JAMES A. GRAASKAMP COLLECTION OF TEACHING MATERIALS

I. MANUSCRIPTS

B. Unpublished Books

1. A Statistical Handbook for Appraisers (Incomplete) by James A. Graaskamp and Karel J. Clettenberg

Chap 1 - Some Basic Stuff about Statistics
Chap 2 - A Histogram is Not a Major Operation
Chap 3 - Population Control with Parameters
Chap 4 - Unwritten - Exercises
Chap 5 - A Sampler Can Embroider a Report:
Rough draft and includes correspondence with the publisher, the American Institute of Real Estate Appraisers, and working notes

Index

A STATISTICAL HANDBOOK FOR APPRAISERS

A Preliminary Draft

For the Exclusive Use of the

Research Committee of the

American Institute of Real Estate Appraisers

Karel J. Clettenberg

James A. Graaskamp

2nd Draft

CHAPTER ONE

SOME BASIC STUFF ABOUT STATISTICS

I. What Are A Statistic?

   A. Introduction

   B. A Definition of Statistics

   C. The Practical Value of Statistics

II. Who Are the Variates in a Statistical Population?

   A. Elementary Units

   B. Measured or Described Observations

   C. Quantitative Observations or Variates

   D. Qualitative Observations or Attributes

   E. Population Samples

III. Frequency Distributions Are Not Electrical Phenomena

   A. Frequency Distribution Defined

   B. Frequency Distribution Applied to Attributes

   C. Frequency Distribution Applied to Variates

CHAPTER ONE

SOME BASIC STUFF ABOUT STATISTICS

I. What Are A Statistic?

A. Introduction

Many real estate professionals remember statistics as a course describing little known facts about red and white balls and see no reason to bust theirs on the subject. After all, nobody can know everything about anything so there is little certainty in this world, let alone mathematical accuracy. However, that simple fact of life is precisely why real estate practitioners need statistics as they must constantly make decisions on the basis of incomplete or sample information. Certainty is only possible when one has information on all the objects or alternatives in a specific situation. Real estate decisions are generally made under conditions of uncertainty but these conditions can often be measured or defined as to probability or reliability. Uncertainty is a manageable condition and a far cry from "doubt" or the prevailing real estate theory that ignorance is bliss.

B. A Definition of Statistics

Lo and behold, statistics is a field of knowledge that permits the researcher to derive and evaluate conclusions about a population (the big picture) from sample data (some snapshots). Specifically, the subject of statistics includes the theorems, methods, tools, and techniques which are used to discipline processes for:

1. gathering, selecting, and classifying data;
2. interpretation and analysis of data;
3. deriving and defining the reliability of conclusions based on sample data.

A parameter is a single number that describes a total population. A statistic is a single number that describes a sample. An array of statistics sometimes provides a more informative description. 422,636 is a single number. 42-24-36 is a statistic because it communicates a description.

C. The Practical Value of Statistics

Now, every appraisal is a research assignment and if such assignments are to be done in a professional manner there ought to be some discipline in the process of slinging facts around to derive a defendable solution. Real estate is too complex a subject to say much about the dynamics of value with absolute certainty. The appraiser may say that market value on a home is $25,000 when he means the $25,000 is the central tendency around which possible alternative transaction prices would group. However, how uncertain should he be in his prediction? How reliable is his answer? Statistics can provide in many situations evaluation of possible error and reliability as a by-product of careful data gathering and interpretation.

A great eastern fakir has written, "He who knows not and knows not that he knows not gives the opinions of foolish arrogance while he who knows not but knows the degree of error in his ways shall bask in the sunshine of professional respect and enjoy the inner peace that passeth all bounds of significance."

II. Who Are the Variates in a Statistical Population?

A. Elementary Units

A set of data becomes a statistical population only if the analyst can be absolutely certain that he can make a decision or statement without chance that another observation may create data which is the exception. The risk of omission and error is completely eliminated if all the possibilities are known to the analyst. However, most of the time the analyst must work with only a part of the population, and conclusions based on any part of a population with unknown exceptions or limits could be very misleading.

A set of data consists of either measured or described characteristics of certain objects within a statistical population. These objects are called the elementary unit of the population. For example, consider:

1. To determine the average assessed value of a single family home in a community of 5,000 single family houses one begins with a population of 5,000 numbers ($6,280, $11,500, $9,850) with each number measuring the total assessed value of a particular house. The elementary unit is a single family house and the characteristic to be measured is its total assessed value.
2. When a characteristic cannot be measured, statistics can only recognize that a certain elementary unit has or does not have that attribute. A single family lot may or may not have shoreline on a lake or stream - and it is then necessary to describe a population of newly subdivided lots in a given county relative to this

attribute by the percent of such lots having the attribute of shoreline.

B. Measured or Described Observations

If a measured or described characteristic has been specified on each and every elementary unit (such as an apartment or household), the set of data represents a statistical population.

Note the words "measured or described." Measurable observations are quantitative factors such as number of rooms or number of lots in a subdivision. On the other hand if the term "lots" as number of lots is qualified as those with trees or the number of lots which have been sold, these conditions can only be described rather than measured and are qualitative factors. Three are no differing degrees of "soldness"; it is or it ain't. A plump woman can be measured as so many pounds "plump" or so many inches too hippy - an issue of degree - but a woman cannot be just a little pregnant. That is a non-measurable "yes-no" choice and thus a qualitative observation.

C. Quantitative Observations or Variates

The values measured by quantitative observations are each called a variate. If the price of a house is \$22,450, the number \$22,450 is termed a variate. Quantitative observations can be further subdivided as either discreet or continuous. A discreet variate can assume only a limited number of values when using some specific measuring unit. For example, the number of persons in the family can be two or six but cannot be 2.5 or 6.3 when you are doing a head count. On the other hand some quantitative observations can be described as a continuous variate. The floor area in the sample of rooms or buildings could be, at least theoretically, 200 sq.ft., 200.001 sq. ft. or on up to thousands of sq. ft. Time can be defined days, hours, minutes, and micro-seconds, virtually an infinite number of values depending on the accuracy of available measuring equipment. Such is the nature of continuous quantitative observations.

D. <u>Qualitiative Observations or Attributes</u>

When discussing qualitative observations, it was shown that an adjective was necessary to qualify the elementary unit as to status, condition, or characteristic. These observations which describe are called <u>attributes</u>, in contrast to quantitative observations defined as <u>variates</u>. The lady's variates may be 38-23-34 but her virtues are attributes.

There is an important basic reason for distinguishing between a data set of variates (measured observations) and a data set of attributes (described observations). For a measurable variate it is possible to compute an arithmetic means or simple average, such as the average age of appraisers or the mean sale price per room for certain types of apartment buildings. However, when combining attributes, it is necessary to calculate ratios or percentages; the percentage of women who are pregnant, not of an individual woman who ~~which~~ is pregnant or of apartments which are vacant and not the vacancy within a single apartment.

D. Population Samples

Whether calculating averages or ratios, a data set usually represents only a portion of the statistical population and a complete set of observations is either impossible or too expensive to obtain. Thus the conclusions to be reached from statistical analysis of a data set are generally based on sample information and the process of collecting a sample of a total population and manipulating the sample data set is a pervasive issue in statistics. When there are 30 or less observations in a data set, it is termed a small sample. Small samples are processed with special adjustments in many statistical problems and small samples are very common in real estate statistics.

III. Frequency Distributions Are Not Electrical Phenomena

A. Frequency Distribution Defined

Since data may collect into a large number of observations almost before you know it, much needs to be done to simplify the communication problem. For certain purposes the data may be provided in a summarized form which reveals the general magnitude of certain relationships by subclassification of data. A frequency distribution is a summary of a large number of observations, generally in some tabular form.

B.. Frequency Distribution Applied to Attributes

It is best to discuss the frequency distribution of attributes first because it is easier than any other kind.

One method of presenting a summary is to list by subclassification with the frequency of occurrence of that particular attribute. Consider the thirty students in a class to learn real estate statistics. Each reported his major field of interest in real estate as follows:

Brokerage, counseling, brokerage, property management, property management, mortgage banking, counseling, brokerage, mortgage banking, brokerage, brokerage, brokerage, counseling, property management, brokerage, mortgage banking, property management, property management, brokerage, mortgage banking, mortgage banking, brokerage, property management, brokerage, brokerage, brokerage, mortgage banking, mortgage banking, mortgage banking, mortgage banking

The above set of attributes can be summarized and more efficiently displayed in a frequency distribution table like the one below:

Table 1-1

Major Activity	Number of Students
Brokerage	12
Counseling	3
Property Management	6
Mortgage Banking	9
Total	30

The tabular summary is easier to comprehend than the list. Tabular summaries may be improved as pictures, such as bar charts or pictograms, but those are the subject of another story reserved for Chapter 2. For now there is no avoiding the subject of the variates among us.

C. <u>Frequency Distribution Applied to Variates</u>

Summarizing a data set of variates for a frequency distribution requires the creation of artificial sub-classes which are called <u>class intervals</u>. A class interval is not a coffee break. Suppose you had collected a sample set of data on monthly rents for two-bedroom unfurnished apartments in a certain neighborhood. Each apartment might have a different rent per month, such as \$105, \$123, \$132 and so on which would produce a list of 100 possibly different prices, an unreadable chunk of data.

If these 100 items of data in the set were grouped into seven class intervals (as in Table 1-3) the data begins to take form for further analysis. Each of the subclassifications has a lower and a upper limit for purposes of sorting the data. Class limits should not overlap or it will be ambiguous for the analyst as to which class is intended to receive rental items which fall on the borderline. Notice that in Table 1-3 the upper limit of the first class is 109 while the lower limit is 110 for the next class. The exact division or mid-point between these two successive class limits is 109.5 and this mid-point is called the <u>upper boundary</u> for the first class interval and the <u>lower boundary</u> for the second class interval. All variates (rents in Table 1-3) in a given class must be larger than the lower boundary and smaller than the upper one. To anticipate each possible borderline case, class boundaries must be carried to one more decimal place than that of the numbers to be grouped. For example, if rents were stated as a $132.50 the upper boundary would need to be defined as 109.505.

The mid-point between the limits of the boundaries of a given class is called a class mark. For example, in Table 1-3 are sorted in the selected class intervals, the number in each class is called the frequency of that class and the array of the frequencies represents a distribution of the sample data.

Table 1-3

Distributions of Rents
For Two-Bedroom Apartments Unfurnished on Ticky-Tacky Lane

Class Interval	Lower Boundary	Upper Boundary	Class Mark	Frequency
100-109	99.5	109,5	104.5	10
110-119	109.5	119.5	114.5	13
120-129	119.5	129.5	124.5	15
130-139	129.5	139.5	134.5	24
140-149	139.5	149.5	144.5	16
150-159	149.5	159.5	154.5	13
160-169	159.5	169.5	164.5	9
				100

Some judgment and experience with data will aid in defining the number of classes to be used when constructing a frequency distribution of variates so that the result does not conceal significant patterns or gaps in the data nor do the classes become so numerous that the available data is too thinly distributed to reveal any differences in the magnitude of frequency among classes. Thus, the number of classes should generally not exceed 20 or be less than 5 and one method of estimating a convenient number of classes for a given set of data is called Sturges' Rule, after a mathematician named Rule. Old Sturges claimed that one is off to a fast start on a frequency distribution table where there are N observations in the data set if:

$$\text{number of classes} = 1 + 3.3 \log N$$

Thus it was no accident that there are 7 classes in Table 1-3 as Sturges' Rule suggested:

$$\begin{aligned}\text{number of classes} &= 1 + 3.3 \log 100 \\ &= 1 + (3.3)2 \\ &= 7.6 \text{ classes}\end{aligned}$$

The rule approximates 7 or 8 classes but to assure having an adequate number of items in those classes it is generally desirable to use the smaller number, in this case, 7 classes.

Once one knows the number of classes, its all down hill in determining the desirable class width. The approximate class width is indicated by the simple expedient of dividing the difference between the highest variate and the lowest variate in the data set by the desired number of classes. For example, in Table 1-3 the lowest rent might have been \$100 per month while the highest was \$168. This range of 68 divided by 7 would suggest a class width of 9.6 or for simplicity and clarity, a class width of \$10. To establish the lower boundary as one decimal beyond the nature of the variate of monthly rent, a lower boundary was established at 99.5 since the lowest item of rent $100 must fall within the boundary. The upper class boundary is then determined by adding the class width of $10 to the lower boundary of 99.5 to determine 109.5. That operation is repeated for the 7 classes until a distribution table begins to take shape as in Table 1-3. In reporting this to the client it is only necessary to provide the frequency distribution table as shown in Table 1-4.

Table 1-4

Frequency Distributions of Rents
For Two-Bedroom Units on Ticky-Tacky Lane
as of May 1, 1970

Class Interval	Frequency
100-109	10
110-119	13
120-129	15
130-139	24
140-149	16
150-159	13
160-169	9
	100

Always take care to label the contents of your distribution table with care as in Table 1-4. The title specifies the type of apartment, the fact that all are unfurnished and therefore comparable, the location, and the date for which the data is true.

Communication requires care and precision in such descriptions or chart titles so that the reader will know exactly what data has been summarized and conveyed by means of the table.

There are some other things you need to know about statistics but a thorough understanding of Chapter 1 will place you in the top 2% in the population - the world population. If you are discouraged, try Chapter 2. It is mostly pictures.

CHAPTER 2

A HISTOGRAM IS NOT A MAJOR OPERATION

I. Happiness is Simple Data Presentation

   A. A Problem in Communication

   B. Art to Conceal Art

   C. Forget the Pictogram

II. A Good Table Setting Aids Digestion

   A. Design of Table Presentations

   B. Comparisons and Emphasis

   C. Stubs and Captions

   D. Details of Table Construction

III. Bar Charts Mean Nothing to a Sailor

   A. A Simple Bar Chart

   B. The Histogram

   C. Component Part Charts

IV. Straight Line Graphs for the Novice

   A. Simple Line Chart

   B. Silhouette Charts

   C. A Z-Chart

V. Statistical Maps

   A. Data With Spatial Coordinates

   B. Hatched Maps

   C. Dot Maps

VI. Coloring Book Exercises

CHAPTER 2

I. Happiness is Simple Data Presentation

A. A Problem in Communication

One thing worse than actually doing statistical analyses is the responsibility of reading a statistical study with the understanding and a favorable impression of the author. A real estate report should communicate an understanding of much detail and many conclusions with a single glance. Three methods of data presentation are available:

1. Inclusion in a paragraph of text material
2. Construction of a table
3. Summary in graphic form

Inclusion of data in text material has the advantage of permitting the author to single out certain comparisons of important or express attitudes about the facts. However, such statistical presentations like the sample paragraph below quickly dissipate the attention and therefore understanding of the reader.

When only a few statistical items are involved, it is possible to increase the effectiveness of communication by formatting the typewritten report to resemble a brief table, such as:

> The number of passenger seat miles flown by shceduled domestic airlines was
>
> xx,xxx,xxx in 1969
> xx,xxx,xxx in 1970
> xx,xxx,xxx in 1971

However, clients do not like to read appraisal reports, particularly those parts of the report which are providing background information or demonstrating how well the analyst has done his home work before giving an opinion. A report should communicate quickly and with variety to stimulate the reader. Tabular presentations and graphic summaries of data can serve that purpose and this chapter is designed to provide some simple and basic methods of improving your reports with tables and graphic presentation. Of course, the real estate professional should take care that his good art work doesn't sell a poor investment.

B. Art to Conceal Art

One of the Chinas has persisted in advocating one picture is worth a thousand words, with no deviation permitted. Data distilled into a table or some graphic chart form communicates like a picture. There are two types of data which may be "pictures". First, there are general or reference items which are used as repeated references sources - as data banks for the statments in the report. This type of data can be very extensive or precisely detailed and may best be presented in tabular form. The primary function of a table is to present the data so that individual items may be found readily and labeled correctly for the reader. On the other hand if the purpose of the data is to communicate one relationship or a few closely related items, a simple graphic summary may be best. Since the author expects the reader to divert his attention from a running commentary to such a graph, it is important that the graphic statement be simple and not subject to visual musunderstanding or distortion. It is intended to communciate one idea at a time.

Every table and every graphic presentation should be noted in the prose of the report text and if it is too minor or too irrelevant for the text it should be omitted from the report altogether. To digest a big glob of information into a useful table or concise graphic statement conceals a great deal of analysis. If titles on the table or graphic statements are labored and confusing, the result is to advertise the confusion and lack of understanding of the analyst. It takes skill to conceal the analytical effort just as the most complex football play looks effortless when the pass is completed on TV.

C. Forget the Pictogram

Statistical comparisons often appear in newspapers, government reports, and department store windows in the form of pictograms, as typified by Chart 2-1. The essential piece of information is the average number of students enrolled as undergraduates for each of the periods indicated (height "H" of figurine) but the relationship is exaggerated by the area or volume of the art work used to dipict the level of enrollment which is a one dimensional concept (H-1, H-2, H-3). Cubes and circles may also be used to exaggerate one dimensional relationships. Cubes and circles are for squares. This chapter is not concerned with this type of graphic summary. Forget the pictogram. Do not associate with newspaper or government statisticians.

II. A Godd Table Setting Aids Digestion

A. Design of Table Presentation

To construct a good table the author must first consider the purpose, comparisons, if any, and data limitations which he must serve. The purpose may be as a general reference data bank where the only objective is to permit a quick search by the reader and there is no reason to arrang e columns in rows to permit comparison or explain the derivation of one row from several which preceded it. However, if the reader is expected to make comparisons or to review computations, the items must be arranged to assist the eye in making such a comparison or following such a calculation. It is useful for the beginner to write down the purpose of the table and the point in the text where it will be used to nail down an observation or conclusion. Then he should write down the specific comparison, data set, or relevant section of the general data available which makes the point. If its purpose doesn't read with punch or the detail seems off the mark, the chances are the table is not needed. The first thing to learn about tables is not to use them unless words fail, or are excessive and clumsy, or unless the numbers are relevant rather than simply interesting.

PICTOGRAM

PROJECTION OF UNDERGRADUATE ENROLLMENT
In Full Time Student Equivalents
For Biennial Fiscal Periods
Beginning June 1 of year indicated

	Average Annual
1970-1972	20,000
1972-1974	26,000
1974-1976	32,000

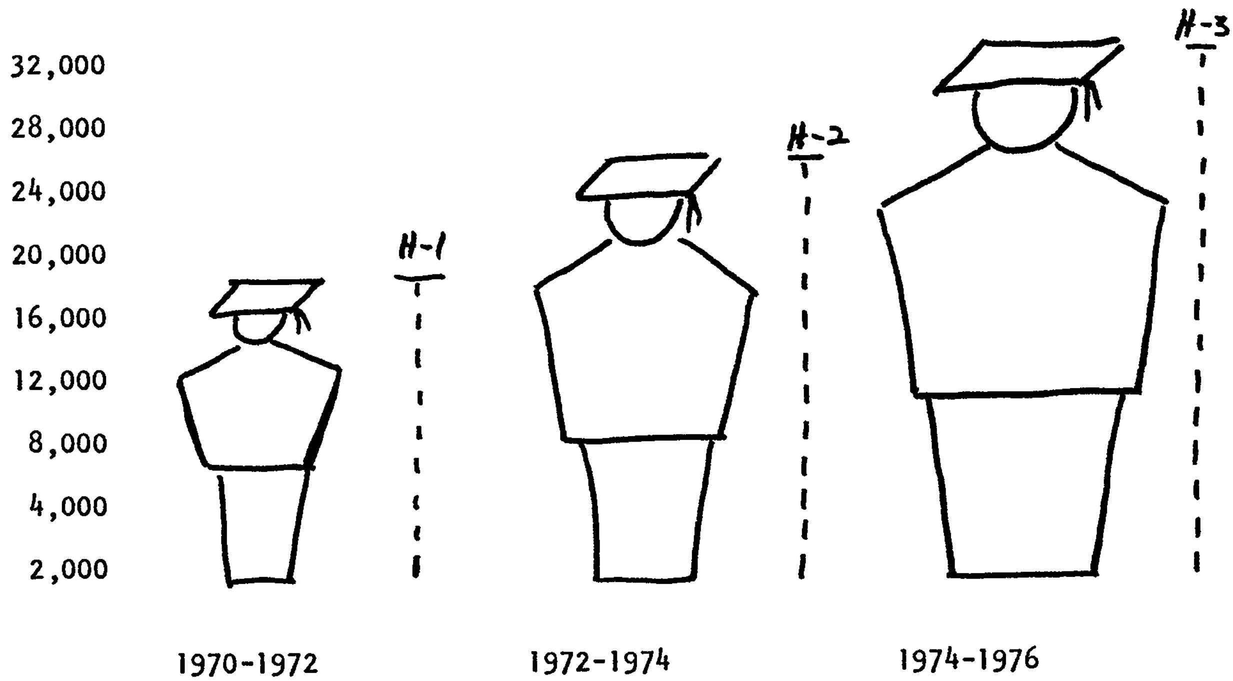


Chart 2-1

Proper credit cannot be given to the university president who presented this to the Legislature to support the new budget.

B. Comparisons and Emphasis

While arrangement of data into columns and rows may facilitate reading of the data, the opportunity for the reader to compare specific items is greatly helped by placing the items to be compared in contiguous columns or rows. Table 2-2 and 2-3 each focus attenti n on a different element of comparison, the first comparing the change in male population from 1960 to 1970 while the second permits the comparison of the relative number of males and females in each census.

The essential relationship in Chart 2-4 is the average price of C-2 land for each year reported but as soon as the reader sees the word average he may wish to know how many transactions were considered in computing that average in order that he might sense the reliability of the numbers reported. Thus logic suggests a table providing time, then price, then an indication of reliability and market activity implicit in the price. Think of the time it would take the analyst to collect the number of sales and transaction information necessary to compute the simple three column Chart 2-4 and the effort that is concealed by a well constructed chart!

If the absolute numbers to be compared are large it is useful for the author to compute for the reader ratios, percentages, averages, or rates of change as in Chart 2-5.

Since most of your reports are not done for the two Chinas it is safe to say that your clients will read from left to right and from top to bottom, if they read at all. For exphasis then, the most prominent position in the table is at the top left, and then the top right, and eventually the least prominent position is at the bottom right of the table. Thus if a total is the most important number to be communicated it should be at the top of the chart separated by a line from the breakdo... its components and the components deemphasized further by indenting to some degree to the right of the left hand column. Consider Chart 2-6; if your point was to emphasize the lilywhite characteristics of xxxx city it would be the first on the list while if the point to be made was the political strength of minorities in xxxx city the chart would be arranged in reverse order.

Key figures, columns or rows of figures may be emphasized by such simple formatting techniques as underlining, extra spacing, or capitalized description stubs and captions. Parentheses () may be used on type written reports much like italics in printed reports to distinguish census returns from other forecasts, to show deficits, to flag items to be subtracted in summing a total or to be omitted from a total. Note that none of these devices for arrangement of the data change the numbers - they only affect the speed with which the significance of the numbers is communicated to the reader and perhaps the willingness with which the conclusion is accepted. Communication with statistics as with women is as much a matter of tone as substance.

C. Stubs and Captions

Surprising as it may seem, there is something to be learned from the accountants in the definition of stub and caption. Stub items are the elements of description in the left hand column or "boxhead" of the chart which apply to each horizontal row of information reading from left to right. Captions refer to the titles on the chart sub-titles, or descriptions at the head of each column. The sequence of items in the stubs or captions may be selected to advance the purpose and emphasis of the chart and a number of systems can be suggested for the beginner:

1. An alphabetical arrangement is well suited for use in a general table to permit quick location of individual items in a series which are already classified geographically or qualitatively (See Chart 2-).
2. Geographical arrangements by state or region are common but the author must be careful to consolidate geographic areas in groups familiar to his readers or compatible with his information sources.
3. Historical arrangements are possible where dats is classified by era, year, or month. When years are listed either the most recent or the earliest may be the first upper left stub, depending which is to be emphasized. Months should be listed in most cases with January first.
4. Arrangement by size or order of magnitude of numbers was already suggested in Chart 2-3.
5. A progressive arrangement is used where each row or column develops logically from those given before as in accounting statement or processing of information as in Chart 2- . It is suitable for general tables or display of specific computations referred to in the text.

6. A numerical arrangement is possible where groups of data are given numerical tags, such as precincts, salesman territories, planning districts, or floor numbers. Obviously this method is quite common for both general reference tables and specific text references.

Unless it is self evident from the topic, the units of measurement for the numbers in a column or row must be indicated in the stub or box head caption as in Chart 2- . A dollar sign or percent sign is necessary only for the first item in each column but square footage or number of bedrooms or degrees of temperature as examples must be specified in the box head.

D. Details of Table Construction

Impressive statistics often are made by the title to the tables rather than by the numbers themselves and the art of good statistical presentation depends on careful titling and identification. Every table must have a number and a title, generally on the top of the page, which should precisely state in this order the what, where, how classified, and when of the data contained. Notice that Chart 2- has all four requirements while Chart 2-1 is less impressive for its failure to state which school or how the data has been classified. Anticipating the grouch in the back row of seminars and the crank letter writer, the authors of this text have carefully arranged both good and bad chart titles throughout this book to polish your critical powers.

In addition to the information contained in the title further explanation for the table may be provided by a prefatory note, a footnote, a supporting schedule, or a source note.

1. A prefatory note may follow a "main title" or subject heading and is placed just below this title with less prominent type. It should provide information about all or a substantial portion of the table as in Chart 2- .
2. Explanations of a specific column caption or row stub should follow the body of the table and be keyed by means of footnote numbers. Footnotes keyed to individual statistical items should be identified by a symbol (*, +, #, &, @, etc.) or by a letter.
3. A supporting schedule should be identified by a letter and used in place of a footnote where the bulk of the information provided is a numerical breakdown of data included in the caption, stub, or statistical item footnoted. For example, the deduction for vacancies might be keyed to a schedule by rental unit vacancy by season, collection losses, and releasing gaps in occupancy. On the other hand heating expense might be footnoted by a simple statement of fuel costs times estimated BTU requirements.
4. Since the data presented in the table often may have been taken from one or more published or unpublished sources, credit should be given in a source note giving author, title, volume, page, publisher, and date. It is best to identify both sources when data is taken from a secondary source as in the example "Source: State of Wisconsin Department of Urban Affairs and Development as quoted in

Statistical Abstract of the United States, 1970, page 532." Source notes generally appear below the footnotes but with some typographical prominence given to the word "source."

5. When percentages are used, the stub or the caption should state to what ratio figures the percentages relate. To avoid confusion the title or caption should state "percent of total" or "percent of increase or decrease" and the column caption might include column numbers and a shorthand formula of the ratio such as "Col. #1/Col. #5". Percentages should be written as whole numbers or as correct to tenths of 1% which may mean the total percentage could exceed 100% because of rounding remainders. Such a rounding error should be noted with a footnote and if the error exceeds .2 of one percent check your arithmetic rather than your footnote.

It is time to have a merry round on the general subject of rounding numbers. Arithmetic produces a large number of digits which may give a false impression of accuracy or which are irrelevant for the comparison or information desired. Everyone should be familiar with the practice of rounding 9.8 to the number 10 or the number 8.578 to 816 or rounding off the pennies to an even number of dollars. But often there is great reluctance to knock it off to the nearest 100,000 or stated to the nearest million. When making tables it is much easier on the reader and the typist to round to the nearest thousand or million to communicate relative magnitude quickly.

There is no need to use the population of the United States in 1970 of 201,896,781 when 202 million will make the point. The stub or caption in a table may note "000,000 omitted" or some mention of the rounding in common for all of the tables can be made in a prefatory note or sub-title. As a general rule six or fewer significant numbers may be accurate enough; however, all the computations such as totals and percentages, should be computed from the original figures before rounding for presentation in the table.

Formatting a table permits a certain degree of creative discretion to avoid the long and narrow, the short and wide or split pages, or crowding by means of cryptic abreviations in the stubs or captions. Avoid large tables with more than 8 or 10 columns and recast such tables into several smaller ones, repeating stub and caption entries on each one, and where possible, placing footnotes at the bottom of the appropriate page. If you put everything you know on one table, you may succeed.

The ruling class of statisticians, that is, those who enclose all their tables with single and double ruled lines, are passing from the scene.(1) The best tables guide the eye by generous spacing or skipping a line every three, four, or five rows. Horizontal lines are generally not used in the body of the table except to set off totals, and where they are necessary to separate major and minor caption heads, the lines do not continue into the stub. Vertical lines should not extend above column captions to complicate title and prefatory note headings at the top of the chart. Zero's are

ABSORPTION RATES FOR NEW AND VACANT APARTMENTS
(Excluding 1-3 Unit Buildings)
In Dallas and Fort Worth, Texas
1961-1970

Year	Number of Apartments		Increase over Previous Year in Units		Percentage Increase Over Previous Year	
	Dallas	Fort Worth	Dallas	Fort Worth	Dallas	Fort Worth
1961	7,000	4,000				
1962	7,500	4,200	500	200	7.1%	5.0%
1963	7,800	4,300	300	100	4.0	2.4
1964	8,000	4,800	200	500	2.6	11.6
1965	9,000	5,000	1,000	200	12.5	4.7
1966	10,000	5,200	1,000	200	11.1	4.0
1967	12,000	5,700	2,000	500	20.0	9.6
1968	15,000	6,000	3,000	300	25.0	5.3
1969	16,000	7,000	1,000	1,000	6.7	16.7
1970	20,000	9,000	4,000	2,000	25.0	28.6

Chart 2-5

nothing to look at either. If there are no items to record or the value is zero the entry should be indicated by a short dash (--). Of course, when there is no entry because the facts are missing or unavailable, an official blush by means of a footnote should be used to acknowledge that gap. If you are still confused on how to construct a table, just leave it with your secretary. She probably knows what <u>she</u> is doing and may be more current on these matters than you.

III. <u>Bar Charts Mean Nothing to a Sailor</u>

A. <u>Simple Bar Chart</u>

For a statistician a bar chart is not a map of saloons between the court house and the real estate office. Just as tables reduce a mess of facts to an organized understandable set of relationships, charts further reduced the amount of detail to gain the advantage of providing a quick picture of a general situation. While exact values can be provided in a table the typical chart or graph presents only approximate values. Since each chart requires an original drawing, the decision to use a table or a chart is an editorial issue of which one will produce the best impact for the cost. Bar chartists have customary procedures or protocol. Vertical bars are generally used for quantitative data or chronological data, variates as you learned in Chapter 1. Data comparisons of attributes or geographical information generally use horizontal bar graphs. Comparisons may be assisted by a two-unit bar chart as illustrated by 2- . Bar charts should make a few comparisons on a few items and if there is similar data over many years or in many sub-classes a line graph or curve will better demonstrate changes over time for area.

The length of the horizontal bar for attributes or the height of the bar for variates is determined by the number of items in the class. Each bar or unit should be uniform in width and the spacing between each bar should be uniform and less than the width of the bar so that it does not distort or mislead the eye as to the relative lengths of the bars to be compared. For greater accuracy and comparison, guidelines are an aid to the eye as in Chart 2- , and generally these guidelines represent bench marks on the scale that is provided.

B. The Histogram

The most common and useful tool in statistics may be the frequency distribution and much will be said about the analytical powers of a frequency distribution in later chapters. A graphic diagram or bar chart of a frequency distribution is called a histogram. The big decisions in doing a frequency distribution of a set of variates involve decisions on how many classes will best represent the relationships in the data. You already understand that problem from Chapter 1. Consider the frequency Table 2-4.

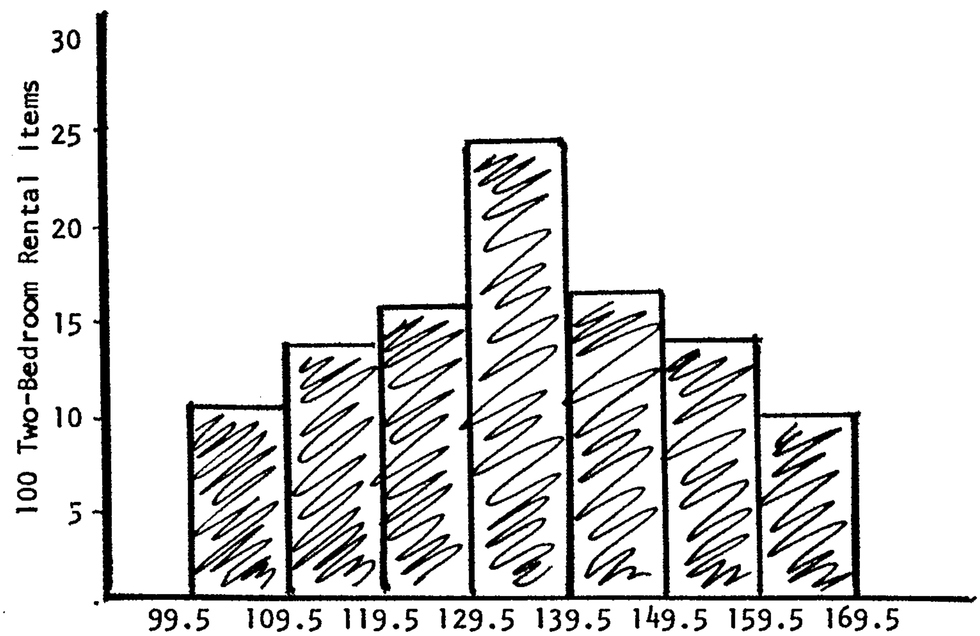
Table 2-4

Frequency Distributions of Monthly Rents

For A Sample Two-Bedroom Units on Ticky-Tacky Land
as of May 1, 1970

Class Interval	Frequency
$100-109	10
110-119	13
120-129	15
130-139	34
140-149	16
150-159	13
160-169	9
	100

The class boundaries could be drawn on the horizontal axis of the graph while the number of rental items in each class could be indicated by the vertical axis. The base of the rectangle corresponds with the class width and the height of the bar which can be drawn with these points represents the frequency of data within each class.



Distribution of Rents for Two-Bedroom Apartments
Unfurnished on Ticky-Tacky Lane

C. <u>Component Part Charts</u>

There have been many dissertations on the use of component part charts, using the device of a bar or a circular pie. This section is not one of them. A simple bar chart as in Chart 2- involves a one dimensional comparison indicated by the height of each section of the bar relative to the total height of the bar. Vertical bar diagrams are simple to draw and much easier to label than multiple bar charts or pie shaped diagrams. Comparison between years are made very easily by a series of component bars as in Chart 2- .

Pie diagrams may have more visual impact because of their analogy to a silver dollar, as in Chart 2- or to a pie such as the tax revenue pie divided among the cities politicians or municipal departments as in Table 2- . A pie diagram communicates the concept 50% (the diameter span of the circle) or 25% (a right angle to the diameter) more effectively than a component bar as in Chart 2- . The pie diagram can involve a two dimensional comparison of area plus arc or central angles so that when area is used it is an exclamation point to the relationship. Pie shaped graphs are generally more confusing than vertical bar graphs when presenting time series changes proportions. Component part depreciation, however, has nothing to do with critical comments on your half-baked pie charts.

IV. Straight Line Graphs for the Novice

A. Simple Line Chart

Data presentation by means of line charts will give you many fits, that is techniques for fitting a line to a set of data. Line charts or curves (a straight line a very s form of a curve!) are often useful for communicating changes over time or frequency distributions of continuous variates and are seldom used for attribute analysis. Much more will be said about graphic devices in the chapters on time series, normal curves, and regression analysis but a few line charts at this point shouldn't throw you any curves.

The easiest line chart to draw is one on which each point on the chart is connected to the next in sequence with a line as in Chart 2- . The critical question is where to put the point on the chart and how to label the scales. Whenever data over time is depicted on a

Average Square Foot Land Costs
For Commercial C-2 Zone Parcels
In Richmond, Vermont 1961-1970

Year (Col. 1)	Average Sq. Ft. Land Costs (Col. 2)	No. of Transactions (Col. 3)
1961	$3.00	7
1962	3.05	9
1963	3.07	10
1964	3.07	8
1965	3.10	11
1966	3.15	14
1967	3.15	13
1968	3.17	13
1969	3.20	15
1970	3.22	16

Chart 2-4

chart, it is customary to use the horizontal axis to indicate the time units for comparison while the vertical axis is used to scale the data series. You must distinguish between period data and point data. Period data refer to aggregates for a period of time, such as deed recordings per month, dollars per quarter, or average price for a year. Point data refer to a specific item at a particular point in time.

Labeling a chart a period data requires that the sequence of dates be placed below the vertical lines. When point data are graphed in a line, spaces rather than vertical lines should be labeled and the data plotted within the space to correspond with the correct point in time, particularly for data on an annual basis. Further questions on technique on time series are deferred to Chapter but consider a chart of period data of annual operating expenses which are increasing at a constant rate as depicted in Chart 2- . The significance of the change may be better communicated by direct comparison with the trend line of effective gross income.

B. Silhouette Charts

Silhouette charts are the stock and trade of many reporting services who analyze trends, such as the Wenzlick Reports or the Dodge Reports. For comparison of operating expenses to effective gross rent trends it is useful to chart on a linear basis effective gross on the same scale and calendar as gross rent as in Chart 2- . The shaded difference reveals by its declining area the significance of expenses increasing at an increasing rate while effective gross rents increase at a decreasing rate. Silhouette charts may be used to show fluctuations above and below a base line or to emphasize the residuals between a greater and a lesser amount which has been charted. Silhouette charts may be used to magnify ranges of price change as in stock market reports by conversion of the data to logarithm or ratios. Charts of this type gain additional impact if "positive" areas are colored blue or black while negative or financial losses are presented in red. Real estate men like publishers may not wish to incur the extra cost.

C. A Z-Chart

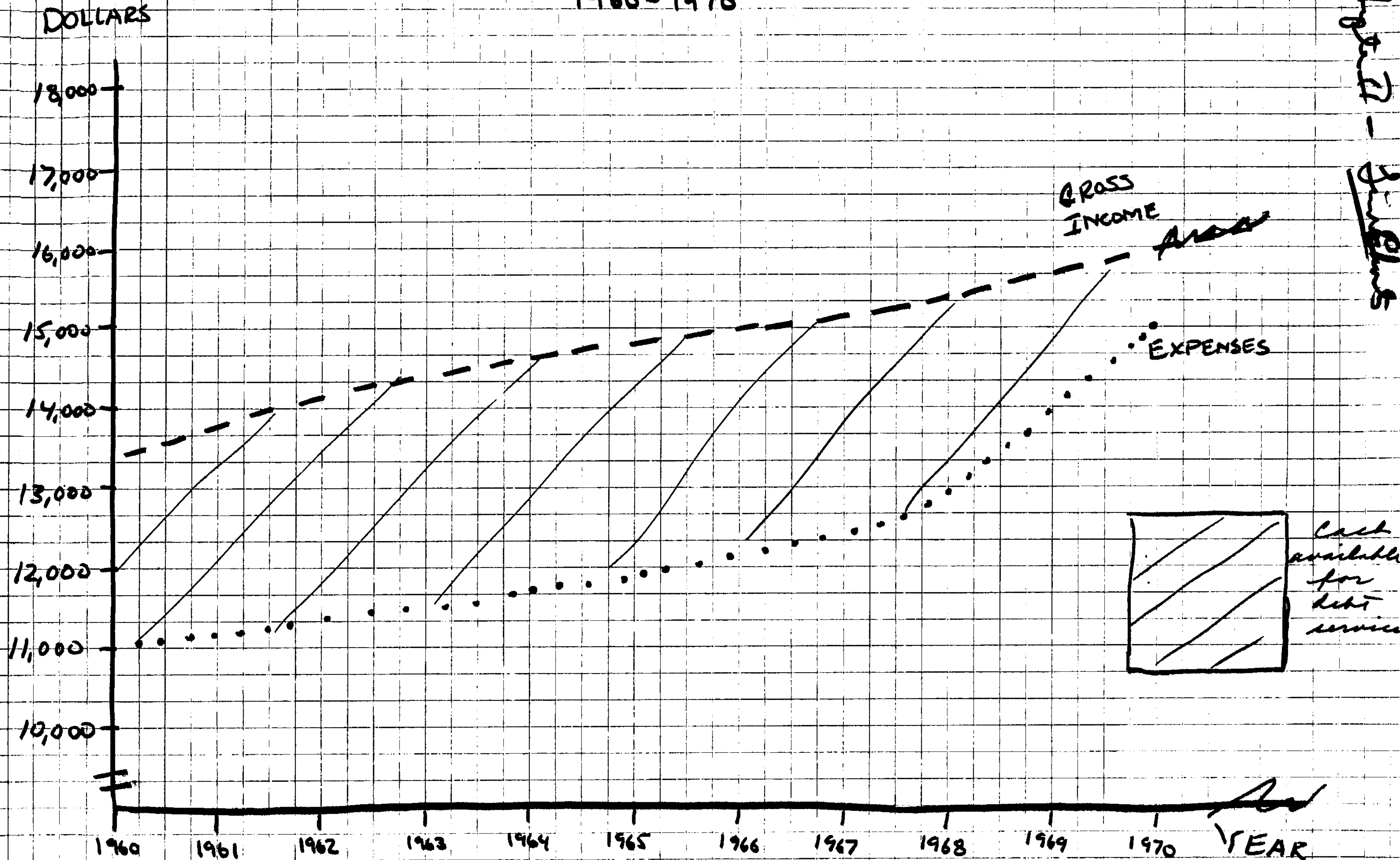
Some imaginative statistician working with his crayola set has termed a three line chart with two vertical scales a Z-chart, because it looks like a Z. Like everything in God's world, it has its uses, particularly in business where the relative of progress of production or sales is of interest. One curve might show monthly sales figures for a recreational lot subdivision, another curve the cumulative figures from the beginning of the year, while the third line may show the moving total for the twelve month period ending with the corresponding one of current sales. Two vertical scales are used so that the monthly data and the cumulative data can be held in close proximity to/avoid a chart three feet high. Since the twelve month cumulative line must intersect with the moving annual total, a projection of the cumulative line at any one time to the right hand vertical scale will indicate if the moving annual rate will be maintained, continue to increase, or fall to the point of intersection. Such a graph is more informative than simple annualized estimates of monthly sales in a seasonal business.

V. Statistical Maps

A. Data With Spatial Coordinates

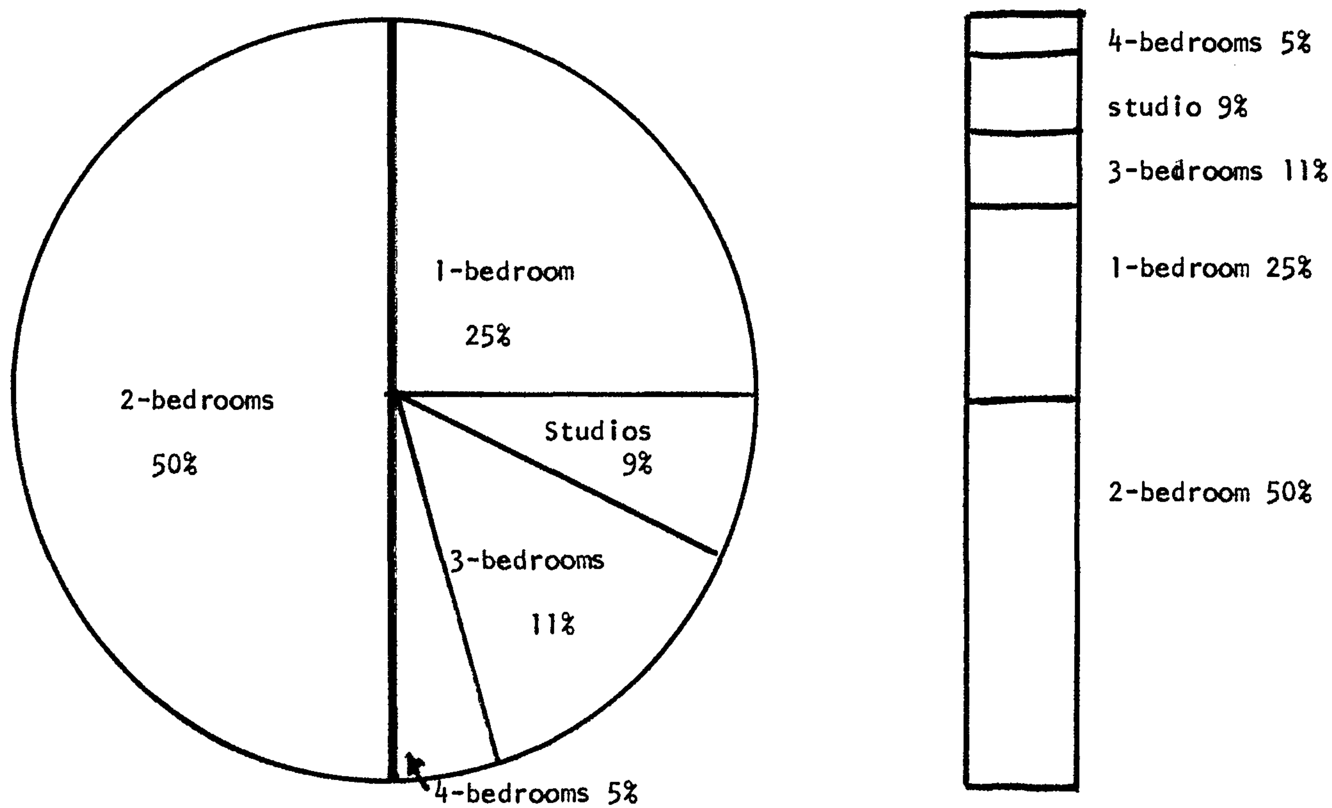
Since the location factor is so critical to successful real estate decisions, it follows that data which can be located in reference to a specific point on the map or in relation to a specific geographic location is of more use than most data which is not assembled by geographic classification. During the next decade the real estate man will be smothered with statistical mapping services based on the 1970 U. S. census,(2) by state agencies,(3) and by many municipalities.(4) A statistical map differs from the typical map as it overlays patterns rather than specific items on an area map with certain geographic reference points, such as political boundaries, major traffic arteries, or other visual reference points familiar to the reading audience. The information is given approximate spatial orientation and aggregated, with each dot representing 1,000 inhabitants, 100 families, or each color representing an average within a class interval. A map is statistical rather than pictorial when it is communicating patterns of data classes rather than locations of each observation as its primary function. Such statistical maps are basic to the analysis of economic or sociological trends in the community or in a trade area for a feasibility study. The techniques are well described in such publications as Guide to Store Location Analysis by Wm Applebaum. The point here is that such mapping is a form of statistical presentation rather than cartography. After all, this chapter has wandered all over the map of graphic presentation, as it is, without wandering into another world of cartography.

GROSS EFFECTIVE INCOME AND OPERATING EXPENSES by QUARTERS
PURE BLUFF ESTATES, BOULDER COLORADO
1960 - 1970



Pie Chart

Glendale, California
Apartment Supply
1971



CHAPTER 3

POPULATION CONTROL WITH PARAMETERS

I. Parameters Ain't Statistics

   A. Definition

   B. Types of Parameters

II. Aggregates in the Wrong Proportion Can Be Abrasive

   A. Aggregates

   B. Proportions

III. What's Average About Average?

   A. The Typical Variate

   B. The Arithmetic Mean

   C. The Median

   D. The Mode

   E. Choosing the Appropriate Average

   F. The Weighted Average

   G. The Group Average

IV. Your Baptism With Full Dispersion

   A. Measures of Dispersion

   B. The Range

   C. The Average Deviation

   D. The Standard Deviation

CHAPTER 3

POPULATION CONTROL WITH PARAMETERS

I. Parameters Ain't Statistics

A. Definition

To further postpone your progress into statistics, it should be noted there are times when the real estate analyst will have data concerning an entire population. Census data as well as tax assessment records for a community are illustrations of the possibility, although eight sales of apartment buildings might represent the total population of apartment sales for a given time and community. A complete data set does not have to be a large data set. A statistic is a single number that describes a sample which may permit some inference about a larger population. A parameter is a single number that describes the population. You should have retained that distinction from Chapter I because this chapter is concerned with methods appropriate to the organization and communication of analysis concerning an entire population rather than an incomplete sub-set or sample. For example, assume you have access to a data set that includes all assessment information for all single family homes in Chicago, Illinois. The average total assessment of 50 single family homes picked at random in Chicago is a statistic. The average total assessment for all the single family homes in Chicago is a parameter. Think parameters.

B. Types of Parameters

For any population one can develop a number of descriptive or analytical parameters. The most common garden variety are the aggregate and the proportion. Other parameters that describe the typical variate are the arithmetic mean, the median, and the mode, which bloom prolifically in appraisal reports. Another lovely bunch of parameters includes the range, the average deviation, and the standard deviation which describe the dispersion of values for units within the population. This chapter is devoted to planting the seeds of these population parameters among the rocks of real estate.

II. Aggregates in the Wrong Proportion Can Be Abrasive

A. Aggregates

You have been using aggregates and proportions for years, although you may not have recognized them in these terms.

The aggregate is the sum of all the variates in a population. For example, assume a population of five shopping centers in a small city with the following sales prices:

$$x_1 = 1{,}200{,}000$$
$$x_2 = 800{,}000$$
$$x_3 = 2{,}000{,}000$$
$$x_4 = 600{,}000$$
$$x_5 = 1{,}000{,}000$$

The aggregate sales price is the sum of the five sales prices, determined as follows:

$$\Sigma x = x_1 + x_2 + x_3 + x_4 + x_5$$

The Greek letter Σ (sigma) indicates the summation of x's, or:

$$\Sigma x = 1{,}300{,}000 + 800{,}000 + 2{,}000{,}000 + 600{,}000 + 1{,}000{,}000$$
$$\Sigma x = 5{,}600{,}000$$

B. Proportions

A proportion is a ratio or percentage with a clumsy title. The proportion is that fraction of a population that has a given characteristic and is denoted as pi (π). For example, in a population of 500 homes, of which 250 have more than one bathroom, the proportion of homes having more than one bathroom is .50, or

$$\pi = \frac{250}{500} = .50$$

The proportion is often converted to a percentage by multiplying by 100. In the above case, 50% of the homes have more than one bathroom, or:

$$.50 \times 100 = 50\%$$

If you were fussy about the use of words, you should note that the term aggregate refers to the sum of values of all variates in a population which means measurable observation whose values can be quantified. On the other hand, a proportion is the parameter most often used in describing a population of attributes as these are non-measurable observations which cannot be summed.

III. What's Average About Average?

A. The Typical Variate

The typical real estate man is generally very sloppy in his use of the word "typical" when he wants to describe a population with a single number such as the typical sales price in a subdivision or the typical operating expense ratio for an apartment building. If he is to speak with professional discipline, the real estate analyst should define his population and the typical variate in that population with care. A common way of describing the population is by means of the average expense ratio or average variate. However, there are three common parameters which might describe the typical or average variate in a population-----the arithmetic mean, the median, and the mode.

B. <u>The Arithmetic Mean</u>

The arithmetic mean is your old friend, the simple average duded up with a fancy name and a Greek letter symbol (u), pronounced "mew" as in music. The formula for the arithmetic mean is the sum of the variates divided by the number of variates, or:

$$\mu = \frac{Ex}{N}$$

where:

μ = the arithmetic mean

Ex = the sum of the variates

N = the number of variates in the population

The appraiser has many applications for the arithmetic mean, a few of which are illustrated in the following examples.

<u>Example 1</u>

Gross Rent Multiples for Five Comparable Apartments

N = 5

	(x)
	5.0
	5.2
	5.2
	6.0
	6.3
Ex =	27.7

$$\mu = \frac{Ex}{N} = \frac{27.7}{5} = 5.54$$

<u>Example 2</u> Operating Expenses for Six Comparable Shopping Centers

N = 6

	(x)
	\$16,000
	16,500
	16,800
	16,900
	17,000
	17,400
Ex =	100,600

$$\mu = \frac{Ex}{N} = \frac{100,600}{6} = 16,767$$

Example 3

Square-Foot Lot Prices for Eight Comparable Lot Sales

N = 9	N = 7
(x)	
.33	.95
.95	.95
.95	.97
.97	.98
.98	1.00
1.00	1.02
1.02	1.03
1.03	
3.05	Σx = 6.9 μ = .9857
Σx = 10.28 μ = 1.14	

Common sense should alert the reader to the distortion of information about the above population which occurs when the average of $1.13 is far beyond the amount the more typical prices of .95 to 1.03. Therefore, when computing the aritmetic mean it is permissible to drop one very unusual value or the highest and lowest values in an array so that only the more typical values are averaged. Remember the purpose of an average is to communicate a typical dimension of the population, not just a fictional number which is correctly computed but portrays nothing about the population or dimension. Dropping the two extremes of 33¢ and $3.05 produces the more representative average of .986. The analyst can modify an average with more self confidence after he has been introduced to measures of dispersion and normal distributions.

The arithmetic mean for a long series of items of a nearly continuous variable may sometimes be computed quickly and accurately by taking the average of the smallest and largest items. For example, the weatherman may have a record of temperatures each half hour of every day, but for many purposes it is not necessary to average 48 readings to arrive at the daily mean temperature. The high and low temperature for the 24 hour period when averaged will be representative of the mean temperature for the day.

C. The Median

The median is the average for the fellow who can count but cannot add or subtract. There is a hitch. Before attempting to locate the median the appraiser must put his data in an array, that is, he must arrange the variates in increasing or decreasing order. If the number of variates in the population is odd, the median is the middle variate. If the number of variates in the population is even, the median is the arithmetic mean of the two middle variates. These procedures are illustrated below:

Example 1

Square-Foot Construction Costs for Five Single-Family Homes

N = 5

(x)
$17.00
19.75
19.85
20.00
26.00

Median = $19.85

Example 2

Heating Expenses for Six Comparable Apartment Buildings

$N = 6$

(x)

$4,000
4,200
4,200
4,300
4,400
4,800

$$\text{Median} = \frac{\$4,200 + \$4,300}{2} = \$4,250$$

Example 3

Gross Rents for Four Two-Bedroom Apartments

$N = 4$

(x)

$2,400
2,500
2,800
2,850

$$\text{Median} = \frac{\$2,500 + \$2,800}{2} = \$2,650$$

D. The Mode

The mode generation is like basic, man. The mode is simply the most frequently appearing variate in the population. The following examples are right on:

Example 1

Assessed Valuation on Eight Comparable Homes

$N = 8$

(x)

$9,000
9,200
9,300
9,300
9,300
9,400
9,400
9,500

Mode = $9,300

Example 2

Sales Prices of Six Comparable Lots

N = 6

(x)
\$8,000
8,200
8,300
8,400
8,400
8,400

Mode = \$8,400

E. Choosing the Appropriate Average

These three measures to describe the typical or average parameter of a population are easy enough to compute. The problem for the analyst is to choose which of these parameters will best communicate the nature of the population characteristic which has been measured or provide the basis for a decision in judging efficiency, or effectiveness, or farm productivity per acre or any of a great variety of quantitative comparisons. The user must decide the most appropriate definition of average and the examples which follow illustrate some guidelines in application of the appropriate measure to different patterns of data.

Example 1

Monthly Gross Rent/Sq. Ft. for Five Apartments

N = 5

(x)	(x)
.25	.21
.23	.23
.26	.25
.21	.25
.25	.26
1.20	1.20

Mean $= \frac{Ex}{N} = \frac{120}{5} = .24$

Median = .25

Mode = .25

In this situation the analyst would probably select the arithmetic mean as the appropriate measure since the mortgage lender is familiar with the technical approach of pennies per month per foot and the arithmetic mean and there is no advantage in simplicity of using the median or the mode when the three parameters have approximately the same value. However, the more natural 130 could be used for a general statement.

<u>Example 2</u>

Condemnation Awards for Duplex Structures in Urban Renewal Project

N = 5

(x)

$8,000
8,200
8,400
8,500
27,000

$$\text{Mean} = \frac{\Sigma x}{N} = \frac{\$60,100}{5} = \$12,020$$

Median = $8,400

Mode = None

In this case the appraiser would probably select the median as the most appropriate measure since the median represents the majority of the variates and the arithmetic mean represents none of the variates.

<u>Example 3</u>

Number of Dwelling Units Per Apartment Building

N = 8

(x)

7
8
8
8
9
9
10
10

$$\text{Mean} = \frac{\Sigma x}{N} = \frac{69}{8} = 8.62$$

$$\text{Median} = \frac{8 + 9}{2} = 8.5$$

Mode = 8

In this case the appraiser would select the mode as the most appropriate measure since the mean and the mode do not represent an actual or typical apartment building. Of course, the analyst might conclude that none of the three averages were truly typical of the impression that he wished to convey, an intuition that would signal that he had chosen the wrong population base for comparison. For example, if his population of sales included buildings with greatly different numbers of apartments in each building, perhaps he should choose an apartment unit as the elementary unit of his population rather than an apartment building. The raw data in terms of number of buildings may better be modified to communicate information with precision by means of weighted average or group average techniques.

IV. Your Baptism With Full Dispersion

A. Measures of Dispersion

Modified forms of averages have already suggested that to describe a population with a single number is to conceal the pattern of items within the population. Averages describe the typical variate while measures of dispersion describe the degree of differences, the disparity, or spread of values of the variates constituting the population. Parameters which describe the degree of differences among and between the variates or which indicate the degree uniformity between the variates are called measures of dispersion. Consider the appraiser who has diligently labored among the files of his friends to develop the following data bank of sales prices for lots in two subdivisions:

Subdivision A	Subdivision B
N = 5	N = 5
(x)	(x)
\$8,000	\$4,000
8,000	5,000
8,100	10,000
8,200	10,000
8,200	11,500

Mean $= \frac{\Sigma x}{N} = \frac{\$40,500}{5} = \$8,000$ Mean $= \frac{\Sigma x}{N} = \frac{\$40,500}{5} = \$8,100$

In this case, the mean sales price in the two subdivisions are identical, \$8,100. In Subdivision A the appraiser may feel very comfortable using the mean sales price for court testimony while in Subdivision B he may not want to use the mean sales price for further analysis. The problem arises because Subdivision B has a greater dispersion in sales price.

There are three parameters that can be used to measure the dispersion in a population - <u>the range</u>, <u>the average deviation</u>, and <u>the standard deviation</u>. These parameters provide statements about the accuracy of portraying a population with a single number, such as the average as well as communicating the internal pattern of items within the population.

B. <u>The Range</u>

To home in on the idea of range, consider it as the difference between the highest and the lowest variate in the population. That concept leads to the "formula" for range which is:

Range = Highest Variate - Lowest Variate

The following examples illustrate the range in several situations.

Example 1

Average Sales Prices of Six Firms

N = 6

(x)

$1,000
1,200
1,300
1,300
1,350
1,375

Range = Highest Variate - Lowest Variate = $1,375 - $1,000 = $375

Example 2

Utilities Expenses for Four Comparable Apartment Buildings

N = 4

(x)
$4,000
4,200
4,200
4,300

Range = Highest Variate - Lowest Variate

Range = $4,300 - $4,000 = $300

In using the range as a measure of dispersion or internal pattern, the analyst should be aware of its major deficiency. The range considers only the highest and the lowest variates in a population while ignoring the rest of the variates.

C. The Average Deviation

The average deviation is a deadend street as a parameter because it does not lead to any other analytical step but is a common measure of dispersion. The average deviation measures the average absolute difference between each variate and the mean of the population. The formula for the average deviation is:

$$AD = \frac{\Sigma|x-\mu|}{N}$$ [1]

Where N = the number of observations in the population and x represents each individual observation. The two vertical bars in the formula indicate that the absolute difference between x and μ is to be considered so that the plus and minus signs should be ignored.

Cash Sales Prices of Five Comparable Single Family
Ranch Homes in Open Meadow Plat
During First Quarter, 1970

N = 5

Sales (x)	μ $18,040	|x-u| |x - $18,040
$18,000		40
17,600		440
16,900		1140
19,000		960
18,700		660
$90,200/5 = μ		E|x-u| = 2240

$$AD = \frac{E|x-u|}{N} = \frac{2240}{5} = \$448$$

Thus, the price of each lot is different from the mean of the population by an average of $448. It is a measure of the average difference of the differences. No need to study average dispersion much for an exam. Concentrate on the standard deviation.

D. The Standard Deviation

The standard devialton is by far the most important parameter for the measure of population dispersion, for it is basic to further statistical analysis as well. It is denoted by the Greek symbol σ which is pronounced Sigma as in the sweetheart of Sigma Chi.

Try and keep your mind on Sigma, rather than a sweetheart. Chi, square, is something altogether different, and reappears in Chapter
The formula for the standard deviation is:

$$\sigma = \sqrt{\frac{E(x-\mu)^2}{N}}$$

For example, consider the investor examining a six month report on vacancy losses for five similar apartment buildings managed by five different brokers. How does the investor know when the vacancy loss is typical and when it indicates exceptionally good or poor performance? He would determine the standard deviation as follows:

Example 1

Vacancy Losses for Five Similar Apartment Buildings
In the Northwest Quarter of Herkimer, Pennsylvania
For the Six Months Ending June 30, 1971

N = 5

Vacancy Losses (x)	(x-μ) (x-\$1,600)	$(x-\mu)^2$
\$ 500	-1100	\$1,210,000
1,200	- 400	160,000
1,500	- 100	10,000
1,800	200	40,000
3,000	1400	1,960,000
8,000/5 = \$1,600 = μ	$E(x-\mu)^2$ =	3,380,000

$$\sigma = \sqrt{\frac{E(x-\mu)^2}{N}} = \sqrt{\frac{3,380,000}{5}} = \sqrt{676,000} = 822.2$$

On the average the vacancy loss for each building differs from the arithmetic mean of the population by \$822. It will be shown that if the average vacancy loss is \$1600 and the standard deviation is \$820, than in 67% of the cases, the normal vacancy loss would be \$1600 plus or minus \$822. While the \$500 vacancy loss is good, it is not far from one standard deviation from the mean but the \$3,000 vacancy is nearly two standard deviation from the mean and far less reasonable to expect. The three property managers for the properties in the center of the cluster are doing an average job within the range of typical expectation. This type of practical application with further refinements of technique in the following chapter would give real estate men the same critical testing power that industry has used for years to evaluate production quantity and quality characteristics.

Gross Income Multipliers for Eight Cabin Resorts
With Less Than Ten Rental Units Each
For the First Five Months of 1969
In Northern Michigan

N = 8

Gross Rent Multiples (x)	$(x-\mu)$	$(x-\mu)^2$
4.0	-.5	.25
4.0	-.5	.25
4.2	-.3	.09
4.6	.1	.01
4.6	.1	.01
4.7	.2	.04
4.8	.3	.09
5.1	.6	.36
		$E(x-\mu)^2 = 1.14$

$$\sigma = \frac{E(x-\mu)^2}{N} = \frac{1.14}{8} = .38$$

Thus, on the average, the gross rent multiple for each resort differs from the mean of the population by .38.or .4 standard deviation.

The appraiser could modify a mean average GRM of 4.5 to 4.9 to reflect a superior property and thus indicate the project was in the top 16% of all units. To use 5.3 or two standard deviations above the mean would indicate the appraiser thought the property in the top 5% of all similar properties. Why so? Read Chapter 4.

Perhaps many of the conclusions drawn from the population parameters in this chapter were obvious to the experienced real estate practioner. However, how does the real estate analyst communicate what he knows to be a reasonable vacancy loss, or gross rent multiplier, or other quantitative characteristic about a population of data? The profess-ional does not have to create a degree of certainty about his con-clusions or information on the basis of reputation. The facts can be made to speak for themselves in most quantitative decision procedures and real estate people must reveal the degree of confidence which can be placed on their quantitative communication. Precision in procedures has always provided the basis by which the true professions have judged the practice of their peers.

Preliminary Draft

(Chapter 4 was to be Exercises)

CHAPTER 5

a. SAMPLER Can EMBROIDER A REPORT

I. The ~~Thought~~ Thread of "[illegible]" Exposition

II. Spinning Some Definitions

- A. Sampling Units
- B. Population Frame

III. Elements of a "Good" Sample

- A. Design of a Survey
- B. Some Angles on Bias
- C. Good May Mean Cheap
- D. THe Non-Response Headache

IV. A Sample of Probability Sampling Techniques ~~Probability Samples Don't Happen Accidentally~~

- A. The Simple Random Sample
- B. The Stratified Sample
- C. The Cluster Sample
- D. The Systematic Sample

V. Non-Probability Samples are Distasteful ~~Can be Tasteful~~

- A. A Convenient Sample
- B. A Judgment Sample

VI. A Condensed Note on Notions About Notations

CHAPTER 5

A SAMPLER CAN EMBROIDER A REPORT

I. The Threadof "Crewel" Exposition

Most real real estate decisions are made without complete information about a particular tenant population, or supply of rental units, or certainty as to the range of expenses. It may be impossible or impractical to gather complete population data. It would not be practical to survey all the buyers of recreation~~l~~ lots to find out why they bought what they did where they did. Moreover it is not necessary to do so as the critical motivations and origins of customers may be equally revealed by a survey ~~of only one hundred~~ in which only 50 or 100 persons reply. The data processing techniques of statistics permit decisions on the basis of only a part or subset of a population and such a subset is called a sample.. This chapter will~~l~~ give you a little sample of sampling, that is, the process by which some of the elementary units in a population are selected for measurement to be used as a representation of the characteristics of the entire population.

To be sure of properly representing the population in any given case, one must select a sampling technique or sample design~~s~~ which is best for the problem at hand. Improbable as it sounds, sample designs ~~are classified into probability samples and non probability samples~~ are either called probability samples or they are not, in which case they are called non-probability samples. A sample is a probability sample if each unit in the population has some chance of selection for inclusion in the sample.

Most important, the researcher must know the probability or chance of selecting any particular elementary unit in the population, and it need not be an equal chance for each unit, despite what some amateur loudmouth might say at the next realtor meeting. The advantage of using a probability sample is that you not only can learn something about the population but also can measure the possibility that you have arrived at the wrong conclusion about the population, ~~can be measured~~ using the theory of probability. You may not be any more correct than you were on a guess-and-by-golly theory, but at least you ~~bookie can give you odds on~~ know how wrong you can be. In a non-probability sample some units, because of the sampling design, do not have any chance of being selected or the probability of selecting any unit cannot be determined in advance. Under the present laws of the land and statistics that situation is unfair, may be discriminatory, and may not give you much real insight into the nature of the population either.

Before you can make use of frequency distributions, standard deviations, normal curves and the other nifty tricks in chapters which follow you need a correct sample to be off on the right foot. This chapter will try to help.

II. Spinning Some Definitions

~~A. Sampling Units~~

Ponder for a moment on two terms, the sampling unit and the population frame. They form the basis for design of a sample. Assume that you wish to obtain a bank charter for a specific location. ~~and~~ In support of the charter application you wish to assert that the typical household is traveling too far for convenience to find bank services or that all households must use the same bank and therefore lack the opportunity of competitive choice. This problem will emerge as a channel marker from time to time in the flow which follows.

A. <u>Sampling Units</u>

You may be ahead of us on this one but sampling units are the elementary units the real estate analyst wishes to study by means of a sample, and in the bank location problem, the household is the basic sampling unit. Of course, a sampling unit can be an inanimate object, such as undeveloped platted lots, square feet of office space, or one-bedroom apartment units. The analyst must be careful to choose a unit which is easily ~~dileanated~~ delineated yet focused on the subject at hand which may require some common sense or ballpark assumptions. For example, before sampling a neighborhood of upper income home owners to see who might consider relocating to a luxury apartment, it was decided to anticipate which household might be ready to move. ~~By defining~~ the sampling unit was defined as home onwers who had lived in the same house for at least 15 years, giving the childred time to fly the nest and maintenance problems time to arrive for what had once been a new home. Definition of a high income neighborhood provided a "population frame".

B. <u>Population Frame</u>

The population frame is a devise for delineating the total ~~population from which~~ number of units from which the sample will be drawn, generally a source which can be reduced to a list of names, license numbers, file folders or ~~other source~~ other access opportunity. In another light the population frame is a device for excluding all of the similar elementary units which are not relevant ~~to~~ in your ~~decision~~ decision situation. Consider the following exqmples of population frames:

1. A primary trade area for a ~~super market is~~ neighborhood bank might be defined as those households within a perimeter representing a maximum of ~~5~~ 10 minutes driving time from any point on the perimeter to the subject site. (1) Using the perimeter line as a "frame", it is then possible to identify virtually all the households in that area from a source such as a street directory, a reverse telephone book, a list of electric meters, or some other source of name and address within the territorial frame. > over

2. In the coming years many more studies will be based on the zip code frame or real estate tax billing address list, a census tract can be a frame, or it can be constructed block by block from an aerial photograph.

3. The building directory in the lobby of each of two or three new office buildings can provide the population frame for a study of tenants to determine where they came from and their primary reasons for moving.

4. ~~For the bank location study the population frame would be a primary trade area defined by past experience as to banking habits and circulation patterns and residential households.~~

~~5.~~ 4. The population frame could be all the cars or all the pedestrians passing a specified point on a certain street during a 24-hour period or some portion of the day.

~~6.~~ 5. A population frame might be all the people who would visit a particular model apartment or model home during a specific period of time, such as a weekend, or a full week, or those appearing as charge customers of a particular store.

A population frame is selected to keep the research focused on relevant sampling units and to permit construction of a sample which will permit measurement and observation by means of telephone, mail, counting, or other required inter action. In addition, the frame should offer some economy and efficientcy appropriate to the budget available for research. While the real estate analyst is permitted considerable attitude and innovation in defining the population frame, it will be shown that he may inadvertently distort his results by the method chosen. For example, he might define his sample unit as single family home sales in a certain high income sector of the city and then choose a population frame such as the record of sales of homes for estate administration purposes by the trust department of his bank. The data would be ~~data was~~ cheap and complete and easily accessible but it might be very ~~was~~ biased by the fact the trust department generally sold for less money then might have been obtained for the property in order to speed the liquidation process of the estate. The design of a "good" sample therefore is concerned with the trade-off between the feasibility of executing the sampl~~e~~ing procedure and the reliability ~~accuracy~~ with which the sample represents the population.

III. Elements of a "Good" Sample

A. Design of a Survey

If you have designs on a survey, the most important aspects of rigorous sampling are the determination of the number of units ~~cases~~ in a satisfactory sample and the method by which the sample is drawn. Answer these two questions and define the unit and the population frame, and your twelve year old can do the heavy work on days he isn't pitching for the Little League. In choosing teh number of observations and the method of selecting a sample, it should be understood that there are a number of things which could casue the sample data to deviate from, to misrepresent the population. These potenitally damaging errors can be ~~termed random error~~ classified random errors or bias.

B. Some Angles on Bias

Any sample will contain some element of deviation due to random errors or bias. Random errors are casued by selection of extreme cases in the sample. For example, in studying average home size in square feet in California, the sampling method might permit the inclusion of a mansion once belonging to William Randolph Hearst which would cause the average area per home in the sample to exceed the mean sales price of all the homes in California. Random errors are mitigated by increasing the number in the sample to increase the land in possibility including an extreme of an opposite sort. For example, the superior oondition of the Hearst mansion might be off-set by the random selection of the governors mansion in California. Although the total distortion of land in error decreases as the sample size incfeases, errors due to bias do not increase significantly with an increase in sample size. A bias is due to a systematic distortion in your sample. Such a bias could result from too narrow a definition of the elementary unit or choice of a population frame which contained an inadvertent exclusion of units presumed to be in the sample. Sales by the trust department contained a bias to the low side while sales prices reported by a single salesman might be consistently exaggerated. Old duplexes might generaly be recorded as land contracts so that a sample based on only warranty deeds ~~might~~ would overlook a substantial market of activity. It is possible that only people with a specific axe to grind or an Atypical interest in the subject may respond to your questionnaire. Tleephone calls made at a certain time of day may only find the housewife or grandmother at home when it is the husband for whom the survey is intended. Almost any population frame, survey technique, or questionnaire format will introduce some bias but the objective is to minimize bias if it cannot be avoided and to eliminate the bias of careless sampling design and testing..(2)

C.. <u>Good May Mean Cheap</u>

A good sample design should provide the maximum amount of well qualified informatiion at the least cost. The dilemma should be clear; the larger the sample, in general the more valid are the conclusions which can be asserted due to a reduction in the random error. Nevertheless, large samples are often expensive and therefore the research design must balance the need for increased level of statistical reliability or error and the additional costs of drawing and processing information from a larger sample. The problem is not much different than that of the appraiser deciding to spend another half a day to locate "one more good comparable." Of course the appraiser, lacking even one good sample unit, can always change the sample unit and adjust rumors. Statistics requires a certain degree of respect for reliability of data at the expense of some marginal dollars lest the real estate statistician be hung on his own population frame in a cross examination or sued for misrepresentation instead of bias.

Chapter 5
Sampling

I. Introduction
II. Some Definitions
III. Elements of a "Good" Sample
IV. Types of Samples
  A. Probability Samples
    1. Simple random sample
    2. Stratified sample
    3. Cluster sample
    4. Systematic sample
  B. Nonprobability samples
    1. Convenience sample
    2. Judgment sample
V. A Note on Notation

Chapter 5
Sampling

I. Introduction

Most real estate decisions are made without complete information. For example, in determining rent levels it would be convenient to know all the current rents for all the apartments in a particular location. Since it is usually impossible ~~to~~ or impractical to gather information that is this extensive, rent levels are generally set after an analysis of a group or a sample of apartments in a given location. This chapter concerns sampling, that is, the process by which some of the data in a population is gathered to be used as a representation of that population.

II. Some Definitions

An understanding of the various sampling procedures requires a brief acquaintance with two terms-- <u>sampling unit</u> and <u>population frame</u>. Sampling units are the elements the real estate analyst wants to study. For example, the sampling units in a rent study could be individual apartment units, apartment buildings, or apartment projects. To further illustrate the sampling units in a land value study could be lots, parcels, acres, & square feet. The population frame is a list of all the sampling units contained within it. Examples of several types of population frames are:

1. The list of tax parcel numbers at the courthouse is the population frame for all the tax parcels (sampling units) in a given area.

2. The list of apartments in a large apartment complex is the population frame for all the apartments ~~in the~~ (sampling units) within the complex.

3. The street directory is the population frame for all the buildings (sampling units) within a given area.

The real estate analyst is afforded considerable latitude in his definition of a population frame. He must, however, use this ~~fact~~ freedom with a degree of caution. For example, an appraiser in his search for comparables could define his population frame as all those homes that sold within the last year in the same neighborhood with the same number of bedrooms. This population frame would be very effective in the selection of good comparables. On the other hand, if the appraiser defined his population frame as all the homes that sold in the past fifty years on the other side of ~~the~~ town, the comparables selected would be of little value.

III. Elements of a "Good" Sample

~~[illegible]~~

~~[illegible]~~. A sample should represent the population from which it is drawn. For example, the mean ~~[illegible]~~ sales price in a sample of single-family homes in Cleveland, Ohio. should represent the mean sales price of all the homes in Cleveland. There are two types of ~~errors~~ things that cause sample data to deviate from population data, random errors and bias. Random errors are caused by selecting an extremely deviant case in the sample. In the above illustration, the selection of an appraiser's home in the sample would cause the mean sales price of the sample to far exceed the mean sales price of the city. Random errors are mitigated by selecting larger samples, that is, sooner or later you will select a professor's home ~~and the mean sales price of~~

to ~~balance off the~~ neutralize the appraiser's home.

Any sample should be selected to avoid bias or a systematic distortion in your sample. Bias can result from a misspecification of the population frame. For example, assume you wanted to determine the average value of single-family homes in a given community. Since it was convenient, you selected as your population frame a list of professor's homes. Since professor's live in meager surroundings, any conclusion you make regarding the mean sales price of single-family homes in this community would be biased toward the lower side.

A good sample should also provide a maximum amount of information at the least cost. In general, the larger the sample the more valid are the conclusions that can be drawn from that sample. Large samples, however, are expensive and the real estate analyst must balance the additional costs of a larger sample against the increased level of accuracy. This problem is not new in the real estate field. All appraisers are familiar with the temptation to find "one more comparable."

IV. Types of Samples

Samples can be distinguished as probability or non-probability samples. Probability samples are those samples in which each sampling unit is given some chance of being included in the sample. The statistician using a probability can calculate the probability of selecting each sampling unit. Since the probability of selecting each sampling unit is known, the risk associated with drawing an incorrect conclusion can be calculated. On the other hand, there is no method of evaluating the risk associated with an incorrect conclusion when a non probability sample is used.

Both probability and nonprobability samples are illustrated below. The reader is cautioned against using any of the nonprobability samples since they can not be used for further analysis.

Probability Samples

Simple ~~Random~~ random sample. In a simple random sample the real estate analyst will select sampling units at random from the population frame until he reaches his desired sample size. The best method of selecting sampling units at random is through the use of a random number table. A random number table is simply a table of digits that has two important properties: (1) each digit will appear approximately the same number of times, and (2) there is no relationship between any of the digits. A sample page of a random number table appears in table 5-1. (Insert)

An example can illustrate the use of the random number table. Assume ~~you~~ we wanted to select, at random, a sample of ~~ten~~ 10 units from a 100 unit apartment complex. The first step would be to assign a number to each of the apartments from 1 to 100. Using the ~~first line~~ first three digits of the first line we would select the ~~following~~ apartments ~~for~~ our sample. The first three digit number is 789 which would be ignored because our population frame consists of only 100 apartments. The following apartments would be selected.

49, 10, 83, 95, 48, 83, 36, 64, 33, 66

The reader should note several things regarding the selection procedure. (1) Although the table is arranged with spaces between the groups of digits for easy reading the selection procedure moved down each row of three digits. For example, the first digit in the first row was 789 and the last digit in the row was 259, the first digit in the second row of three digits was 943 and the last was 036. (2) It makes no difference where you start on the table as long as your movement is systematic, that is, we could have moved from side to side or crosswise. (3) If a number which has already been selected appears the second time it is ignored.

A simple random sample such as the one illustrated above is one of the most highly recommended sampling procedures for further statistical analysis.

Stratified sample[1] In a stratified sample the population is divided into a number of groups or strata. For example, the population of single-family homes could be divided into ranch homes, two-story homes, one-and-one-half story homes, and split levels. From each of these stratum a simple random sample is selected using the selection procedure outlined above. These smaller samples are then combined into one larger sample.

Stratified samples have a twofold advantage. The inferences drawn from a stratified sample will have increased accuracy and the statistician is provided with simple random samples

---

1 The reader should be aware that the formulas for calculating the mean and the variance for a stratified sample are different ~~the~~ from those ~~used~~ for a simple random sample used throughout this book. These are:

~~The mean of the population ... is~~

$$\text{Mean} = \bar{X}_s = \frac{\sum M_i \bar{X}_i}{M}$$

$$\text{Variance} = S^2_{\bar{X}_s} = \sum \left(\frac{M_i}{M}\right)^2 \cdot \frac{M_i - N_i}{M_i - 1} \cdot \frac{S_i^2}{N_i}$$

where:

(over)

k = the number of strata

M_i = the number of individuals in the ith strata

$M = \sum_{i=1}^{k} M_i$ = the number of individuals in the population

N_i = the number of individuals in the sample taken from the ith stratum

$N = \sum_{i=1}^{k} N_i$ = the number of individuals in the sample

$\bar{X}_i = \frac{\Sigma X_i}{N_i}$ = mean of the observations on the N_i individuals in the sample from the ith stratum

$s_i^2 = \frac{\Sigma (X_i - \bar{X}_i)^2}{N_i - 1}$ = variance of the observations on the N_i individuals in the sample from the ith stratum for any of the subgroups that may be of interest. Stratified samples, however, are more expensive in both time and money, a factor ~~which~~ that should be considered.

<u>Cluster Sample</u>:[2] In this sampling procedure the sampling units are clustered into natural groups from which random samples are drawn and combined into a larger sample.

---

[2] Again the reader should be aware that the formulas for calculating the mean and the variance for a cluster are different from those used for a simple random sample used throughout this book. These are.

$$\text{Mean} = \bar{X}_c = \frac{\sum_{i=1}^{m} n_i \bar{X}_i}{\sum_{i=1}^{m} n_i} = \Sigma\Sigma X_i$$

$$\text{Variance} = s^2_{\bar{X}_c} = \frac{M-m}{Mm} \; \frac{\sum_{i=1}^{m} n_i^2 (\bar{X}_i - \bar{X}_c)^2}{(m-1)(\bar{n})^2}$$

where:

(over)

$M =$ the number of clusters in the population

$m =$ the number of clusters in the sample

$n_i =$ the number of individuals in the ith cluster

$\bar{x}_i = \frac{\sum x_i}{n_i} =$ mean of the ith cluster

$\bar{n} = \frac{\sum_{i=1}^{m} n_i}{m} =$ average number of individuals per cluster

$N = \sum n_i = m\bar{n} =$ the number of individuals in the sample

Cluster sampling can be illustrated through the use of census data. Assume we wanted to do a study of rent levels using the value of a block as our sampling unit. The cluster sampling procedure could be as follows. First, several clusters of blocks (census tracts) could selected using a simple random procedure. Second, individual blocks within each of these census tracts could be selected by a simple random sampling procedure. Once selected, these individual samples could be aggregated to represent the entire population.

Cluster sampling techniques can be used in two situations. First, when the cost of measuring an individual unit is small as opposed to the cost of reaching that unit. For example, it is less expensive for the real estate analyst to sample 10 homes within a given neighborhood as opposed to 10 homes apart throughout the city.

Second, when there is a great heterogeneity within each cluster. If all of the sample units were very similar, such as a "look-alike" subdivision there would be no advantage to sampling more than one.

Systematic Sample. A common method of sampling is to select every i th item from the population frame. For example, to evaluate rent levels in a large apartment complex the appraiser might select every 6th apartment so that his sample would include apartment 6, 12, 18 and so on. In using a systematic sampling procedure the real estate analyst must be very careful that his ~~systematic~~ sample is not systematically distorted. In the above illustration, if every 6th apartment was located on the corner of the building with the best view, any conclusion regarding the rent levels for the entire complex could be biased toward higher levels.

Nonprobability Samples

Two of the most common types of nonprobability samples are convenience samples and judgment samples. The reader is again cautioned against drawing conclusions from nonprobability samples. It is the recommendation of the authors that the reader be aware of these nonprobability samples so that he will never use them.

Convenience sample. The sampling units ~~in an~~ this type of sample are selected because they are most convenient. For example, in selecting a sample of office building expense ratios the real estate analyst may go to the local bar and upon seeing four office building owners at the same table ask them what their expense ratios are. This sample could be extremely bias if these four men were ~~~~ meeting to sell their office buildings because their expenses were out of line. To further, illustrate a convient method of sampling the average height of a person in O'hare Field is to walk over to the first plane and record the average height of the first 10 people who get off the plane. This procedure would work well as long as ~~the flight doesn't~~ the Milwaukee Bucks are not on the particular flight!

Judgment sample. A judgment sample is selected on the basis of some "experts" opinion as to what the typical sampling units are. For example, to choose a sample of 10 office buildings we would ask a local real estate expert to select 10 typical buildings in the area. Obviously, the quality of sample selected will vary with the quality of the expert and such a sample can not be used for further analysis.

V A Note on Notation

Chapter III included formulas on the calculation of the mean and standard deviation of a population. When calculating the mean and standard deviation of a sample the real estate analyst should use the following formulas:

For the arithmetic mean of a sample:

$$\bar{x} = \frac{\Sigma x}{n}$$

For the standard deviation of a sample:

$$s = \sqrt{\frac{\Sigma (x - \bar{x})^2}{n-1}}$$

For future reference these differences are summarized in the following Table 5.2

Table 5-2

	Population	Sample
MEAN	$\mu = \dfrac{\Sigma X}{N}$	$\bar{X} = \dfrac{\Sigma X}{n}$
STandard Deviation	$\sigma = \sqrt{\dfrac{\Sigma(X-\mu)^2}{N}}$	$s = \sqrt{\dfrac{\Sigma(X-\bar{X})^2}{n-1}}$

CORRESPONDENCE REGARDING

<u>A STATISTICAL HANDBOOK FOR APPRAISERS</u>

Graduate School of Business 1155 Observatory Drive, The University of Wisconsin, Madison 53706

January 18, 1972

T. C. Hitchings
T. C. Hitchings and Son, Inc.
601 Broadway
Denver, Colorado 80203

Dear Ted:

In response to your recent phone call regarding the Statistics book for the American Institute of Real Estate Appraisers, may we propose the following options:

Option 1

We could finish the book according to our original contract. The completion date in this case would be September, 1972.

Option 2

We could prepare a workbook based upon Fadil Zuwaylif, General Applied Statistics, Addison-Wesley, Publishing Company, (1970). This workbook would be comprised of examples from the real estate industry and would be designed for structuring illustrations and demonstrations in future courses and/or seminars by the AIREA.

In this case our original contract would be modified. We would be willing to prepare the manuscript for this workbook for $6,000 (Six-thousand) dollars and $1 (one) per copy after each sale. The completion date in this case would be June 30, 1972.

We think your suggestion of a workbook correlated to the Zuwaylif text is a good one and we are willing to cooperate with this modification of your original intent. However, as a condition of this redirection we would receive the ownership of the four chapters which we have completed for use as articles. These chapters would be submitted as articles to The Real Estate Journal for consideration and if not accepted within a reasonable period of time, we would feel free to offer these as a serial publication of individual articles wherever they might be received.

Sincerely,

James A. Graaskamp

Karel J. Clettenberg

KJC:ls

ADDISON-WESLEY PUBLISHING COMPANY

READING, MASSACHUSETTS 01867

(617) 944-3700

July 27, 1972

Professor James A. Graaskamp
Graduate School of Business
1155 Observatory Drive
The University of Wisconsin
Madison, Wisconsin 53706

Dear Professor Graaskamp:

Addison-Wesley hereby gives you and Professor Clettenberg and the American Institute of Real Estate Appraisers permission to prepare a workbook to accompany our text, GENERAL APPLIED STATISTICS by Fadil Zuwaylif, which will include identical chapter headings and chapter section headings as they appear in the Table of Contents of our 1970 edition.

We understand there will not be extensive quotations directly from the text, but if the total number of words is greater than one page of text, we will require a separate letter of permission from the authors and the Institute.

We want to thank you for your interest in our book and hope this arrangement better suits the needs of your students.

Sincerely yours,

Roger

Roger L. Drumm
Executive Editor

RLD/kmg

cc: AIRES
Professor Zuwaylif

AMERICAN INSTITUTE OF REAL ESTATE APPRAISERS
OF THE NATIONAL ASSOCIATION OF REAL ESTATE BOARDS
155 EAST SUPERIOR STREET CHICAGO, ILLINOIS 60611 • (312) 664-9700

ROBERT W. FORD, M.A.I.
GENERAL CHAIRMAN
EDUCATION COMMITTEE
P.O. BOX 57
MODESTO, CALIF. 95353

August 2, 1972

Dr. James A. Graaskamp
Graduate School of Business
1155 Observatory Drive
University of Wisconsin
Madison, Wisconsin 53706

Dear Dr. Graaskamp:

Our earlier correspondence indicated that you would have the manuscript for the new book on statistical applications in real estate appraising ready by July 31. I attempted to call last week while I was at the summer meetings of the Institute's Executive Committee, but failed to reach you; I was able to reach Chuck Clettenburg.

Would you please advise as to when we can expect receipt of the manuscript. It is important that this project be completed at the very earliest.

Regards,

Bob Ford

Robert W. Ford, M.A.I.
General Chairman, Education

RWF:ps

Founded 1932—Our 40th Anniversary Year

AMERICAN INSTITUTE OF REAL ESTATE APPRAISERS

OF THE NATIONAL ASSOCIATION OF REAL ESTATE BOARDS

155 EAST SUPERIOR STREET CHICAGO, ILLINOIS 60611 • (312) 664-9700

ROBERT W. FORD, M.A.I.
GENERAL CHAIRMAN
EDUCATION COMMITTEE
P.O. BOX 57
MODESTO, CALIF. 95353

November 20, 1972

Dr. James A. Graaskamp
Graduate School of Business
The University of Wisconsin
1155 Observatory Drive
Madison, Wisconsin 53706

Dear Jim:

At the November meetings of the Education Committee and the Executive Committee we discussed the delayed completion date for the statistics workbook manuscript. As a result of this discussion the Executive Committee approved cancellation of the project.

Therefore, in compliance with your offer contained in your letter to me of October 10, 1972, I will ask Bill Collis of our Chicago office to make arrangements with you and The University of Wisconsin Foundation for return of whatever monies advanced to date.

Sincerely,

Robert W. Ford, M.A.I.
General Chairman
Education Committee

RWF:lk
cc: Mr. Richard Bourguignon, M.A.I.
Mr. T. C. Hitchings, M.A.I.
Mr. William Collis
Mr. Charles Seymour, M.A.I.

T. A. Chores:

Bridgemark

1. 1970 version of Table 3.1, p. 51
   3.3, p. 54 for midwestern states.
2. Table 3.5, p. 56, should be duplicated using cities & % of minority populations.
3. 2-unit bar chart as on page 122 and a 2-direction bar chart as on p. 123.
4. Current example of bar chart on p. 127 and multiple bar chart on center page 129.
5. Provide single line chart of expenses and then a silhouette chart using same expenses to show declining net income available for debt service (have gross income increasing at a decreasing rate while expenses climb at constant rate) (See Ted Hitchings illustration)
6. Find examples of silhouette chart in Wenzlick and the Architectural Record.
7. Absorption rate and sales by month and cumulative and 6 month moving total for recreational land project with 3,000 lots.
8. Statistical maps from 1970 census, the Harvard Sy-Map and municipal map keyed to tax parcel numbers or planning districts using size of dot, cross-hatched and number of dots to suggest patterns. Use customer pattern chart from Applebaum.
9. Bibliography on statistical maps and trade area from Appraisal Journals.