APPLICATIONS, USES, AND ABUSES OF STATISTICS

Edited by
Wilbert Marcellus Leonard II
Professor of Sociology

FOREWORD

The purpose of this Reader is to supplement and complement the standard statistics text. Having taught statistics for two and one-half decades, I have been somewhat disquieted by the amount of time necessary to convey the principles of computation (even with the assistance of computer technology) and interpretation of statistical formulae at the expense of practical application and usage. To fill this void, I have edited a wide range of statistical materials and selected those that would facilitate students' appreciation of the actual use, and sometimes abuse, of quantitative tools. My hope is that these Vignettes stimulate the "statistical imagination" and cultivate "mathematical literacy."

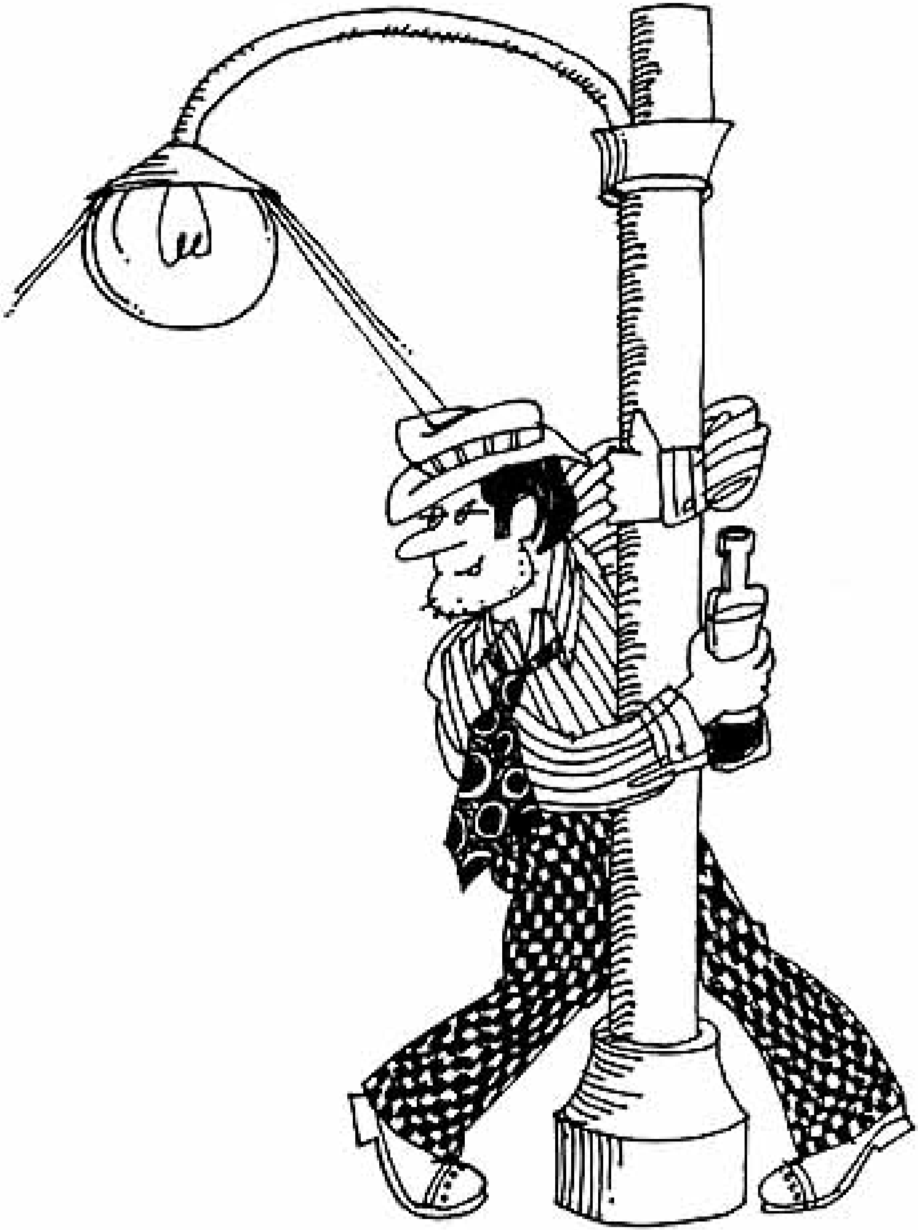
I have organized the "Readings" into three major sections. First, materials are arranged by their use in various fields, e.g., journalism/social policy, business, health, politics, biology and energy. These essays lucidly illustrate some contemporary uses of statistical analysis in several fields. Second, materials are organized by the subtopic within statistics that they most clearly fall into, e.g., Frequency Distributions and Graphic Representation, Measures of Central Tendency, Measures of Dispersion, the Normal Curve and z-Scores, and Measures of Association and Regression Analysis. These topics fall under the rubric of Descriptive Statistics. Examples of Hypothesis Testing and Parameter Estimation are subcomponents of Inferential Statistics and are organized as such. Third, several miscellaneous statistics selections are also contained herein.

This supplement was supported by the Faculty and Instructional Development Program at Illinois State University and the first edition was completed during the summer of 1981. The second edition was produced during the summer of 1983 and the third edition during the fall of 1987. I would like to express my thanks to the authors of the original writings. The essays have been adapted or used in total for inclusion in this booklet. My appreciation to the staff in the Department of Sociology - Pat Walsh, Mickey McCombs, Christine Gruber - and Penny Williams for their assistance on this, and other, projects.

Wilbert Marcellus Leonard, II
Professor of Sociology
Winter 1996

Menu

Table of Contents



It has been said that many people use statistics in much the same way that an inebriate uses a lamp pole: more for support than for illumination.

"Seventy-three percent are in favor of one through five, forty-one percent find six unfair, thirteen percent are opposed to seven, sixty-two percent applauded eight, thirty-seven percent . . ."

IF NUMBERS MAKE YOU NUMB...

...You're not alone. You can count yourself among a multiplying percentage of people who find themselves taxed by the array of figures upon which the smooth functioning of this increasingly complex society is calculated.

Innumeracy, the inability to understand and deal with numbers, is the mathematical equivalent of illiteracy, the inability to read. It afflicts consumers who can't balance their checkbooks or who don't understand credit-card interest, doctors who unwittingly misinform patients about the risks of operations, gamblers who wager money on near-impossible odds. Also included are investors duped by stock-market scams, and fast-food franchises that, for the sake of undereducated, underpaid employees, must place pictures of menu items on cash registers that automatically add appropriate taxes.

What would one megaton of TNT do? (The world's nuclear-weapons stockpiles contain 25,000 megatons of explosive power for every man, woman and child on earth.) What does it mean when a country spends a million dollars, or a billion? (It takes about 11½ days for a million seconds to pass—almost 32 years for a billion seconds to tick off.)

What are your odds of being killed by terrorists? (In a recent year, 17 Americans were killed by terrorists, out of 28 million who traveled abroad. That's one in 1.6 million.)

"I'm distressed by a society which depends so completely on mathematics and science and yet seems so indifferent to the innumeracy and scientific illiteracy of so many of its citizens," writes mathematician John Allen Paulos (*Innumeracy: Mathematical Illiteracy and Its Consequences*, New York: Hill and Wang, 1988, page 134).

Those citizens, says Dr. Paulos, "seem to have no mathematical frame of reference and no basic understandings on which to build. They're afraid. They've been intimidated by officious and sometimes sexist teachers and others who may themselves suffer from math anxiety" (page 88).

Society's failure to effectively educate people about numbers contributes to disparities in employment, among other problems. For example, women often are steered away from heavy mathematics and science curriculums. This denies them better, more lucrative jobs, though they're just as capable as men of understanding math.

"Women, in particular, may end up in lower-paying fields because they do everything in their power to avoid a chemistry or an economics course with mathematics or statistics prerequisites," writes Dr. Paulos. "I've seen too many bright women go into sociology and too many dull men go into business, the only difference between them being that the men managed to scrape through a couple of college math courses" (pages 78-79).

What dosage of medicine should you take? Which investment will produce a better return for you? What, exactly, is a calorie and how many of them do you need? When the media scream about "epidemics" such as AIDS, how scared should you get?

Everyone of us needs to be able to use numbers to reason inductively, to analyze, to make proper judgments, to estimate. Acquiring basic numeracy means overcoming the fear of figures.

Norman L. Sheaf

what do statisticians do?

For many people the word "statistics" brings visions of zig-zag graphs, batting averages, passes completed—all the tabulations on the sports pages or columns of figures in the business section of newspapers. This is only one kind of statistics; the statistician today does far more than construct and examine graphs and tables. He or she must develop a greater understanding of the origins of data, their possible meanings, and especially their accuracy. Knowing that Council Bluffs, Iowa had 60,348 inhabitants in 1970 is not enough. How was the count made? Did it include students temporarily away at school? Or those students temporarily living there during the school year? How about the new-born babies in the hospital? Did the enumerators do their job well? Important decisions, such as who gets what amount of Federal funds in revenue sharing, depend upon the answers to these questions.

U.S. Government statisticians conduct Censuses of Population, Housing, Manufactures, and Agriculture. Compilations are made of sales, production, inventories, payrolls, and other internal industrial and business data. These statistics tell the alert business manager how his industry is growing, how his company is growing in relation to the industry, and what plans he should make for future expansion. Government officials could not function without this information; planning would come to a standstill.

The statistician also helps the manager make wise decisions. When a food chain, for example, considers opening a new branch of a supermarket, it turns to the statistician to help find the best location. Sample surveys help determine the prospects for success, by measuring such items as population concentrations, income levels of potential customers, the availability of transportation, the needs of the community and existing competition. The expert interpretation of these data puts the statistician right at the heart of the final decision.

Central to the work of many statisticians is the use of computers. Statistical calculations that once took weeks of labor can now be done on a high-speed electronic computer in a few seconds. Some statisticians use the computer to analyze data. Others use it to help solve thorny statistical problems whose mathematical complexity might otherwise be overwhelming. Almost all agree that what the test tube is to the chemist, the modern computer is to the statistician.

Statistics is a changing field with new methods being generated constantly. The user of statistical techniques regularly has more and better tools at hand to help solve his problems.

There seems to be no limit to the areas of challenge to the statistician. Let's look at some specific examples of statistical problems, in widely ranging fields.

<u>Careers in Statistics</u>, American Statistical Association, p. 5.

SIR RONALD A. FISHER

The ideas and methods that we study as "statistics" were mostly invented in the nineteenth and twentieth centuries by people working on problems that required analysis of data. Astronomy, biology, social science, and even surveying can claim a role in the birth of statistics. But if anyone can claim to be "the father of statistics," that honor belongs to Sir Ronald A. Fisher (1890–1962).

Fisher's writings helped organize statistics as a distinct field of study whose methods apply to practical problems across many disciplines. He systematized the mathematical theory of statistics and invented many new techniques. But the randomized comparative experiment is perhaps Fisher's greatest contribution.

Like many statistical pioneers, Fisher was driven by the demands of practical problems. Beginning in 1919, he worked on agricultural field experiments at Rothamsted in England. How should we arrange the planting of different crop varieties or the application of different fertilizers to get a fair comparison among them? Because fertility and other variables change as we move across a field, experimenters used elaborate checkerboard planting arrangements to obtain fair comparisons. Fisher had a better idea: "arrange the plots deliberately at random."

This chapter explores statistical designs for producing data to answer specific questions like "Which crop variety has the highest mean yield?" Fisher's innovation, the deliberate use of chance in producing data, is the central theme of the chapter and one of the most important ideas in statistics.

Garbage In, Garbage Out

Several years ago Time magazine took a hard look at the American fascination with statistics -- and the doubtful quality of some of those statistics. In an editorial essay entitled "The Science and Snares of Statistics," Time suggested:

> This dedication to numbers has created its own pitfalls for the innocent -- and opportunities for the purveyors. There is an air of certainty about the decimal point or the fractionalized percentage -- even in areas where the measurement is statistically absurd or the data basically unknowable. A classic example is a survey made some years ago, which solemnly reported that 33-1/3% of all the coeds at Johns Hopkins University had married faculty members. True enough, Johns Hopkins had only three women students at the time, and one of them married a faculty member. The American Medical Association announces not that very few people dream in color, but that "only 5% of Americans" dream in color. New York City has 8,000,000 rats. How does anybody know? Statisticians have a phrase for this, borrowed from the computer industry on which they now rely. The phrase is "garbage in, garbage out" — meaning that the result that comes out is only as good as the material that is fed in.
>
> For the sake of drama or publicity, numbers are slapped on nearly everything -- and the bigger the number the better . . . . Newsmen during the Detroit race riots pressed a harried fire chief for damage estimates. His guess: $500 million. So far, in the cooling aftermath of riot, insurance companies are processing only $84 million worth of damage claims, and the overall loss is now put at $144 million. For newsmen, the National Safety Council issues forecasts of expected highway deaths over holiday weekends usually with a prediction tacked on of "record fatalities". What the forecast never says is that the record is due to population increases and wider use of automobiles, and that the fatality rate is usually just as high proportionately on other weekends -- holiday weekends are just a bit longer . . . .
>
> Since it is obviously impractical to poll the nation on anything less important than the selection of a President, one cherished statistical tool is the sample. Not even statisticians can agree on how big or good a sample can be relied upon as representing the whole. Dr. Alfred C. Kinsey's celebrated reports were criticized by statisticians not so much for their moral implications but because they made sweeping presumptions on the basis of too small a sample (in the male study, only 5,300 men provided data). The Nielsen ratings, by which television programs live or die, have been justly attacked because Nielsen recorders are necessarily hooked to the sets of those viewers willing to have a recorder -- a special class by definition, whose tastes may or may not correspond with those of the unpolled millions of the total TV audience.
>
> Still the state of the art of statistics has come a long way since 1661, when its founding father, London Haberdasher John Graunt, began a careful count and found that more boy babies died in infancy than girls, and concluded that therefore, there must be more women than men in Britain. Today's scientists, who no longer believe that anything is absolutely certain, also believe that many things are predictably probable. And it is the computer, fed with vast amounts

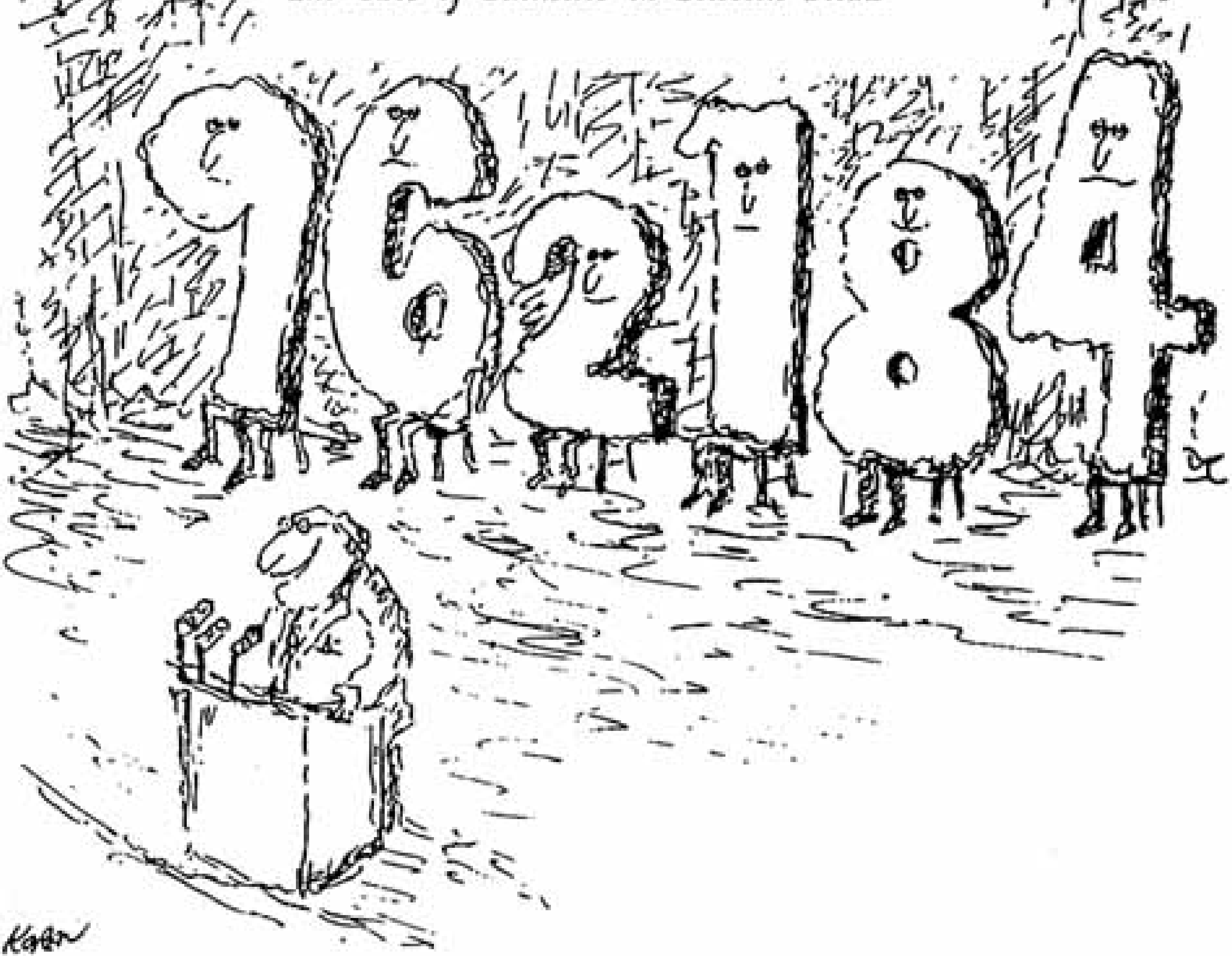
of past data, that can project or at least outline the alternatives of several possible futures. "The computer has enshrined statistics," says M.I.T.'s Professor Harold A. Freeman. "Without it, statistics would still be a grubby business." Where once all they had to do was count, and perhaps draw graphs, statisticians are now "programmers," with a mystique all their own. Unquestionable, for the moment, numbers are king. But perhaps the time has come for society to be less numerically conscious and therefore less willing to be ruled by statistics.[1]

Since it doesn't seem likely that society will become "less numerically conscious," at least in the near future, a better suggestion might be that we try to become more statistically sophisticated, more familiar with proper usage of statistics and more aware of their limitations. We might move in this direction by including statistics in the standard high school -- or even elementary school -- curriculum. Maybe the time has come for Reading, 'Riting, and Regression analysis.

[1]Condensed from "The Science and Snares of Statistics," Time (September 8, 1967): 29.

Understanding and Using Statistics-Basic Concepts, pp. 37-38, Marty J. Schmidt (Second edition), 1979.

The Uses of Statistics in Selected Fields



"Tonight, we're going to let the statistics speak for themselves."

From the Literature

Sociology is a rather new discipline. In fact, so new that many modern-day sociologists regard Emile Durkheim (1858–1917) as one of its founders. Durkheim set the tone for sociology as a research-oriented science of society that uses various statistical techniques for obtaining, summarizing, and interpreting data in order to acquire knowledge.

Emile Durkheim is especially noted for several major works, including *Le Suicide*, which was written in French and published in 1897. It was translated into English in 1951.

Durkheim's *Suicide* is clearly a classic. Many of the ideas contained in it are presented in almost every basic sociology text. We often remember Durkheim for his typology of suicides: altruistic, anomic, and egoistic. He should also be remembered for publishing the first empirical study in which a sociologist used a variety of simple yet appropriate statistical tools (such as rates, percentages, and proportions) to summarize and interpret data.

Among the data described and summarized in *Suicide* are the following:

suicides per 100,000 inhabitants

deaths per 100,000 inhabitants

percentage differences in suicide rates between two consecutive years

suicide rate per million inhabitants in 11 European countries

number of insane persons per 1000 inhabitants of each religious faith by country

Based on these and other research findings, Durkheim highlighted the relationship between group conditions and suicide. Specifically, he showed that individuals are integrated into social groups and that the risk of suicide depends, in part, on the extent of that integration. In addition, he noted that suicide depends on the quality of group life, which can range from stable and supportive to unstable and nonsupportive. In this context his data suggest that suicide becomes a viable option for persons weakly integrated into the group. It also becomes a viable option among highly integrated individuals when the group itself supports self-destruction. Finally, it is a viable option when individuals are caught up in a disorganized environment, an environment that does not provide clear-cut guidelines for meaningful actions.

Lyman Ott, Richard F. Larson, & William Mendenhall, Statistics (Boston: Duxbury Press), 1987: 20.

When we laugh at monkeys in the zoo, we may be laughing at ourselves. Statistics have shown a correlation between animal and human behavior and between animal behavior and natural phenomena.

Because Man is a primate, scientists are carefully observing monkeys, chimpanzees, gorillas and other primates to learn more about human behavior. Jane Goodall, famous for studying chimpanzees in the African wilds, gathers data on such chimp behavior as lone play, group play, mating, grooming, restlessness, aggressiveness, and closeness to others.

At the Stanford Outdoor Primate Facility (SOPF) at Stanford University, California, the nature, frequency, and duration of these behaviors have been sampled on a statistically random basis and the data fed into a computer. Correlational statistics are being compiled on development behavior and hormonal changes during puberty. Scientists want to learn if social behavior in adolescent chimps is related to body biology.

Because they had the data available, Stanford researchers were able to study another interesting correlation. Reports from China after the catastrophic earthquake on February 4, 1975, indicated that striking changes in animal behavior had been observed before the quake. The researchers decided to go back to look at their data to compare chimp behavior at SOPF before California quakes.

The researchers correlated dates of shocks registered on nearby U.S. Geological Survey equipment with chimp behavior. They discovered that the chimps had been significantly more restless and group oriented (statistically speaking) for 12 hours preceding seismic shocks in June and July of 1975.

Since the behavior data were collected and stored in the computer before the previously unpublished seismic activity was made available, it is assumed no observer bias colored the data. The researchers, however, caution that such a small animal group (15) and limited events (2) make too small a sample for generalizing from their statistical analysis. Nonetheless, speculation on the possibilities of using animals to predict earthquakes is fun. And certainly this study demonstrates the value of quantitative recording of even commonplace events.

Source: Ukena, Ann Seymour, Statistics Today, Harper & Row, Publishers, 1978, p. xiii-xix.

STATISTICS TODAY IN BUSINESS

When you are in business, you are always trying to outguess your customers. You have to. Unless you can judge ahead of time how many of any item customers will buy, you cannot know how many to order if you are a retailer, or how many to manufacture if you are a manufacturer. It's a tough problem.

Statistics help a lot. People in business gather statistical data to learn customer preferences and attitudes. This is called market research and helps companies to judge what turns consumers on and off.

The Coco-Cola Co., for instance, confirmed through market research that the famous "Coke" bottle does indeed induce customers to "reach for a Coke." That contour shape is known throughout the world. But Coca-Cola executives were not sure the bottle shape would play any significant role in sales of a new 16-ounce size the company planned to introduce several years ago.

A straight-walled bottle instead of the traditional contour one would save glass and speed the production line when the bottles were being filled. But executives want to be sure no sales would be lost.

So, a market test was run to check the hypothesis that potential customers identify with the "Coke" bottle shape sufficiently strongly to be lured into purchase more often.

Researchers interviewed about 800 people in shopping centers in some five cities. Shoppers were shown one bottle or another, then asked questions on intention to buy "Coca-Cola" that day. Later, answers were cross-tabulated and checked with known buying patterns from previous years at each shopping center.

When the data were analyzed, the contour bottle indeed proved to have extra selling power in the 16-ounce size. The traditional shape was retained.

Market research takes many forms. Sometimes businesses want to evaluate the potential effects of promotional programs before spending large sums of money. Then computer studies may be used. The computer compares past history with expected changes in customer habits. From these projected results, marketing executives decide whether the contemplated campaign is likely to produce statistically significant improvements in company sales.

You can be sure that whatever you intend to merchandise as a business person, you will look at statistics before making a major decision.

Source: Ukena, Ann Seymour. Statistics Today, Harper & Row, Publishers, 1978, pp. 350-351.

Energy. Where is it going to come from, and how are we going to use it in the future? Researchers are busy gathering statistics trying to forecast answers to both those questions.

Many persons suppose the sun is likely to be a leading energy source by the year 2000, while others believe wind will provide us considerable power. Researchers, however, predict no more than 6 percent of the United State's total energy needs are apt to be supplied by solar power by the year 2000. Wind is rated even lower.

These figures come from research being conducted on solar and wind energy as potential future power sources in Arizona, California, and New York, among other states. Similar research is going on for other energy forms: nuclear, geothermal, hydro(water), oil, gas, and coal.

Accurate forecasting depends on reliable statistics. Utility companies use statistical data to forecast what our main energy sources will be so that they can decide now how to invest in equipment for future power production. Manufacturers of autos, appliances, steel, and other products also must know what fuels will be available. And the general public needs to know in order to make intelligent policy decisions about environment that may change with varying fuel choices.

With statistics we can forecast our future energy habits now so we can tabulate statistically our future needs against possible energy sources. Determining need is a complicated process. Social customs, technology, climate, economics, and environment must all be considered.

Researchers must find answers to questions like: How many people are in a household? What new appliances have caught consumer fancy?

Will technical breakthroughs make solar power more widespread than predicted? How often are furnaces or air conditioners turned on? What's the price of energy? What pollutants might various energy sources create?

Researchers reduce all these environmental, social, and economic variables to statistical data. These data are converted to computer models that play-off statistics, one against another, in changing patterns to simulate actual behavior. Sometimes these computer-projected patterns are checked out in live situations.

For example, 8-hour workdays determine when most residential and industrial customers make peak demand on utilities. Were utilities to price power based on time of day--cheaper when demand is low, more expensive when demand is high--would consumers change their habits? Computer models say maybe. A utility company in Connecticut is conducting real-life experiments to find out.

Interpreting the statistical data from the experiments being conducted now will help shape how we use energy in the future.

Source: Okeun, Ann Seymour, <u>Statistics Today</u>, Harper & Row, Publishers, 1978, pp. 68-69.

Statistics help health professionals provide quality care to patients everyday. As nurses make their hospital rounds, they are constantly on the lookout for patients' vital signs: temperature, blood pressure, respiratory and pulse rates. They have to understand the standard deviation for each of these vital signs so they can take action if they spot a statistically significant variation.

Health professionals who work in special-care areas, such as the cardiac unit, rely on statistics even more. In these units, patients who have had heart attacks are hooked to monitoring machines that record heart rhythms, pressure within heart chambers, blood pressure, and other important body functions, which give clues to the likelihood of repeat heart attacks. Nurses, medical technicians, and physicians keep tabs on the machines that display precise information. These health professionals recognize when statistically significant changes in information from any body function is displayed.

Nurses, for example, who observed abnormal heart rhythms might administer medicine or help the patient relax, as measures to restore normal rhythms. Arrhythmias, as abnormal heart rhythms are called, are the number one cause of death after the most common type of heart attack, myocardial infarction. In the 15 years since arrhythmia monitoring was first introduced, nurses have reduced arrhythmia deaths within hospitals by about 20 percent.

The same statistical approach to caring for all patients is taking place in many hospitals around the country, though monitoring machines are not used. In a new auditing technique, diseases are reduced to a standard list of nursing-care criteria. These criteria are derived from statistical analysis of patient charts. Stroke patients, as one example, might require help in bathing, turning, eating, physical therapy; they might have to be watched carefully for speech difficulties, pupil changes, reduced bodily movements.

Once the criteria are established, a minimum desirable compliance percent is set; perhaps 98 percent for one disease, 90 percent for another. A compliance record compiled mainly from nurses' notes is audited regularly. Nursing quality control committees meet regularly, too, to analyze compliance deficiencies. The committees then make recommendations on what needs to be done to bring the quality of care up to the statistically desired rate.

Thus valid statistical procedures provide objective, consistent decisions in situations for which the chief guideline formerly was intuition. Health professionals now combine quantitative data with qualitative concern to bring top care to sick people.

Source: Ukena, Ann Seymour, Statistics Today, Harper & Row, Publishers, 1978, pp. 286-287.

Journalists would be surprised to be told they are walking statisticians since most hate math. They have, however, been trained to gather observations, which they process and reduce to readable articles.

Observations analyzed and reduced to quantitative terms are called data. Data organized in summary forms are called statistics. Thus, in a sense, a journalist is a statistician who summarizes data in a way the general public can understand.

Journalists also consciously use statistics in developing meaningful stories. They interpret masses of public data gathered everyday by government and other organizations. From this data they write exposes of illegal campaign contributions, for example, or why suicide rates are climbing among suburban youth.

Typical of modern journalism was an important statistical analysis made several years ago by two reporters on the Philadelphia Inquirer. They decided to scientifically measure the administration of justice in the local court system.

The reporters were interested in answers to such questions as: Are blacks and whites treated alike in the judicial system? How similar are sentences for the same crime among judges in Philadelphia? Do certain types of crimes result in harsher sentences even though legally they are of the same seriousness as other crimes?

The reporters studied 1034 persons indicted during 1971 for one of four major crimes: murder, rape, aggravated robbery, and aggravated assault and battery. These 1034 defendants represented a statistica--, random 39 percent sample of all persons indicted for those crimes during 1971.

More than 10,000 court documents and some 20,000 pages of notes of testimony concerning trials and other legal proceedings were used to feed data into the computer for analysis.

The Inquirer story of the study disclosed fascinating information, interesting to courts around the country as well as to Philadelphia residents. For example, only 30 percent of all persons convicted under one judge actually served prison sentences of any sort. What is the implication of such a statistic? Does it mean, as many claim, that criminals are being released back into society?

Noticeable statistical differences among blacks and whites also showed up. Black criminals were sent to jail for 7 months or longer 20 percent more often than white criminals, for example, while black judges imposed longer sentences 6 percent more often than white judges.

These are just a few of many facts the journalists wrote about after their statistical analysis. But these facts indicate the useful sociological information buried in dry statistics, information journalists and other social-problem researchers can present to society at large. Then it is up to society to decide what use to make of this knowledge.

Source: Ukena, Ann Seymour. Statistics Today, Harper & Row, Publishers, 1978, pp. 14-15.

Political opinion polling before elections is an excellent example of predicting results from a small sample. Pollsters need check only a small fraction of the voting population to come close to what actually happens on any given election night.

Often the forecasts are within five percent above or below the voted preference. That may appear to be a wide margin of error. But when you are assessing independent choices made by millions of individuals, that tolerance is sufficiently realistic to prove useful.

Pollsters achieve reasonable predictions by careful statistical probability principles. The goal is for evey eligible voter to stand an equal chance to be included in a sample. If this would be achieved, a truly random sample would result.

However, true random samples seldom work out. There are too many difficulties: finding the right people at home to give an accurate cross section of socioeconomic status, sex, age, race, and other variables, including honest answers to pollster questions.

Instead of a true random sample, pollsters try for artificial randomization that gives a representative sampling of the voting population. Various techniques are used, among them stratification.

For example, California has 10 million registered voters, out of 15 million eligible adults, of who some 8 or 9 million vote in any election. Poll sampling in that state stratifies, or divides the state into population concentrations. Then a sample of about 1000 persons is distributed throughout the state. That is, since Los Angeles county represents 38 percent of the state's voting population, 38 percent, or 380, will come from that county. And so it goes.

Within each stratification, census tract information provides the breakdown on the variables necessary to reflect an accurate picture of that voting community.

Usually those to be polled are put into distributive groups based on information collected from census data. Pollsters begin at some random start then pick, say, every one-thousandth individual until they have sufficient people. Extra persons are included to allow for those who cannot be reached and for balancing variables.

The above explanation is overly simplified and does not explain correction procedures for biases that inevitably creep in: if interviews are during the day, the sample is biased against many employed people; if in the evening, movie-goers are neglected. It's difficult to overcome them all.

But reputable pollsters know how to compensate for many of these biases. Nowadays we have fewer major blunders, such as Truman's unexpected win over Dewey, that made laughingstocks of opinion polls in years past.

Source: Ukena, Ann Seymour. Statistics Today, Harper & Row, Publishers, 1978, pp. 212-213.

From the Literature

Public policy is often based on social research findings. For example, one of the major policy interpretations set by the U.S. Supreme Court was based, in fact, on a series of research findings. That policy interpretation put an end to legally sanctioned segregation in the public schools, and it came about as a direct result of the Supreme Court desegregation decision (*Brown v. Board of Education*) of 1954. For many experts in the field of race relations, the 1954 Supreme Court ruling was the beginning of the militant civil rights activities.

Prior to the Court's decision, the seventeen Southern states and the District of Columbia had two complete sets of school systems (at the elementary and secondary levels). One set was for whites, the other for blacks. Previous Court rulings had upheld the doctrine of "equal but separate" school systems. Social research data, however, systematically revealed that the two systems, while separate, were not equal.

Sample sizes for the social research studies were rather large in several instances and required considerable summarization. One, for example, involved data collected on nearly 15,000 Southern blacks. Some of the data summarized and made available to the Court during its deliberations included the following:

> A U.S. Commission on Civil Rights study reported the median number of school years completed by blacks who were 25 years of age and over was 8.1, while the median for whites was 11.4.
>
> Performance tests showed that blacks in Northern schools (Illinois, New York, Ohio, and Pennsylvania) scored higher, on the average, than whites in Southern schools (Arkansas, Georgia, Kentucky, and Mississippi).
>
> These same performance tests showed white Southerners scored, on the average, higher than black Southerners.

Lyman Ott, Richard F. Larson, & William Mendenhall, Statistics (Boston: Duxbury Press), 1987: 104.

FLORENCE NIGHTINGALE

Florence Nightingale (1820–1910) won fame as a founder of the nursing profession and as a reformer of health care. As chief nurse for the British army during the Crimean War, from 1854 to 1856, she found that lack of sanitation and disease killed large numbers of soldiers hospitalized by wounds. Her reforms reduced the death rate at her military hospital from 42.7% to 2.2%, and she returned from the war famous. She at once began a fight to reform the entire military health care system, with considerable success.

One of the chief weapons Florence Nightingale used in her efforts was data. She had the facts, because she reformed record keeping as well as medical care. She was a pioneer in using graphs to present data in a vivid form that even generals and members of Parliament could understand. Her inventive graphs are a landmark in the growth of the new science of statistics. She considered statistics essential to understanding any social issue and tried to introduce the study of statistics into higher education.

In beginning our study of statistics, we will follow Florence Nightingale's lead. This chapter and the next will stress the analysis of data as a path to understanding. Like her, we will start with graphs to see what data can teach us. Along with the graphs we will present numerical summaries, just as Florence Nightingale calculated detailed death rates and other summaries. Data for Florence Nightingale were not dry or abstract, because they showed her, and helped her show others, how to save lives. That remains true today.