or percentages may prove effective for business planning and control, and yet be easy to understand. It is, however, necessary that three persons know precisely what is being measured and with what purpose: these are the collector of the raw data, the person who computes the average, and the person using the resulting figure. As Dr. Paul H. Nystrom says in *Marketing Handbook:* "Without this knowledge and familiarity with the conditions that give rise to the data the statistician holding sharp analytical tools is not unlike a small boy holding a sharp knife. Each can do very serious damage."

We have a choice of several averages, and we pick out the one which is most appropriate to our data and most meaningful for our purpose. They are mentioned here just as a reminder that there is more than one sort. The arithmetic mean, the most common, is obtained by adding a series of numbers and dividing the sum by the number of items added. The median divides a series equally with the same number of items above as below. The mode is the value in a series that occurs most often. The geometric mean sounds more difficult than it is: the *n*th root of the product of *n* items. If there are three items, you multiply them together and extract the cube root of the product. The harmonic mean, used principally in certain cases of averaging time rates, or when dealing with rates and prices, is a special kind of arithmetic mean, working through reciprocals.

The difference between two averages may be illustrated in an example. The arithmetic average of 2, 4, and 8, obtained by adding them together and dividing the sum by 3, is 4.6; the geometric mean of the same figures, obtained by multiplying them and taking the cube root, is 4.

No one type of average can be considered the best, but only the best for a purpose. Each of them has characteristics that are favourable or unfavourable under the circumstances; the statistician selects the type which, under the circumstances and for the purpose, is representative.

Whatever average is used, it is necessary that the crude data be of the same sort. If we scramble together several distinct classes of individuals and take an average of whatever sort we shall get a meaningless figure. For example, if we put into one table the heights and weights of men and women, the average height and weight would represent neither men nor women. Similarly, the average of hourly earnings in a factory would be meaningless if wages of those working at hourly rates and of supervisory employees working on a monthly basis were put into the same table.

Index numbers

The index number is a statistical device for measuring changes in groups of data, such as employment, prices, academic grades, and so forth. In order to measure the changes in a large number of varying items, it is necessary to resort to some relative averaging device that will serve as a yardstick of comparative measurement. The comparisons may be between periods of time, between places, or between things that fall into the same category, such as articles of merchandise, persons or factories.

Index numbers must be based upon valid representative samples that are homogeneous. The base period, perhaps a year or several years, is assigned the value of 100 per cent. Index numbers are computed as relative to the base period. An index of 125 shows that business, or whatever is being measured, is 25 per cent greater than in the base period, while an index of 85 reflects a 15 per cent decline.

Not only business people, but all who buy, are interested in price indexes. Consumer prices, formerly called the "cost of living", cover prices of goods and services normally purchased by families of wage earners and moderate income workers in cities. This index is widely misunderstood. It is merely a barometer of prices for a fixed bill of goods considered as necessary to life. It is not a measure of living standards, which vary from income to income and according to personal whims and desires.

Canada's consumer price index, based on 1949 prices equal to 100, has been hovering around 116 for many months. This means that it takes $1.16 to buy the same amount of the specified goods as could have been bought for $1 in 1949.

To ascertain the purchasing power of your earnings today as compared with any year in the past, divide your earnings now by the current consumer price index and your earnings in the earlier year by the consumer price index of that year. This is called bringing the amounts into dollars of equal value.

It must be kept in mind that ordinary index numbers do not relate to quality, but to quantity. A price index may show, for example, that an automobile tire costs

twice as much as it did thirty or forty years ago, but it doesn't show that the current tire lasts about ten times as long as its ancestor. This might be overcome by building an index that showed the cost per mile of a tire then and now. Judged by utility, the cost might be only 80 per cent of the old price.

Keep it simple

How exact must we be in statistical work? It is easy and fatal to think that the detail of our statistics is equivalent to the accuracy of our knowledge about the problem in hand. In many cases of computation we shall find that much meticulous work may be eliminated without significant change. Consider the percentage we find by dividing 375,541, 940 by 5,847,159,678. It is 6.4 The same result would have been reached by dividing 376 by 585, and almost the same result (except for a decimal) by dividing 38 by 58.

Some confusion exists because of our way of expressing big numbers. We on this side of the Atlantic commonly refer to a thousand millions as a billion, but the British billion is a thousand times as big. The *Economic Digest* suggests that, in the interests of international understanding, we should speak merely of thousand millions. It makes no difference in the present state of misunderstanding that the British, clinging to the original sixteenth-century sense of the word billion (as the second power of a million) are historically and etymologically in the right. For mutual understanding of statistics expressed in billions of whatever sort, something will have to give way.

Intelligence needed

It is not enough to have honest statistics placed in front of us; we need to be intelligent in our interpretation and use of them. Statistical methods cannot be relied upon to take the risk element out of enterprise, nor to create certainty of judgment, nor to predict future events. They are a base upon which to formulate sound business judgment: that is all.

Among the best criteria by which to judge a statistical statement is the simple test of reasonableness. Is the conclusion consistent with other known data?

Care is needed in determining the existence and extent of relationship between facts reported statistically. Cause and effect are not always to be deduced because two factors move together. For example: we see in a report that 90 out of every 100 bus drivers have gastritis some time between the ages of thirty and forty. We are not justified in concluding that there is any special connection between driving a bus and having gastritis, if this is all the information we have.

About predictions

It is important to all business enterprises to be able to estimate with at least fair accuracy what the demand will be for products and what the supply will be of components. Such an estimate enables production and inventory to be adjusted to the probable level of sales.

Statistics will not do all this work, but conclusions based upon statistics can usually be applied to the future, under guidance of the user's common sense and experience.

The fact given statistically is the trend-to-now. When we go on to estimate the trend-from-now we must include in our educated guess "everything else being equal" and "present trends continuing." A nonsense illustration by Darrell Huff in *How to Lie with Statistics* (W. W. Norton Inc., New York) will serve as a warning against accepting trend-to-now without challenge as an indicator of the future. He is writing about the trend of television. "The number of sets in American homes increased around 10,000 per cent from 1947 to 1952. Project this for the next five years and you find that there'll soon be a couple billion of the things, Heaven forbid, or forty sets per family."

Qualified by the two phrases given above, trend study may pay off well. Too frequently enterprisers are deceived by surface appearances and move into the future without detecting the ground swell that would be revealed if they studied wider and deeper aspects of their own business and surrounding businesses.

For those interested in following up this general and scanty survey, there are excellent books available, including the following: Croxton and Cowden: *Applied General Statistics* (Prentice-Hall Inc., New York); Greendlinger: *Financial and Business Statements* (The Alexander Hamilton Institute Modern Business Texts, New York); Neiswanger: *Elementary Statistical Methods* (The Macmillan Company, New York and Toronto); Riegel: *Elements of Business Statistics* (Appleton, New York); and Arkin and Colton: *An Outline of Statistical Methods* (Barnes and Noble, Inc., New York). A list of statistical reports published by the Government of Canada may be obtained from The Queen's Printer, Office of the Supervisor of Government Publications, Ottawa (the *Annual Catalogue* costs $1).

In concluding we quote from Mr. Moroney's *Facts from Figures:* "If you are young, then I say: Learn something about statistics as soon as you can. Don't dismiss it through ignorance or because it calls for thought . . . If you are older and already crowned with the laurels of success, see to it that those under your wing who look to you for advice are encouraged to look into this subject. In this way you will show that your arteries are not yet hardened, and you will be able to reap the benefits without doing overmuch work yourself. Whoever you are, if your work calls for the interpretation of data, you may be able to do without statistics, but you won't do so well."

Authorized as second class mail,
Post Office Department, Ottawa

PRINTED IN CANADA
by The Royal Bank of Canada