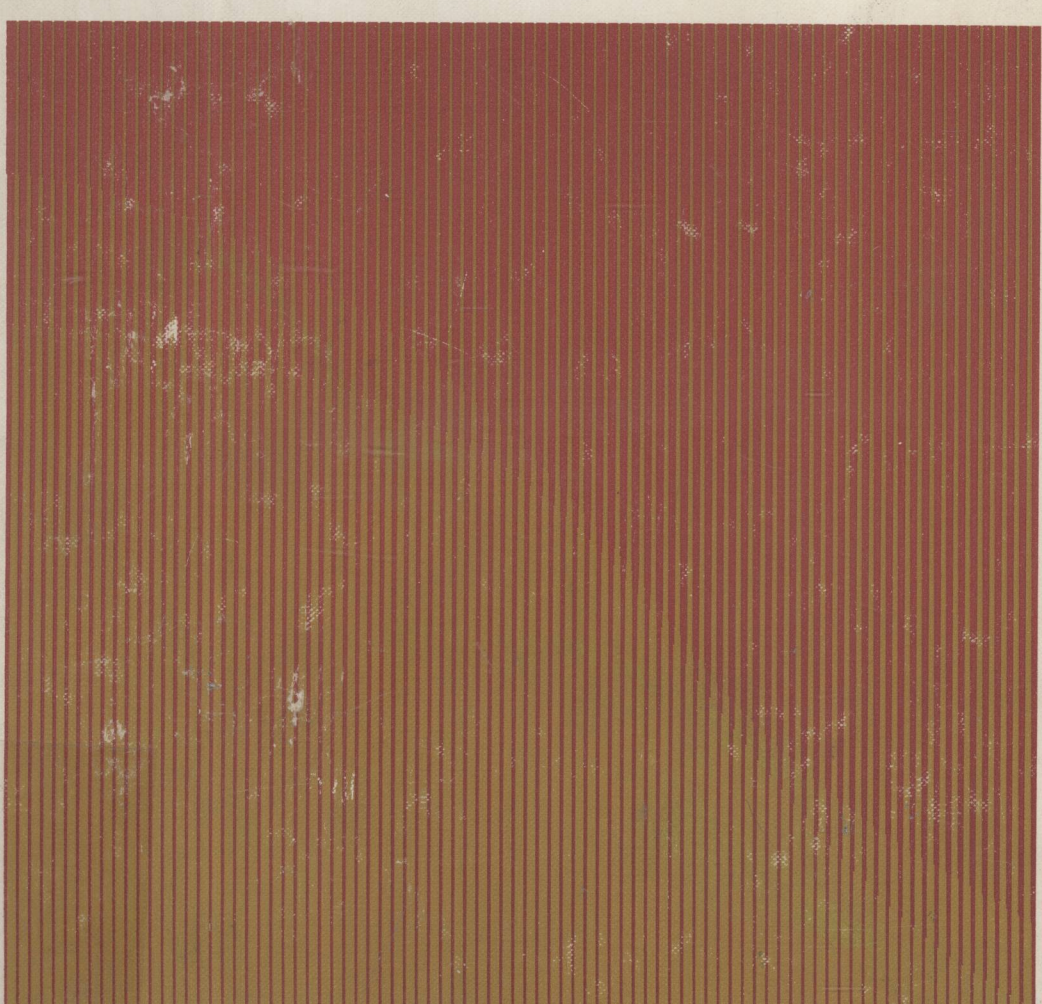
HOW TO USE STATISTICS

Joe D. Megeath



HOW TO USE STATISTICS

Joe D. Megeath
Metropolitan State College

With the assistance of
Allen T. Peck
Arapahoe Community College

CANFIELD PRESS  **SAN FRANCISCO**
A Department of Harper & Row, Publishers, Inc.
New York Evanston London

Sponsoring Editor: Jerry Papke
Production Editor: Pat Brewer
Copy Editor: Don Yoder
Designer: Penny L. Faron
Illustrator: Paulette Hanson
Line Artist: Carl Brown

Library of Congress Cataloging in Publication Data

Megeath, Joe D 1939–
How to use statistics.

Includes bibliographies.
1. Statistics. I. Title.
HA29.M45 519.5 74–28237
ISBN 0–06–385445–7

HOW TO USE STATISTICS

75 76 77 10 9 8 7 6 5 4 3 2 1

HOW TO USE STATISTICS

PREFACE

THIS TEXT HAS been written as a basic statistics primer for people who have no prior knowledge of the field. The user may be a formal student or may be someone on the job needing a reference book at work. The material has been generally directed toward business applications, but persons in other disciplines will find it relevant and useful. The contents are well suited to a two-quarter course, a one-semester course, or with proper selection of topics, a one-quarter course.

The material assumes no more than a working knowledge of basic algebra. There are no mathematical derivations. The emphasis throughout is on introducing topics on an intuitive level and then developing these topics into formal statistical concepts. Purely academic presentations have been studiously avoided whenever possible.

Most people I have encountered use statistics as a drunk uses a light pole—for support rather than illumination. This could be due in part to the human disposition, but it is due also to a lack in our training. In most institutions, the statistics course has one of the highest "fatality" rates of the courses offered. Students drop the course, fail the course, and have been known to change their major to avoid the course. And among those who have successfully "passed" their statistics requirements, there appears to be a woeful lack of understanding of what they studied. I remember in particular a colleague who said, "Oh, yeah, I took a statistics course in graduate school. It had something to do with a bell-shaped curve." Good Lord! A graduate course in statistics and he remembers "it had something to do with a bell-shaped curve"—what a waste!

Each chapter in this book begins with a set of general objectives that will alert the reader to what he or she should be looking for in the chapter. The problems are carefully coordinated with these objectives and cover a broad range of difficulty. There are problems for stu-

dents who need a helping hand with basic concepts, and there are problems that are geared to push the reader to transfer his or her understanding of concepts to entirely new situations. At the end of each chapter is a summary of information and equations to aid the reader in reviewing the topics covered. Some chapters also contain an autopsy, which is designed to point out particular pitfalls or important subjective areas. The objectives, summaries, autopsies, and the motivating light-hearted examples and problems make this book very appropriate for self-paced courses.

The reader of this book will also be challenged to go beyond the cut-and-dried syndrome that can plague statistics. The fact of the matter is that the *use* of statistics is not cut and dried; there are many subjective decisions to be made by the analyst. I fully realize that it is far easier to ignore these subjective decisions for both the student and the instructor. However, practical experience has indicated that the easiest route is not necessarily the best route.

I have taught statistics to students of business, engineering, education, behavioral sciences, and mathematics. It has been my experience that *more* students get a *better* understanding of the subject if it is heavily seasoned with realistic examples and lively problems. Consequently, you will occasionally find that you don't know whether to laugh or to cry as you read. This is a noticeable improvement, though, over definitely knowing that you want to cry.

I would like to acknowledge the many students and friends who helped in a hundred different ways in the preparation of this text. Particularly, I would like to thank Paulette Hanson for her initial work without promise of pay; Howard Flomberg for the major share of the computer programming needed; and my wife, Donna, for putting up with a lot during this writing. This book might not have happened without the inspiration provided by my friends Irv Forkner, R. E. D. Woolsey, and John Rushton.

Joe D. Megeath
Denver, Colorado
January 1975

CONTENTS

PART I
HOW TO DESCRIBE DATA

CHAPTER 1
STATISTICAL AWARENESS

OBJECTIVES: After studying this chapter and working the problems, you should be able to:

Demonstrate with a current example your optimistic skepticism toward statistical data.

Explain why figures don't necessarily tell all the truth.

IN TODAY'S WORLD, including both our work world and our private world, we are besieged with numbers. We are exposed to long lists of sales figures, charts of costs, and such statements as:

"Our profits are 1 percent."

"A survey shows 55 percent of the voters support the bond issue."

"The average income of our salesmen is $20,000."

Statements like these can be heard or read almost every day. These three comments have two things in common: they all involve numbers and not one of them means anything standing by itself. Take a look at the statement, "We couldn't possibly be contributing to inflation. Our profits are only 1 percent." What does that mean? One percent of what? Is it a 1 percent return on investment? That's hardly likely in view of the fact that you can invest your money safely in a bank for far more than a 1 percent return. Perhaps the man meant 1 percent of sales. If that's the case, then we'd better look more closely. If I

loan you 99 cents and at the end of the day you pay me back $1, then my profit of 1 cent is 1 percent of the amount you paid me. If we do this every day, however, I'll make a penny each day and at the end of one year I will have made $3.65. Putting that in terms of a return on my investment of 99 cents, I will have made 3.65/.99 = 369 percent return on my investment. The 369 percent profit (percentage of return on investment) looks a lot different than the 1 percent profit (percentage of sales), doesn't it? The point is that by looking at "1 percent profit" the amount of information you received was totally nothing. You still have no idea whether or not the profit was relatively high or relatively low.

Suppose in your community there is a controversial bond issue that would be used to build a new ten-lane, fiber glass bridge to replace the current footbridge over Little Dry Gulch at the edge of town. One morning you open the newspaper and the headline screams, "Survey Shows 55 percent of Voters Support Bond Issue." You might think, "Well, people seem to support this issue. I guess it must be a good deal." You might think, though, "Says who?" Perhaps the last sentence of the lengthy article states that the survey of 100 voters was conducted by the National Bridge Builders Association. That should lead you to conclude that the survey results, by themselves, tell you nothing. The Bridge Builders Association has a heavy interest in the project, and a survey can be gerrymandered in a lot of ways—either consciously or subconsciously. For instance, the people surveyed could have included all major suppliers of bridge-building materials in the area. Or the question could have been leading, such as "Do you support the new bridge bond or do you want to see your child drown in Little Dry Gulch?" These examples may seem extreme, but you should be wary of surveys if they are not conducted by a neutral party.

In the statement "The average income of our salesmen is $20,000," the implication to the lay reader is that most of the salesmen make about $20,000. If one salesman makes $100,000 and four salesmen make $0, however, the average income for the five is $20,000. Such general statements don't tell you anything by themselves because the word "average" can imply something different than it actually means.

If you get nothing else out of a study of statistics, you should learn to be optimistically skeptical of quoted numbers. This attitude of optimistic skepticism and an increased awareness of numbers is what I call *statistical awareness.*

Once you develop statistical awareness you will find that most

numbers you hear or read tell you very little without further investigation. Numbers are like the Bible in one sense. When they are quoted, they add credence to whatever is being said. But also, like the Bible, they are often quoted erroneously or out of context.

People often talk about football statistics or about something becoming a statistic, referring to actual counts and figures that are compiled for various reasons. This usage of the term has led many people to assume that the field of statistics deals with counting the number of births, recording the number of traffic accidents, adding the total yardage gained, and so forth. Such numbers are *statistical data.*

The term *statistics* refers here to a much larger scope, including methods of analyzing such data. These methods encompass techniques for summarizing and describing statistical data (descriptive statistics) and techniques for making inferences from sample data (inferential statistics).

The businessman today, large or small, receives statistical data in huge amounts. If he cannot get the correct information from those data, his business will suffer. To be statistically illiterate is to be at the mercy of any number (or lack of numbers) you happen to come across.

To protect yourself against "tell you nothing" data and to reach the proper decisions and actions from data, you must be able to handle data almost like a witness in a courtroom. A witness is sworn to tell the truth, the whole truth, and nothing but the truth. You must develop the ability to place statistical data on the witness stand and find out what information they really tell you, all the information they contain, and nothing more than the information they contain.

This ability requires you to ask pertinent questions of the number in the witness chair. If you read that "the unemployment rate is 5 percent," for instance, you should ask:

> Percentage of what?
>
> Does this include people who are seeking temporary employment?
>
> Does this include people who have obtained employment but haven't begun work yet?
>
> What was this percentage a year ago?
>
> Is this a national or a local figure?

Are there any extenuating circumstances that could cause this figure to be so high (or low)?

Who compiled the data?

The first step in putting statistical data on the witness stand is to develop a statistical awareness. Ask yourself what the numbers really mean and "says who?" The second step is to develop some technical skill and learn definite procedures for cross-examining the data in the witness chair. These procedures to be learned constitute the field of statistics.

The major point, though, is making numbers "talk" to you. Many people think that a number is a number is a number is a number. But that isn't necessarily true. Take the number 2. If I say, "I have two cents," the listener will yawn and think so what. But if I say, "I have two wives," the listener is immediately going to raise an eyebrow . . . at least! It is the same number in both statements, yet it obviously carries a different impact. (Haven't you ever been in a conversation where someone says, "Do you realize we sold ten purple thingamajigs yesterday?" And you don't know whether to say "That's too bad" or to say "That's great"?)

At other times we become immune to numbers. For instance, the cost to the United States of the war in Indochina was reported as $108 billion.* How big is 108 billion? That is, written out, 108,000,000,000. We often accept such figures because we don't grasp their magnitude. If a baby started counting one digit each second at the moment of his birth and passed this task on to his children, it would take more than 48 generations to count to 108 billion!

As you develop your ability with formal statistical techniques, you will also be (probably subconsciously) developing your statistical awareness. Start right now by looking more closely at the numbers being used around you. Some people claim that figures don't lie, but liars figure. You will begin to find that figures don't necessarily tell all the truth either.

PROBLEMS

1. Watch for a quoted or written figure that is misleading or unsubstantiated. Analyze why the figure tells you nothing.

* "The Fighting Finally Stops for the U.S.," *Time,* 27 August 1973, p. 34.

2. Watch for a quoted or written figure that you think is not misleading and is totally substantiated. Discuss it with friends to see if they agree with you.

3. Assume you are the police commissioner in your community. A newspaper headline proclaims that there were 1250 thefts during the last six months.

a. Not a pleasant figure, but what does it tell you?

b. What if there were 1100 thefts during the previous six months?

c. Would your reaction be any different if your town's population was 10,000 or 500,000 people? Why?

d. How would crime rate figures from a similar city help you decide whether or not to fire the police chief?

e. Suppose your city could get a federal grant of $10,000 to put a special antitheft squad on the street. To get the grant your city would have to put up $5000 and show a theft reduction of 20 percent. What questions would you ask before you made a decision?

SELECTED REFERENCES

Campbell, Stephen K. *Flaws and Fallacies in Statistical Thinking.* Englewood Cliffs, N.J.: Prentice-Hall, 1974.

Huff, Darrell. *How to Lie with Statistics.* New York: Norton, 1954.

CHAPTER 2
MAKING DATA USABLE

OBJECTIVES: After studying this chapter and working the problems, you should be able to:

Explain why the proper presentation of data is important.

Explain how and why ratios are used.

Discuss when to use a table and when to use a chart to display data.

Prepare examples of charts and tables (bar chart, line chart, cross-classification table, and others).

Illustrate why scale selection and area proportioning for charts are important.

DATA DON'T ALWAYS come as a single number or a pair of numbers. Often they appear as a crowd. When this happens, we need to organize them into a usable format. By usable format, I mean that the full impact of the numbers becomes apparent. This normally requires a table, graph, or chart. A look at the construction of tables, graphs, and charts will benefit the would-be analyst in two ways: he will learn how to prepare them and how to read them.

Preparing good tables, graphs, and charts requires a mixture of general technical knowledge about what you should and shouldn't do with them, common sense, and practice. There are two cardinal rules in preparing them. Observance of these two rules will help keep you out of trouble when you are displaying data: